

Joint Congress of the Canadian Society for Mechanical Engineering and CFD Society of Canada

Proceedings of the 2019 CSME International Congress



June 2 – June 5, 2019

**Department of Mechanical and Materials Engineering
Western University
London, Ontario, Canada**



Message from the 2019 CSME International Congress Chair



Welcome to the 2019 CSME International Congress. The Congress is hosted by the Department of Mechanical and Materials Engineering at the Western University. We have the privilege to host the CSME Congress again after 15 years. We are very pleased to co-host this year's Congress with the CFD Society of Canada. This joint Congress provides a unique opportunity to bring researchers and students whose interests align with both societies to a common platform that facilitates the dissemination of results from cutting-edge research as well as networking opportunities among academia, government agencies and industry. We hope that this joint Congress will further strengthen the ties between the two societies.

This two and a half day Congress has an extensive technical program featuring plenary lectures from internationally-recognized distinguished speakers, technical symposia, CSME Student Paper Competition, CSME Student Design Competition, NSERC Workshop on grants and scholarships, Annual General Meeting of the CSME Board, meetings of various Technical Committees of CSME and tours of various state-of-the-art research facilities.

CSME is sponsoring three of the five plenary lectures covering a broad range of mechanical engineering topics. Dr. Peter Walker, Professor at the New York University in the area of Biomechanics will speak about the recent advancements in artificial knees. Dr. Andy Sun, Professor at the Western University in the area of Nano-materials will talk about novel applications of Nano-materials for energy storage. Dr. Ruxu Du, Professor at South China University of Technology in the area of Advanced Manufacturing will talk about a new idea of resilient manufacturing.

The CSME Congress has 17 technical symposia covering all key areas in the mechanical engineering discipline. These symposia are comprised of 282 technical presentations including eight Keynote presentations, all run in 10 parallel sessions. The Congress has also organized tours of various state-of-the-art research facilities at Western and collaborating institutes including Fraunhofer Project Centre, the WindEEE Dome, the Boundary Layer Wind Tunnel, National Research Council, Canadian Centre for Product Validation, Wolf Orthopedic Biomechanics Lab and Energy Materials Lab.

A large number of CSME-CFDSC Congress participants are also attending the co-located Industry 4.0 Symposium. This symposium provides an opportunity to learn about new trends in Industry 4.0, and meet and network with relevant academic and industry people.

I would like to thank all individuals who have contributed to the organization of the Congress, especially, the members of the Congress Organizing Committee, Western Conference Services, symposia organizers, student volunteers and session chairs.

We also gratefully acknowledge the generous sponsorships from the Faculty of Engineering, Western University, High Speed Imaging, FLIR, Delta Photonics, Bombardier and Canadian Science Publishing.

I hope you have a great time at the Congress

A handwritten signature in dark ink, appearing to read 'Kamran Siddiqui'.

Kamran Siddiqui
2019 CSME Congress Chair



Congress Organizing Committee



Kamran Siddiqui
Chair
CSME International
Congress and Technical
Program
Western University



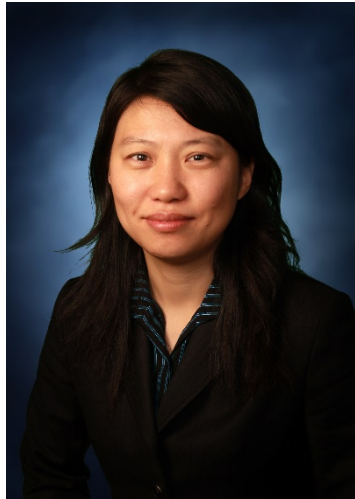
Christopher DeGroot
Co-Chair
CFDSC Annual
Conference
Western University



Chao Zhang
Co-Chair
CFDSC Annual
Conference
Western University



Ryan Willing
Communications Chair
Western University



Liying Jiang
Publications Chair
Western University



Remus Tutunea-Fatan
Exhibitions Chair
Western University

CSME Symposia Organizers

Symposium	Organizers
Advanced Manufacturing	Alex Czekanski, Evgueni Bordatchev, Remus Tutunea-Fatan
Additive Manufacturing	Jun Yang, Tengyuan Zhang
Advanced Energy Systems	Xianguo Li
Biomechanics	Ali Ahmadi, Ryan Willing
Biomedical Engineering	George Knopf, Ali Ahmadi
Energy Materials	Yang Zhao, Keegan Adair
Engineering Analysis and Design	Kamran Behdinin
Environmental Engineering	Jubayer Chowdhury
Fluid Mechanics	Martin Agelin-Chaab, Dana Grecov
Heat Transfer	Dominic Groulx
Materials Engineering	Frank Cheng, Ali Nasiri
Mechatronics, Robotics and Automation	Farrokh Janabi-Sharifi, Mehrdad Kermani
Micro- & Nano-Technology	Carlos Escobedo
Nuclear Materials	Hamid Abdolvand
Solid Mechanics	Hamid Akbarzadeh, Liying Jiang
Transportation Systems	Yuping He
Wind Energy and Engineering	Jubayer Chowdhury, Horia Hangan

Plenary Speakers

Plenary Speaker (Sponsored by CSME and Western's Bone & Joint Institute)

Artificial Knees: Can They Benefit from New Technologies?



Peter S. Walker, PhD

Professor, New York University Langone Orthopaedic Hospital
Director, Laboratory for Orthopaedic Implant Design
Professor, Dept. of Mechanical & Aerospace Engineering, New York University

ABSTRACT

The volume of total knee replacements in North America has reached about 700,000 per year and several billion dollars. The early designs of total knee were produced in the 1970's based on collaboration between surgeons and bioengineers, with companies providing the manufacturing technology. The basic engineering principles of the laxity and stability, the kinematics, the fixation of the components to the bone, and the factors affecting wear, were developed in parallel. Since 1980, there has been an ever-increasing amount of research and clinical evaluations, including new design concepts, but 90% of the total knees used by surgeons today are still of two types designed in the 1970's, the posterior cruciate-retaining and the posterior-stabilized designs. The main application of the research since then has mainly been directed to refinements with the result that survivorship has reached a level of around 90 percent at 20 years. Since about the year 2000, a computer-assisted approach has produced several methods for improving the accuracy and consistency of the surgical technique, yet the uptake has been only around 10 percent. An important question today is to what extent technology can further improve the treatment of the arthritic knee and in what ways bioengineers can have a significant further impact. Two areas will be discussed as possible approaches. The first is to develop new methodologies for evaluating the performance of total knees in a way which will identify the optimal result achievable for each individual patient. The second is to formulate implant designs, manufacturing technologies, and surgical techniques, which will deliver those optimal results on an individual basis at a cost affordable to the health care system.

BIOGRAPHY

Dr. Peter Walker has completed a PhD at the University of Leeds. Since then he has been an active participant in the field of biomechanics of joints and the design of joint replacements, with a major emphasis on the knee. He has collaborated with surgeons and bioengineers at renowned institutions, including the Hospital for Special Surgery, Howmedica (Stryker), Brigham & Women's Hospital, Royal National Orthopaedic Hospital, University College London, and currently at New York University Langone Orthopaedic Hospital. He has co-authored over 300 articles with almost 12,000 citations. The basic research has contributed to innovative ideas for total knee design, while many of the test machines and methodologies have been emulated by other labs. Dr. Walker has been a co-designer of several successful total knee systems, the most notable being the Total Condylar, Graduated Knee System, Kinematic & Kinemax, Smiles Bone Tumor System, NexGen and Persona. In 2018, Dr. Walker was chosen as the recipient of the Lifetime Achievement award by The Knee Society, a high honor only given every two years.

Monday, June 3rd (8:30 am – 9:30 am), Ivey BMO Auditorium

Plenary Speaker (Sponsored by CSME)

Advanced Materials for Energy Storage and Conversion: From Nano Scale to Single Atoms



Andy (Xueliang) Sun, PhD

Professor and Canada Research Chair (Tier I)
Department of Mechanical and Materials Engineering
Western University

ABSTRACT

In this talk, Dr. Sun will talk about application of nanotechnology such as atomic layer deposition (ALD) to address challenges in both lithium ion batteries and low temperature fuel cells. The advanced rechargeable batteries have attracted intensive research attention as one promising solution to solve the global energy and environment problems due to high energy density and long working life. A few successful examples from his work will be given including lithium-sulfur (Li-S) batteries, alkali metal-oxygen (Li-O₂, Na-O₂) batteries and all-solid-state batteries. In addition, fuel cells still have challenges ahead which are hindering the market implementation of low temperature fuel cell technology, mainly high cost of materials and the durability during fuel cell life-time operation. The high cost is primarily associated with precious metal catalysts (Pt or Pt alloys). His group used ALD to control the size of Pt down to single Pt atom catalysts. Dr. Sun will also discuss some future perspectives.

BIOGRAPHY

Dr. Andy (Xueliang) Sun is a Professor and senior Canada Research Chair (Tier I) for the development of nanomaterials for clean energy, at the Western University. Dr. Sun received his Ph.D degree in Materials Chemistry at the University of Manchester, UK, in 1999. Dr. Sun's research is focused on advanced nanostructured materials for energy conversion and storage including fuel cells and Li batteries. Dr. Sun was named as Web of Science "Highly Cited Researchers" in 2018 by Clarivate Analytics. Dr. Sun is an author and co-author of over 370 refereed-journals with citations of over 20,000 times and H-index of 72. He edited 3 books and published 16 book chapters as well as filed 18 US patents. Dr. Sun is actively collaborating with industries and government labs such as Ballard Power Systems, General Motors, Lithium Phostech Inc., 3M, and GLABAT Solid-State Battery Inc. He also serves as an Editor-in-Chief of "Electrochemical Energy Review" under Spring-Nature and an Associate Editor for Frontier of Energy Storage (2013-present). Dr. Sun received various awards such as Early Researcher Award, Canada Research Chair, University Faculty Scholar Award, Western Engineering Prize for Achievement in Research, Fellow of Royal Society of Canada, Fellow of the Canadian Academy of Engineering, Professional Achievement Awards from Cross-cultural Professionals Association of Canada, Award for Research Excellence in Materials Chemistry Winner from Canada Chemistry Society and Award of Merit of the Federation of Chinese Canadian Professionals.

Tuesday, June 4th (1:00 pm – 2:00 pm), Ivey BMO Auditorium

Plenary Speaker (Sponsored by CSME and Industry 4.0)

Resilient Manufacturing System



Ruxu Du, PhD, FCAE, FSME, FASME

Professor

S. M. Wu School of Intelligent Manufacturing

South China University of Technology

Guangzhou, China

ABSTRACT

In the past two years, the world's manufacturing environment has been completely changed from in-favor of globalization to that of nationalization. As a result, companies large or small will have to face ever increasing government regulation, resources restriction and market limitation. To survive such a harsh environment, companies will have to make their manufacturing system resilient. In this talk, we introduced the idea of resilient manufacturing system. It consists of four parts: (1) Product design and optimization, (2) Manufacturing system modeling and optimization, (3) Advanced manufacturing methods, and (4) Market and supply-chain management. The talk includes a couple of examples, which demonstrates the uses of reconfigurable machines and buffers as well as prioritizing resources can improve the resilience of manufacturing systems .

BIOGRAPHY

Dr. Ruxu Du has received his Ph.D. degree from the University of Michigan in 1989. He has taught at the University of Windsor, in Windsor, Ontario, Canada, the University of Miami, in Coral Gables, Florida, USA, and the Chinese University of Hong Kong, in Hong Kong SAR. He has built several institutes, including the Institute of Precision Engineering of in the Chinese University of Hong Kong, and Guangzhou Chinese Academy of Sciences Institute of Advanced Technology. Currently, he is a Professor and the founding Dean of S. M. Wu School of Intelligent Engineering in the South China University of Technology. His area of research include: design and manufacturing, as well as robotics and automation. He has published over 450 papers in various academic journals and international conferences. He is the associate editor / the members of editorial board of six international journals. He has received a number of awards including Fellow of Canadian Academy of Engineering; Fellow of SME (Society of Manufacturing Engineers); Fellow of ASME (Society of American Mechanical Engineers); Fellow of HKIE (Hong Kong Institute of Engineers).

Wednesday, June 5th (8:30 am – 9:30 am), Ivey BMO Auditorium

June 3, 2019 (Monday) - CSME TRACK

8:00 AM - 4:00 PM	Conference Registration (Amit Chakma Engineering Building Atrium)									
8:30 AM - 9:30 AM	CSME-CFDSC Congress Opening Ceremony Plenary Session (Ivey BMO Auditorium): "Artificial Knees: Can They Benefit from New Technologies?" Peter Walker (New York University)									
9:30 AM - 10:00 AM	Coffee Break (Amit Chakma Engineering Building Atrium)									
	Session A-1 (Advanced Manufacturing I)	Session B-1 (Advanced Energy Systems I)	Session C-1 (Biomechanics I)	Session D-1 (Biomedical Engineering I)	Session E-1 (Energy Materials I)	Session F-1 (Heat Transfer I)	Session G-1 (Fluid Mechanics I)	Session H-1 (Micro-Nano-Technologies I)	Session J-1 (Materials Engineering I)	Session K-1 (Solid Mechanics I)
	Session Chair: Alex Czekanski	Session Chair: Xianguo Li	Session Chair: Ali Ahmadi	Session Chair: George Knopf	Session Chair: Yang Zhao	Session Chair: Syeda Tasnim	Session Chair: Paul Henshaw	Session Chair: Michael Boutilier	Session Chair: Ali Nasiri	Session Chair: Hamid Akbarzadeh
	Room: ACEB 1410	Room: ACEB 2440	Room: SEB 1056	Room: SEB 1059	Room: ACEB 1450	Room: SEB 2099	Room: SEB 2100	Room: ACEB 2435	Room: ACEB 1415	Room: ACEB 1420
10:00 AM - 10:20 AM	Laser Polishing and Structuring of Tooling and Functional Surface (Keynote Presentation) Evgeni Bordatchev (National Research Council of Canada)	Thermodynamic Modeling and Optimization of Ejector Refrigeration Systems Using Alternative Working Fluids Aggrey Mwesigye (Ryerson University); Amir Kiamari (Ryerson University); Seth Dworkin (Ryerson University)	Computational Modeling of Passive Ankle Exoskeleton Biomechanics Jin Zhu (University of Ottawa); Marc Doumit (University of Ottawa)	Multiscale Fluid Flow Analysis of Decellularized Mouse Lungs to Improve Lung Regeneration Eric Chadwick; Mohammadali Ahmadipour; Takaya Suzuki; Golnaz Karoubi; David Romero; Cristina Amon; Thomas Waddell; Aimy Bazylak; (University of Toronto);	Phase Inversion: A Universal Method to Create High-Loading Porous Electrodes for Energy Storage Devices Xiaofei Yang (Western University); Xueliang Sun (Western University)	Heat Flux Analysis of a Li-Ion Cell during Charging and Discharging Periods Abu Raihan Siddique (University of Guelph); Shohel Mahmud (University of Guelph); S. Andrew Gadsden (University of Guelph); Bill Van Heyst (University of Guelph)	Mechanics of Active Fluids: Fluid Mechanics Refashioned by Microorganisms (Keynote Presentation) Hassan Peerhossaini (Western University)	Development of BioMEMS 3D Stress/Strain Sensing Mechanism Mohammed Kayed (University of Alberta); Amr Balbola (University of Alberta); Walled Moussa (University of Alberta)	Novel Welding Processes and Joint Evaluation Techniques in Pipeline Construction (Keynote Presentation) Adrian Gerlich (University of Waterloo)	Modelling of Creep Recovery in Carbon Fiber Reinforced Polymer Composites Spencer Cabel (University of Manitoba); Raghavan Jayaraman (University of Manitoba)
10:20 AM - 10:40 AM		Exergy Analysis of Renewable Resources for Energy Production in Remote Communities Mudit Nijhawan (University of Windsor); Ofelia Jianu (University of Windsor)	Computational Investigation into Materials with Lower Stiffness for Hemiarthroplasty Carolyn Berkmortel (University of Western Ontario); Daniel Langohr (University of Western Ontario); James Johnson (University of Western Ontario); Graham King (Saint Joseph's Healthcare)	3D Bio-Printing of Starch-Chitosan Scaffolds for Neuron Cells Haley Butler; Emad Naseri; Debra MacDonald; Andrew Tasker; Ali Ahmadi; (University of Prince Edward Island)	Sulfide Solid-State Electrolyte Protection Layer In-Situ Formed on the Surface of Metallic Li for High-Performance Li Batteries Jianwen Liang (Western University); Xueliang Sun (Western University)	The Impacts of Exterior Surface Convective Heat Transfer Coefficients on High-Rise Building Energy Consumption Meseret Kahsay (Western University); Girma Bitsuamlak (Western University); Fitsum Tariku (British Columbia Institute of Technology)		Microfluidic Oil-Water Separation Using a Hydrophobic-Oleophilic Stainless Steel Membrane Tartela Alkayyali (University of Prince Edward Island); Bradley Halti (University of Prince Edward Island); Russell Kerr (University of Prince Edward Island); Ali Ahmadi (University of Prince Edward Island)		Using Lime/Hemp Concrete as a Multi-Function Material E. Fotsing (Polytechnique Montréal); A. Ross (Polytechnique Montréal); T. Lecompte (Université Bretagne Sud)
10:40 AM - 11:00 AM	Innovative Strategies for Assessing the Metrological Performance of Optical 3D Scanners Farbod Khameneifar (Polytechnique Montréal)	A Data-Driven Thermal Fault Detection Method in Lithium-Ion Batteries Olaoluwa Ojo (University of Ontario Institute of Technology); Xianke Lin (University of Ontario Institute of Technology); Haoxiang Lang (University of Ontario Institute of Technology)	Artificial Intelligent Examination of DXA Image Quality for Automatic Finite Element Assessment of Hip Fracture Risk Wanshi Yu (University of Manitoba); Yunhua Luo (University of Manitoba)	Carbon-based pH Biosensors for Monitoring Infections Dogan Sinar (Western University); George Knopf (Western University)	Organic/Inorganic Composite Electrolyte for Dendrite-Free Solid-State Lithium Batteries Hanyu Huo (Western University); Xueliang Sun (Western University)	Experimental Investigation of a Loop Heat Pipe with a Two-Way Pressure Regulated Valve Hooman Jazebizadeh (Carleton University); Tarik Kaya (Carleton University)	Effect of Background Turbulence on the Evolution of and Entrainment into a Turbulent Jet Rana Sahebjam (McGill University); Susan Gaskin (McGill University)	Design of a High Density MEMS Membrane-Rod based Hydrophone Tara Ahmadi (University of Windsor); Mohammed Jalal Ahamed (University of Windsor); Shahpour Alirezaee (University of Windsor)	Experimental Characterization of Anisotropic Mechanical Properties of Extruded AA6061-T6 Matthew Bondy (University of Windsor); William Altenhof (University of Windsor); John Magliaro (University of Windsor); Morten Jensen (CertaSIM LLC)	Molecular Dynamics Simulations of Nanoindentation - The Effect of Force-Field Choice on the Predicted Elastic Modulus Douglas Pratt (Memorial University); Sam Nakhla (Memorial University)

June 3, 2019 (Monday) - CSME TRACK

	Session A-2 (Additive Manufacturing I)	Session B-2 (Advanced Energy Systems II)	Session C-2 (Engineering Analysis & Design)	Session D-2 (Biomedical Engineering II)	Session E-2 (Energy Materials II)	Session F-2 (Fluid Mechanics II)	Session G-2 (Fluid Mechanics III)	Session H-2 (Micro-Nanao-Technologies II)	Session J-2 (Materials Engineering II)	Session K-2 (Solid Mechanics II)
	Session Chair: Mohsen Mohammadi	Session Chair: Thomas Cooper	Session Chair: Eric Lanteigne	Session Chair: Aimy Bazylak	Session Chair: Chandra Veer Singh	Session Chair: Martin Agelin-Chaab	Session Chair: Amir Aliabadi	Session Chair: Carlos Escobedo	Session Chair: Adrian Gerlich	Session Chair: Liying Jiang
	Room: ACEB 1410	Room: ACEB 2440	Room: SEB 1056	Room: SEB 1059	Room: ACEB 1450	Room: SEB 2099	Room: SEB2100	Room: ACEB 2435	Room: ACEB 1415	Room: ACEB 1420
2:30 PM - 2:50 PM	Multi-Element Slicing Approach for Intelligent Additive Manufacturing Preprocessing Dylan Bender (University of Ontario Institute of Technology); Ahmad Barari (University of Ontario Institute of Technology)	Natural Gas Engine Driven Heat Pump System – A Case Study of an Office Building Farshad Kimiaghalam (Ryerson University); Altamash Baig (Ryerson University); Rakesh Kumar (Ryerson University); JuanLi Ma (Ryerson University); Alan Fung (Ryerson University)	Dynamic Analysis and Design Optimization of an Outdoor Fitness Machine Victor Hatai (University of Ottawa); Eric Lanteigne (University of Ottawa)	A Novel Method to Improve Estimation of Tissue Displacement Field in Ultrasound Elastography Niusha Kheirkhah (Western University); Ali Sadeghi-Naini (York University); Abbas Samani (Western University)	Local and electronic structure study of black phosphorus by XAFS Wei Han Li (Western University); Minsi Li (Western University); Tsun-Kong Sham (Western University); Xueliang Sun (Western University)	Multilayer Viscoplastic Fluid Flows in Axially Rotating Pipes Seyed Mohammad Taghavi (Université Laval); Shan Lyu (Université Laval)	PIV Analysis of Flow Around a Simplified Ship Hull Jennie Flaig (University of Manitoba); Baafour Nyantekyi-Kwakye (Memorial University); Xingjun Fang (University of Manitoba); Mark Tachie (University of Manitoba)	Quantifying the Influence of Strain Engineering on Piezoresistivity Using a Six Elements Sensing Rosette Amr Balbala (University of Alberta); Mohammed Kayed (University of Alberta); Walied Moussa (University of Alberta)	Moisture Effect on Mechanical Performance of Out-of-Autoclave Composite Material at Different Void Level Afshin Bayatpour (Concordia University); Mehdi Hojjati (Concordia University)	A New Anisotropic Elastic Metamaterial (Keynote Presentation) Xiaodong Wang (University of Alberta)
2:50 PM - 3:10 PM	Prospects and Challenges of Reinforcement Methodologies for 3D Printable Plastics Aman Preet Singh; Basim El Essawi; Ali Daneji; Salman Pervaiz; (Rochester Institute of Technology - Dubai Campus); Sathish Kannan (American University of Sharjah)	Sizing, Modeling and Analysis of a Solar Seasonal Energy Storage System for Space Heating in New Foundland Rabbani Rasha (Memorial University); Habibur Rahaman (Memorial University); Tariq Iqbal (Memorial University)	Experimental and Numerical Modelling of a One-Degree of Freedom Non-Smooth Mechanical System Solène Kojtych (Polytechnique Montréal); Yann Colaitis (Polytechnique Montréal); Elsa Piollet (Polytechnique Montréal); Alain Batailly (Polytechnique Montréal)	Modelling of Flow-Induced Shear Stress to Predict Targeted Delivery of Cells in the Decellularized Lung Jason Chan; Eric Chadwick; Takaya Suzuki; Golnaz Karoubi; David Romero; Cristina Amon; Thomas Waddell; Aimy Bazylak; (University of Toronto)	Pt Single Atom Electrocatalysts by Atomic Layer Deposition for Fuel Cell Related Catalytic reactions Lei Zhang (Western University); Xueliang Sun (Western University)	Aspect Ratio Effects on Turbulent Flow over Forward-Backward-Facing Steps Heath Chalmers (University of Manitoba); Baafour Nyantekyi-Kwakye (Memorial University); Xingjun Fang (University of Manitoba); Mark Tachie (University of Manitoba)	Three-Dimensional Characterization of Bursting and Sweeping Phenomena in the Turbulent Boundary Layer Kadeem Dennis (Western University); Kamran Siddiqui (Western University)	Quantitative Analysis and Predictive Engineering of Self-rolling of Nanomembranes Jun Song (McGill University)	On the Influence of the Environmental Factors on Corrosion Behavior of 13Cr Stainless Steel Using Box-Behnken Design Mostafa Kazemipour (Memorial University); Salar Salahi (Memorial University); Ali Nasiri (Memorial University)	
3:10 PM - 3:30 PM	3D Printed Metamaterials of Dual/Negative Poisson's Ratio Behaviors Induced by Hierarchical Architecture Junfeng Xiao (Western University); Dongxing Zhang (Western University)	Investigating n-Heptane Spray Characteristics Inside the Ignition Quality Tester (IQTTM) Chamber through Temperature Variation Measurements Osama Hmood; Edgar Matida; (Carleton University); Dave Gardiner; Luc Menard; Omar Ramadan; Aaron Wilcox; Gary Webster; (Advanced Engine Technology Ltd.)	Experimental Measurements of Thermal Distribution of Belt Drive System: Application to Vehicle Front Engine Accessory Drives Xingchen Liu (University of Toronto); Kamran Behdinan (University of Toronto)	Prosthetic Foot Design – A Simplified Analytical Approach Stephanie Gibbons (Memorial University); Sam Nakhla (Memorial University); Adam Dubrowski (Memorial University)	3D-Printed Sulfur Cathode for Lithium Sulfur Batteries Xuejie Gao (Western University); Xueliang Sun (Western University)	An Exact Solution to the 3D Navier-Stokes and Applications to the Ranque-Hilsch Vortex Tube Nolan Dyck (Western University); Anthony Straatman (Western University)	Measuring Dynamic Contact Angles of a Droplet on Inclined Hydrophobic Surfaces with Irregular Roughness Kewei Shi (Memorial University); Xili Duan (Memorial University); Kristin Poduska (Memorial University)	Electroadhesive Carbon Nanotube Surfaces Michael Boutilier (Western University); Changhong Cao; Nigamaa Nayakanti; Sanha Kim; Seyedeh Mohadeseh Taheri Mousavi; A. John Hart; (Massachusetts Institute of Technology)	Impact of Fibre Modification on the Properties of Natural Fibre Composites Nabeel Syed (University of Ontario Institute of Technology); Utkarsh (University of Ontario Institute of Technology); Ghaus Rizvi (University of Ontario Institute of Technology); Remon Pop-Iliev (University of Ontario Institute of Technology)	Manipulating the Architecture of Beams for High Toughness and Strength Ahmed Dalaq (McGill University); Francois Barthelat (McGill University)

June 4, 2019 (Tuesday) - CSME TRACK

8:00 AM - 4:00 PM	Conference Registration (Amit Chakma Engineering Building Atrium)									
8:30 AM - 9:30 AM	Plenary Session (Ivey BMO Auditorium): "Advances in Numerical Modelling of Flow, Heat, and Mass Transfer in Heterogeneous Media" Marcelo de Lemos (Instituto Tecnológico de Aeronáutica)									
9:30 AM - 10:00 AM	Coffee Break									
	Session A-3 (Advanced Manufacturing II)	Session B-3 (Robotics, Mechatronics, Automation I)	Session C-3 (Biomechanics II)	Session D-3 (Environmental Engineering I)	Session E-3 (Energy Materials III)	Session F-3 (Heat Transfer II)	Session G-3 (Fluid Mechanics IV)	Session H-3 (Transportation Systems I)	Session J-3 (Materials Engineering III)	Session K-3 (Solid Mechanic III)
	Session Chair: Farbod Khameneifar	Session Chair: Mehrdad Kermani	Session Chair: Hossein Rouhani	Session Chair: Animesh Dutta	Session Chair: Xianke Lin	Session Chair: Jubayer Chowdhury	Session Chair: Ali Tarokh	Session Chair: Yuping He	Session Chair: Ali Nasiri	Session Chair: Hamid Akbarzadeh
	Room: ACEB 1410	Room: ACEB 2440	Room: SEB 1056	Room: SEB 1059	Room: ACEB 1450	Room: SEB 2099	Room: SEB 2100	Room: ACEB 2435	Room: ACEB 1415	Room: ACEB 1420
10:00 AM - 10:20 AM	Implementations and Applications of Elliptical Vibration Micro/Nano-Cutting Nikolai Farrus (Western University); O. Remus Tutunea-Fatan (Western University); Nicolas Milliken (National Research Council of Canada); Evgueni Bordatchev (National Research Council of Canada)	Nonlinear Model Predictive Control of Path Following by Waterloo's Autonomous Car Mohammad Hassanpour (University of Waterloo); John McPhee (University of Waterloo); Krzysztof Czarnecki (University of Waterloo)	Comparing Fixation Constructs for Scapula Insufficiency Fractures Following Reverse Shoulder Arthroplasty Jason Lockhart; David Axford; G. Daniel Langohr; James Johnson; (Western University); George Athwal; Francis Ting; (St. Joseph's Health Care Foundation)	An Airborne Thermal Imaging Methodology for Mapping Land Surface Temperature (LST) with a High Spatiotemporal Resolution Ryan Byerley (University of Guelph); Mohammad Biglarbegian (University of Guelph); Amir Aliabadi (University of Guelph)	Computational Design of 2D Nanomaterials for Sustainable Energy (Keynote Presentation) Chandra Veer Singh (University of Toronto)	Numerical simulation of heat transfer and insulation performance of honeycomb structure Hao Liu (University of Guelph); Syeda Tasnim (University of Guelph); William Lubitz (University of Guelph)	A New Multivariate Correlation for Liquid/Liquid Separation Weiwei E (Memorial University); Kevin Pope (Memorial University); Xili Duan (Memorial University)	Multibody Dynamics in Julia Bruce Minaker (University of Windsor)	Additive Manufacturing of Aluminum Alloys: Opportunities for Extreme Environments (Keynote Presentation) Mohsen Mohammadi (University of New Brunswick)	Non-Fourier Heat Conduction and Heat Wave Interference in CNT Reinforced Composites (Keynote Presentation) Zengtao Chen (University of Alberta)
10:20 AM - 10:40 AM	Application of EWMA Control Chart on Volumetric Errors Change Recognition Kanglin Xing (Polytechnique Montréal); Rene Mayer (Polytechnique Montréal); Achiche Sofiane (Polytechnique Montréal)	Model Predictive Control, Identification and Implementation Design for a Pendubot: A Microcontroller-Based Architecture Ricardo Bautista Quintero (University of New Brunswick); Rickey Dubay (University of New Brunswick)	Design of a Biaxial Mechanobioreactor for Engineering Pediatric Aortic Valves Edwin Wong (University of Toronto); Craig Simmons (University of Toronto)	A Novel Option for City Water Recycle with Carbon Capture Bashnu Acharya (University of Prince Edward Island); Prabir Basu (Dalhousie University); Akash Kulashreshtha (Dalhousie University)		Performance Evaluation of Siding Materials Subjected to Radiant Heat Loads Eder Villa Coronel (University of Alberta); Razim Refai (FPInnovations); André McDonald (University of Alberta)	Influence of Trailing Edge Flap Pitching on Dynamic Stall in Pitching Airfoils Farid Samara (University of Waterloo); David Johnson (University of Waterloo)	SDP-Based Energy Management Strategy with Lithium Plating Prevention for PHEVs Siyang Wang (University of Ontario Institute of Technology); Xianke Lin (University of Ontario Institute of Technology)		
10:40 AM - 11:00 AM	Orthogonal Machining of Pure Magnesium-Silica Syntactic Foam Abdulla Alhourani (American University of Sharjah); Sathish Kannan (American University of Sharjah); Salman Pervaiz (American University of Sharjah)	A Point Cloud Generation Scheme Shuwei Qiu (Western University); Mehrdad Kermani (Western University)	Energy Loss of Aortic Wall in Aneurysms Predicts Tissue Delamination Edwin Wong; Mingyi Tang; Craig Simmons; Jennifer Chung; Thomas Forbes; Jagdish Butany; Maral Ouzounian; (University of Toronto)	Sea Lettuce to Biocarbon and Biogas with Recovery of Nutrients Ankita Shrestha (University of Prince Edward Island); Bishnu Acharya (University of Prince Edward Island)	Interface Engineering of Sulfide Electrolyte-Based All-Solid-State Lithium Batteries Changhong Wang (Western University); Xueliang Sun (Western University)	Thermal Characteristics of Steel and Composite Risers for Flow Assurance Samar Aly (Memorial University); Sam Nakhla (Memorial University)	System Simulation of Engine Cooling Water Jacket Elle Mistruzzi (University of Windsor); Shankar Natarajan (FCA Canada Inc.); Ronald Barron (University of Windsor); Ram Balachander (University of Windsor)	A Fault-Tolerant Steering Control Algorithm Using MPC for Driving Safety of All-Terrain Cranes Kwang-seok Oh (University of Ontario Institute of Technology); Jaho Seo (University of Ontario Institute of Technology)	Microstructural Evolution of Adiabatic Shear Bands in AZ31 Magnesium Alloy During High Strain Rate Impact Francis Tetteh (York University); Solomon Boakye-Yiadom (York University)	Investigating the Accuracy of Various Many-Body Potentials in Estimating Mechanical Properties of a Nanoscale Aluminum Sabir Subedi (Memorial University); Sam Nakhla (Memorial University)

June 4, 2019 (Tuesday) - CSME TRACK

1:00 PM - 2:00 PM	Plenary Session (Ivey BMO Auditorium): "Advanced Materials for Energy Storage and Conversion: From Nano Scale to Single Atoms" Andy Sun (Western University)									
2:00 PM - 2:30 PM	Coffee Break (Amit Chakma Engineering Building Atrium)									
	Session A-4 (Additive Manufacturing II)	Session B-4 (Advanced Energy Systems III)	Session C-4 (Materials Engineering IV)	Session D-4 (Biomedical Engineering III)	Session E-4 (Environmental Engineering II)	Session F-4 (Fluid Mechanics V)	Session G-4 (Fluid Mechanics VI)	Session H-4 (Transportation Systems II)	Session J-4 (Materials Engineering V)	Session K-4 (Solid Mechanics IV)
	Session Chair: Sathish Kannan	Session Chair: Xianguo Li	Session Chair: Sam Nakhla	Session Chair: Reza Fotouhi	Session Chair: Bishnu Acharya	Session Chair: Prabir Basu	Session Chair: Xili Duan	Session Chair: Bruce Minaker	Session Chair: William Altenhof	Session Chair: Annie Ross
	Room: ACEB 1410	Room: ACEB 2440	Room: SEB 1056	Room: SEB 1059	Room: ACEB 1450	Room: SEB 2099	Room: SEB 2100	Room: ACEB 2435	Room: ACEB 1415	Room: ACEB 1420
2:30 PM - 2:50 PM	Accurate Estimation of Process Parameters in Fused Deposition Modeling Hossein Gohari Bahabadi (University of Ontario Institute of Technology); Ahmad Barari (University of Ontario Institute of Technology); Hossam Kishawy (University of Ontario Institute of Technology)	Efficient Non-Contact Steam Generation and Superheating Using Sunlight Thomas Cooper (York University)	Performance of Metals in Space: A Molecular Dynamics Simulation of Atomic Oxygen Impacts on Silver Liam Morrissey (Memorial University); Stephen Handrigan (Memorial University); Sam Nakhla (Memorial University); Ali Rahnamoun (John Hopkins University)	The Effect of Stem Fit on the Radiocapitellar Contact Mechanics of a Metallic Axisymmetric Radial Head Hemiarthroplasty: Is Loose Fit better than Rigidly Fixed J. Szmit (Western University); G. J. W. King (Western University); J.A. Johnson (Western University); G.D.G. Langohr (Western University)	WRF Simulations of Urban Climate: Do We Need Low or High Topographic Resolution? Md. Rafsan Nahian (University of Guelph); Amir Nazem (University of Guelph); Mohsen Moradi (University of Guelph); Shohel Mahmud (University of Guelph); William Lubitz (University of Guelph); Amir Aliabadi (University of Guelph)	Numerical Investigations of a Comb-Like Film-Cooling Scheme Hao-Ming Li (Concordia University), Wahid Ghaly (Concordia University), Ibrahim Hassan (Texas A&M University at Qatar)	Design of Passive Wind-Tunnel Grids using a Prediction Model for Turbulence Intensity Dwaipayan Sarkar (Western University); Eric Savory (Western University)	Design Analysis of an Active Trailer Steering System for Multi-Trailer Articulated Heavy Vehicles Mutaz Keldani (University of Ontario Institute of Technology); Yuping He (University of Ontario Institute of Technology); Ghaus Rizvi (University of Ontario Institute of Technology)	Investigative Experimental Study of the Wetting Behavior of Transparent Coatings of PMMA Blends with PVC/PS/PVDF/PDMS Wing Yi Pao (University of Ontario Institute of Technology); Remon Pop-Iliev (University of Ontario Institute of Technology); Ghaus Rizvi (University of Ontario Institute of Technology)	Mechanical Properties of 3D Printed Primitive Shellular Materials with Metasurface Jiahao Shi (McGill University); Hamid Akbarzadeh (McGill University); Armin Mirabolghasemi (McGill University); Gilles Desharnais (Axis Prototypes)
2:50 PM - 3:10 PM	Effect of Printing Parameters on the Porosity of Stainless Steel Alloy 316L Using Direct Metal Laser Sintering Ali Eliasu (York University); Aleksander Czenkanski (York University); Solomon Boakye-Yiadom (York University)	Detailed Characterization of Novel Radiant Floor Heating and Cooling Systems S. Shukla (Ryerson University); R. Daneshazarian (Ryerson University); A. Mwesigye (Ryerson University); SB. Dworkin (Ryerson University); J. Swift (Grip Metal Manufacturing NUCAP Industries Inc.)	On Reliability of Thermally-Sprayed Alumina Dielectric Layers in a Multi-layered Coating-based Heating System Morvarid Bajirani (Concordia University); Milad Rezvani Rad (University of Alberta); André McDonald (University of Alberta); Christian Moreau (Concordia University)	Haptic Force Feedback for a Teleoperated Sonography System Atieh Najafi Semnani (University of Saskatchewan); Reza Fotouhi (University of Saskatchewan); QianWei Zhang (University of Saskatchewan); Haron Obaid (University of Saskatchewan); Scott Adams (University of Saskatchewan)	Separation of Carbon dioxide from Biogas by Adsorption of Activated Carbon Eniola Adewakun (University of Guelph); Animesh Dutta (University of Guelph)	Mechanics of Active Fluids at a Solid-Liquid Interface: Diffusion of Synechocystis SP. PCC 6803 in Dilute Regime Zahra Habibi (Western University); Thomas Vourc'h (Institut Curie Physico-Chimie Curie); Julien Léopoldès (Institut Langevin, Université Paris-Est); Malihe Mehdi-zadeh Allaf (Western University); Hassan Peerhossaini (Western University)	Experimental Simulation of a Density-Driven Downburst Translating within an Atmospheric Boundary Layer Shivani Jariwala (Western University); Kyle Graat (Western University); Eric Savory (Western University)	Control-Oriented Dynamic Model of a Three-way Catalytic Converter Utilizing Orthogonal Collocation Amer Keblawi (University of Waterloo); John McPhee (University of Waterloo)	On Fabrication and Durability of Superhydrophobic Electrodeposited Zn Coating on Stainless Steel: the Effect of Substrate's Surface Roughness Mona Amiriafshar (Memorial University); Xili Duan (Memorial University); Ali Nasiri (Memorial University)	Time-Fractional Dual-Phase-Lag Heat Conduction and Transient Crack Problem in Viscoelastic Materials under Thermal Shock Wenzhi Yang (University of Alberta); Zengtao Chen (University of Alberta)
3:10 PM - 3:30 PM	Thermal Distribution of a Stainless Steel 316 Body-Centered Cube Packing Particle at Sintering Temperatures Carlos Abel Rojas Dorantes (York University); Aleksander Czekanski (York University)	Optimal Design of Li-Ion Batteries for Health-Aware Fast Charging Xianke Lin (University of Ontario Institute of Technology); Youngki Kim (University of Michigan Dearborn)	Determining the Liquid Window of Al-Si Coated 22MnB5 during Austenitization using Reflectivity Measurements Cameron Klassen (University of Waterloo); Kyle Daun (University of Waterloo)	Validation of a Temperature Sensor for Implantable Infection Telemetry Michael Lavdas (Western University); David Holdsworth (Western University); Matthew Teeter (Western University); Brent Lanting (London Health Sciences Center)	Eggshell-Derived Carbon Dioxide Sorbent: Calcination and Carbonation Kinetic Studies Shakirudeen Salaudeen (University of Guelph); Animesh Dutta (University of Guelph); Bishnu Acharya (University of Prince Edward Island)	Numerical Study on Gas-Solids Circulating Fluidized Bed Downer Reactors Zeneng Sun (Western University); Chao Zhang (Western University)	A Very Large-Eddy Simulation (VLES) Model for Investigation of the Thermally-Stratified Atmospheric Boundary Layer: Reductionist and Practical Amir Aliabadi (University of Guelph); Gonçalo Pedro (RWDI)	A Review of Active Aerodynamic Control for Increasing Safety of High-Speed Road Vehicles Mohammed Hammad (University Of Ontario Institute Of Technology); Yuping He (University Of Ontario Institute Of Technology)	Manufacturing Processes Over the Final Residual Stress State of Crankshafts Luiz Fonseca (Instituto Tecnológico de Aeronáutica); Mario Batalha (Instituto de Pesquisas Tecnológicas - IPT); Alfredo de Faria (Instituto Tecnológico de Aeronáutica); Hamid Jahed (University of Waterloo)	Investigating the Effect of Porosity and Temperature on Modulus of Nanoscale Aluminium Sabir Subedi (Memorial University); Sam Nakhla (Memorial University)

June 5, 2019 (Wednesday) - CSME TRACK

8:30 AM - 9:30 AM	CSME-CFDSC Congress Closing Ceremony Plenary Session (Ivey BMO Auditorium): "Resilient Manufacturing System" Ruxu Du (South China University of Technology)									
9:30 AM - 10:00 AM	Coffee Break									
	Session A-5 (Advanced Manufacturing III) Session Chair: Evgueni Bordatchev Room: ACEB 1410	Session B-5 (Robotics, Mechatronics, Automation II) Session Chair: Farrokh Janabi-Sharifi Room: ACEB 2440	Session C-5 (Biomechanics III) Session Chair: Ryan Willing Room: SEB 1056	Session D-5 (Biomedical Engineering IV) Session Chair: Ali Ahmadi Room: SEB 1059	Session E-5 (Energy Materials IV) Session Chair: Lei Zhang Room: ACEB 1450	Session F-5 (Heat Transfer III) Session Chair: Sunny Ri Li Room: SEB 2099	Session G-5 (Fluid Mechanics VII) Session Chair: Tony Straatman Room: SEB 2100	Session H-5 (Nuclear Materials) Session Chair: Hamid Abdolvand Room: ACEB 2435	Session J-5 (Materials Engineering VI) Session Chair: Ben Jar Room: ACEB 1415	Session K-5 (Wind Energy and Engineering) Session Chair: Jubayer Chowdhury Room: ACEB 1420
10:00 AM - 10:20 AM	Axial Strategy for Ultraprecise Single Point Cutting of V-grooves with Constant Cross-Sectional Cutting Area Delfim Joao (Western University); O. Remus Tutunea-Fatan (Western University); Nicolas Milliken (National Research Council); Evgueni Bordatchev (National Research Council)	Design of the Mechanical Transmission for 5 DOF Robot with Distributed Active Semi-Active Actuation Sergey Pisetskiy (Western University); Mehrdad Kermani (Western University)	Performance of QCT-Derived Scapula Finite Element Models in Predicting Local Displacements Using Digital Volume Correlation Jonathan Kusins; Nikolas Knowles; Louis Ferreira; (Western University); Melissa Ryan; Enrico Dall'Ara; (University of Sheffield)	Microfluidic Droplet Transport by Photoelectrowetting Khaled Al-Arife (Abu Dhabi University); George Knopf (Western University)	Interfacial Engineering of Solid-State Batteries using Atomic Layer Deposition (Keynote Presentation) Neil P. Dasgupta (University of Michigan, Ann Arbor)	Development of a Numerical Model for Hydrothermal Conversion of Biomass Mohammad Heidari (University of Guelph); Animesh Dutta (University of Guelph); Bishnu Acharya (University of Prince Edward Island)	Drag Reduction in Conduits with Vibrating Walls Josuel Rogenski (Western University); Maciej Floryan (Western University)	Orientation Dependent Irradiation Hardening in Pure Zr Studied by Nanoindentation, TEM, and Crystal Plasticity Finite Element Modeling Qiang Wang; Christopher Cochran; Travis Skippon; Zhouyao Wang; Mark Daymond; (Queen's University); Hamid Abdolvand (Western University)	Crystal Orientation Effect on the Incipient Plasticity of the Thin Film Nanoindentation Mahdi Bagheripoor (Western University); Robert Klassen (Western University)	Calibration of Icing Severity Measurement on a Wind Turbine Nacelle using Image Analysis of Ice Accretions Patrice Roberge (Université Laval); Jean Lemay (Université Laval); Jean Ruel (Université Laval); André Bégin-Drolet (Université Laval)
10:20 AM - 10:40 AM	Effect of Tool Paths on Fine & Ultrafine Particles Emission & Distribution During Polishing of Medium Silica Content Granite Brice Kouamou (École de technologie supérieure); Victor Songmene (École de Technologie Supérieure); Jules Kouam (École de technologie supérieure)	Pneumatic Hyperelastic Actuators for Grasping Soft Organic Objects Alexandre Galley (Western University); George K. Knopf (Western University)	Material Mapping of QCT-Derived Scapular Models: A Comparison with Micro-CT Loaded Specimens using Digital Volume Correlation Nikolas Knowles; Jonathan Kusins; Mohammadreza Faiegi; Louis Ferreira; (Western University); Melissa Ryan; Enrico Dall'Ara; (University of Sheffield)	A Simulation on Multi-Pixel Carbon Nanotube Field Emitter for Medical Imaging Jiayu Liu (University of Waterloo); Elahe Cheraghi (University of Waterloo); Yonghai Sun (University of Waterloo); John Yeow (University of Waterloo)		Experimental Study of Natural Convection in a Horizontal Uniformly Heated Cylinder Kyle Teather (Western University); Kamran Siddiqui (Western University)	Measurement of Ice Adhesion on Stainless Steel Hydrophobic Surfaces Sadman Rhythm (Memorial University); Kewei Shi (Memorial University); Xili Duan (Memorial University)	A Non-Local Crystal Plasticity Finite Element Model for Simulating Irradiation Growth in α-Zirconium Polycrystals Omid Sedaghat (Western University); Hamid Abdolvand (Western University)	Effect of Heat Treatment on the Micro-structural Evolution and Mechanism of Wear of 3D Printed Ti6Al4V Alloys Niyousha Azgomi (York University); Solomon Boakye-Yiadom (York University)	A Large Eddy Simulation of Atmospheric Boundary Layer for Vertical Axis Wind Turbine Application Maryam Zabarjad Shiraz (Concordia University); Marius Paraischivou (Concordia)
10:40 AM - 11:00 AM	Online monitoring of built-up edge formation in turning stainless steel using acoustic emission signals Yassmin Ahmed (McMaster University); A.F.M. Arif (McMaster University); S.C. Veldhuis (McMaster University)	Kalman Filtering and PID Control of an Inverted Pendulum Robot Claire Bourque (University of Guelph); Andrew Lee (University of Guelph); Elyse Hill (University of Guelph); S. Andrew Gadsden (University of Guelph); Alex Bardelcik (University of Guelph); Mohammad Biglarbegian (University of Guelph)	Experimental Study of the Effect of Organic and Mineral Content on Bone Mechanical Properties Ogheneroborue Amromanoh (University of Manitoba); Yunhua Luo (University of Manitoba); Xinyi Wu (University of Manitoba)	Wireless Telemetry Load Sensor for Orthopaedic Applications William Anderson (Western University); Sydney Wilson (Western University); David W Holdsworth (Western University)	Organic Polyurea Thin Film for Stable and Long-Life Lithium Metal Anode Yipeng Sun (Western University); Tsun-Kong Sham (Western University); Xueliang Sun (Western University)	Comparing Artificial Neural Networks (ANN) with REFPROP Software in Predicting the Carbon Dioxide (CO2) Properties as a Working Fluid for the 10 MW Gas Turbine Power Plant Ibrahim Ali (Carleton University); Henry Saari (Carleton University)	Analysis of Gas Dispersion in Non-Newtonian fluids with a Coaxial Mixer through Tomography and CFD and Response Surface Methodology Maryam Jamshidzadeh (Ryerson University); Farhad Ein-Mozaffari (Ryerson University); Ali Lohi (Ryerson University)	On the State of Thermal Residual Stresses in Zirconium: Modeling and Experimentation Abdulla Alawadi (Western University); Hamid Abdolvand (Western University)	Characterization of Long-Term Mechanical Performance of Polyethylene (PE) and its Pipe Ben Jar (University of Alberta); Na Tan (University of Alberta); Yi Zhang (University of Alberta); Souvenir Muhammad (University of Alberta)	Experimental Simulations of Large Scale Tornado-Like Vortices Arash Ashrafi (Western University); Horia Hangan (Western University)

Authors Index (CSME Track)

Abdolvand, Hamid (Session H-5)
Abdullah, Abdullah (Session B-3)
Abdul-latif, Akrum (Session A-1)
Abhi, Trishan Deb (Session F-5)
Abid, Najmul (Session J-1)
Aboshosha, Haitham (Session K-5)
Aboutabikh, Moustafa (Session K-5)
Acharya, Bishnu (Session D-3, E-4, F-5, J-2)
Adair, Keegan (Session E-3)
Adams, Scott (Session D-4)
Adewakun, Eniola (Session E-4)
Agelin-Chaab, Martin (Session B-2)
Agrawal, Sumit (Session C-3)
Agur, Anne (Session C-3)
Agyapong, Joseph (Session J-3)
Ahamed, Mohammed Jalal (Session B-3, H-1)
Ahmadi, Ali (Session D-1, D-4, D-5, F-5, H-1)
Ahmadi, Tara (Session H-1)
Ahmadian, Niloufar (Session C-1)
Ahmadipour, Mohammadali (Session D-1)
Ahmed, Marya (Session D-4)
Ahmed, Mohamed (Session B-5)
Ahmed, Yassmin (Session A-5)
Akbarzadeh, Hamid (Session K-1, K-4)
Al Khatib, Fadi (Session H-4)
Al-Arife, Khaled (Session B-3, D-5)
Alatrash, Abubaker (Session D-1)
Alawadi, Abdulla (Session H-5)
AlBajian, Hamoud (Session B-5)
AlFailakawi, Eissa (Session B-5)
Alhourani, Abdulla (Session A-3)
Ali, Ibrahim (Session F-3, F-5)
Aliabadi, Amir (Session D-3, E-4, F-2, G-4)
Alirezaee, Shahpour (Session H-1)
Al-Kaabi, Zainab (Session D-3)
Alkayyali, Tartela (Session H-1)
AlKhudher, Ahmad (Session B-5)
Allaf, Malihe Mehdizadeh (Session F-4)
AlSalahi, Dalal (Session B-5)
Altenhof, William (Session J-1)
Aly, Samar (Session F-3)
Amirafshar, Mona (Session J-4)
Amirmaleki, Maedeh (Session E-3)
Amon, Cristina (Session D-1, D-2)
Amromanoh, Ogheneriobororue (Session C-5)
Anderson, William (Session D-5)
Angelin-Chaab, Martin (Session F-4)
Arif, A.F.M. (Session A-5)
Arku, Precious (Session F-3)
Armstrong, Jerrold (Session D-5)
Ashgriz, Nasser (Session G-3)
Ashrafi, Arash (Session K-5)
Asokanthan, Samuel (Session K-3)
Atefrad, Pedram (Session B-4, F-1)
Athwal, George (Session C-3)
Axford, David (Session C-3)
Azangbebil, Hayford (Session B-2)
Azgomi, Niyousha (Session J-5)
Bagheripoor, Mahdi (Session J-5)
Bahabadi, Hossein Gohari (Session A-4)
Baig, Altamash (Session B-2)
Bajgiran, Morvarid (Session C-4)
Baker, Mitchel (Session G-2)
Balachander, Ram (Session G-3)
Balbola, Amr (Session H-1, H-2)
Balsdon, Megan (Session C-2)
Banik, Sudhan (Session K-5)
Bannigan, Niall (Session F-2)
Barari, Ahmad (Session A-2, A-3, A-4)
Bardelcik, Alex (Session B-5)
Barrera, Jacob (Session D-4)
Barron, Ronald (Session G-3)
Barthelat, Francois (Session J-1, K-2)
Basu, Prabir (Session D-3)
Batailly, Alain (Session C-2)
Batalha, Mario (Session J-4)
Bayatpour, Afshin (Session J-2)
Bazylak, Aimy (Session D-1, D-2, E-1, E-2)
Bégin-Drolet, André (Session D-1, K-5)
Behdinan, Kamran (Session C-2)
Bender, Dylan (Session A-2)
Berkmortel, Carolyn (Session C-1)
Berry, Cody (Session A-3)
Bian, Kewei (Session C-1, D-2)
Biglarbegian, Mohammad (Session B-5, D-3)
Birmingham, Trevor (Session C-5)
Bitsuamlak, Girma (Session F-1)
Boakye-Yiadom, Solomon (Session A-4, J-3, J-5)
Bognash, Mohamed (Session K-3)
Bondy, Matthew (Session J-1)

Authors Index (CSME Track)

Bordatchev, Evgueni (Session A-1, A-3, A-5)
Boulama, Kiari Goni (Session B-1)
Bourque, Claire (Session B-5)
Boutilier, Michael (Session H-2)
Brahimi, Salim (Session J-2)
Brailovski, Vladimir (Session A-4)
Brochu, M. (Session K-3)
Bryant, Dianne (Session C-5)
Busmann, Markus (Session G-2)
Butany, Jagdish (Session C-3)
Butler, Haley (Session D-1, D-4)
Byerlay, Ryan (Session D-3)
Cabel, Spencer (Session K-1)
Cameron, Tiffany (Session D-5)
Campbell, Robert (Session H-1)
Campigotto, Angelica (Session H-1)
Cao, Changhong (Session H-2, E-3)
Carassale, Luigi (Session G-4)
Chadwick, Eric (Session D-1, D-2)
Chaieb, Oussama (Session A-1)
Chakrabarty, Rohan (Session A-3)
Chalmers, Heath (Session F-2)
Chan, Jason (Session D-2)
Chan, Sarah (Session H-1)
Charter, Edward (Session F-5)
Chatterjee, Abhijit (Session A-4)
Chauhan, Shakti (Session A-1)
Chen, Siyu (Session G-3)
Chen, Zengtao (Session K-3, K-4)
Cheraghi, Elahe (Session D-5)
Chew, Lup Wai (Session E-4)
Chiang, Yi Wai (Session F-2)
Choukir, Sahar (Session K-2)
Chowdhury, Jubayer (Session K-5)
Chung, Jennifer (Session C-3)
Clare, Lisa (Session C-2)
Cochrane, Christopher (Session H-5)
Colaitis, Yann (Session C-2)
Collin, Simon (Session D-1)
Cooper, Thomas (Session B-4)
Coronel, Eder Villa (Session F-3)
Cui, Teng (Session E-3)
Cvijanovic, Srdjan (Session A-5)
Czarnecki, Krzysztof (Session B-3)
Czekanski, Alex (Session J-3)
Czenkanski, Aleksander (Session A-4)
Dahal, Raj (Session J-2)
Dalaq, Ahmed (Session K-2)
Dall'Ara, Enrico (Session C-5)
Daneji, Ali (Session A-2)
Daneshazarian, R. (Session B-4, F-1)
Das, Tuhin (Session J-2)
Dasgupta, Neil P. (Session E-5)
Daun, Kyle (Session C-4)
Daymond, Mark (Session H-5)
de Bruyn, John (Session D-1)
de Faria, Alfredo (Session J-4)
De Snoo, Bob (Session J-3)
DeGroot, Christopher (Session B-2)
Deiab, Ibrahim (Session A-1, C-4)
Dellinger, Guilhem (Session G-4)
Deng, Sixu (Session E-5)
Dennis, Kadeem (Session G-2)
Derny, Alexander (Session F-2)
Desharnais, Gilles (Session K-4)
Dixon, S. Jeffrey (Session D-1, D-5)
Djokoto, Sylvester (Session B-2)
Dong, Zuomin (Session H-4)
Dorantes, Carlos Abel Rojas (Session A-4)
Doumit, Marc (Session C-1)
Dragašius, Egidijus (Session B-2)
Duan, Xili (Session F-1, G-2, G-3, G-5, J-4)
Dubay, Rickey (Session B-3)
Dubrowski, Adam (Session D-2, D-4)
Dudalski, N. (Session G-5)
Duntu, Solomon (Session J-5)
Dutta, Animesh (Session D-3, E-4, F-3, F-5)
Dworkin, Seth (Session B-1, B-4, F-1)
Dyck, Nolan (Session F-2)
Dyer, Benjamin (Session D-3)
E, Weiwei (Session G-3)
Ein-Mozaffari, Farhad (Session G-1, G-5)
El Essawi, Basim (Session A-2)
Elfasi, Osama (Session A-3)
Eliasu, Ali (Session A-4)
Elmaghraby, Hossam (Session F-2)
Elruby, Ahmed (Session J-1)
Eltaggaz, Abdelkrem (Session A-1, C-4)
Faber, Kenneth (Session C-5)
Faiegi, Mohammadreza (Session C-5)

Authors Index (CSME Track)

Fan, Ying (Session J-2)
Fang, Xingjun (Session F-2, G-2)
Farnese, Jeff (Session H-2)
Farrus, Nikolai (Session A-3)
Fayed, Eslam (Session A-4)
Ferreira, Louis (Session C-5, D-5)
Filleter, Tobin (Session E-3)
Fiorenze, Frank (Session D-1)
Flaig, Jennie (Session G-2)
Floryan, Maciej (Session G-5)
Flynn, Morris (Session C-2, G-2)
Fonseca, Luiz (Session J-4)
Forbes, Thomas (Session C-3)
Fotouhi, Reza (Session D-4)
Fotsing, E. (Session K-1)
Fradette, Julie (Session D-1)
Fung, Alan (Session B-2)
Gadallah, Aly (Session G-5)
Gadsden, S. Andrew (Session B-5, F-1)
Galley, Alexandre (Session B-5)
Gao, Xuejie (Session E-2)
Gardiner, Dave (Session B-2)
Gaskin, Susan (Session G-1)
Ge, Nan (Session E-2)
Gebrel, Ibrahim (Session K-3)
Gerlich, Adrian (Session J-1)
Getgood, Alan (Session C-3, C-5)
Ghaffari, Mahya (Session A-2)
Ghaly, Wahid (Session F-4)
Ghazal, Tarek (Session K-5)
Ghobeity, Amin (Session B-1)
Gibbons, Stephanie (Session D-2, D-4)
Gordon, Karen (Session C-1, C-3)
Graat, Kyle (Session G-4)
Greco, Dana (Session D-4)
Habibi, Zahra (Session F-4)
Haltli, Bradley (Session H-1)
Hammad, Mohammed (Session H-4)
Handrigan, Stephen (Session C-4, J-4, J-5)
Hangan, Horia (Session G-4, K-5)
Hanson, Ronald (Session G-1)
Hart, A. John (Session H-2)
Hassan, Ibrahim (Session F-4)
Hassanpour, Mohammad (Session B-3)
Hatai, Victor (Session C-2)
Hayward, Cindy (Session D-1)
He, Yuping (Session H-4)
Hebert, Marie (Session H-1)
Hegab, H. (Session A-1, C-4)
Heidari, Mohammad (Session F-5)
Henshaw, Paul (Session B-2)
Hijazi, Khaled (Session D-5)
Hill, Elyse (Session B-5)
Hmood, Osama (Session B-2)
Hojjati, Mehdi (Session J-2)
Holdsworth, David (Session D-1, D-4, D-5)
Holland, Sara (Session D-5)
Holness, Benjamin (Session A-5)
Hopkins, C. (Session A-5)
Hosein, Yara (Session D-5)
Hosseini, A. (Session A-5)
Hu, Huan (Session K-1, K-2)
Huang, Xiao (Session J-3)
Huculak, Greg (Session E-4)
Huissoon, Jan (Session H-1)
Huo, Hanyu (Session E-1)
Idrissa, Abdoul Kader Mossi (Session B-1)
Imad, M. (Session A-5)
Iqbal, Tariq (Session B-2)
Isaza, Pedro (Session G-2)
Islam, Mohammad (Session J-5)
Jahazi, Mohammad (Session A-1, A-4)
Jahed, Hamid (Session J-4)
Jamshidzadeh, Maryam (Session G-5)
Jar, Ben (Session J-5)
Jariwala, Shivani (Session G-4)
Jayaraman, Raghavan (Session K-1)
Jazebizadeh, Hooman (Session F-1)
Jegatheeswaran, Sinthuran (Session G-1)
Jennings, Morgan (Session C-5)
Jensen, Morten (Session J-1)
Jiang, Liying (Session K-2, K-4)
Jianu, Ofelia (Session B-1)
Joao, Delfim (Session A-5)
Johnson, David (Session G-1, G-3)
Johnson, James (Session C-1, C-3, D-4)
Jūrėnas, Vytautas (Session B-2)
Kahnamouei, Jalal Taheri (Session H-3)
Kahsay, Meseret (Session F-1)
Kannan, Sathish (Session A-2, A-3)

Authors Index (CSME Track)

Karami, Mohammad (Session G-4)
Karoubi, Golnaz (Session D-1, D-2)
Kassab, Aya (Session K-5)
Kaya, Tarik (Session F-1)
Kayed, Mohammed (Session H-1, H-2)
Kazemipour, Mostafa (Session J-2, J-3)
Keblawi, Amer (Session H-4)
Keldani, Mutaz (Session H-4)
Kermani, Mehrdad (Session B-3, B-5)
Kerr, Russell (Session H-1)
Khameneifar, Farbod (Session A-1)
Khanafer, Khahlil (Session A-1)
Khanjari, Hossein (Session G-1)
Khayat, Roger (Session F-5, G-3)
Kheirkhah, Niusha (Session D-2)
Kiamari, Amir (Session B-1)
Kim, Sanha (Session H-2)
Kim, Youngki (Session B-4)
Kimiaghalam, Farshad (Session B-2)
King, G. J. W. (Session D-4)
King, Graham (Session C-1)
Kishawy, A. (Session A-1, A-4)
Kizhakkeniyil, Manoj (Session D-3)
Klassen, Cameron (Session C-4)
Klassen, Robert (Session H-5, J-5)
Knopf, George (Session B-5, D-1, D-5)
Knowles, Nikolas (Session C-5)
Kojtych, Solène (Session C-2)
Kouam, Jules (Session A-5)
Kouamou, Brice (Session A-5)
Krayenhoff, E. Scott (Session D-3)
Kulashreshtha, Akash (Session D-3)
Kumar, Rakesh (Session B-2)
Kusins, Jonathan (Session C-5)
Kuska, Mitchell (Session B-2)
Ladak, Hanif (Session C-3)
Lai, Yongjun (Session H-1)
Lalone, Emily (Session D-5)
Lang, Haoxiang (Session B-1)
Langohr, Daniel (Session C-1, C-3, C-5, D-4)
Lanteigne, Eric (Session C-2)
Lanting, Brent (Session D-4)
Lavdas, Michael (Session D-4)
Lavoie, Philippe (Session G-1)
Lecompte, T. (Session K-1)
Lee, Andrew (Session B-5)
Lee, Jason (Session E-1)
Lemay, Jean (Session K-5)
Léopoldès, Julien (Session F-4)
Levy, Yanir (Session D-2)
Li, Hao-Ming (Session F-4)
Li, Minsi (Session E-2)
Li, Ri (Session F-3)
Li, Weihang (Session E-2)
Li, Xianguo (Session B-4)
Li, Xiaona (Session E-1)
Li, Yuanping (Session K-2)
Li, Zhengwei (Session K-1)
Liang, Jianneng (Session E-1)
Liang, Jianwen (Session E-1)
Lin, Xianke (Session B-1, B-4, H-3)
Lin, Xiaoting (Session E-2)
Liu, Hao (Session F-3)
Liu, Hugh (Session B-3)
Liu, Jiayu (Session D-5)
Liu, Kefu (Session F-2)
Liu, Liyan (Session H-3)
Liu, Xingchen (Session C-2)
Liu, Xingxing (Session C-4)
Lockhart, Jason (Session C-3)
Lohi, Ali (Session G-5)
Lorusso, Daniel (Session D-1)
Louca, Karim (Session H-5)
Lu, Lihong (Session D-2)
Lubitz, William (Session D-3, E-4, F-3, G-4)
Luo, Yiming (Session K-3)
Luo, Yunhua (Session C-1, C-5)
Lyons, Murray (Session G-4)
Lyu, Shan (Session F-2)
Ma, JuanLi (Session B-2)
Ma, Zhe (Session H-3)
MacCallum, Ben (Session D-5)
MacDonald, Debra (Session D-1)
MacNevin, Wyatt (Session D-4)
Magliaro, John (Session J-1)
Mahmud, Shohel (Session E-4, F-1)
Mao, Haojie (Session C-1, D-2, D-5)
Martin, Elia (Session F-5)
Matida, Edgar (Session B-2, D-1)
Matsuura, Kazuo (Session G-2)

Authors Index (CSME Track)

Mayer, Rene (Session A-3)
Mazyan, Walid (Session F-5)
McDonald, André (Session C-4, F-3)
McKay, Gordon (Session D-3)
McPhee, John (Session B-3, H-4)
Medraj, Mamoun (Session A-4)
Mehri, Rym (Session D-1)
Menard, Luc (Session B-2)
Mhoammadi, Mohsen (Session J-3)
Milliken, Nicolas (Session A-3, A-5)
Milner, Jaques (Session D-1)
Minaker, Bruce (Session H-3)
Mirabolghasemi, Armin (Session K-4)
Mistruzzi, Elle (Session G-3)
Mohamed, A. (Session G-5)
Montesano, John (Session J-2)
Moradi, Mohsen (Session D-3, E-4)
Moreau, Christian (Session C-4)
Morrissey, Liam (Session C-4, J-4, J-5)
Moslemian, Alireza (Session C-3)
Mousavi, Seyedeh Mohadeseh Taheri (Session H-2)
Moussa, Walied (Session H-1, H-2)
Mubareka, S. (Session G-5)
Muhammad, Souvenir (Session J-5)
Mwesigye, Aggrey (Session B-1, B-4, F-1)
Nahian, Md. Rafsan (Session D-3, E-4)
Nakhla, Sam (Session C-4, D-2, D-4, F-3, J-1, J-4, J-5, K-1, K-3)
Naseri, Emad (Session D-1, D-4, D-5)
Nasiri, Ali (Session A-2, A-4, J-2, J-3, J-4)
Natarajan, Shankar (Session G-3)
Naveen, Easwaran (Session J-5)
Nayakanti, Nigamaa (Session H-2)
Nazarahari, Milad (Session C-1)
Nazem, Amir (Session D-3, E-4)
Nec, Yana (Session E-4)
Nemani, Alireza Vahedi (Session A-2)
Nguyen, Hiep (Session B-4, F-1)
Nguyen, Thu (Session H-1)
Nicholson, Sarah (Session B-1)
Nijhawan, Mudit (Session B-1)
Nikolov, Hristo (Session D-1)
Norford, Leslie (Session E-4)
Nyantekyi-Kwakye, Baafour (Session F-2, G-2)
Obaid, Haron (Session D-4)
Ocran, Emmanuel (Session C-3)
Oh, Kwang-seok (Session H-3)
Ojo, Olaoluwa (Session B-1)
Oliver, Michele (Session C-1, C-3)
Olufayo, Oluwole (Session A-1, A-3)
Orf, Leigh (Session F-2)
Ouzounian, Maral (Session C-3)
Pao, Wing Yi (Session J-4)
Paraischivou, Marius (Session K-5)
Park, James (Session B-4)
Parker, Mark (Session C-2)
Patnaik, Kondavalasa (Session H-4)
Pedro, Gonçalo (Session G-4)
Peerhossaini, Hassan (Session F-4, G-1, G-4)
Pepler, Daniel (Session A-2)
Pervaiz, Salman (Session A-2, A-3)
Pessard, E. (Session K-3)
Pickering, J. Geoffrey (Session D-1)
Piollet, Elsa (Session C-2)
Pisetskiy, Sergey (Session B-5)
Pjontek, Dominic (Session F-4, G-5)
Poduska, Kristin (Session G-2)
Poepping, Tamie (Session A-5, D-1)
Pope, Kevin (Session G-3)
Pop-Iliev, Remon (Session J-2, J-4, K-1)
Pradhan, Ranjan (Session D-3)
Pratt, Douglas (Session K-1)
Price, Aaron (Session A-5)
Qian, Longhao (Session B-3)
Qinghui, Zhou (Session H-3)
Qiu, Shuwei (Session B-3)
Qu, Tianliang (Session K-3)
Quintero, Ricardo Bautista (Session B-3)
Rad, Milad Rezvani (Session C-4)
Rafieazad, Mehran (Session A-2, A-4)
Rahaman, Habibur (Session B-2)
Rahnamayan, S. (Session A-1)
Rahnamoun, Ali (Session C-4)
Ramadan, Omar (Session B-2)
Ramanan, N. (Session J-5)
Rans, Calvin (Session J-3)
Rasha, Rabbani (Session B-2)
Ratzlaff, Tim (Session H-1)
Refai, Razim (Session F-3)
Ren, Carolyn (Session H-1)
Rhythm, Sadman (Session G-5)

Authors Index (CSME Track)

Ricketts, Amanda (Session D-2)
Rizkalla, Amin (Session D-5)
Rizvi, Ghaus (Session J-2, J-4, K-1)
Roberge, Patrice (Session K-5)
Rogenski, Josuel (Session G-5)
Rohani, Seyed Alireza (Session C-3)
Rohiith, G.B (Session J-5)
Romero, David (Session D-1, D-2)
Ross, A. (Session K-1, K-3)
Rouhani, Hossein (Session C-1)
Ruel, Jean (Session D-1, K-5)
Ryan, Melissa (Session C-5)
Saari, Henry (Session F-3, F-5)
Sadeghi-Naini, Ali (Session D-2)
Saeedi, Mohammad (Session F-2)
Sahebjam, Rana (Session G-1)
Salahi, Salar (Session J-2, J-3)
Salaudeen, Shakirudeen (Session E-4)
Saleem, Junaid (Session D-3)
Salem, A. (Session A-1)
Salt, David (Session B-4, F-1)
Samani, Abbas (Session D-2)
Samara, Farid (Session G-3)
Sarkar, Dwaipayan (Session G-4)
Savory, Eric (Session F-2, G-4, G-5)
Secanell, Marc (Session C-2)
Sedaghat, Omid (Session H-5)
Semnani, Atieh Najafi (Session D-4)
Sentissi, M. (Session K-3)
Seo, Jaho (Session H-3)
Shahriari, Davood (Session A-4)
Sham, Tsun-Kong (E-2, E-5)
Shamoon, Matthew (Session B-3)
Shi, Jiahao (Session K-1, K-4)
Shi, Kewei (Session G-2, G-5)
Shiraz, Maryam Zabarjad (Session K-5)
Shojaei-Kavan, Loabat (Session H-5)
Shrestha, Ankita (Session D-3)
Shrestha, Pranay (Session E-2)
Shukla, S. (Session B-4)
Siddique, Abu Raihan (Session F-1)
Siddiqui, Kamran (Session B-2, F-1, F-5, G-2, G-5)
Siddiqui, Naseeb (Session F-4)
Simmons, Craig (Session C-3)
Simmons, Scott (Session G-4)
Sinar, Dogan (Session D-1)
Singh, Aman Preet (Session A-2)
Singh, Chandra Veer (Session E-3, J-4, K-2, K-4)
Skippon, Travis (Session H-5)
Sofiane, Achiche (Session A-3)
Soltanmohammadi, Pendar (Session C-5)
Song, Jun (Session A-3, H-2, J-2)
Songmene, Victor (Session A-1, A-3, A-5)
Soon, Kayla (Session D-1)
Srabet, Saleh (Session F-4, G-5)
Steeves, Craig (Session A-2)
Straatman, Anthony (Session C-2, F-2)
Subedi, Sabir (Session K-3, K-4)
Sudharshan, Raghunathan (Session J-5)
Sun, Hao (Session J-4)
Sun, Xueliang (Session E-1, E-2, E-3, E-5)
Sun, Yipeng (Session E-5)
Sun, Yonghai (Session D-5)
Sun, Yu (Session E-3)
Sun, Zeneng (Session F-4)
Suratkar, Aaditya (Session J-2)
Suzuki, Takaya (Session D-1, D-2)
Swift, J. (Session B-4)
Syed, Nabeel (Session J-2)
Szmit, J. (Session D-4)
Tachie, Mark (Session F-2, G-2)
Taghavi, Seyed Mohammad (Session F-2)
Tam, Jaosn (Session E-3)
Tan, Na (Session J-5)
Tang, Mingyi (Session C-3)
Tariku, Fitsum (Session F-1)
Tarokh, Ali (Session F-2)
Tasker, Andrew (Session D-1)
Tasnim, Syeda (Session F-3)
Tavakoli, Amir (Session C-5)
Teather, Kyle (Session F-1, F-5)
Teeter, Matthew (Session D-4)
Tepyllo, Nicholas (Session J-3)
Tetteh, Francis (Session J-3)
Thapa, Devi (Session F-4)
Thevathasan, Naresh (Session D-3)
Ting, David (Session B-2)
Ting, Francis (Session C-3)
Tutunea-Fatan, O. Remus (Session A-3, A-5)
Umapathy, Jagathish (Session J-5)

Authors Index (CSME Track)

Ungarish, Marius (Session G-2)
Utkarsh, Utkarsh (Session J-2)
Van Heyst, Bill (Session E-4, F-1)
Varty, Robert (Session G-3)
Vatao, Muhammad Tariq (Session K-1)
Veldhuis, S.C. (Session A-5)
Vickers, Brad (Session C-2)
Vickery, Peter (Session K-5)
Vogt, Anja (Session F-5)
Vourc'h, Thomas (Session F-4)
Waddell, Thomas (Session D-1, D-2)
Wang, Baohua (Session H-4)
Wang, Biqiong (Session E-3)
Wang, Changhong (Session E-3)
Wang, Ligang (Session K-3)
Wang, Qiang (Session H-5)
Wang, Siyang (Session H-3)
Wang, Xi (Session B-2)
Wang, Xiaodong (Session K-1, K-2)
Wang, Yunpeng (Session F-5, G-3)
Wang, Zhouyao (Session H-5)
Webster, Gary (Session B-2)
Whitaker, Stephen (Session G-2)
Whittaker, Jackie (Session C-1)
Wilcox, Aaron (Session B-2)
Willing, Ryan (Session C-3, C-5)
Wilson, Sydney (Session D-5)
Wong, Edwin (Session C-3)
Wood, David (Session E-4)
Wood, Jeffrey (Session C-2, J-2)
Wu, Xinyi (Session C-5)
Xia, Yakang (Session F-3)
Xiao, Junfeng (Session A-2)
Xing, Kanglin (Session A-3)
Yang, Jianming (Session H-3)
Yang, Jun (Session A-4, C-4)
Yang, Wenzhi (Session K-4)
Yang, Xiaofei (Session E-1)
Yeh, Han Hung (Session D-4)
Yeow, John (Session D-5)
Yifan, Xi (Session H-3)
Yim, Evelyn (Session H-1)
Yongyuan, Su (Session H-3)
Young, Calvin (Session C-1)
Younis, Adel (Session B-5, H-4)
Yu, Wanshi (Session C-1)
Yue, Stephen (Session J-2)
Zhang, Chao (Session F-4)
Zhang, Daqian (Session F-1)
Zhang, Dongxing (Session A-2)
Zhang, Guangde (Session H-4)
Zhang, Lei (Session E-2)
Zhang, QianWei (Session D-4)
Zhang, Yi (Session J-5)
Zhao, Feipeng (Session E-5)
Zhao, Huiyong (Session H-4)
Zhao, Jian (Session B-4)
Zhao, Yang (Session E-2, E-3)
Zheng, Mingyue (Session A-4)
Zhou, Jianyou (Session K-4)
Zhu, Jesse (Session F-4, G-5)
Zhu, Jin (Session C-1)
Zilstra, Alison (Session G-1)

Technical Abstracts & Papers

Additive Manufacturing

Anisotropy in Mechanical Properties of a Wire Arc Additive Manufactured Low Carbon Low Alloy Steel

Mahya Ghaffari, Alireza Vahedi Nemani, Mehran Rafieazad, Ali Nasiri

Faculty of Engineering and Applied Science,
Memorial University of Newfoundland,
St. John's, NL, Canada, A1B 3X5
mghaffari@mun.ca

ABSTRACT

In this study, ER70S-6 low carbon low alloy steel wall was 3D-printed by Wire Arc Additive Manufacturing (WAAM) method using a gas metal arc welding (GMAW) torch mounted on a six-axis robotic arm and employing advanced surface tension transfer (STT) mode. Microstructural characterization of the manufactured component was carried out utilizing optical microscopy (OM), field emission scanning electron microscopy (FESEM), X-ray diffraction (XRD), and electron backscatter diffraction (EBSD). Mechanical properties of the component were studied at different orientations relative to the building direction utilizing micro-hardness testing, and uniaxial tensile testing followed by fractography of the fractured surfaces. The dominant microstructure of the fabricated material contained polygonal ferrite grains and the pearlite lamellae formed at the ferrite grain boundaries, which were observed at a high volume fraction among all layers of the wall. A small volume fraction of bainite and acicular ferrite was also distinguished along the melt pool boundaries, where the material is encountered a higher cooling rate during solidification in comparison with the rest of the melt pool. The mechanical properties characterization of the fabricated part revealed a comparable tensile strength in deposition (horizontal) and building (vertical) directions, measured at ~400 MPa and ~500 MPa for the yield and ultimate tensile strengths, respectively. However, the elongation percentage in horizontal samples was nearly three times higher than that of the vertical samples. The measured anisotropy in ductility was found to be ascribed to the higher density of the interpass regions and the melt pool boundaries in the vertical direction, containing heat affected zones with coarser grain structure, brittle martensite-austenite constituent, and possibly a higher density of discontinuities.

Effect of printing parameters on the porosity DMLS AISI 316L Alloy

Ali Eliasu

Alex Czenkanski

Solomon Boakye-Yiadom

Mechanical Engineering Department,
Lassonde School of Engineering, York University
Toronto, ON, Canada, M3J 1P3

ABSTRACT

Powder bed fusion technology is the most common metal 3D printing technique due to its ability to create dense parts that require little to no post processing. The capability of controlling porosity level in printed parts has started gaining traction in the Additive manufacturing research field. This gives the opportunity to fine tune the stiffness of parts based on its applications

This study is to investigate how the printing parameters affect pore formation in 3D printed 316L Stainless steel using Direct Metal Laser Sintering technique. Two types of pores were investigated; the complete spherical pores that occur because of gas atomization of the feed material (metal powder) and the pores that occur during the printing process itself. The former is spherical in shape while the latter has irregular shape. Fifteen samples with varied 3D printing parameters were produced with EOS M280 DMLS Printer and these samples were metallographically prepared and studied extensively under the optical microscopy to characterize the nature of the porosity as the laser energy changes from sample to sample. A very common way of comparing printing parameters is the use of Energy Density (E.D) which is representative of the laser power per unit volume delivered to the part during the printing process. This quantity combines the varying parameters used in the experiment, namely laser power(P), scan speed(v) and the hatch spacing(h). The energy density is calculated using the formula; $\text{Energy Density (J/mm}^3\text{)} = P/(h \cdot v \cdot \delta)$. The porosity levels are determined with the image processing software, ImageJ, where 10 different micrographs per samples were used.

Generally, the porosity level begins to reduce as the energy density increases because more energy is fed into the material for fusion. Hence as the energy increase more of the particles fuse together create less pores in the part. The study showed that for AISI 316L parts, when the E.D is lower than 20 J/mm³, the porosity is approximately 50% and higher. E.D within 30-40 J/mm³ gives a porosity level between 1%-20% and having the energy density higher than 40 J/mm³ bring the pore counts further to a minimum value (less than 1%). It is very crucial to not that as the E.D goes beyond a certain threshold (~120 J/mm³ for AISI 316L) the pore begin to re-emerge. The pores reappear is attributed to the fact that at that high energy level the material starts to evaporate, and the laser penetrates deeper than is expected which results in keyhole formation (a site with affinity to pore formation). The energy density does not correspond to an exact porosity level but a range. There are samples with the same energy density values which show a slight variation in porosity because the individual parameters are entirely different from each other. Comparing the individual parameters also shed more light on the matter. For samples with just the scanning speed or hatching spacing changing had a linear relationship with the level of porosity. Samples with varying scan speed showed a direct relation to porosity level, as the laser scan speed increases so does the level of porosity in the part, same can be said for increase in hatch spacing and porosity level. However, the laser power follows the same trend as the E.D against porosity for the same reasons.

3D Printing of Vitrimer with Recycable and Precise In-Situ Self-Healing Properties

Mingyue Zheng

Department of Mechanical and Materials Engineering
Western University
London, Ontario, Canada
mzheng46@uwo.ca

Jun Yang

Department of Mechanical and Materials Engineering
Western University
London, Ontario, Canada
jyang@eng.uwo.ca

ABSTRACT

Three-dimensional (3D) Printing, which is also known as the Additive Manufacturing (AM), refers to the manufacturing processing of a three-dimensional product from a computer-driven digital model. This remarkable invention promotes the fourth industrial revolution development. Comparing with traditional manufacturing, AM has a great advantage in green economic which can achieve by two aspects. First, this computer controllable manufacturing technology can combine with cloud computing and achieve personal small-scale manufacturing, which can greatly reduce the industry equipment investment and transport energy costs. Second, additive manufacturing does not have residual parts, so it quite reduces the raw materials consumption and can achieve zero materials waste.

Whereas 3D printing technology still remains a number of challenges to overcome before it is accepted as a standard manufacturing method and widely employed for production use. The worldwide 3D printing industry market forecast will be over \$20 billion in 2020, and over half of the 3D printing materials will be the thermosets. Thermosetting materials are ideal 3D printing materials for their thermal stability, good mechanical properties, and high solvent resistance ability. However, due to their permanently crosslinked network, 3D printing thermoset polymers have problems in damage repairing, waste recycling, and the printing methods limitation. The thermoset 3D printing parts recycling will become a serious issue for 3D printing industry development and environment problem. Vitrimers is a derivative of thermosetting polymers. With their covalent adaptable networks (CAN), vitrimers behave like thermosets at service temperature, are insoluble at all temperature, but can flow when heated. Based on the dynamic covalent bond mechanism, we developed the new extrusion base 3D printable thermoset materials. These printable materials can be repaired under heat treatment and can achieve 100% recyclable. This technology can be widely used in different types of applications.

3D printed metamaterials of dual/negative Poisson's ratio behaviors induced by hierarchical architecture

Junfeng Xiao
Mechanical and Materials Engineering
Western University, UWO
London, Canada
jxiao47@uwo.ca

Dongxing Zhang
Mechanical and Materials Engineering
Western University, UWO
London, Canada
zhangdx1113@gmail.com

ABSTRACT

Metamaterials with artificially designed architectures can achieve unique and even unprecedented physical properties, which show promising applications in actuators, amplifiers and micromechanical controls. An initiator-integrated 3D printing technology (i3DP) was applied in this study to create scalable, metal/polymer meta-mechanical materials. By increasing structure's hierarchy, the materials can gradually achieve negative Poisson's ratio, high strength and ultralow density, as well as high compressive and super-elastic behavior. A class of metamaterials with mechanically tunable Poisson's ratio are fabricated. Furthermore, by inducing hierarchically hybrid structure, the materials can be endowed with dual Poisson's ratios.

Topology Optimization for Layered Material

Daniel Pepler

University of Toronto Institute for Aerospace Studies
 Toronto, Canada
d.pepler@mail.utoronto.ca

Craig Steeves

University of Toronto Institute for Aerospace Studies
 Toronto, Canada
csteeves@utias.utoronto.ca

Abstract—Topology optimization is a powerful optimization technique because it efficiently generates optimal structures that can be very intricate, far beyond what a human designer could conceive. The number, size and shape of the holes in the structure are not inherently limited, allowing designs of near infinite complexity. Such highly complex designs often cannot easily be manufactured through conventional approaches. Structures with internal lattices are an example. Additive manufacturing is one of the few viable manufacturing approaches to construct these complex designs. However, additive manufacturing has distinct characteristics imposed by the layering process, such as anisotropy. If the material properties associated with additive manufacturing are not taken into consideration during the design process, heavy post processing is required to ensure a robust design. In this paper the characteristics of fused deposition modelling additive manufacturing are considered during the topology optimization process using a new layered element approach. The ensuing optimal designs are compared with corresponding designs using conventional anisotropic elements. The results show that both anisotropic and layered elements give unique topologies with advantages and disadvantages. Most notable is that a layered element description allows for stochastic properties associated with the additive manufacturing to be directly measured and modelled.

Keywords- *Additive Manufacturing; Topology Optimization; Layered Material; Robust Optimization; Stochastic Models*

I. INTRODUCTION

Topology optimization has the freedom to give designs of near infinite complexity, but which often cannot be manufactured through conventional approaches. Additive manufacturing is one of the few viable manufacturing techniques, with its layer by layer production approach [1]. Additive manufacturing has some ideal properties: it can manufacture complex shapes as easily as simple ones, there is less material waste, and there is no overhead cost associated with constructing a mold [2]. These properties make it ideal for fabricating the complex structures obtained using topology optimization. However, the layering method associated with all

additive manufacturing processes affects material properties. When a new layer is fused to a previous one there is a transition phase or adhesion layer where the material properties are changed [3]. Often this adhesion layer is weaker, more brittle and has more voids compared with bulk material properties [3]. On the macroscopic scale this results in anisotropic material properties [3]. Two different methods were used to model the layered material to determine its effect on topology optimization.

A. Additive Manufacturing

For this paper the focus will be on the fused deposition modelling (FDM) method of additive manufacturing. In FDM a molten filament is deposited in a raster-like fashion one layer at a time creating the 3D structure [4]. Figure 1 gives a visualization of the FDM process. The extruder tip is moved while the molten material is being extruded to deposit each layer. As layers stack on top of one another a 3D shape is created.

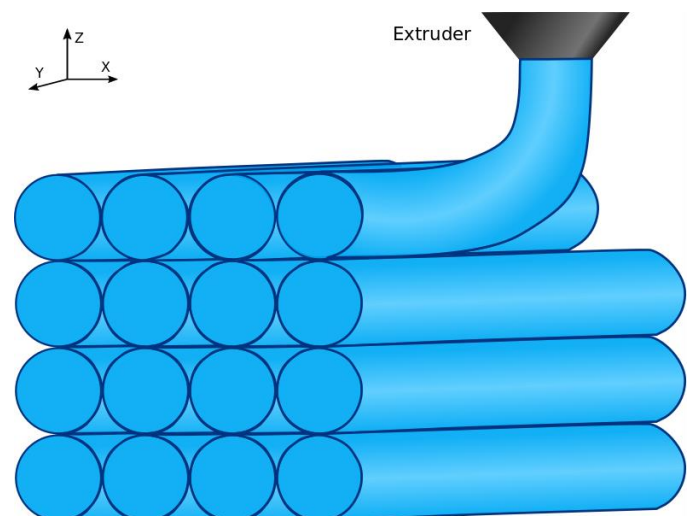


Figure 1 A depiction of the fused deposition modeling process. The molten material is being extruded onto/ next to the previously deposited filaments.

As the molten filament is placed on the previous layer its heat locally re-melts the previous layer allowing the filaments

to fuse [4] by allowing the polymer chains in the previous layer to move and entangle with the polymer chains in the molten filament. This entanglement process is referred to as 'fusion' [5]. The amount of fusion between the new filament and the previous layer can vary greatly depending on ambient temperature, air gap and melt temperature [4]. The balance of these parameters is difficult to optimize and often the adhesion layer has weaker material properties compared to bulk material [5]. This filament fusion process is shown in figure 2 where three levels of adhesion and polymer chain entanglement are shown. The dashed boxes show how thick the adhesion layer is in each case. As the amount of fusion increases the thickness of the adhesion layer increases which can be measured physically using the surface roughness [6,7]. This weaker adhesion layer is why parts manufactured through FDM possess anisotropic material properties: stronger along the length of the polymer filament (the y direction in figure 1) and weaker in the bond between filaments (the x and z directions in figure 1) [5].

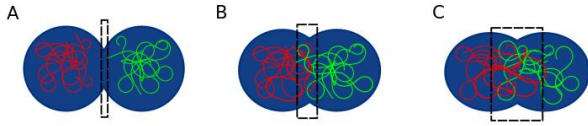


Figure 2 The process of two filaments fusing together through polymer chain entanglement. Images A, B, and C show how the filaments go from no entanglement to thorough entanglement. The dashed boxes show the adhesion layer.

B. SIMP Topology Optimization

One popular technique for topology optimization is the Solid Isotropic Material with Penalization (SIMP) method. This method allows densities to take intermediate values, which enables gradient-based optimization [8,9]. In most cases the intermediate densities have no physical meaning and therefore are penalized and filtered for final solutions [10,11]. Codes employing the SIMP formulation are available in Sigmund's 99 and 88 line Matlab codes for academic purposes [8,9]. Equation (1) gives the formulation for the standard compliance minimization topology optimization with a volume constraint, where C is the compliance, d is the displacement, F is the applied force and K is the stiffness matrix. This equation is solved using finite element analysis. Compliance and displacement are both a function of material density ρ , which in SIMP can vary between 0 and 1. V_e gives the volume of each element and V_{max} is the total volume constraint.

$$\begin{aligned} \min_{\rho} : \quad & C(\rho) = Fd(\rho) \\ \text{subject to :} \quad & F = K(\rho)d(\rho) \\ & \sum_e \rho_e V_e \leq V_{max} \\ & \rho_{min} \leq \rho_e \leq \rho_{max}. \end{aligned} \quad (1)$$

C. Robust Topology Optimization

Robust topology optimization takes into consideration the stochastic nature of a system. To perform robust topology optimization uncertainties need to be quantified and the Monte-Carlo method was chosen for this study. The Monte-Carlo method can handle the large variance expected in the FDM and it is easy to implement [10]. The Monte-Carlo method requires many simulated systems, based on the statistics of the model system, to give representative samples. For the layered material a gaussian random variable was assigned to each adhesion layer and evaluated to produce M simulated systems. The objective function for these M simulated systems was calculated which allowed the expected value and variance of the objective function to be determined. The expected value and variance of the original objective function are used in a new weighted objective function. The expressions for a weighted objective function and its sensitivity with respect to density are [10]:

$$\hat{C} = \alpha_E E[C] + \alpha_S Std[C], \quad (2)$$

and

$$\frac{\partial \hat{C}}{\partial \rho} = -\frac{1}{M} \sum_i \left(\alpha_E + \frac{\alpha_S (\hat{C}_i - E[C])}{Std[C]} \right) \tilde{d}_i^e \frac{\partial \tilde{K}_i^e}{\partial \rho^e} \tilde{d}_i^e. \quad (3)$$

The constants α_E and α_S are the weight factors for the expected value ($E[\bullet]$) and standard deviation ($Std[\bullet]$) respectively. The weight factors are set at the beginning of the simulation such that the contribution of the expected value and the standard deviation of compliance are equal. This balances the weight factors but can cause problems in systems with very high or low variance. To mitigate this problem the values for these constants are checked at the beginning of each simulation. \tilde{K}_i^e denotes the elemental stiffness matrices, while $\frac{\partial \tilde{K}_i^e}{\partial \rho^e}$ denotes their sensitivity with respect to the element density. The elemental displacements are given by \tilde{d}_i^e .

II. PROBLEM FORMULATION

A. Methods for Modeling the Layered Material

A simple method for modeling the FDM process is to change from an isotropic to anisotropic stiffness matrix. The integral definition of the stiffness matrix is,

$$K^e = \int_{\Omega_e} B^T D B. \quad (4)$$

Where D is the elastic matrix, B is the gradient of the nodal shape functions, K^e is the element stiffness matrix and Ω_e is the element domain. An anisotropic stiffness matrix takes separate moduli for each different direction modifying the elastic matrix. An anisotropic elastic matrix allows for a unique modulus and Possion's ratios for the different directions as well as requiring a unique shear modulus. An isotropic material is described by a single Young's modulus (E) and Possion's ratio (ν), and the shear modulus is a function of E and ν . An

anisotropic material has a unique Young's modulus (E_1 and E_2) for each direction, two Poisson's ratio (ν_{12} , ν_{21}) for conservation of energy and a unique shear modulus (G_{12}). The matrix of elastic constants for an anisotropic matrix is:

$$D = \begin{bmatrix} \frac{E_1}{(1-\nu_{12}\nu_{21})} & \frac{\nu_{21}E_1}{(1-\nu_{12}\nu_{21})} & 0 \\ \frac{\nu_{12}E_2}{(1-\nu_{12}\nu_{21})} & \frac{E_2}{(1-\nu_{12}\nu_{21})} & 0 \\ 0 & 0 & G_{12} \end{bmatrix} \quad (5)$$

An alternative method for modeling the FDM process is to use a layered meshing scheme with six nodes and two Young's moduli E_b and E_a for the bulk and adhesion material respectively. Using six nodes allows for two local element stiffness matrices to be created, K_b^e and K_a^e , representing the bulk material and the adhesion material respectively. The two local element stiffness matrices are then assembled into an elemental stiffness matrix, $K^e = K_b^e \cup K_a^e$. In SIMP the density value was assigned to K^e to ensure that adhesion layer material was always associated with deposited material. These elements were used in all but the bottom most elements, the bottom layer is not fused to anything below it so does not need a layered element. This meshing structure, shown in figure 3, allows for a more physically accurate description of the additive manufacturing system. The drawback is that this method is more computationally expensive, almost doubling the number of nodes.

A benefit to using layered elements is that it allows for the stochastic nature of the process to be associated with only the adhesion layer, allowing for stochastic anisotropy to be modelled. This is not possible with the anisotropic elastic matrix as the degree of anisotropy cannot be taken out of the integral definition of the stiffness matrix. The stochastic properties associated with the adhesion layer are physically measurable. For example, it is feasible to measure the difference in layer thickness or number of voids in the adhesion layer. It is very difficult to measure variation in local Young's modulus. Typically, stochastic bulk properties are assumed to hold true locally, which can fail to describe what is happening physically.

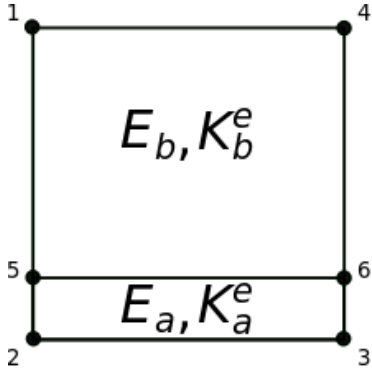


Figure 3 Linear layered mesh structure using six nodes and two elements. Each element has a unique Young's Modulus and stiffness matrix.

B. Model Systems

Two model systems were used to show the affect that introducing anisotropic and layered elements have on the final topology. These systems can be found in figure 4. System 1 is an MBB beam; for all simulations the dimensions of the MBB-beam were 600mm by 100mm with a mesh size of 2mm. The symmetry of this system requires that only half of the system be solved. This plane of symmetry is shown in the figure 4 by a dashed line. The results of this system will be shown as the half solution, which would be mirrored in a final design. System 2 is an L-bracket. For all simulations the length of the L-bracket arms was 200mm and the width were 80mm with a mesh size of 2mm.



Figure 4 System 1: The MBB beam problem with a 6:1 length to width ratio and load F . The grey region Ω will be taken as the design space for topology optimization, the two red circles show the supports. System 2: The L-Bracket problem with each arm being 200mm in length and 80mm in width.

III. RESULTS

A. Isotropic Topology Optimization

Each system was first optimized using an isotropic material description for reference purposes. The topology optimization simulations were run using a 30% volume fraction. Sigmund's modified sensitivity filter was used with a 3mm filter radius [10,11]. This filter removes known SIMP problems such as checkerboarding but still results in a well-defined boundary with very little intermediate densities [10,11]. The method of moving asymptotes was used as the optimizer [13-15]. These simulation parameters will be used for all simulations.

System 1

System 2

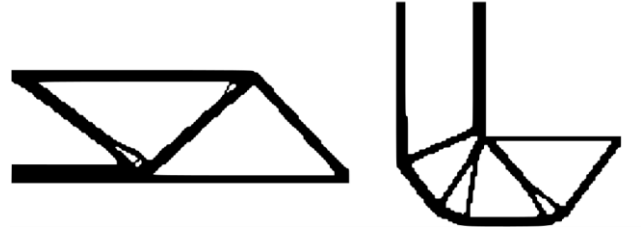


Figure 5 Isotropic solutions for system 1 and 2

B. Anisotropic Element Topology Optimization

An anisotropic elasticity matrix, as defined in equation 5, was used to model the bulk anisotropy found when using the

FDM process. Three degrees of anisotropy (da) were used (10%, 20%, 30% and 40%) where E_1 , E_2 , ν_{12} and ν_{21} were determined as follows:

$$E_1 = E, \quad (6)$$

$$E_2 = E(1 - da), \quad (7)$$

$$\nu_{12} = \nu, \quad (8)$$

$$\nu_{21} = \nu \frac{E_2}{E_1}. \quad (9)$$

The directional Young's moduli and Possion's ratio were determined such that E can still be separated from the problem allowing SIMP to be used. E_1 and E_2 were taken as the x and y directions respectively, as defined in figure 4. Figure 6 gives the results for both systems at the four different anisotropy levels. A comparison is made between isotropic solutions (in blue) and anisotropic solutions (in black) by overlapping the results. This overlapping allows not only topological differences but also the structural changes to be more apparent.

In system 1 the inner struts are split into multiple thin struts as the degree of anisotropy is increased. This allows the right most strut to take a shallower angle, orienting the strut toward the x direction. In system 2 the inner struts which are split near the outer support beam are joined so that a single

thicker strut can be used. Both systems vary greatly from the isotropic solution. This shows the importance of modeling the material well during the optimization process.

C. Layered Element Topology Optimization

For implementation of the layered mesh the adhesion layer thickness was varied while the relative Young's modulus was held at a constant 5% of the bulk Young's modulus. This value was exaggerated so that topological and structural changes are more apparent. The adhesion layer thickness was expressed as a percentage of the total element height, the values chosen were 10%, 20%, 30%, and 40%. Figure 7 gives the results for both systems using layered elements. A comparison is made between the layered mesh solutions (in black) and the isotropic solutions (in blue) by overlapping the solutions.

The topological and structural changes are much more consistent, and a few trends can be noted. Introducing the adhesion layer elements shortens and thickens the struts oriented in the y direction. This can be seen in all cases but is most noticeable in system 2 with 40% adhesion layer thickness, where the bottom support does not extend to the bottom of the design space allowing the struts to be thicker. System 1 is structurally rearranged compared with the isotropic case. This can be seen best in 40% adhesion layer thickness solution, where the top strut is extended and is

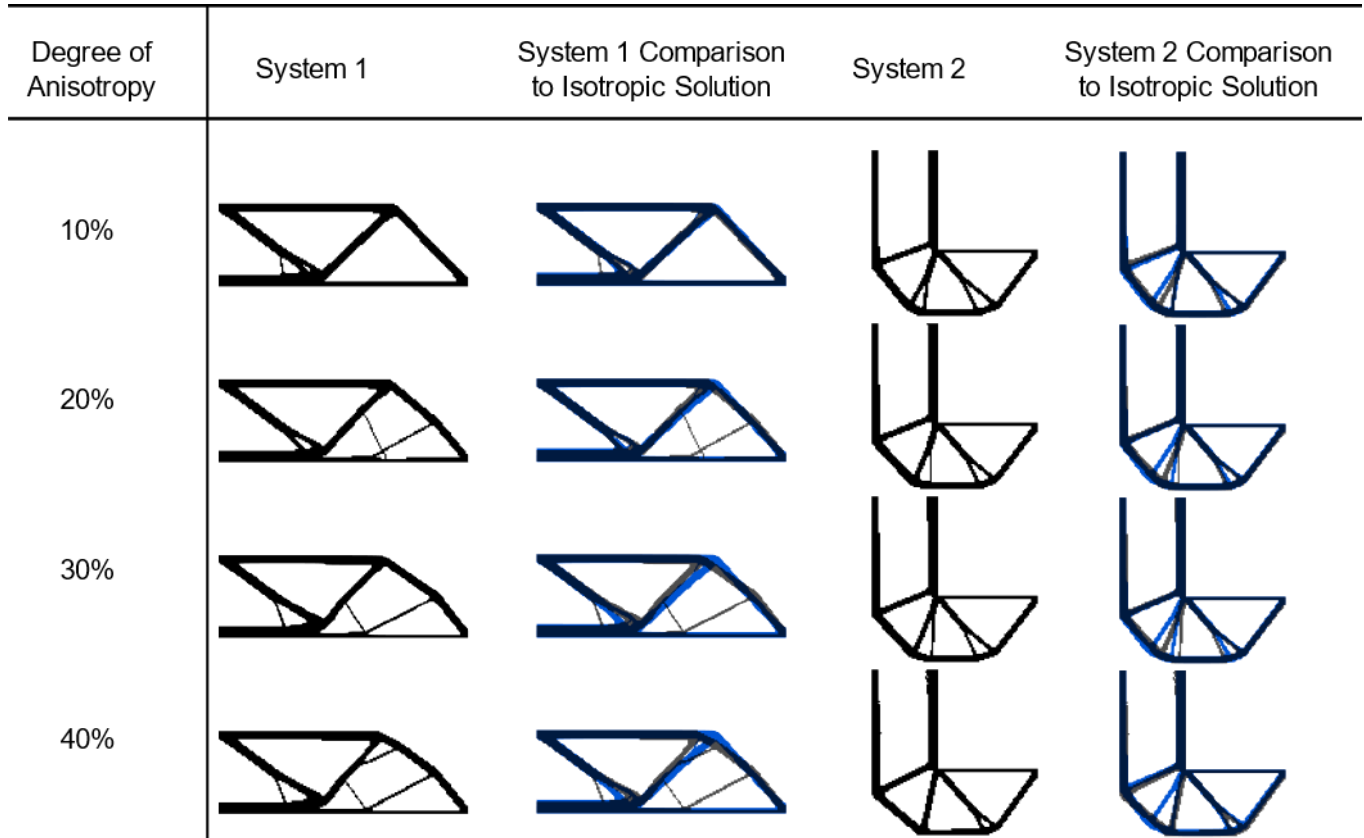


Figure 6 Comparison of the effect that varying degrees of anisotropy have on topology optimization solutions. The solutions for 10%, 20%, 30% and 40% degrees of anisotropy are presented for both system 1 and system 2. There is also a column comparing anisotropic to isotropic solutions. In these comparisons the isotropic solution is blue and anisotropic is black. The two solutions are overlapped so that topological and structural changes become easier to identify.

angled such that the vertical strut can be made shorter. This rearrangement also allows the two inner struts to take shallower angles allowing more of the load to be transferred in the x direction

D. Discussion of Anisotropic vs Layered Elements

There are some major differences between using anisotropic and adhesion layer elements. In system 2 the number of inner struts is increased when using layered elements and the number of inner struts decreases when using anisotropic elements. For system 1 the two different element types resulted in shorter and thicker struts as the degree of anisotropy or adhesion layer thickness was increased. In system 1 both element types cause a reorientation of the inner struts so that more load is transferred in the x direction. However, the topology is quite different when comparing the element types. The adhesion layer elements resulted in thicker support beams and anisotropic elements used thinner support beams for the reorientation. Overall there are advantages and disadvantages to using either element type. The adhesion layer elements describe the physical system better than both the isotropic and anisotropic elements but use twice the amount of CPU time. Anisotropic elements give a better description of the system compared with isotropic elements, without the extra CPU time and without the ability to describe stochastic anisotropy.

E. Robust Layered Element Topology Optimization

To show how the layered mesh enables new description of stochastic nature of FDM a Monte-Carlo simulation was performed on both systems. For generating each simulated system an independent random Gaussian variable was used to represent the relative adhesion layer Young's modulus. Some of the common problems associated with FDM that this can be used to model are: voids or poor fusion between layers, misalignment of layers and varying layer thicknesses. All these random properties can also be directly measured. For example, misalignment of layers, which can be measured using surface roughness, is modeled very well by a weaker Young's modulus. When layers are misaligned, the effective cross-sectional area decreases, which is proportional to the Young's modulus. As discussed previously the stochastic anisotropy cannot be modeled using anisotropic elements without very complex descriptions of the sensitivity. The effect that introducing a stochastic Young's modulus in the adhesion layer was explored for both system 1 and system 2 and the results are shown in figure 8. For these results adhesion layer thickness is taken as 5% of the element, the mean relative Young's modulus was taken as 50% of the bulk Young's modulus, and the standard deviation was taken as 10% of the bulk Young's modulus. It is important to note that the minimum Young's modulus value (taken at three standard deviations) for the random adhesion layer is greater than the highest Young's modulus value used for the isotropic layered element simulations.

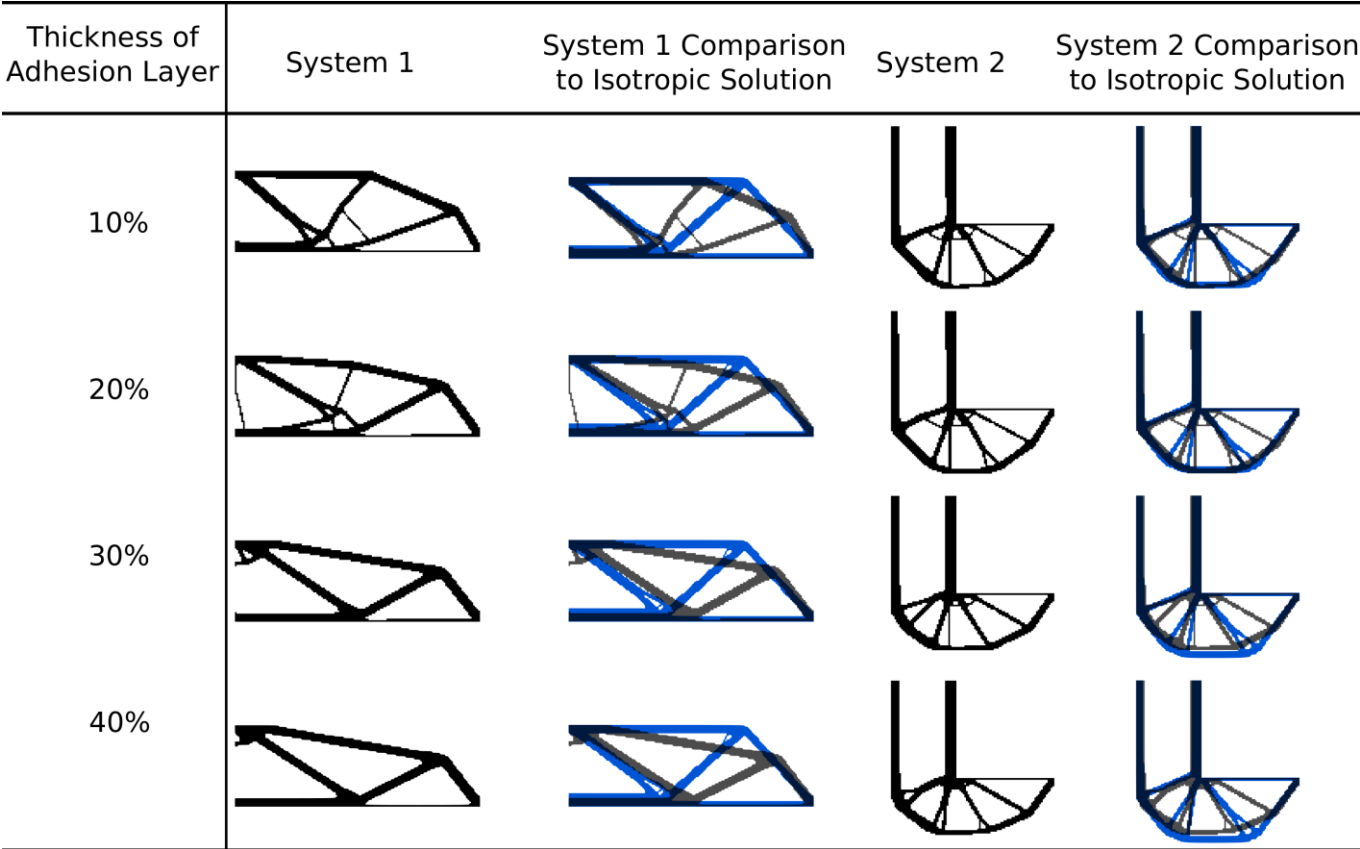


Figure 7 Comparison on the effect that varying degrees of adhesion layer thickness have on topology optimization solutions. The solutions for 10%, 20%, 30% and 40% adhesion layer thickness are presented for both system 1 and system 2. There is also a column comparing anisotropic to isotropic solutions, in these comparisons the isotropic solution is in blue and anisotropic is in black. The two solutions are overlapped so that topological and structural changes become easier to interpret.

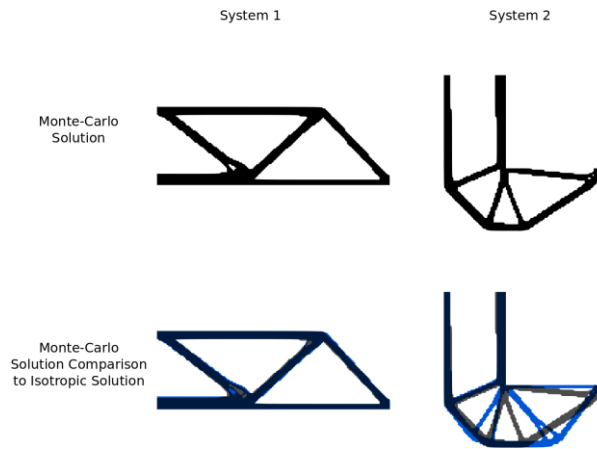


Figure 8 Monte-Carlo solution for both system 1 and 2 using a mean of 50% relative Young's modulus and a standard deviation of 10% relative Young's Modulus. A comparison is also made to the isotropic solution (shown in blue) by overlapping the two results.

Using a random variable to describe the Young's Modulus of the adhesion layer provided very different results for system 1 and 2. System 1 had no major difference topologically or structurally when compared with the isotropic solution. However there is a difference when comparing to the uniformly weak adhesion layer solutions. In the uniformly weak adhesion layer solutions the number of inner struts was increased, greatly contrasting the removal of inner struts found with the introduction of a stochastic Young's modulus. System 2 resulted in a more simplified structure with some of the support beams being removed when compared to the isotropic solution. Interestingly the Monte-Carlo solutions are more topologically different compared with the uniformly weak adhesion layer solutions or with the isotropic solution. In the uniformly weak adhesion layer element solutions there was a reorientation of the support beams not found in the Monte-Carlo solution. There is no obvious reason why system 1 does not differ from the isotropic solution while system 2 does. These results show the importance of including the stochastic properties present in the system because it can greatly affect the topology of the solution.

IV. CONCLUSIONS

In the case of modeling FDM for topology optimization the use of adhesion layer elements is better than using anisotropic elements for accuracy of the system description and it has been shown that this increased accuracy results in different topologies. Adhesion layer elements also allow for the randomness associated with FDM to be more accurately modeled. If planning to use FDM or other additive

manufacturing methods for fabrication of topology optimization solutions the layering effect on material properties should be taken into consideration during the optimization process.

REFERENCES

- [1] X. Wang, S. Xu, S. Zhou, W. Xu, M. Leary, P. Choong, M. Qian, M. Brandt, and Y. Xie, "Topological design and additive manufacturing of porous metals for bone scaffolds and orthopedic implants: A review," *Biomaterials*, vol. 83, pp. 127-141, 2016.
- [2] G. Kim and Y. Oh, "A benchmark study on rapid prototyping processes and machines: quantitative comparisons of mechanical properties, accuracy, roughness, speed, and material cost," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 222, no. 2, pp. 201-215, 2008.
- [3] A. K. Sood, R. K. Ohdar, and S. S. Mahapatra, "Experimental investigation and empirical modelling of fdm process for compressive strength improvement," *Journal of Advanced Research*, vol. 3, no. 1, pp. 81-90, 2012.
- [4] B. Mueller, "Additive manufacturing technologies-rapid prototyping to direct digital manufacturing," *Assembly Automation*, vol. 32, no. 2, 2012.
- [5] C. Bellehumeur, L. Li, Q. Sun, and P. Gu, "Modeling of bond formation between polymer filaments in the fused deposition modeling process," *Journal of Manufacturing Processes*, vol. 6, no. 2, pp. 170-178, 2004.
- [6] S. Rahmati and E. Vahabli, "Evaluation of analytical modeling for improvement of surface roughness of fdm test part using measurement results," *The International Journal of Advanced Manufacturing Technology*, vol. 79, no. 5-8, pp. 823-829, 2015.
- [7] G. Budzik, "Geometric accuracy of aircraft engine blade models constructed by means of the generative rapid prototyping methods fdm and sla," *Advances in Manufacturing Science and Technology*, vol. 34, no. 1, pp. 33-43, 2010.
- [8] O. Sigmund, "A 99 line topology optimization code written in Matlab," *Structural Multidisciplinary Optimization*, no. 21, pp. 120-127, 2001.
- [9] E. Andreassen, A. Clausen, M. Schevenels, B. Lazarov, and O. Sigmund, "Efficient topology optimization in Matlab using 88 lines of code," *Structural Multidisciplinary Optimization*, vol. 43, no. 1, pp. 1-16, 2011.
- [10] O. Sigmund and J. Petersson, "Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima," *Structural optimization*, vol. 16, no. 1, pp. 68-75, 1998.
- [11] O. Sigmund, "Morphology-based black and white filters for topology optimization," *Structural and Multidisciplinary Optimization*, vol. 33, no. 4-5, pp. 401-424, 2007.
- [12] A. Asadpoure, M. Tootkaboni, and J. Guest, "Robust topology optimization of structures with uncertainties in stiffness - application to truss structure," *Computers and Structures*, no. 89, pp. 1131-1141, 2011.
- [13] K. Svanberg, "MMA and GCMMA - two methods for nonlinear optimization," tech. rep., KTH Royal Institute of Technology, Sweden, 2007.
- [14] K. Svanberg, "The method of moving asymptotes - a new method for structural optimization," *International Journal for Numerical Methods in Engineering*, vol. 24, pp. 359-373, 2002.
- [15] K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," *SIAM Journal of Optimization*, vol. 12, pp. 555-573, 2002.

Thermal distribution of a stainless steel 316 body-centered cube packing particle at sintering temperatures

Carlos Abel Rojas Dorantes
Department of Mechanical Engineering
York University
Toronto, Canada
carlosr1@yorku.ca

Aleksander Czekanski
Department of Mechanical Engineering
York University
Toronto, Canada
alex.czekanski@lassonde.yorku.ca

Abstract— Additive manufacturing, or 3D printing, is gaining in popularity across many industries. Despite the advantages of 3D printing, the properties of manufactured parts differ from those of parts produced by more traditional means. The primary reason for this difference is that additive manufacturing is based on building a part layer by layer. In fact, nonuniform connections between layers result in a decrement in the values of the ideal properties. In the powder bed fusion method of 3D printing, the interaction between the powder particles is a critical parameter determining the reliability of the part. A uniform exchange of heat is required for the part to have higher mechanical strength. In this study, we simulated different packing densities in a face-centered cube arrangement. Higher packing densities and areas of contact resulted in faster changes in temperature. Furthermore, minor changes in packing density had a significant effect on the time to reach the sintering and melting temperature.

Keywords- *component; additive manufacturing; powder bed; packing density; sintering;*

I. INTRODUCTION

Additive manufacturing, or 3D printing, is a manufacturing technique that produces objects through the layer-by-layer deposition of material by a 3D printer according to a design defined using CAD software. Recently, the number of metallic materials available for 3D printing have multiplied, giving users the ability to manufacture different types of parts; nevertheless, library of metallic materials is still limited. In addition, manufactured prototypes can be produced for running tests and analysis, and to apply them to the real world as well.. Even if the initial cost is greater, adding features to the internal geometry of a 3D printed part can improve its final design [1].

In the field of aerospace engineering, printed parts are gaining in popularity. Reducing the weight of an aircraft by using different internal structures leads to improved aircraft performance [2]. For instance, using 3d printing technology it is possible to reduce the number of parts to assembly one object,

such as a fuel nuzzle developed by GE which replace a 20 parts whit only one. Moreover, the GE fuel nozzle is 25% lighter[3].

Additive manufacturing is not only used to produce prototypes but also for industrial applications and many fields where custom models are required, such as in medicine. A major challenge in medicine is that all patients are different and thus require customization solutions, a problem that 3D printing is intrinsically able to address. For instance, for many years researchers have been working on creating 3D parts for use in prostheses [4]. However, 3D printing technology can be used for more than producing mechanical elements for prosthetic implants. 3D printed parts have been used to craniofacial reconstruction, porosity in printed parts is a positive factor because it allows bone ingrowth [5].

While 3D printing offers many unique advantages, its main limitation continues to be the mechanical and metallurgical properties of the fabricated part. Other common inconveniences related to 3D printing are the cost of the material as well as the production time. Once manufactured, fabricated parts still need to be postprocessed to improve their properties. Furthermore, to change some materials, hardware and software themselves must be modified. Hence, additional studies are required to increase the number of materials, reduce cost, and improve the properties of the fabricated part. For these reasons, a number of materials for use in 3D printing is still in development[6].

The powder bed fusion (PBF) method of 3D printing consists of sintering layers of powder in a defined geometry, adding new layers above the sintered ones. A laser beam is used to sinter the layers of powder, with the laser's energy being absorbed by powder particles in the top layer. After sintering one layer, the recipient moves down, and using a roller, more layers of raw material are added to be sintered in the laser's path. This process is repeated until the full model is done[7]. However, the process is not limited to sinter particles, but in some conditions with higher energy melting point is reached.

A few manufacturing techniques are based on the PBF method, which uses low-power lasers to sinter particles of high-power beams to melt the powder particles.

A common technique for sintering is known as selective laser sintering (SLS), while common melting processes are called selective laser melting (SLM) and electro-beam melting (EBM) [8]. SLS systems are also able to melt the powder completely depending on exposure time and temperature.

Unfortunately, as with other additive manufacturing techniques, printed parts made using SLS are susceptible to having defects that impact the mechanical properties of the fabricated part. To address some of those problems, postprocessing is commonly used, which requires material removal as part of the finishing process. However, if the part cannot meet the required specifications, the 3D printing process would be unsuccessful, thus the resources would be wasted. For that reason, it is critical to define the parameters of the process to achieve successful printing [8].

The quality properties that are controlled by process parameters [7] are as follows:

- Surface temperature
- Residual stress
- Geometry verification
- Dimensional accuracy
- Surface quality
- Mechanical properties
- Porosity
- Density

The properties of the powder are a significant factor to control the quality of printed parts. Powder solidification depends on most of the parameters. It is not only the properties of the powder that must be considered, but also the properties of the reused powder. For that reason, it is crucial to study each part of the process [8]. Studying the temperature distribution in the different particle densities is vital to determine the properties of the layers of powder. Obtaining reliable results about the possible temperature distributions will help to improve the behavior of the materials and the accuracy of the PBF method.

A uniform powder arrangement is required for the process to be successful. The different methods of producing powders result in different properties. In the atomized process, gas is used to generate more spherical particles; however, particles atomized using water have a wider size distribution as well as varied shapes. Moreover, some research findings indicate that parts manufactured with gas-atomized powders have a higher density [9]. There are different arrangements of powder particles in a powder bed, which can be represented in unit cells, which have different relative densities [10].

II. METHOD

Using a face-centered cube geometry, with a radius particle of 21 μm , the initial volume of the geometry represents 68% of the entire cube. Initially, the contact area is geometrically

correct. The area of contact is established as the exact point where the spheres are touching. For that reason, the area of contact must be changed to contrast the exchange of temperature in the representative geometries when the area of contact increases.

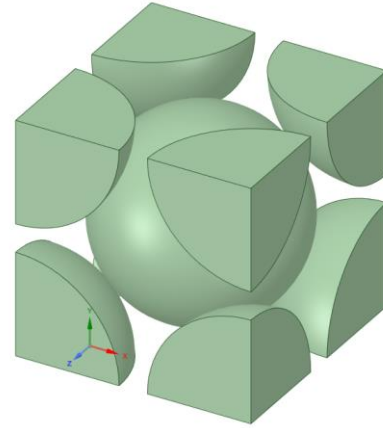


Figure 1. Cubic equivalent of face-centered powder particles.

The dimensions of the packing cube, shown in Figure 1, are $50 \mu\text{m} \times 50 \mu\text{m} \times 50 \mu\text{m}$, resulting in a volume of $1.25 \times 10^{-13} \text{ m}^3$. A temperature corresponding to the melting temperature of stainless steel 316 is applied to the top of the surface. The melting temperature is $1370\text{--}1400^\circ\text{C}$, but the sintering temperature is lower. Some authors recommends a sintering temperature around 60% or 70% of the melting temperature. However, the density of the final part will vary — to achieve a density of stainless steel $>80\%$, a sintering temperature of $\sim 82\%$ is recommended [11].

The initial packing density is 68% of the full cube, but to increase the area of contact, the proportion is raised, increasing the radius in steps of 1% until 110% (Figure 2) packing density is reached, when the portion of the cube's volume is 77.76% (Table 1). In fact, in the real process a roller is in charge to compact the powder bed. By increasing the area of contact, it is expected that the time required to reach the melting point will be reduced. The thermal conductivity is higher when the heat moves in a conduction mechanism.

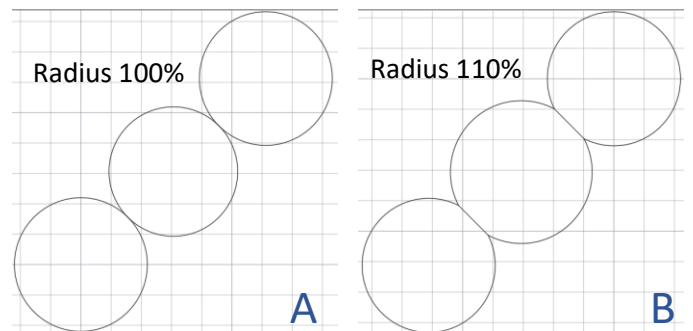


Figure 2, Examples of changes in contact area in two-dimensional representation of the particles. In the picture A, an ideal arrangement of 3 particles is shown. On the other hand, in picture B the radius was increased

TABLE 1. INCREMENT IN PARTICLE RADIUS AND ITS EFFECTS ON VOLUME.

Radius (10^{-6} m)	Equivalent radius	Total volume (10^{-14} m ³)	Contrast to initial volume	Packing density	Contact area .3 (10^{-12} m ²)	Contact %
21.303	100.01%	8.5095	1.000129284	0.68076	0.146845582	0.0026
21.513	101%	8.629	101%	69.03%	14.43634534	0.2482
21.726	102%	8.7504	103%	70.00%	29.00371935	0.4890
21.939	103%	8.8717	104%	70.97%	43.69691017	0.7224
22.152	104%	8.993	106%	71.94%	58.43298515	0.9476
22.365	105%	9.1143	107%	72.91%	73.23673999	1.1651
22.578	106%	9.2356	109%	73.88%	88.10579122	1.3754
22.791	107%	9.3585	110%	74.87%	103.0377315	1.5786
23.004	108%	9.4779	111%	75.82%	118.0300305	1.7749
23.217	109%	9.5991	113%	76.79%	133.0815949	1.9647
23.43	110%	9.7202	114%	77.76%	148.1891484	2.1481

The following is a list of assumptions we made about this process:

- The original packing structure is modified in order to increase contact areas. Perfect cube arrangement was
- not part of the study because of the impact infinitesimal point contrasted to other results. Initial
- cube was set with a radius of 100.01% of the perfect geometry. Random sizes of powder or different shapes were not part of the simulation.
- Maximum temperature was fixed to 1400°C, increasing the temperature the results of the process change, as well as pre-heating process can decrease sintering and melting times. Sintered temperature (80% of melting temperature) was established as 1120°C. Initial temperature was set as 25°C
- The dimensions of the packing cube are fixed; 50 microns x 50 microns x 50 microns having a volume of $1.25 \times 10^{-13} \text{ m}^3$. Only a unit cell was considered, thus

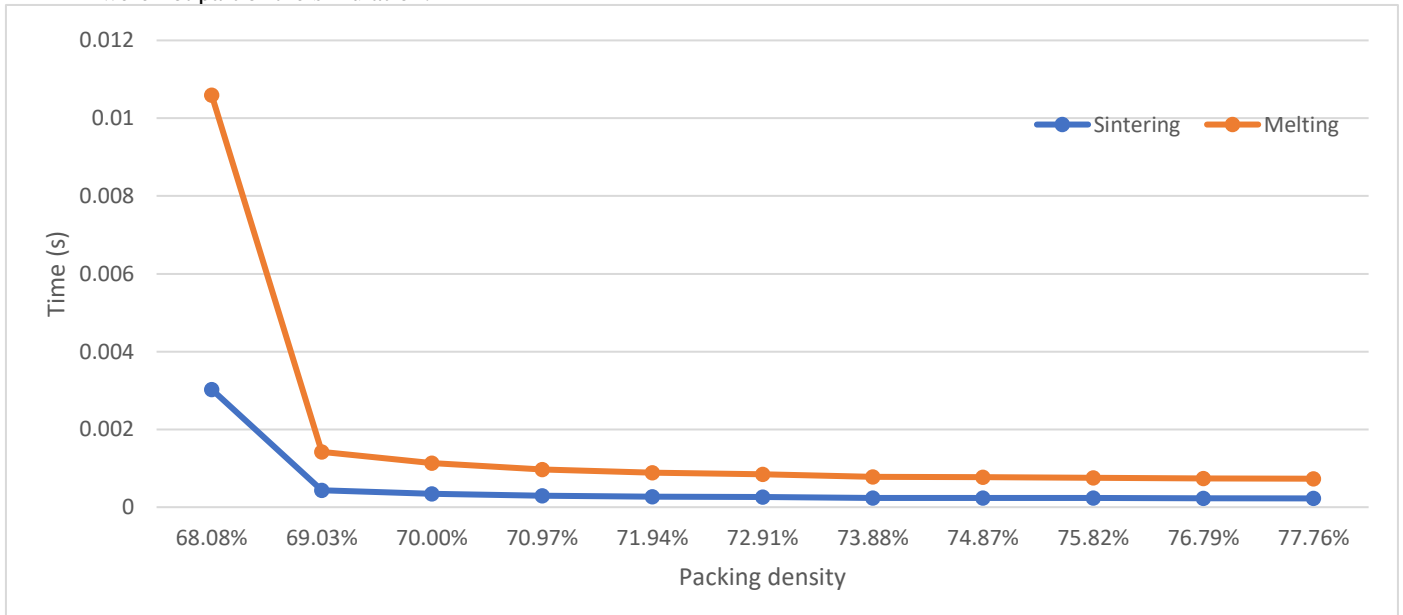


Figure 3. Increasing the packing density (x-axis) drastically decreases melting and sintering time (y-axis).

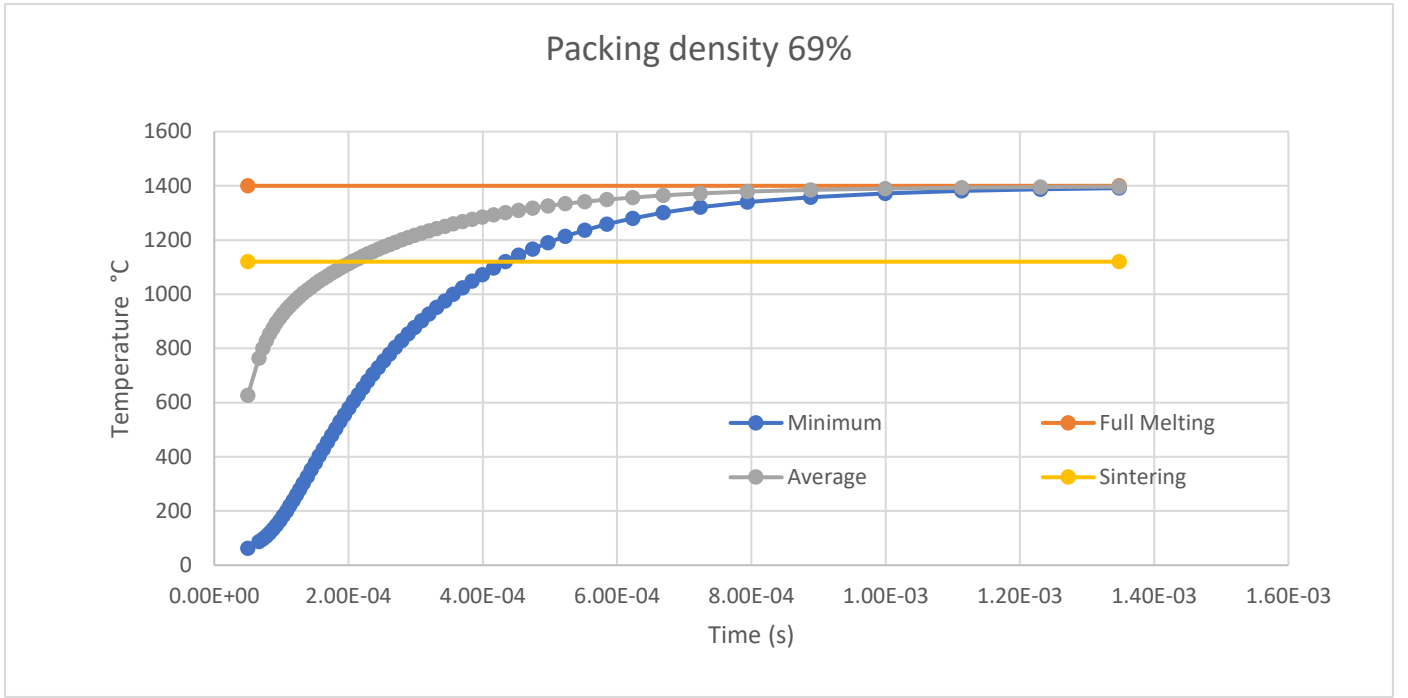


Figure 4. Temperature increasing with time at a packing density of 69%. interaction with other particles was not part of the simulation. More complex studies required more computational time and special hardware [12]. Thermal behaviors were not considered

III. RESULTS

The result was obtained by contrasting the difference in temperature when the area of contact between particles was increased simulating the compression process.

The first result, which has a minimum area of contact, required 0.003 s to reach the melting point. Other packing densities led to significantly less time to reach the melting point (Figure 3). Less time was also required to reach the sintering point; for instance, at a packing density of 77.76% (Figure 5). However, real contact points are bigger than theoretical structures, for that reason is better to compare the subsequent points starting in a packing density of 69%. There is an outstanding difference when comparing 68% and 69%, Figure 3, Table 2. Even more, when 69% and 70% are contrasted, the difference is not excessive.

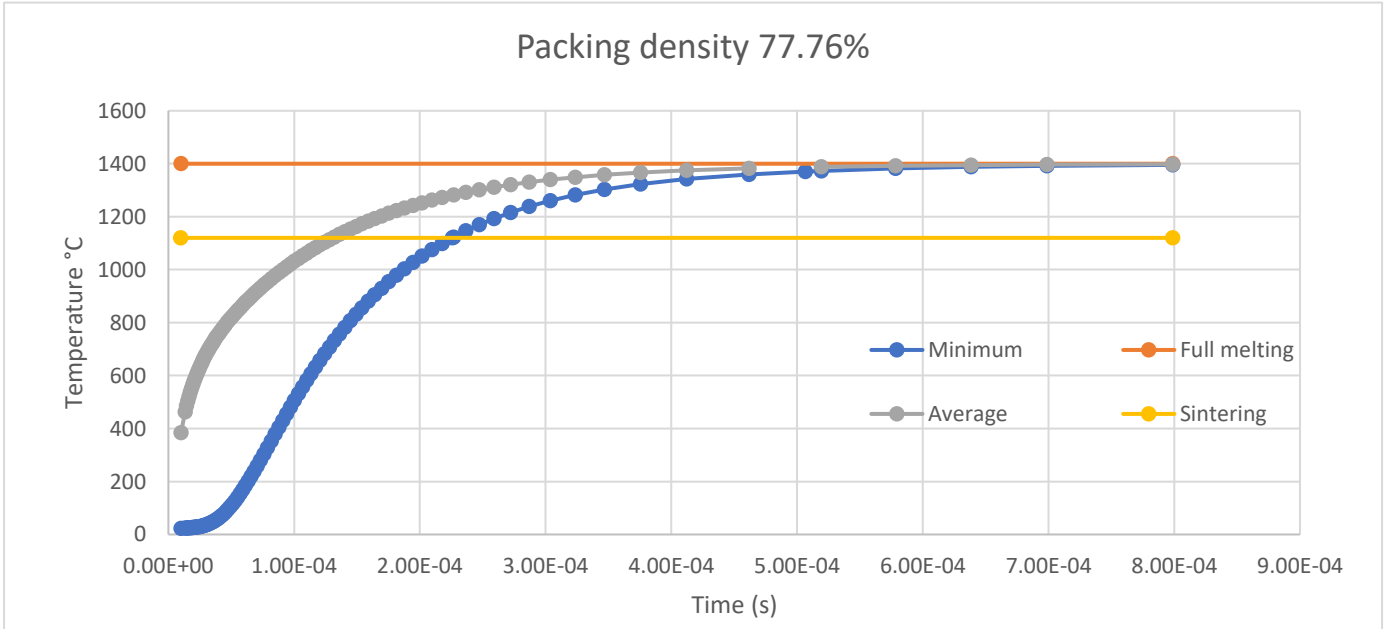


Figure 5. Temperature increases with time at a packing density of 77.76%.

Furthermore, after 69% the time to reach the melting and sintering point continued decreasing with increasing packing density following a trend. Moreover, when time to reach the sintering point is contrasted with time to reach the melting temperature, the trend is similar. Comparing the results from 77.76%, the time required to reach the sintering point is considerably lower.

Figure 4 and 5, are examples to compare the average temperature of the particles as well as the minimum temperature which is located in the bottom of the cubic cell.

IV. DISCUSSION

Even if the real powder is not 100% spherical and the area of contact is larger, it is remarkable to contrast the effect of the contact area under ideal circumstances. The results obtained support the importance of packing density as a parameter in 3D printing. Compacting processes in PBF are crucial for effective manufacturing.

Some experiments using a thermal camera and a multilayer process [13] show that the results are lower than the simulated values obtained using numerical calculations. The top temperatures in the process, were measured between 1626.85°C and 1526.85°C, and subsequent scans measured less than 1326.85°C.

Other experiments confirms the importance of the sintering temperature. [14]. Powder particles which received more heat, formed parts with higher density and much more tensile strength.

To conclude, this study demonstrate the importance of the powder compacting process. Thermal conductivity is improved when the contact area and packing density increase. Benefits of higher thermal conductivity are not limited to a scanning speed ratio but also to improve properties of the parts.

REFERENCES

- [1] Bandyopadhyay, A., Bose, S., & ENGnetBase - Additive manufacturing. CRC Press.
- [2] Wong, K. V., & Hernandez, A. (2012). A review of additive manufacturing. ISRN Mechanical Engineering, 2012.
- [3] <https://www.ge.com/additive/blog/new-manufacturing-milestone-30000-additive-fuel-nozzles>
- [4] Ventola C. L. (2014). Medical Applications for 3D Printing: Current and Projected Uses. P & T: a peer-reviewed journal for formulary management, 39(10), 704-11.
- [5] Sunpreet Singh, Seeram Ramakrishna, Rupinder Singh, Material issues in additive manufacturing: A review, Journal of Manufacturing Processes, Volume 25, 2017, Pages 185-200, ISSN 1526-6125, <https://doi.org/10.1016/j.jmapro.2016.11.006>.
- [6] Parthasarathy J. 3D modeling, custom implants and its future perspectives in craniofacial surgery. Ann Maxillofac Surg. 2014;4(1):9-18.
- [7] Mani, M., Feng, S., Lane, B., Donmez, A., Moylan, S., & Fesperman, R. (2015). Measurement science needs for real-time control of additive manufacturing powder bed fusion processes.
- [8] Gillespie, L., & Access Engineering. Design for Advanced Manufacturing: Technologies, and Processes. McGraw-Hill Education.
- [9] Li, R., Shi, Y., Wang, Z., Wang, L., Liu, J., & Jiang, W. (2010). Densification behavior of gas and water atomized 316L stainless steel powder during selective laser melting. Applied surface science, 256(13), 4350-4356.
- [10] de Moraes, Diego & Czekanski, Aleksander. (2018). Parametric Thermal FE Analysis on the Laser Power Input and Powder Effective Thermal Conductivity during Selective Laser Melting of SS304L. Journal of Manufacturing and Materials Processing. 2. 47. 10.3390/jmmp2030047.
- [11] Klar, E., & Samal, P. K. (2007). Powder metallurgy stainless steels: processing, microstructures, and properties. ASM international.
- [12] Johannes A. Koepf, Martin R. Gotterbarm, Matthias Markl, Carolin Körner, 3D multi-layer grain structure simulation of powder bed fusion additive manufacturing, Acta Materialia, Volume 152, 2018, Pages 119-126, ISSN 1359-6454, <https://doi.org/10.1016/j.actamat.2018.04.030>.
- [13] Chivel, Y. (2013). Optical in-process temperature monitoring of selective laser melting. Physics Procedia, 41, 904-910.
- [14] Gusarov, A & Yadroitsev, Igor & Bertrand, Ph & Smurov, I. (2009). Model of Radiation and Heat Transfer in Laser-Powder Interaction Zone at Selective Laser Melting. Journal of Heat Transfer. 131. 072101. 10.1115/1.31

TABLE 2. CONTRASTING DIFFERENT PACKING DENSITIES WITH MELTING AND SINTERING TIMES.

Packing density	Sintering time (s)	Melting time (s)	Proportional sintering time	Proportional melting time
68.08%	0.00303	0.007567	100.00%	100.00%
69.03%	0.000434	0.000985	14.31%	13.01%
70.00%	0.000348	0.000787	11.48%	10.41%
70.97%	0.000296	0.00067	9.78%	8.86%
71.94%	0.000272	0.000616	8.99%	8.14%
72.91%	0.000259	0.000585	8.56%	7.73%
73.88%	0.000241	0.000542	7.97%	7.16%
74.87%	0.000237	0.000535	7.83%	7.07%
75.82%	0.000235	0.000523	7.76%	6.91%
76.79%	0.000228	0.000509	7.53%	6.73%
77.76%	0.000226	0.000507	7.45%	6.69%

Post-Printing Heat Treatment of a Wire Arc Additively Manufactured Low Carbon Low Alloy Steel

Alireza Vahedi Nemani, Mahya Ghaffari, Mehran Rafieazad, Ali Nasiri

Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL, Canada, A1B 3X5
avahedineman@mun.ca

Abstract—Wire arc additive manufacturing (WAAM) is a novel technology capable of producing large-scale engineering components with medium to simple geometries. Analogous to other additive manufacturing processes, WAAM induce a considerable anisotropy in mechanical properties of the fabricated part, which can be potentially minimized either by optimizing processes parameters or adopting appropriate post heat treatment processes. In this study, the effects of two heat treatment cycles, including hardening (austenitizing followed by water quenching) and normalizing (austenitizing followed by still-air cooling) on the microstructure and mechanical properties of a WAAM fabricated low carbon low alloy steel (ER70S-6) are studied. The microstructural analysis was carried out using optical microscopy (OM) and field emission scanning electron microscopy (FESEM). The microstructure in the melt pools of the as-printed sample contains low volume fraction of lamellar pearlite formed along the grain boundaries of polygonal ferrite as the predominant constituents. Heat affected zone (HAZ) grain coarsening was also detected at the periphery of each melt pool boundary, as an adverse effect of the thermal cycle associated with each layer on its previously deposited ones, leading to a noticeable microstructural inhomogeneity in the as-fabricated sample. In order to modify the nonuniformity of the microstructure, a normalizing heat treatment process was employed to promote a homogenous microstructure with uniform grain size throughout the melt pools and HAZs. Differently, the hardening heat treatment process contributed to the formation of two non-equilibrium micro-constituents, *i.e.* acicular ferrite and bainite, primarily adjacent to the lamellar pearlite phase. Mechanical properties of the as-printed and heat-treated components were also studied at different orientations relative to the building direction using the Vickers microhardness method and the uniaxial tensile testing. The results of microhardness testing revealed that the normalizing heat treatment slightly decreases the microhardness of the as-printed sample from 160 HV to 154 HV; however, the formation of non-equilibrium phases during hardening process significantly increased the microhardness of the component up to 260 HV. Tensile testing of the as-printed part in the building (vertical) and deposition

(horizontal) directions revealed an anisotropy in ductility. Although normalizing heat treatment did not contribute to the tensile strength improvement of the component, it suppressed the observed anisotropy in ductility. On the contrary, the hardening heat treatment raised the tensile strength, but intensified the anisotropic behaviour of the component.

Keywords- WAAM; ER70S-6; anisotropy; microstructure; mechanical properties; heat treatment

I. INTRODUCTION

Wire arc additive manufacturing (WAAM) is a novel technology capable of producing metallic components utilizing an arc welding process to additively fabricate an engineering part [1]. Different from the metal powder-based additive manufacturing processes, such as direct metal laser sintering (DMLS) and selective laser melting (SLM), wire arc additive manufacturing uses a consumable metallic wire as the feedstock material [2]. In WAAM, the entire consumable wire is continuously fed into an adopted electric arc or plasma, leading to an extremely high deposition rate as compared to that of the powder-based AM systems [3]. Therefore, wire-based systems are generally suitable for producing large-scale components with less complexity in geometry and design, in contrast to the powder-bed systems, which typically fabricate small and high-definition parts [4]. From another perspective, powder-bed additive manufacturing techniques are limited to a build envelope, but in wire-based systems, the torch is usually mounted on a robotic arm having more freedom of movement, implying that the component's size is not confined by a chamber.

During wire arc additive manufacturing process, the feedstock material is melted and deposited in the form of weld beads layer by layer on the previously solidified tracks. As the consecutive layers fuse into the previous ones, the material is built up until the near-net-shape component is completed. Since the process is involved with sequential melting and solidification, each region of the component is subjected to periodic fast heating and cooling cycles by the deposition of upper layers. Such complex localized thermal cycles lead to

heterogeneous and anisotropic microstructures, and mechanical properties in the AM fabricated components [5][6]. This is one of the main drawbacks of the WAAM technique as compared to the conventional methods of manufacturing. Suryakumar *et al.* [7] studied the mechanical properties of a mild steel built through hybrid layered manufacturing (HLM) and reported anisotropic tensile properties in different locations including torch, stepover and vertical directions. They concluded that the anisotropy in mechanical properties is related to the number of heating cycles experienced by each layer of the component [7]. According to their experimental and thermal analysis data, the thermal effect of each deposited layer affects only 5 previous layers, which means that the thermal history of the final 5 layers at the top of the part defers from the other layers [7]. Haden *et al.* [8] also investigated wire arc additive manufacturing of 304 stainless steel and reported graded wear and hardness properties in both directions of deposition and building. Their findings showed that this anisotropy is due to the fluctuation in localized thermal histories, consequently leading to the formation of a variety of microstructures, from austenitic to solidification structures owning mixed ferrite morphology with abrupt texture changes at different regions of the sample [8].

In addition to the microstructural inhomogeneity, the formation of internal defects between the deposited layers as a result of high heat removal capacity from the inter-pass regions may deteriorate the mechanical properties of the additively manufactured components [9][2][10]. The inter-pass defects being formed in the fusion boundaries commonly include lack of fusion, entrapped gas, and porosities [11][12]. Such defects can readily facilitate the brittle fracture by providing easy potential sites for crack initiation and growth since they can act as a strong stress concentrator during tensile loading [13]. Among these defects, the lack of fusion is of particular interest and highly probable to form during multi-pass deposition based processes, such as WAAM, resulted from the special heat transfer condition between the layers in the sample along the building direction [9].

The anisotropy in microstructure and mechanical properties can be minimized or even be eliminated either by optimizing the process parameters or adopting appropriate post-fabrication heat treatment processes [10][11]. Yang *et al.* [14] investigated the influence of different heat treatment cycles on the anisotropy in microstructure and mechanical properties of A357 aluminum alloy fabricated by selective laser melting. They concluded that a long-time solution heat treatment on the as-printed component generates a homogenous microstructure containing evenly distributed Si precipitates, leading to the elimination of the anisotropy in both yield strength and ductility [14].

In this study, with the aim of homogenizing the microstructure and diminishing the induced anisotropy in an as-printed WAAM-ER70S-6 low carbon low alloy steel part, two heat treatment cycles, including normalizing (austenitizing followed by still-air cooling) and hardening (austenitizing followed by water quenching), were conducted on the as-printed samples. Microstructural and mechanical properties

characterizations were carried out on both as-printed and heat-treated samples in different orientations, including deposition (horizontal) direction and building (vertical) direction.

II. EXPERIMENTAL PROCEDURE

A. Fabrication Process

In the present study, a thin wall of low carbon low alloy steel (ER70S-6) was fabricated using the wire arc additive manufacturing method utilizing a Lincoln Electric GMA machine with a torch mounted on a 6-axis Fanuc robot as the power source. To minimize the heat input of the WAAM process and be able to adjust the heat independent from the wire feed speed, an advanced current controlled surface tension transfer (STT) process was employed for fabrication. Utilizing the STT can further contribute to reducing the surface irregularities and spattering during the building process [15]. In order to smoothly feed the wire to the melt pool, the stand-off distance was held constant at 3 mm between the tip of the filler wire and the surface of the previous layer. Fig. 1 schematically represents the set-up of the WAAM process.

The nominal chemical composition of the filler wire and the WAAM process parameters are listed in Table 1 and Table 2, respectively. ASTM A36 mild steel with a thickness of 12 mm was used as the substrate, which was attentively wire brushed and then degreased by acetone to prevent contamination of the melt pools and the formation of gas pores during the solidification process. The whole part contained 50 consecutive layers, and each layer was comprised of six beads with a length of 135 mm and 3 mm center-to-center overlap, resulting in a wall with the a total width of 22 mm and a height of 150 mm. Two heat treatment cycles were applied to the as-printed component, including (i) normalizing (austenitizing followed by still-air cooling), and (ii) hardening (austenitizing followed by water quenching). For initial austenitizing in both cycles, the samples were heated up to 900 °C for 1 hour. The purpose of the normalizing process was to homogenize the microstructure by producing a uniform grain size along the melt pools, fusion boundaries, and heat affected zones. The intention of the hardening heat treatment was also to increase the hardness and tensile strength of the component.

B. Microstructural Characterization

To prepare the samples for microstructural characterizations, a Tegramin-3 Struers auto-grinder/polisher was employed, then the samples were etched chemically using a 5 vol.% Nital reagent for 20 s as suggested by ASM Metals Handbook [16]. The microstructure of the fabricated component was characterized at different regions from the bottom to the top of the wall to detect any microstructural changes throughout the whole part. To perform the microstructural characterization at different magnifications, an optical microscope (Nikon Eclipse 50i) and a field emission scanning electron microscope (FEI MLA 650F) were employed.

TABLE I. THE NOMINAL CHEMICAL COMPOSITION OF THE ER70S-6 FEEDSTOCK WIRE (WT. %)

C	Mn	Si	S	P	Cr	Ni	Mo	V	Cu
0.06-0.15	1.40-1.85	0.80-1.15	0.04 max	0.03 max	0.15 max	0.15 max	0.15 max	0.03 max	0.50 max

TABLE II. THE PROCESSING PARAMETERS IN WIRE ARC ADDITIVE MANUFACTURING OF THE LOW CARBON LOW ALLOY STEEL (ER70S-6)

Arc Current	Arc Voltage	Wire Feeding Rate	Scanning Rate	Argon Flow Rate
320 A	28 V	104 mm/s	5 mm/s	45 L/min

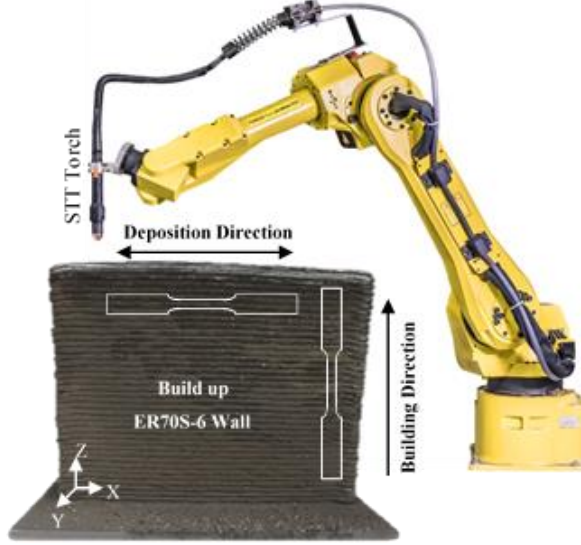


Figure 1. Schematic illustration of the WAAM process set-up.

C. Mechanical Properties Measurement

Microhardness distribution was measured and plotted along a line covering five successive layers through the building (vertical) direction on different zones including the center of the melt pools, fusion boundaries, and heat affected zones (HAZs), using a Buehler Micromet hardness test machine with the applied load of 300 g and an indentation time of 45 s. It should be noted that the reported data of microhardness are the average of five different measurements. The indentations were done on the polished and etched surfaces in order to distinguish the position of each indentation relative to the melt pool's geometry. Moreover, microhardness profiles were produced by subsequent indentations with 300 μm intervals (approximately five times more than the diagonal of each indent) to avoid the work hardening effect.

The tensile test specimens of as-printed and heat-treated samples were machined along both parallel and perpendicular to the building direction based on ASTM E8m-04 standard sub-size specimen [17] with dimensions of (100 mm \times 25 mm \times 5 mm). The room temperature uniaxial tensile tests were carried out using an Instron load frame equipped with an extensometer at the crosshead speed of 8 mm/min. Each tensile test was repeated five times under the same conditions to obtain a reliable average value.

III. RESULTS AND DISCUSSION

A. Microstructural Characterization

Fig. 2a illustrates a low magnification OM micrograph of the as-printed sample showing the transition from the center of the melt pool to the melt pool boundary and then the heat affected zone. The dominant microstructure in the center of the melt pool consists of a low volume fraction of lamellar pearlite (P) primarily formed at the grain boundaries of polygonal ferrite (PF) (Fig. 2b), in agreement with the results of a recent study on the WAAM of ER70S-6 [8]. Fig. 2c depicts the SEM micrograph taken from the melt pool boundary region (denoted as the fusion boundary shown in Fig. 2a) at higher magnification, revealing the formation of acicular ferrite (AF) and bainite (B) due to the faster cooling rate along the boundaries of each deposited bead as compared to its center. The aforementioned transition in the microstructure during 3D-printing of ER70S wire is also reported by Haselhuhn *et al.* [18]. In another investigation, Lee *et al.* [19] also studied the microstructure of the welded low carbon low alloy AH36 steel and similarly reported the formation of acicular ferrite and bainite near the fusion line.

It is well established that the presence of acicular ferrite and bainite constituents in the microstructure of steels can promote the mechanical properties of the component. This is primarily resulted from the finer structure of both phases, more uniform distribution of carbide and higher dislocation density and internal stresses in the bainite phase, contributing to a higher hardness/strength and ductility in the alloy [20][21][22]. However, it should be noted that since the volume fraction of acicular ferrite and bainite constituents are negligible as compared to the dominant ferritic and pearlitic microstructure of the alloy, the presence of acicular ferrite and bainite cannot have a significant contribution to the mechanical properties of the WAAM-ER70S sample. On the other hand, according to Fig. 2a, the microstructure of the HAZ consists of coarser grains of polygonal ferrite in comparison with the interior of the melt pool, as the thermal cycle associated with each depositing track stimulates the grain growth in the previously deposited bead. The grain coarsening in the HAZ can potentially cause a substantial softening in this region, consequently leading to a lower localized hardness and a reduced strength in a sample that accommodates this region [23]. Formation of such microstructural inhomogeneity from the center of the melt pool to its boundaries and to HAZ is attributed to the overlapping scanning strategy associated with multi-layer deposition nature of the WAAM process. Consequently, this process evokes various thermal cycles in different regions of the sample [24].

Post printing heat treatment is commonly being used to modify the microstructure and consequently the mechanical properties of an additively manufactured component [10][1].

Fig. 3a-d depict the microstructure of the WAAM-ER70S-6 wall after applying different heat treatment cycles, including normalizing (Figs. 3a and b) and hardening (Figs. 3c and d) at different magnifications. As indicated from Figs. 3a and b, the normalizing heat treatment from the austenitizing temperature of 900 °C, has not altered the pre-existing constituents of the microstructure of the as-printed sample. However, the grain size became more uniform and homogenous from the center of each melt pool toward the heat affected zone in the adjacent track. In other words, the grain coarsening in the heat affected zone was eliminated after normalizing heat treatment.

For the initial austenitizing step, the sample was heated up to 900 °C, where γ is the only stable phase, since the A_{c3} temperature of the alloy was calculated to be at 883.6 °C according to the empirical equation developed by Andrews [25] to predict austenite formation temperatures A_{c1} and A_{c3} for low alloy steels with less than 0.6 wt.% C. Subsequently, uniformly distributed austenite grains nucleate and grow evenly in any region of the material during austenitization. Following the full austenitization of the microstructure, the sample is subjected to a relatively slow cooling by exposing the sample to the room temperature. The slow cooling rate associated with the normalizing heat treatment hinder the formation of unstable or metastable phases, such as bainite or martensite, during the normalizing heat treatment. Therefore, the microstructure of the sample in terms of the formed constituents was analogous to the dominant microstructure of the as-printed sample, whilst after the normalizing heat treatment, the grain size distribution became more homogeneous and uniform. Natividad *et al.* [26] also performed the normalizing heat treatment on Grade X70 pipeline steel, possessing a similar chemical composition to ER70S-6, and reported the formation of polygonal ferrite and pearlite areas with more homogenous and uniform microstructure in comparison with the as-received material. Figs. 3c and d illustrate the microstructure after hardening heat treatment, which is a mixture of pearlite, bainite and acicular ferrite. In the case of hardening heat treatment cycle, the sample was exposed to an extremely high cooling rate (water quenching), resulting in the formation of the meta-stable bainite and acicular ferrite phases besides the lamellar pearlite phase. It should also be mentioned that similar to the scenario of the normalizing heat treatment, the homogeneity and uniformity of the microstructure throughout the sample after hardening both in terms of the formed micro-constituents and grain size are promoted. Formation of the acicular ferrite and bainite phases by quenching of the sample at higher cooling rates from the austenite stability region has also been reported in other low carbon low alloy steels, such as API X70 and X80 [26][20][21].

B. Mechanical Properties

Fig. 4 shows the microhardness profile of the as-printed and heat-treated samples along a line covering five consecutive layers through the building (vertical) direction on different zones including the center of the melt pools, fusion boundaries, and heat affected zones (HAZs). The overall microhardness of the as-printed sample was 160 ± 7 HV, which showed a relatively significant fluctuation from a minimum of 150 ± 1

HV to 160 ± 2 HV, and then to the maximum of 175 ± 2 HV. The observed fluctuation was ascribed to the presence of different phases along the melt pool center, the fusion boundaries, and the heat affected zone. In particular, the lowest amount of microhardness (150 ± 1 HV) corresponded to the HAZ containing coarser polygonal ferrite grains than the rest of the fusion zone, and the maximum microhardness (175 ± 2 HV) was correlated to the fusion boundaries, where acicular ferrite and bainite constituents exist. The center of the melt pool, owning the dominant microstructure of the component (lamellar pearlite and polygonal ferrite), revealed the microhardness of 160 ± 2 HV. The fluctuations in the microhardness values were considerably lower in both hardened (water-quenched) and normalized samples as compared to the as-printed component due to the homogeneity of the microstructure in the heat-treated samples.

The normalized sample with a microstructure analogous to the dominant microstructure of the as-printed sample showed the microhardness of 154 ± 3 HV, comparable to that of the center of the melt pools in the as-printed sample. It is also worth noting that the microhardness of the normalized sample was slightly decreased after the heat treatment, primarily due to (i) the release of the residual stresses of the as-printed sample, (ii) reduction of lattice defects generated during the WAAM, (iii) grain growth [27], and (iv) omitting the acicular ferrite and bainite phases from the microstructure of the fusion boundaries. Contrarily, the microhardness of the hardened (water-quenched) sample was 260 ± 3 HV, drastically higher than the other samples, ascribing to its microstructure, including pearlite, bainite, and acicular ferrite as the predominant micro-constituents in its structure. It has been reported elsewhere that the presence of bainite along with a finer microstructure can increase the microhardness of low carbon steels [27]. However, it should also be noted that a higher microhardness is not always beneficial to the overall mechanical performance of the material since the ductility and toughness of the alloy could potentially be degraded. The adverse effect of existing hard micro-constituents can be more crucial particularly in the case of samples manufactured by a welding process, which are usually prone to the presence of welding defects, discontinuities, and residual stresses. Such discontinuities can readily propagate into a brittle microstructure and form internal micro-cracks during tensile loading of the sample.

Fig. 5 shows the engineering tensile stress-strain curves for the as-printed and heat-treated WAAM-ER70S-6 samples in the building (vertical) and deposition (horizontal) directions. In the as-printed component, the vertical and horizontal tensile strength were approximately similar (503 ± 21 MPa). However, the ductility (elongation percentage) of the vertical sample only reached to $12 \pm 3\%$, whereas the horizontal sample showed a significantly higher ductility at $35 \pm 2\%$, indicating a large plastic deformation prior to the fracture with an obvious necking phenomenon. Therefore, the results of tensile testing of the as-printed part revealed anisotropy in ductility. Wang *et al.* [28] investigated the anisotropy in the mechanical properties of the additively manufactured 304 stainless steel parts, and

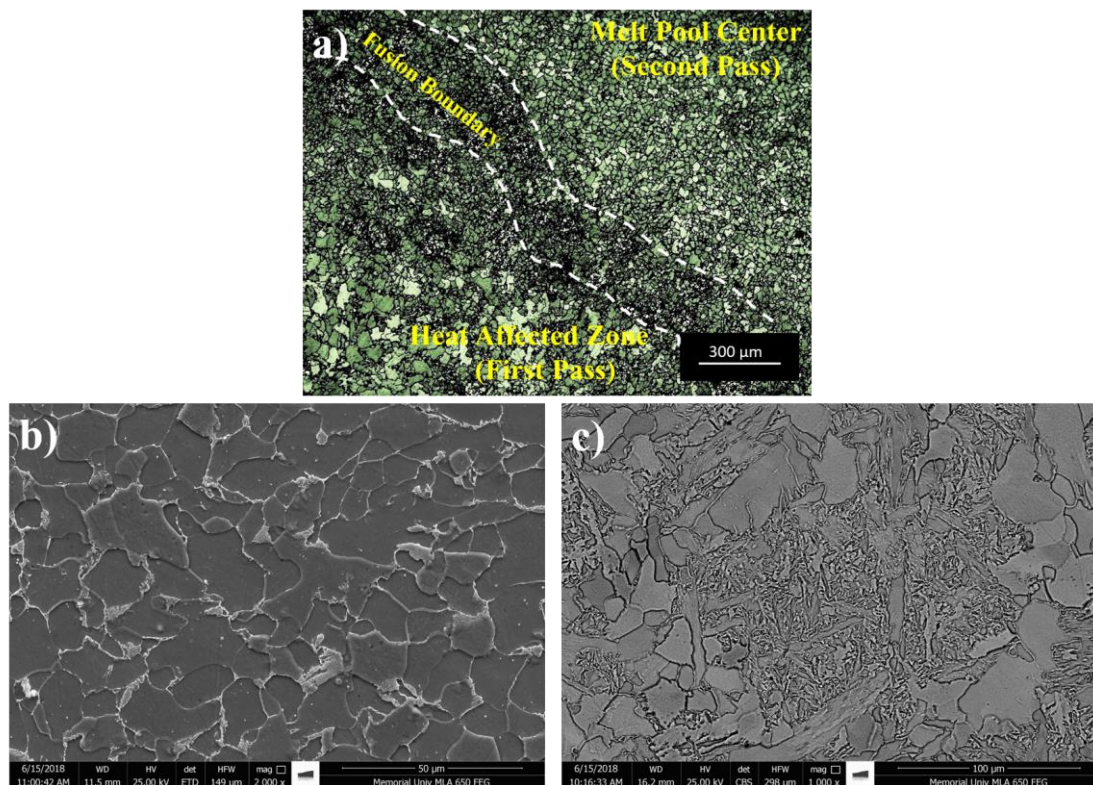


Figure 2. (a) Low magnification optical micrograph of the as-printed sample shows the transition from a melt pool center to the fusion boundary and to the heat affected zone. Higher magnification SEM micrograph of (b) melt pool center reveals lamellar pearlite and polygonal ferrite and (c) fusion boundary including acicular ferrite and bainite.

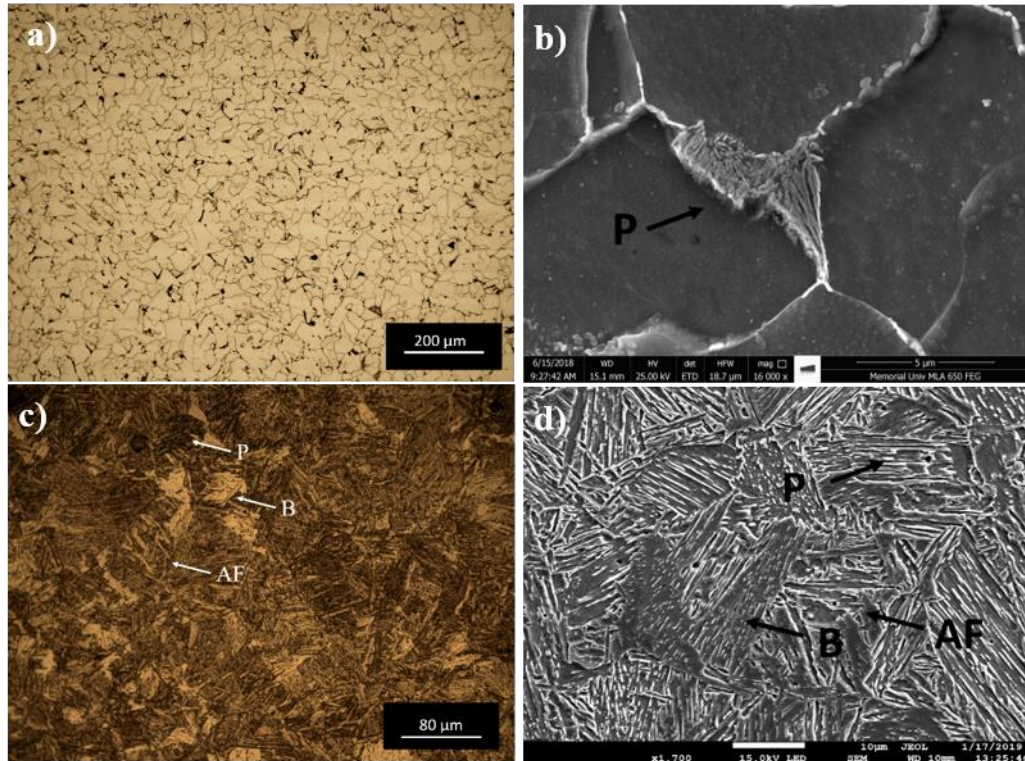


Figure 3. The microstructure of the normalized sample (a and b), and hardened (water-quenched) sample (c and d) at different magnifications.

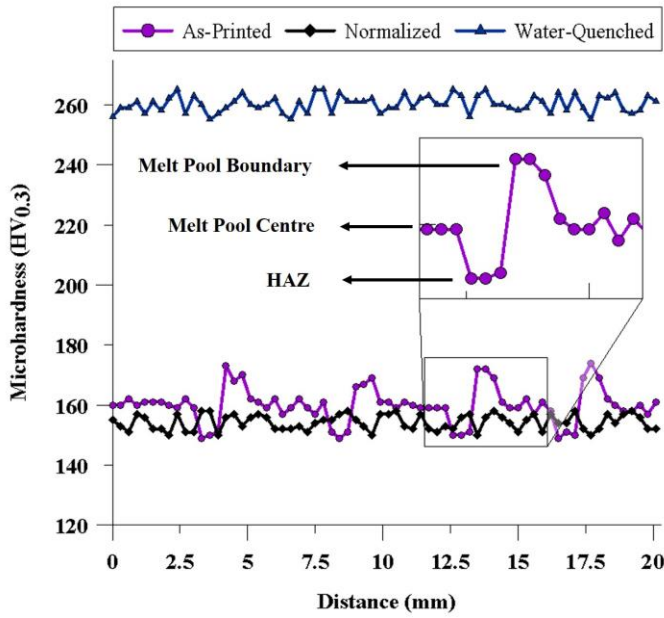


Figure 4. Vickers microhardness profile of the as-printed and heat-treated samples along a line covering five successive layers through the building (vertical) direction.

reported that the ductility in the longitudinal direction was less than that of the transverse direction, while there was no clear anisotropy in the strength. The drastic reduction in the ductility value from the horizontal sample to the vertical one herein can be rationalized by a combination of (i) the presence of manufacturing discontinuities and defects, such as inter-pass lack of fusion or porosity, and (ii) the grain coarsening in the heat affected zone resulting in the HAZ softening. In the vertical samples since the long axis of the inter-pass lack of fusions is perpendicular to the loading direction, the sharp edges of these defects can serve as stress concentration sites, causing the propagation of the discontinuity in the vertical direction, but not in the horizontal samples. A similar observation was reported by Wang *et al.* [29]. In another study, Lopez *et al.* [11] demonstrated the formation of manufacturing defects in the wire arc additive manufacturing of ER70S-6 using different nondestructive examination methods including radiographic testing (X-ray), liquid penetrant inspection (LPI), and ultrasonic testing (UT), and reported the presence of lack of fusion between deposited layers. Another major factor that can further contribute to the observed mechanical anisotropy in the as-printed sample is the grain growth phenomena in the HAZ leading to the formation of coarse ferrite grains along with the relatively low dislocation density, facilitating the decohesion and the propagation of cracking in this zone [30].

The anisotropy in the mechanical properties has been reported as a common issue in various metals and alloys [31][32][33] produced by additive manufacturing processes, which can be minimized by applying proper post-printing heat treatment cycles [10][1]. Hardening treatment could increase the tensile strength of the WAAM-ER70S steel from around 500 MPa for the as-printed sample to 640 ± 14 MPa and 624 ± 13 MPa in vertical and horizontal samples, respectively. This improvement in tensile strength is attributed to its bainitic,

acicular ferritic, and pearlitic microstructure with around 62% higher microhardness compared to the as-printed sample. However, the ductility of the hardened samples was reduced by 4% and 7% for the vertical and horizontal samples, respectively. Although, the horizontal sample with the UTS value of 624 ± 13 MPa and elongation of $28 \pm 2\%$ plausibly satisfies the mechanical integrity requirement for the service conditions of this alloy, the anisotropic mechanical behaviour of the component cannot be diminished since the sample revealed a severe brittle fracture in the vertical direction with only $8 \pm 1\%$ elongation.

It should be mentioned that in the case of the as-printed vertical sample, the formation of some defects, flaws, and also possible weak metallurgical bonding between the layers (lack of fusion) can potentially reduce the degree of plastic deformation that the material can accommodate before its failure. This scenario can be intensified when the microhardness increases from 160 ± 7 HV for the as-printed sample to 260 ± 3 HV for the hardened sample. Consequently, as a result of hardening cycle, the ultimate tensile strength increased at the expense of a reduction in ductility. Overall, it can be inferred that the hardening heat treatment exhibited a positive effect on the mechanical properties in the horizontal direction but was not found beneficial to the vertical sample, leading to its brittle fracture during the uniaxial tensile loading.

On the other hand, the normalizing process that contributed to the formation of a homogenized microstructure, characterized by a uniform grain size along the center of the melt pool, fusion boundary, and the heat affected zone could not increase the tensile strength neither in horizontal nor in the vertical samples. Similar to the slight reduction in the microhardness of the normalized sample (~ 10 HV lower than that of the as-printed sample), its tensile strength was expected to be slightly lower (460 ± 12 MPa) as compared to the as-printed sample (503 ± 21 MPa). Moreover, a closer look at the stress-strain curves of the normalized samples revealed that there is not a huge difference between the elongation of the component in the vertical ($29 \pm 2\%$) and horizontal ($34 \pm 3\%$) directions. It should be mentioned that the purpose of normalizing cycle was to eliminate the inhomogeneous microstructure resulted from the complex thermal cycles associated with the layer-by-layer deposition nature of the wire arc additive manufacturing process. As a consequence of heating the sample up to a temperature above the upper critical temperature (A_{c3}), and formation of new austenite grains, the inhomogeneous microstructure including coarse grains of HAZ was eliminated. Accordingly, during the cooling process in still-air, the whole part experiences a similar cooling rate leading to a uniform microstructure at different zones of the sample. Additionally, all residual stresses induced by the rapid solidification during the WAAM process are released [27]. Therefore, the anisotropy in the elongation can be eliminated or weakened through modifying the microstructure from an inhomogeneous one to a homogenized microstructure with a uniform grain size.

Fig. 6 demonstrates the reduction in area (RA) for the as-printed and heat-treated samples in both vertical and horizontal directions. The stereomicroscope images of the fractured surfaces are also attached to each point of the plot in order to

clarify the brittle or ductile nature of the fracture in different samples. As clearly shown in Fig. 6, there is a considerable difference between the RA values of the horizontal and vertical tensile samples in both as-printed and hardened conditions, implying a significant anisotropy in ductility of the component. However, the RA values of the normalized sample in the vertical and horizontal directions are relatively close to each other, indicating a negligible anisotropy in ductility. Therefore, normalizing treatment can be utilized as a post-printing cycle to minimize the anisotropic behaviour of the wire arc additively manufactured low carbon low alloy steel (ER70S-6) by homogenizing the grain size and eliminating the inhomogeneous microstructure through the melt pool center, fusion boundary, and the heat affected zone.

IV. CONCLUSIONS

This study aimed to investigate the effect of two post-printing heat treatment cycles, including normalizing and hardening, on the microstructure and mechanical properties of a wire arc additively manufactured low carbon low alloy steel (ER70S-6). The following are the key conclusions of the study:

1) The dominant microstructure of the as-printed WAAM-ER70S-6 component consisted of polygonal ferrite grains along with a small volume fraction of lamellar pearlite formed at the ferrite grain boundaries. In addition, formation of acicular ferrite and bainite constituents were detected as the primary phases along the melt pool boundaries. Furthermore, a heat affected zone comprised of coarser polygonal ferrite grains adjacent to each deposited track in the previously solidified bead also formed associated with the layer-by-layer deposition nature of the process, inducing multiple heating cycles on each deposited track.

2) Normalizing heat treatment eliminated the meta-stable constituents, *i.e.* acicular ferrite and bainite, from the as-printed microstructure, leading to a more uniform and homogenous ferritic/pearlitic microstructure within the melt pool center, fusion boundaries, and the heat affected zone. On the contrary, the hardening heat treatment altered the microstructure of the as-printed part to a combination of pearlite, bainite, and acicular ferrite.

3) Microhardness of the as-printed sample slightly decreased from 160 ± 7 HV to 154 ± 3 HV after the normalizing heat treatment, while the hardening treatment could increase the microhardness to 260 ± 3 HV.

4) Uniaxial tensile testing of the as-printed samples indicated a comparable tensile strength in horizontal and vertical samples, while a considerable anisotropy in the ductility was apparent in horizontal and vertical samples with $35 \pm 2\%$ and $12 \pm 3\%$ of elongation, respectively.

5) Hardening heat treatment could increase the tensile strength of the component by around 20%, but intensified the anisotropy in the ductility of vertical and horizontal samples.

6) The anisotropy in ductility was minimized by normalizing heat treatment due to the removal of the coarse grain regions in the HAZ and the resultant uniformity and homogeneity of the microstructure.

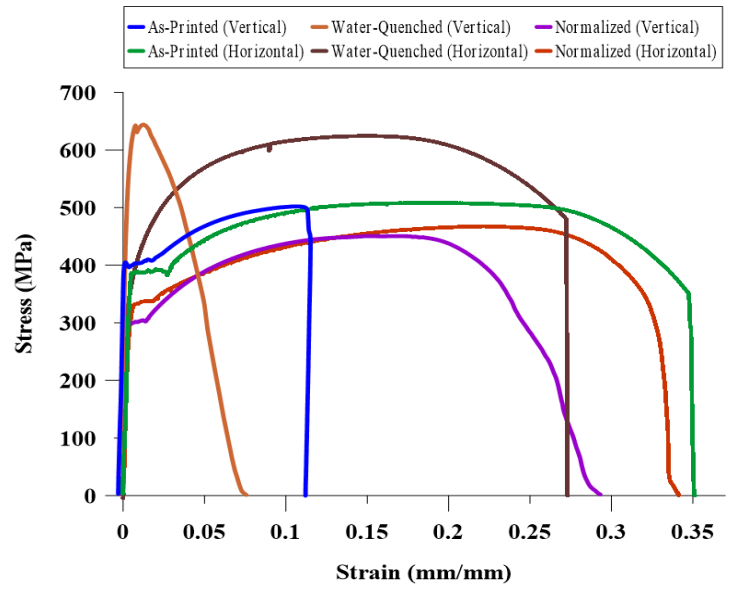


Figure 5. The stress-strain curves for the as-printed and heat-treated samples in the building (vertical) and deposition (horizontal) directions.

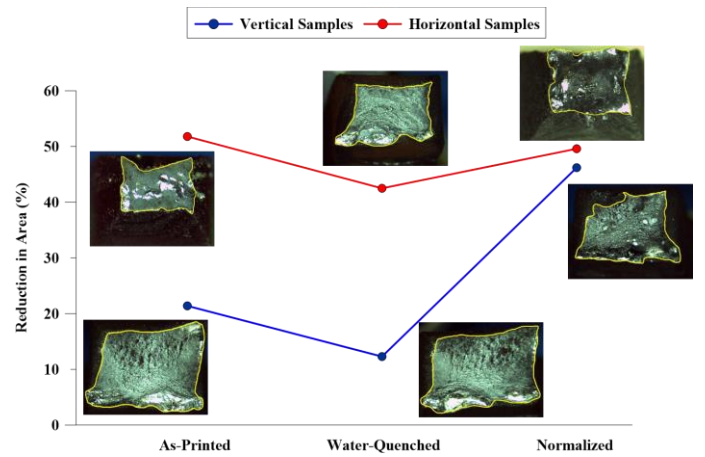


Figure 6. The reduction in area (RA) for the as-printed and heat-treated samples in both vertical and horizontal directions.

ACKNOWLEDGMENT

The authors would like to thank the support of Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number RGPIN-2017-04368] for sponsoring this work.

REFERENCES

- [1] E. Herderick, "Additive manufacturing of metals: A review," *Mater. Sci. Technol.*, no. MS, p. 1413, 2011.
- [2] W. E. Frazier, "Metal Additive Manufacturing: A Review," *J. Mater. Eng. Perform.*, vol. 23, no. 6, pp. 1917–1928, Jun. 2014.
- [3] J. Xiong, Z. Yin, and W. Zhang, "Closed-loop control of variable layer width for thin-walled parts in wire and arc additive manufacturing," *J. Mater. Process. Technol.*, vol. 233, pp. 100–106, 2016.

- [4] J. Xiong, Y. Lei, H. Chen, and G. Zhang, "Fabrication of inclined thin-walled parts in multi-layer single-pass GMAW-based additive manufacturing with flat position deposition," *J. Mater. Process. Technol.*, vol. 240, pp. 397–403, 2017.
- [5] Z. Wang, A. D. Stoica, D. Ma, and A. M. Beese, "Diffraction and single-crystal elastic constants of Inconel 625 at room and elevated temperatures determined by neutron diffraction," *Mater. Sci. Eng. A*, vol. 674, pp. 406–412, 2016.
- [6] A. M. Beese and B. E. Carroll, "Review of mechanical properties of Ti-6Al-4V made by laser-based additive manufacturing using powder feedstock," *Jom*, vol. 68, no. 3, pp. 724–734, 2016.
- [7] S. Suryakumar, K. P. Karunakaran, U. Chandrasekhar, and M. A. Somashekara, "A study of the mechanical properties of objects built through weld-deposition," *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, vol. 227, no. 8, pp. 1138–1147, 2013.
- [8] C. V. Haden, G. Zeng, F. M. Carter III, C. Ruhl, B. A. Krick, and D. G. Harlow, "Wire and arc additive manufactured steel: Tensile and wear properties," *Addit. Manuf.*, vol. 16, pp. 115–123, 2017.
- [9] N. Sridharan, M. W. Noakes, A. Nycz, L. J. Love, R. R. Dehoff, and S. S. Babu, "On the toughness scatter in low alloy C-Mn steel samples fabricated using wire arc additive manufacturing," *Mater. Sci. Eng. A*, vol. 713, no. November 2017, pp. 18–27, 2018.
- [10] P. Zhang, J. Liu, and A. C. To, "Role of anisotropic properties on topology optimization of additive manufactured load bearing structures," *Scr. Mater.*, vol. 135, pp. 148–152, 2017.
- [11] A. Lopez, R. Bacelar, I. Pires, T. G. Santos, J. P. Sousa, and L. Quintino, "Non-destructive testing application of radiography and ultrasound for wire and arc additive manufacturing," *Addit. Manuf.*, vol. 21, no. January, pp. 298–306, 2018.
- [12] B. A. Szost *et al.*, "A comparative study of additive manufacturing techniques: Residual stress and microstructural analysis of CLAD and WAAM printed Ti-6Al-4V components," *Mater. Des.*, vol. 89, pp. 559–567, 2016.
- [13] A. Kudzal *et al.*, "Effect of scan pattern on the microstructure and mechanical properties of Powder Bed Fusion additive manufactured 17-4 stainless steel," *Mater. Des.*, vol. 133, pp. 205–215, 2017.
- [14] K. V. Yang, P. Rometsch, C. H. J. Davies, A. Huang, and X. Wu, "Effect of heat treatment on the microstructure and anisotropy in mechanical properties of A357 alloy produced by selective laser melting," *Mater. Des.*, vol. 154, pp. 275–290, 2018.
- [15] B. D. DeRuntz, "Assessing the benefits of surface tension transfer welding to industry," *J. Ind. Technol.*, vol. 19, no. 4, pp. 55–62, 2003.
- [16] A. O. Benscoter and B. L. Bramfitt, "Metallography and microstructures of low-carbon and coated steels," *Metallogr. Microstruct.*, vol. 9, pp. 588–607, 2004.
- [17] ASTM Int., "Standard Test Methods for Tension Testing of Metallic Materials 1," *Astm*, vol. i, pp. 1–27, 2015.
- [18] A. S. Haselhuhn, B. Wijnen, G. C. Anzalone, P. G. Sanders, and J. M. Pearce, "In situ formation of substrate release mechanisms for gas metal arc weld metal 3-D printing," *J. Mater. Process. Technol.*, vol. 226, pp. 50–59, 2015.
- [19] H. K. Lee, K. S. Kim, and C. M. Kim, "Fracture resistance of a steel weld joint under fatigue loading," *Eng. Fract. Mech.*, vol. 66, no. 4, pp. 403–419, 2000.
- [20] Y. M. Kim, H. Lee, N. J. Kim, and J. Y. Yoo, "Microstructural characteristics of acicular ferrite in linepipe steels," in *The Fifteenth International Offshore and Polar Engineering Conference*, 2005.
- [21] X. Yang, M. Salem, and E. J. Palmiere, "The effect of processing conditions and cooling rate on the microstructure and properties of API X-70 and API X-100 steels," *Mater. Manuf. Process.*, vol. 25, no. 1–3, pp. 48–53, 2010.
- [22] B. Hwang, Y. M. Kim, S. Lee, N. J. Kim, and S. S. Ahn, "Correlation of microstructure and fracture properties of API X70 pipeline steels," *Metall. Mater. Trans. A*, vol. 36, no. 3, pp. 725–739, 2005.
- [23] T. Mohandas, G. M. Reddy, and B. S. Kumar, "Heat-affected zone softening in high-strength low-alloy steels," *J. Mater. Process. Technol.*, vol. 88, no. 1–3, pp. 284–294, 1999.
- [24] X. Shi *et al.*, "Selective laser melting-wire arc additive manufacturing hybrid fabrication of Ti-6Al-4V alloy: Microstructure and mechanical properties," *Mater. Sci. Eng. A*, vol. 684, no. December 2016, pp. 196–204, 2017.
- [25] K. W. Andrews, "Empirical formulae for the calculation of some transformation temperatures," *J. Iron Steel Inst.*, pp. 721–727, 1965.
- [26] C. Natividad, R. García, V. H. López, L. A. Falcón, and M. Salazar, "Mechanical and Metallurgical Properties of Grade X70 Steel Linepipe Produced by Non-conventional Heat Treatment," in *Characterization of Metals and Alloys*, Springer, 2017, pp. 3–11.
- [27] S. Bordbar, M. Alizadeh, and S. H. Hashemi, "Effects of microstructure alteration on corrosion behavior of welded joint in API X70 pipeline steel," *Mater. Des.*, vol. 45, pp. 597–604, 2013.
- [28] Z. Wang, T. A. Palmer, and A. M. Beese, "Effect of processing parameters on microstructure and tensile properties of austenitic stainless steel 304L made by directed energy deposition additive manufacturing," *Acta Mater.*, vol. 110, pp. 226–235, 2016.
- [29] Z. Wang and A. M. Beese, "Effect of chemistry on martensitic phase transformation kinetics and resulting properties of additively manufactured stainless steel," *Acta Mater.*, vol. 131, pp. 410–422, 2017.
- [30] A. J. Haq, K. Muzaka, D. P. Dunne, A. Calka, and E. V. Pereloma, "Effect of microstructure and composition on hydrogen permeation in X70 pipeline steels," *Int. J. Hydrogen Energy*, vol. 38, no. 5, pp. 2544–2556, 2013.
- [31] B. E. Carroll, T. A. Palmer, and A. M. Beese, "Anisotropic tensile behavior of Ti-6Al-4V components fabricated with directed energy deposition additive manufacturing," *Acta Mater.*, vol. 87, pp. 309–320, 2015.
- [32] L. Thijs, M. L. M. Sistiaga, R. Wauthle, Q. Xie, J.-P. Kruth, and J. Van Humbeeck, "Strong morphological and crystallographic texture and resulting yield strength anisotropy in selective laser melted tantalum," *Acta Mater.*, vol. 61, no. 12, pp. 4657–4668, 2013.
- [33] Q. Zhang, J. Chen, Z. Zhao, H. Tan, X. Lin, and W. Huang, "Microstructure and anisotropic tensile behavior of laser additive manufactured TC21 titanium alloy," *Mater. Sci. Eng. A*, vol. 673, pp. 204–212, 2016.

Additively Manufactured DMLS- AlSi10Mg using Recycled Powder: the Impacts on Microstructure and Corrosion Properties

M. Rafieezad*, A. Chatterjee, and A.M. Nasiri

Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL, Canada, A1B 3X5

(*Corresponding author: mrafieezad@mun.ca)

Abstract— This study examines the impact of using recycled powder on microstructures and the corrosion performance of the Direct Metal Laser Sintered (DMLS) AlSi10Mg alloy. Microstructural analysis of the DMLS- AlSi10Mg alloy using recycled powder revealed that the use of recycled powder in the fabrication of the DMLS- AlSi10Mg alloy primarily impacts the degree of Si networks coarsening by stimulating the precipitation of eutectic Si phase during solidification of the alloy. Field emission scanning electron microscope (FESEM) and X-ray diffraction (XRD) analysis were utilized to characterize the evolution of the microstructure in samples fabricated using the recycled powder in comparison with the ones fabricated using the virgin powder. Additionally, to investigate the impact of recycled powder on the corrosion resistivity of DMLS- AlSi10Mg , the anodic polarization testing and electrochemical impedance spectroscopy were performed in aerated 3.5 wt.% NaCl solution. The results confirmed a slight degradation of the corrosion properties of the recycled powder fabricated samples, ascribed to further coarsening of Si network along the melt pool boundaries. Furthermore, the intergranular corrosion testing demonstrated formation of superficial micro-cracks at the melt pool boundaries of the recycled powder fabricated sample, whereas the virgin powder fabricated part did not reveal superficial cracking after the intergranular corrosion testing.

Keywords- Additive manufacturing (AM); Direct laser metal sintering (DMLS); AlSi10Mg ; Recycled powder; Microstructure; Corrosion properties

I. INTRODUCTION

The continuously rising demands for advanced engineering components with complex designs and exceptional mechanical performance requires the implementation of innovative manufacturing technologies. In this regard, additive manufacturing (AM) technology is gaining enormous attention from industrial manufacturing sectors, owing to its malleability to design complicated, near-net shape engineering components mostly without any geometrical constraint, and its affordability due to minimal waste of the feedstock material, leading to its efficient processing route for fabrication of engineering materials for low volume manufacturing [1].

Aluminum alloys containing Silicon (Si) and Magnesium (Mg) as the primary alloying elements have conceived a lightweight alloy family with adequate mechanical properties and acceptable corrosion performance [2]. Among a wide variety of alloys in this family, AlSi10Mg has been mostly adapted by the AM industry, in particular through Direct Metal Laser Sintering (DMLS) process. As a hypo-eutectic alloy, AlSi10Mg has a multiplicity of applications in the aerospace and automotive industries due to its distinctive properties [3].

Regardless of the economic impacts of the AM technology on manufacturing of intricate engineering parts, there is still a major drawback of adopting this technology by many industries due to the high final cost of the printed components, particularly for parts having medium to simple designs [4]. One way to increase the affordability of the final product is to minimize wasting of the feedstock metal powder. In DMLS process, 90% of the feedstock powder is commonly collected after each building process and reused [5]. However, the high temperature of the powder bed in the preheating stage and more severely the interaction between the virgin powder and the laser can cause possible microstructural and morphological changes in the virgin powder close to the parts that are being built [6].

As powder reuse cycle increases, many properties of the powder are expected to change [4]. This can adversely affect the build quality. To suppress all the side effects associated with implementing the used powder in fabrication of new parts without compromising on the build quality, many AM industries have incorporated a sieving process into the manufacturing cycle of the AM products, in which all collected powder after completion of each building cycle will pass through a sieving step to separate and discard partially melted or highly heated condensate powder from the remaining unmelted virgin powder, denoted as the recycled powder, and reusing them after the sieving process. Nonetheless, it has been shown that the recycled powder will not be identical to the virgin powder in terms of shape, size distribution, and the microstructure [1,7,8]. However, the manufactured part using the recycled powder should have comparable properties with parts produced using virgin powder.

So far, some efforts have been made by the research communities to investigate the effect of using recycled powder on porosity level, density, and mechanical properties of the AM components and auspicious results have been obtained [1,7,8].

For instance, Ardila *et al.* [8] reported a comparable mechanical properties for the selective laser melting (SLM) fabricated Inconel 718 using virgin powder and the recycled powder even after 14 times of recycling. In a recent study, Asgari *et al.* [1] extensively characterized the AlSi10Mg recycled powders after the DMLS process, and reported a noticeably different particles' size, morphology, and chemical composition for the recycled powder relative to the virgin powder [1]. Such differences were ascribed to the high tendency of fine powders for sintering and agglomeration [9,10]. The sintering of the AlSi10Mg metal powder during the DMLS process and the subsequent agglomeration was reported to cause approximately 12% increase in the average size of the particles [1]. Furthermore, it has been reported that the recycled AlSi10Mg powder is slightly deformed and elongated as compared to the spherical shape of the virgin powder, resulting in a lower apparent density as compared to that of the spherical virgin powder [7].

The change in the morphology and the size of metal powder has been also shown to impact the density and the volume fraction of the internal defects in the AM fabricated parts. Abd-Elghany and Bourell [11] have reported that an increase in 304L stainless steel particle size on the layer thickness of 30 μm to 70 μm can lead to a 6 % decrease in the density of manufactured parts due to a reduction in the contact area of each powder layer. It is also known that the densification process has a significant impact on the formation of pores in the samples. As Maamoun *et al.* [7] reported, two types of pores have been commonly detected in the DMLS-AlSi10Mg samples fabricated using the recycle powders, including Spherical pores and Keyhole pores. Spherical pores are small sized pores (around 10 μm diameter), which form due to the large interparticle spacing in the powder samples during the laser-melting process. Such interparticle voids can potentially act as gas entrapment sites, contributing to the formation of the Spherical pores [7]. On the other hand, Keyhole pores mainly form due to the improper melting of the powders. These Keyhole pores are almost five times larger than the Spherical ones and generated from the lack of thermal conduction during the laser melting process [7]. Contrarily, the virgin powder with a smaller particle size and larger particle packing, has a higher thermal conductivity. Accordingly, formation of Keyhole pores can be suppressed when virgin powder is used as the feedstock material [12].

The reported mechanical properties of the fabricated DMLS-AlSi10Mg parts using recycled powder with larger average size than that of the virgin powder confirmed that implementing the recycled powder in manufacturing of the DMLS-AlSi10Mg does not affect the mechanical integrity of the final product adversely [1]. Contrarily, Liu *et al.* [13] studied the impact of particle size distribution on the mechanical properties of the SLM-316L stainless steel and reported that using 316L powders with a narrower distribution range of particle size instead of a wider distribution range can contribute to a greater powder flowability, and thermal conductivity, resulting in a higher ultimate tensile strength (UTS) in the fabricated part [13]. Analogously, Spierings *et al.* [14] reported that mechanical properties of the SLM-316L

parts are highly dictated by the feedstock powder particle size distributions.

In the context of corrosion properties, although Revilla *et al.* [15] reported no significant influence of the recycled powder on the corrosion potential of SLM-AlSi10Mg using anodic polarization testing, the observed slight corrosion performance difference between the recycled powder fabricated samples and the virgin ones was not properly correlated to the microstructural variations between the samples. Furthermore, albeit the increased size of the recycled powder were detected in the previous studies [1,7,15], the effect of the powder size on the final microstructure of the samples were not clearly discussed, and no elucidation for the slight microstructural variations between the recycled powder and the virgin powder fabricated samples has been provided.

Therefore, despite the existing few studies on the microstructure and mechanical properties of the DMLS fabricated AlSi10Mg alloy using the recycled powders [1,7], the impact of using recycled powder on corrosion properties of the parts is still unknown. This knowledge is crucial for AM fabricated AlSi10Mg alloy since one of the major concerns with using DMLS-AlSi10Mg components in marine or aerospace applications is their corrosion performance, and high temperature induced microstructural modifications of the printed AlSi10Mg have been shown to deteriorate the corrosion properties [16,17]. This study investigates the impact of powder reuse times on the microstructural evolution and the resulting corrosion properties of the DMLS-AlSi10Mg parts produced from recycled powders vs. its virgin powder fabricated counterpart.

II. EXPERIMENTAL PROCEDURE

A. Material

DMLS-AlSi10Mg cubic samples with a dimension of 15 \times 15 \times 15 mm were fabricated using an EOS M290 metal additive manufacturing machine. The samples were printed under a 400 W Yb-fiber laser irradiation with 100 μm spot size. Three types of AlSi10Mg feedstock powder were utilized for fabrication, *i.e.* virgin powder, recycled powders after four times and five times of reuse cycles. The fabricated samples were denoted as Virgin, 4X and 5X samples, respectively. For the case of the recycle powders, after each building cycle the collected powder was sieved prior to the next building cycle. Therefore, the 4X samples were built using a recycled powder that was passed through 4 building cycles along with 4 subsequent sieving steps before being used for fabrication of 4X samples, whereas the used recycled powder for fabrication of 5X sample was passed through one additional building and its subsequent sieving cycle as compared to the 4X sample. The nominal chemical composition of the virgin powder used herein is given in Table 1.

TABLE I. THE NOMINAL CHEMICAL COMPOSITION OF ALSi10MG VIRGIN METAL POWDER USED IN THIS STUDY (WT. %)

Si	Mg	Fe	Mn	Ti	Zn	Cu	Al
9.00-11.00	0.20-0.45	≤ 0.55	≤ 0.45	≤ 0.15	≤ 0.10	≤ 0.05	Bal.

B. Microstructural characterization

For the microstructure characterization of the samples, all DMLS-samples were sectioned along both the building plane and the building direction, and prepared following the standard grinding and polishing procedures for aluminum alloys utilizing a Tegramin-30 Struers auto-grinder/polisher. The polished surfaces were then chemically etched using Keller's reagent (2.5 vol. % HNO₃, 1 vol. % HF, 1.5 vol. % HCl, and 95 vol. % H₂O) to reveal the microstructure. The microstructures of all samples were characterized using an optical microscope (Nikon Eclipse 50i) and a field emission scanning electron microscope (FESEM) (JEOL JSM 7100F). Phase characterization of different samples was carried out using a Rigaku Ultimate IV X-ray diffraction (XRD) with Cu-K_α source at 40 kV and 44 mA at the diffraction angle range of 5°-90° with a step size of 0.02°.

C. Anodic polarization testing

The anodic potentiodynamic polarization (APP) testing was conducted using an IVIUM CompactStat™ Potentiostat computer-controlled instrument with a three-electrode cell setup according to the ASTM G5 standard. The three-electrode setup consisted of a graphite rod as the counter electrode, saturated Ag/AgCl as the reference electrode, and the DMLS-AlSi10Mg sample as the working electrode. All electrochemical corrosion tests were performed in simulated sea water environment (an aerated 3.5 wt.% NaCl solution) in a temperature-controlled water bath at constant temperature of 25±0.5 °C. Prior to each polarization test, the open circuit potential (OCP) was monitored for 1 h to ensure samples reach to an electrochemical stability before the test was run. The APP tests were conducted in the potential range of -0.02 V_{Ag/AgCl} to +0.3 V_{Ag/AgCl} with respect to OCP value and with a scanning rate of 0.15 mV/s.

D. Electrochemical Impedance Spectroscopy (EIS) testing

Additionally, the EIS tests were carried out on polished samples after 1 h and 120 h of immersion times in an aerated 3.5 wt. % NaCl solution at 25 °C. Signals with 0.01 V amplitude were applied over the open circuit potential with a frequency range between 100 kHz and 10 mHz. All corrosion tests were carried out on the polished surfaces to eliminate effects of surface roughness on corrosion properties.

E. Intergranular corrosion testing

The intergranular corrosion testing was performed on the polished surfaces of the side view (aligned with the building direction) of different samples as per ISO 11846 standard, where samples were immersed in 30 g/L of sodium chloride solution containing 10 mL/L of HCl at 25 °C for 24 h. After one day of immersion, samples were rinsed with distilled water and investigated using FESEM.

III. RESULTS AND DISCUSSION

A. Microstructure Characterization

The optical micrographs taken from the Virgin, 4X, and 5X DMLS-AlSi10Mg samples are shown in Fig. 1. The 3D representation of the optical micrographs of DMLS-AlSi10Mg samples, from both side and top views (parallel and perpendicular to the building direction, respectively) confirmed a proper overlapping and densification between melt pools. The morphology of the melt pools from the side view of the fabricated samples revealed a semi-circle shape, which its size is primarily dependent on the volumetric energy density used in manufacturing of the samples. The top view of all three fabricated samples shows irregular melt pool geometries and directions, attributed to the 67° rotation of the laser scan between the consecutive layers [18]. The optical microscopy investigation also confirmed that employing recycling powder does not alter the overall macrostructure and morphology of the melt pools noticeably in the fabricated samples. This was found to be in agreement with the results reported in previous studies [1,7,15], confirming a consistent overall macrostructure for both virgin powder and recycled powder fabricated samples.

One important difference between the Virgin sample and the recycled ones to note is the volume of internal porosities in the samples. A closer look at the optical microscopic images of the built samples (shown in Fig. 2) reveals that using recycled powders in fabrication of DMLS-AlSi10Mg (Figs. 2b and 2c) has resulted in the formation of larger and higher quantities of Spherical and Keyhole porosities in the as-printed samples compared to the Virgin sample. The higher volume fraction of internal porosities in both 4X and 5X samples can be correlated to the reduced effective thermal conductivity of the recycled powders, as it depends on the size, shape, and void fraction between the powders.

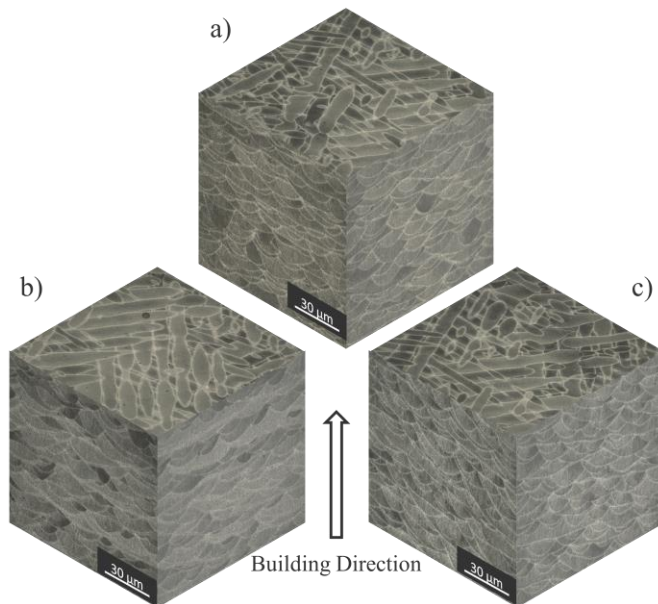


Figure 1. 3D representation of optical micrograph of DMLS-AlSi10Mg sample for a) Virgin, b) 4X recycled, and c) 5X recycled samples, demonstrating the side view along the building direction, and the top view perpendicular to the building direction.

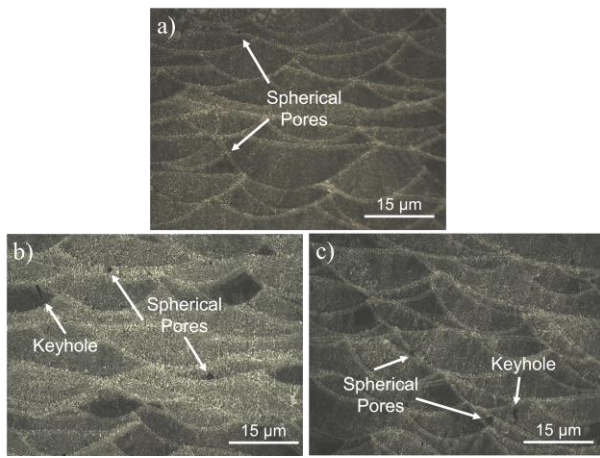


Figure 2. Optical micrographs of the DMLS-AISi10Mg samples fabricated from the a) Virgin, b) 4X recycled powder, and c) 5X recycled powder.

The increased size and the distorted shapes of the recycled powder particles relative to the virgin particles would increase the volume fraction of internal voids between the particles, leading to a lower thermal conductivity between the powders [10]. In other words, in the case of virgin powder, the sintering kinetics are more accelerated than that of the recycled powders, ascribed to the higher active contact surface area of small particles and smaller gaseous pockets in the inter-particle spacing. It has been shown by Olakanmi [4] that the irregular shapes of the metal powders as a result of the agglomeration can intensify the formation of voids and porosity in the SLM fabricated Al-Mg, and Al-Si parts. In addition, the great impact of the particle size on the increased porosity level of the iron parts manufactured by direct laser sintering (DLS) method has been reported [10]. Thus, the irregular shape [7] and larger particle size [1] of the recycled powders contribute to the entrapment of larger gaseous pores between the powder layers, leading to a reduced thermal conductivity between the particles, and consequently lowering the sintering kinetics during the fabrication process, leading to the increased volume fraction of porosities. Furthermore, the larger size of the porosities in the recycled powder fabricated samples than the Virgin sample was attributed to the decreased packing density of the recycled powders as a result of their agglomeration.

To further investigate the microstructure of the fabricated DMLS-AISi10Mg samples, scanning electron microscopy analysis was employed. Fig. 3 shows the SEM images taken from the side views of the DMLS-AISi10Mg alloy fabricated using Virgin, 4X, and 5X recycled powders. Similarly, all samples revealed formation of an extremely fine cellular dendritic structure of aluminum containing a continuous network of eutectic Si phase primarily formed along the interdendritic regions. This microstructure is commonly reported for the additively manufactured AISi10Mg alloy [17,19,20]. Over a melt pool area, there are three distinguishable regions with distinct microstructures formed by the moving heat source, *i.e.* fine cellular structure, coarse cellular structure, and transition heat affected zone (HAZ) structure that forms due to the overlapping scanning lines and layer-by-layer solidification of material [21].

It should be noted that critical solidification parameters, *i.e.* temperature gradient (G) and solidification rate (R), control the microstructure for a given composition [22]. These values vary from the melt pool borders towards its center, resulting in a fine cellular-dendritic morphology in the melt pool center, where G is the maximum and R is the minimum during solidification. On the other hand, an elongated and coarser dendritic structure along the melt pool boundaries forms, due to the lowest temperature gradient and the highest cooling rate [20].

The microstructure along the melt pool boundaries adjacent to the HAZ regions in Virgin, 4X and 5X recycled samples are shown in Fig. 3. A noticeably coarser and more expanded Si network was detected along the melt pool boundary of the recycled samples, particularly in the 5X sample, in comparison with the Virgin sample. The coarsening of the microstructure in the recycled samples can be correlated to the different thermal conductivity between the recycled powders and the virgin powders. As reported by Simchi [10], the effective thermal conductivity of the particles varies with their size, void space between them, arrangement, and active surface contact. Therefore, the recycled powders, having a larger size [1] and elongated shapes [7], form a greater void space and a smaller contact area between them, leading to a lower thermal conductivity between them. Thus, a smaller portion of laser's volumetric energy density is dissipated by conduction through the metal powders into the previously solidified tracks. Consequently, a larger portion of laser beam energy is consumed for melting of the metal powders and plausibly raising the temperature of the melt pools to higher temperature. This can potentially cause slower solidification rate and cooling rate of the melt pool, leading to further coarsening of the solidified structure and excessive solute segregation (Si particles) from the liquid during solidification. The observed coarser Si network along the interdendritic regions and further precipitation of this phase at the interior of the dendrites and even more expanded HAZ confirm the slower cooling rate that melt pools experience during solidification in the case of 4X and more noticeably 5X sample.

Interestingly, in the coarse region of the recycled samples along the melt pool boundary, especially in the 5X sample, the formation of the solidification micro-cracks primarily in the areas where Al-Si lamellar eutectic structure had formed, was detected (shown in Figs. 3b and 3c by arrows). It is well established that materials with high coefficient of thermal expansion, such as aluminum alloys, excessive solute segregation during solidification, *e.g.* Si phase in the case of AISi10Mg alloy, in the presence of high degree of residual stresses (common to DMLS fabricated parts, as material experiences extremely high cooling rates (10^6 - 10^8 °C/s)) can enhance solidification cracking susceptibility of the material. Excessive solute segregation in the case of 4X and 5X recycled samples as compared to the Virgin sample were detected along the melt pool boundaries. Consequently, the samples fabricated using the recycled powder seem to be more susceptible to solidification cracking than the virgin powder fabricated ones.

Although, in a recent study, comparable mechanical properties were reported for the DMLS-AISi10Mg parts fabricated using virgin and recycled powders [1], their investigated recycled powder was only one time re-used,

implying that their recycled powder could not have experienced a significant change in its shape, size distribution, and morphology, only after one building cycle. Accordingly, formation of solidification cracks were not detected in their study [1]. However, the presented results in this study confirmed that after four or five times of powder re-use cycle, a noticeable number of solidification micro-cracks formed in the as-printed samples, which can potentially deteriorate the mechanical properties of the fabricated part. Despite the vital importance of mechanical properties in the build quality, the current study is mainly focused on the microstructure and corrosion properties of the parts, and the study of mechanical properties of the fabricated samples is subjected to a future work by the authors.

The XRD spectra of the DMLS-AISi10Mg samples are presented in Fig. 4, revealing a co-existence of both Al and Si phases for all three different samples. The diffraction peaks for Si and Al phases were identified according to the JCPDS patterns of 01-071-4631 and 98-001-3659, respectively. Noticeably, the Si peaks at 28° , 47° , and 56° in the XRD patterns of the samples intensify from the Virgin to 4X, and 5X recycled samples, corresponding to the increased content of Si in precipitation form from the Virgin to the 5X recycle samples. This is consistent with the microstructural observations results presented in Fig. 3, revealing the coarsening of intercellular Si network in the recycled samples compared to that of the Virgin sample

B. Open Circuit Potential Variations and Anodic Polarization Results

Fig. 5a shows the evolution of the open circuit potential (OCP) of all samples over 3600 s in an aerated 3.5 wt.% NaCl solution. Since the entire composition of all samples were approximately the same, the OCP values for all fabricated samples using different powders were stabilized around -0.70 ± 0.01 V_{Ag/AgCl}, which was found to be consistent with previous studies [15]. However, the 5X recycled sample revealed a slightly lower and more unstable E_{OCP} , as compared to the other samples. The fluctuation of OCP values has been ascribed to simultaneous localized dissolution and re-formation of the passive film on the surface of alloys with self-passivating properties, causing the electrochemical instability of the surface [23]. Meanwhile, the more negative E_{OCP} value of the 5X recycle sample, indicates the higher activity and less electrochemical stability of its surface. This can be associated with the impact of recycled powder on the microstructure and the resulted higher volume fraction of porosities in the fabricated samples, in addition to further coarsening and breakage of Si network into idiomorphic crystals in the HAZ, as compared to those in the Virgin sample [15].

Fig. 5b shows anodic polarization curves of the fabricated samples using virgin and recycled powders in fully polished condition. Previous studies have shown a lower corrosion resistivity on the side view (parallel to the building direction) of the DMLS-AISi10Mg than on top view, which has been associated to the higher density of the melt pool boundaries containing an increased content of Si, and the breakage of the Si network along the HAZ on the side views of the sample [15,24].

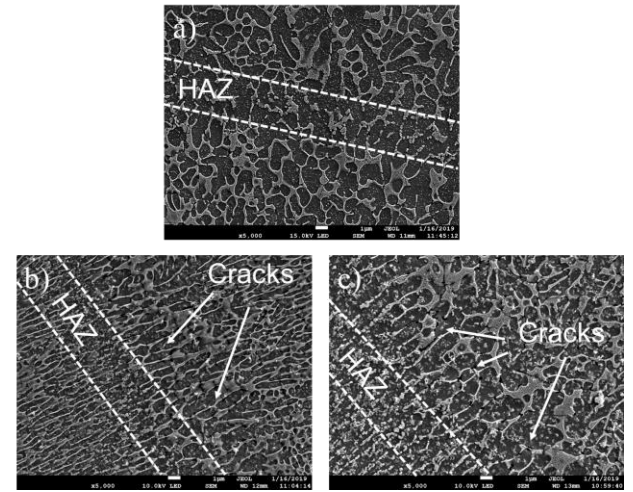


Figure 3. SEM micrographs from the side view of the DMLS-AISi10Mg fabricated by the a) virgin, b) 4X recycled, and c) 5X recycled powder.

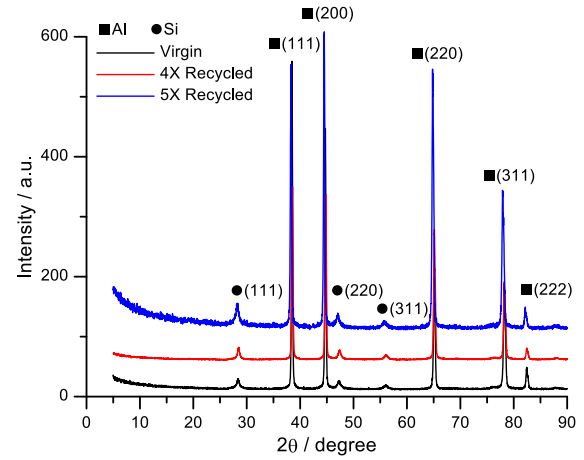


Figure 4. X-ray diffraction spectra of the DMLS-AISi10Mg samples fabricated using virgin and recycled powders.

For this reason, in this study, electrochemical properties of the samples were only measured on the side view planes. As depicted in Fig. 5b, the anodic current was rapidly increased by applying potential, resulting in the metal dissolution at an elevated rate. This represents an active-like behavior for all polished surfaces, and in agreement with the results reported in previous studies [15,26]. Therefore, regardless of the type of the employed metal powder, formation of a passive region was not detected on the polarization graphs of all samples, confirming that surface pitting can immediately happen once the corrosion potential is reached.

Table 2 also shows all the results extracted from the anodic polarization plots shown in Fig. 5b, including the corrosion potential ($E_{Corr.}$), and corrosion current density ($I_{Corr.}$), of each of the fabricated DMLS samples.

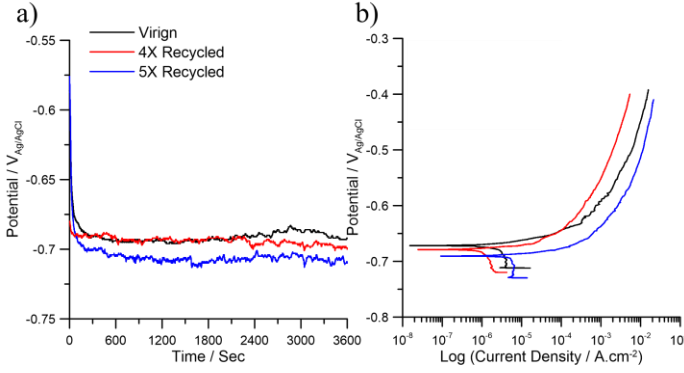


Figure 5. a) Open circuit potential measurement over time, b) anodic polarization curves comparing the corrosion behavior of fabricated DMLS-AlSi10Mg samples using recycled powders versus the virgin powder.

TABLE II. ANODIC POLARIZATION PARAMETERS OF FABRICATED DMLS SAMPLES USING VIRGIN AND RECYCLED POWDERS IN AERATED 3.5 WT.% NaCl

	Virgin	4X	5X
$E_{\text{Corr.}} (V_{\text{Ag/AgCl}})$	-0.674 ± 0.02	-0.679 ± 0.01	-0.693 ± 0.03
$I_{\text{Corr.}} (\mu\text{A}/\text{cm}^2)$	0.642 ± 0.01	0.828 ± 0.02	1.262 ± 0.02

As a general trend, a higher corrosion potential and a lower corrosion current density represent greater corrosion properties and a higher electrochemical stability [20]. The corrosion potential of all three samples were in the same range: -0.674 ± 0.02 , -0.679 ± 0.01 , and -0.693 ± 0.03 $V_{\text{Ag/AgCl}}$ for the Virgin, 4X, and 5X recycled samples, respectively, which is potentially ascribed to the uniformity of the chemical composition on the surface for all samples. However, when the corrosion current density ($I_{\text{Corr.}}$) values are compared, a noticeably higher (~ 2 times) corrosion current density was measured for the 5X recycled sample versus the Virgin sample, and the 4X recycled sample showed an intermediate value. This indicates increased severity of corrosion attack on the surface of the samples in the order of Virgin < 4X recycled < 5X recycled sample, confirming the deterioration, although not significant, of the corrosion performance when recycled powder is used as the feedstock material. The degraded corrosion properties of the recycled powder samples was attributed to the coarser microstructure and more expanded HAZ along the melt pool borders of the recycled samples containing a higher content of eutectic Si phase and lower content of solute Si in α -Al solid solution.

This consequently results in an increased potential difference between the cathodic Si precipitates and anodic α -Al matrix, leading to a greater susceptibility of melt pool boundaries of the recycled samples to galvanic selective attack [15]. To further investigate the stability and protectiveness of the passive layer formed on the fabricated samples over time, EIS tests in aerated 3.5 wt.% NaCl solution were carried out at different immersion times.

C. Electrochemical Impedance Spectroscopy (EIS) Results

Fig. 6 shows the modification of the EIS spectra for all samples over two different immersion times, *i.e.* 1 h and 120 h. The constant high-frequency impedance values in the Bode plots follow the ohmic drop in the electrolyte. However, the impedance differences in middle to low frequency ranges are more evident for different samples. As a general trend, at lower frequency range, the higher modulus of impedance for the Virgin sample confirms a slower kinetic for the corrosion reactions (Fig. 6a). This is in agreement with the anodic polarization results (shown in Fig. 5 and Table 2), corroborating the lowest corrosion current density ($I_{\text{Corr.}}$) for the Virgin sample. The impedance plots in just immersed condition (after 1 h) for both 4X and 5X recycled samples confirmed the lower resistivity of the protective passive film on their surfaces as compared to the Virgin sample. Collectively, all DMLS-AlSi10Mg samples showed one broad capacitive peak in the phase angle *vs.* frequency diagram at the early stage of immersion (shown in Fig. 6a). A closer look at the Nyquist plots of the recycled samples (shown in Fig. 6b) reveals that the broad peak is composed of two different peaks, one in low frequencies and the other in intermediate frequencies, with non-evident time constants [16], which is a typical behavior of passive aluminum.

After a longer immersion time (120 h) the phase angle *vs.* frequency diagrams (Fig. 6c) revealed two distinct capacitive peaks with well-defined time constants in the range of low to intermediate frequencies. The low-frequency peak is assigned to the diffusion within the passive layer and through localized corroded zones, while the high-frequency peak corresponds to the sealing effect of the corrosion products inside surface porosities or other active areas, such as α -Al/Si interface [24]. It is also worth mentioning that the frequencies of the distinct peaks were constant, independent from the sample type. Similarly, two peaks were also formed on the Nyquist plots after 120 h of immersion time (shown in Fig. 6d). Nyquist curvature radius is an indication of the improvement in the corrosion properties of the sample. The first curvature in the Nyquist plot represents the high-frequency peak, and the second one corresponds to the low-frequency peak shown in the phase angle plots of Fig. 6c.

Analogous to the initial immersion time, even after 120 h of immersion time, the Bode plot of the virgin powder sample (Fig. 6c) showed a higher absolute value of impedance than the recycled powder samples at a lower frequency range. This implies the superior protective nature of the passive layer on the Virgin sample against the corrosion degradation even after longer immersion times. Likewise, comparing the Nyquist responses of all three samples after 120 h of immersion time (shown in Fig. 6d) revealed a larger capacitive arc for the Virgin sample, confirming the higher stability of the protective passive film on its surface, leading to its greater corrosion resistance compared to the 4X and 5X recycled samples.

Therefore, although the microstructural variations resulted from applying recycled powders in fabrication of DMLS-AlSi10Mg were not significant, the corrosion testing results presented herein confirmed that even a slight coarsening of interdendritic Si network along the melt pool boundaries of

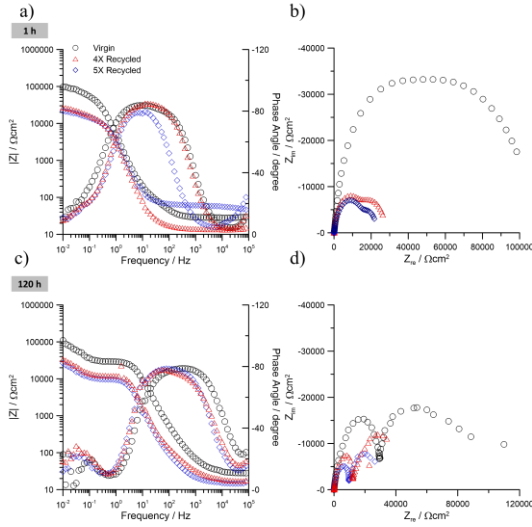


Figure 6. The EIS spectra of the DMLS-AM-ALSi10Mg fabricated using virgin and recycled powders, a) Z modulus, Bode phase angle plot and b) Nyquist spectra after 1 h of immersion time, c) Z modulus, Bode phase angle plot and d) Nyquist spectra after 120 h of immersion time.

DMLS fabricated AM-ALSi10Mg can noticeably affect the corrosion performance of the part, leading to a significant deterioration of its electrochemical stability in chloride containing environments.

D. Intergranular Corrosion Testing

To investigate the morphology of the corrosion on recycled samples and compare them with the Virgin sample, the intergranular corrosion test was conducted, and surfaces of all three samples were studied using a field emission scanning electron microscope. Fig. 7 shows the FESEM images from the exposed surfaces of the Virgin, 4X, and 5X recycled samples after the intergranular corrosion test (as per ISO 11846). As shown in Fig. 7a, a selective corrosion attack preferably along the melt pools borders was detected on the Virgin sample. The selective corrosion attack along the melt pool boundaries of DMLS/SLM-AM-ALSi10Mg is commonly reported in previous studies [17,19,20,26] and is attributed to the enrichment of Si phase in that region, stimulating the micro-galvanic corrosion of the anodic Al matrix along the melt pool boundaries and in particular in HAZ, where coarsening and breakage of Si network into idiomorphic crystals were observed. On the other hand, the recycled samples (Figs. 7b and 7c) behaved differently and revealed formation of several superficial cracks on the Al matrix and particularly in the HAZ and close to the melt pool boundaries after the intergranular corrosion testing. Formation of these corrosion mediated surface cracks in AM-ALSi10Mg products has been also reported in a recent study, and was ascribed to the existence of high level of residual stresses in the as-printed part along with the microstructural evolutions near the melt pool boundaries [25]. Therefore, further coarsening of Si network along the melt pool regions in the recycled samples relative to the Virgin sample and the resulted accelerated dissolution of anodic α -Al matrix from those regions have contributed to the surface cracking of the 4X and 5X recycled samples after intergranular corrosion testing.

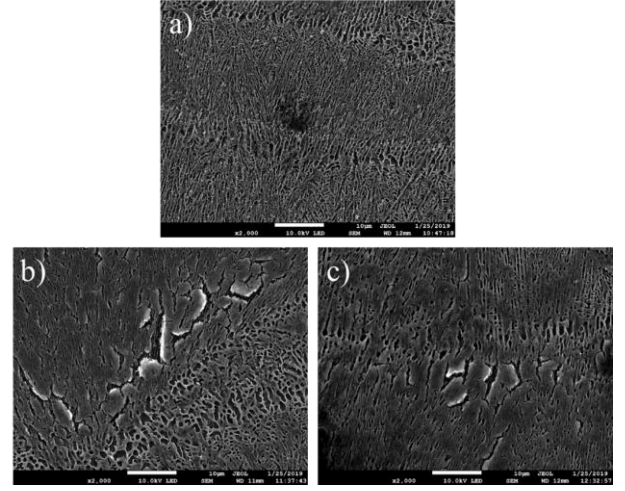


Figure 7. SEM micrographs of the DMLS-AM-ALSi10Mg fabricated from the a) virgin, b) 4X recycled, and c) 5X recycled powder after the intergranular corrosion test.

IV. CONCLUSIONS

With the aim to improve affordability of metal additive manufacturing technology by using recycled powder as the feedstock material instead of highly expensive virgin powder, the effects of using AM-ALSi10Mg recycled powder after 4 times (4X) and 5 times (5X) of re-use cycles on the microstructure and electrochemical properties of the DMLS-produced AM-ALSi10Mg were investigated. Microstructural analysis results confirmed that using the recycled powder in fabrication of the alloy leads to an increased content of eutectic Si and coarsening of the intercellular Si network in the structure, mostly along the melt pool boundaries, accompanied by the formation of higher density of internal discontinuities and defects, including porosities and micro-cracks. Such microstructural variations in the recycled samples were correlated to the larger size and irregular shape of the recycled powders compared to the virgin powders, leading to the reduced thermal conductivity of the recycled powders. The lower thermal conductivity associated with the recycled powders was found to lower the cooling rate of the melt pools during fabrication, causing excessive Si solute segregation during solidification, which was found to increase the solidification cracking acceptability of the alloy. The electrochemical measurements, *i.e.*, anodic polarization and EIS tests, confirmed that the increased eutectic Si content along the melt pool boundaries and more expanded HAZ in the 4X and 5X samples deteriorated the corrosion resistivity in recycled powder samples. The intergranular corrosion test confirmed an accelerated preferential corrosion attack combined with the formation of superficial micro-cracks along the melt pool boundaries in recycled samples dominated by the microstructural evolutions and in particular coarsening of the Si phase in the recycled samples. Therefore, the slight microstructural variations in recycled samples were found to play a key role in dictating the material's corrosion behavior.

ACKNOWLEDGMENT

The authors wish to acknowledge the support of Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number RGPIN-2017-04368], for sponsoring this work.

REFERENCES

- [1] H. Asgari, C. Baxter, K. Hosseinkhani, and M. Mohammadi, "On microstructure and mechanical properties of additively manufactured AlSi10Mg_200C using recycled powder," *Mater. Sci. Eng. A*, vol. 707, pp. 148–158, Nov. 2017.
- [2] Y. Bai, Y. Yang, Z. Xiao, M. Zhang, and D. Wang, "Process optimization and mechanical property evolution of AlSiMg0.75 by selective laser melting," *Mater. Des.*, vol. 140, pp. 257–266, 2018.
- [3] B. A. Fulcher, D. K. Leigh, and T. J. Watt, "Comparison of AlSi10Mg and Al 6061 processed through DMLS," in *Proceedings of the Solid Freeform Fabrication (SFF) Symposium, Austin, TX, USA*, 2014, vol. 46.
- [4] E. O. Olakanmi, "Selective laser sintering/melting (SLS/SLM) of pure Al, Al–Mg, and Al–Si powders: Effect of processing conditions and powder properties," *J. Mater. Process. Technol.*, vol. 213, no. 8, pp. 1387–1405, 2013.
- [5] W. J. Sames, F. A. List, S. Pannala, R. R. Dehoff, and S. S. Babu, "The metallurgy and processing science of metal additive manufacturing," *Int. Mater. Rev.*, vol. 61, no. 5, pp. 315–360, 2016.
- [6] N. T. Aboulkhair, I. Maskery, C. Tuck, I. Ashcroft, and N. M. Everitt, "On the formation of AlSi10Mg single tracks and layers in selective laser melting: Microstructure and nano-mechanical properties," *J. Mater. Process. Technol.*, vol. 230, pp. 88–98, 2016.
- [7] A. H. Maamoun, M. Elbestawi, G. K. Dosbaeva, and S. C. Veldhuis, "Thermal post-processing of AlSi10Mg parts produced by Selective Laser Melting using recycled powder," *Addit. Manuf.*, vol. 21, pp. 234–247, 2018.
- [8] L. C. Ardila *et al.*, "Effect of IN718 Recycled Powder Reuse on Properties of Parts Manufactured by Means of Selective Laser Melting," *Phys. Procedia*, vol. 56, pp. 99–107, 2014.
- [9] G. Nichols *et al.*, "A Review of the Terms Agglomerate and Aggregate with a Recommendation for Nomenclature Used in Powder and Particle Characterization," *J. Pharm. Sci.*, vol. 91, no. 10, pp. 2103–2109, 2002.
- [10] A. Simchi, "Direct laser sintering of metal powders: Mechanism, kinetics and microstructural features," *Mater. Sci. Eng. A*, vol. 428, no. 1, pp. 148–158, 2006.
- [11] K. Abd-Elghany and D. L. Bourell, "Property evaluation of 304L stainless steel fabricated by selective laser melting," *Rapid Prototyp. J.*, vol. 18, no. 5, pp. 420–428, 2012.
- [12] A. C. Hoffmann and H. J. Finkers, "A relation for the void fraction of randomly packed particle beds," *Powder Technol.*, vol. 82, no. 2, pp. 197–203, 1995.
- [13] B. Liu, R. Wildman, C. Tuck, I. Ashcroft, and R. Hague, "Investigation the effect of particle size distribution on processing parameters optimisation in selective laser melting process," *22nd Annu. Int. Solid Free. Fabr. Symp. - An Addit. Manuf. Conf. SFF 2011*, 2011.
- [14] A. Spierings, N. Herres, and G. Levy, "Influence of the particle size distribution on surface quality and mechanical properties in AM steel parts," *Rapid Prototyp. J. - RAPID Prototyp. J.*, vol. 17, pp. 195–202, 2011.
- [15] R. I. Revilla, J. Liang, S. Godet, and I. De Graeve, "Local Corrosion Behavior of Additive Manufactured AlSiMg Alloy Assessed by SEM and SKPFM," *J. Electrochem. Soc.*, vol. 164, no. 2, pp. C27–C35, Jan. 2017.
- [16] M. Cabrini *et al.*, "Effect of heat treatment on corrosion resistance of DMLS AlSi10Mg alloy," *Electrochim. Acta*, vol. 206, no. Supplement C, pp. 346–355, 2016.
- [17] M. Rafieazad, M. Mohammadi, and A. Nasiri, "On Microstructure and Early Stage Corrosion Performance of Heat Treated Direct Metal Laser Sintered AlSi10Mg," vol. Submitted, 2019.
- [18] L. Thijs, K. Kempen, J. Kruth, and J. V. Humbeeck, "Fine-structured aluminium products with controllable texture by selective laser melting of pre-alloyed AlSi10Mg powder," *Acta Mater.*, vol. 61, no. 5, pp. 1809–1819, 2013.
- [19] P. Fathi, M. Mohammadi, X. Duan, and A. Nasiri, "Effects of Surface Finishing Procedures on Corrosion Behavior of DMLS-AlSi10Mg_200C Alloy Versus Die-Cast A360.1 Aluminum," *JOM*, 2019.
- [20] P. Fathi, M. Mohammadi, X. Duan, and A. M. Nasiri, "A Comparative Study on Corrosion and Microstructure of Direct Metal Laser Sintered AlSi10Mg_200C and Die Cast A360.1 Aluminum," *J. Mater. Process. Technol.*, vol. 259, pp. 1–14, 2018.
- [21] X. P. Li *et al.*, "A selective laser melting and solution heat treatment refined Al–12Si alloy with a controllable ultrafine eutectic microstructure and 25% tensile ductility," *Acta Mater.*, vol. 95, no. Supplement C, pp. 74–82, 2015.
- [22] W. Kurz and D. J. Fisher, *Fundamentals of solidification*, vol. 1. trans tech publications Aedermannsdorf, Switzerland, 1986.
- [23] B. Wu, Z. Pan, S. Li, D. Cui, D. Ding, and H. Li, "The anisotropic corrosion behaviour of wire arc additive manufactured Ti-6Al-4V alloy in 3.5% NaCl solution," *Corros. Sci.*, vol. 137, pp. 176–183, 2018.
- [24] M. Cabrini *et al.*, "Evaluation of corrosion resistance of Al–10Si–Mg alloy obtained by means of Direct Metal Laser Sintering," *J. Mater. Process. Technol.*, vol. 231, no. Supplement C, pp. 326–335, 2016.
- [25] T. Rubben, R. I. Revilla, and I. De Graeve, "Influence of heat treatments on the corrosion mechanism of additive manufactured AlSi10Mg," *Corros. Sci.*, 2018.
- [26] M. Cabrini *et al.*, "Corrosion Behavior of Heat-Treated AlSi10Mg Manufactured by Laser Powder Bed Fusion," *Materials (Basel)*, vol. 11, no. 7, 2018.

Prospects and Challenges of Reinforcement Methodologies for 3D Printable Plastics

Aman Preet Singh, Basim El Essawi, Ali Daneji,
Salman Pervaiz

Department of Mechanical and Industrial Engineering,
Rochester Institute of Technology – Dubai Campus,
Dubai, United Arab Emirates
sxcad@rit.edu

Sathish Kannan

Department of Mechanical Engineering, College of
Engineering, American University of Sharjah
Sharjah, United Arab Emirates
skannan@aus.edu

Abstract— Three-dimensional (3D) printing is a prosperous and emerging field which has the potential to boost development related activities by offering freedom in designing a complex engineering product. Besides that, it offers opportunities of waste minimization and product customization at the same time. With the advent of additive manufacturing-based technology, the conventional manufacturing failed in providing the design flexibility as compared to 3D printing technology. 3D printing has gained significant popularity in the biomedical, aerospace, automotive, and construction related industries. However, there are serious concerns about the mechanical behavior of the 3D printed components due to its anisotropic behavior and void formations etc. In case of 3D printable plastics, there are several approaches of providing reinforcements. These reinforcement methodologies can improve the mechanical behavior of 3D printable plastics significantly and make them usable for different demanding applications. This paper reviews the conventional techniques for 3D printing, methods of reinforcements, and the prospects and challenges.

Keywords— 3D printing, Reinforcements, Polymer, Plastics, FDM

I. INTRODUCTION

With the advancement of technologies came a demand for parts, and suitable equipment for advancement. While welding and other novel methods of machining still are relevant and show massive impact on the manufacturing industry, studies done on 3D printing as an alternative have been on the rise and gaining popularity with the passage of time. 3D printing, also called additive manufacturing through many literatures, works precisely like a printer where entries from a 3D imaging or CAD software are sent to the printer where the material gets extracted onto a surface producing an actual product, the extrusion usually takes place on an X-Y CNC gantry [1], which allows the printer to print in layers, based on X-Y co-ordinates, the layers then move down, or the apparatus moves based on

the specifications of the 3D printer, this allows the formation of complex parts with great accuracy. The ease of which a printed part can be made is incomparable to one machined and while machined parts have a tendency to create wasted material 3D printing is concise which does not include any losses making it more cost efficient as well as energy and time efficient [2], but while demand rises the limitations of materials formed by additive manufacturing has become more and more expressed.

Various 3D printing technologies are accessible to us today. Thermoplastics and thermosetting polymer materials are widely used to produce products through existing 3D printing techniques. These printed polymers are used in various fields such as automotive, architectural, medical etc. [3]. The parts printed are limited to conceptual prototypes because these polymeric materials do not provide sufficient mechanical strength and other desired properties to achieve functional parts that can be used in the above fields. This opens up gates to introducing polymer composites to 3D printing for improving performance. As an alternative to classical materials used for 3D printing, new methods of reinforcements can improve material's mechanical, thermodynamic and chemical properties [4]. The reinforcements are summarized in three main fields of polymer matrix composites: particle-composites, fiber-composites, and nanomaterial reinforcements. Adding carbon fiber to the novel PLA allows the formation of carbon fiber reinforced PLA which increases the mechanical and flexural properties of the material and taking into consideration various parameters increases the strengths to exceed certain aluminum alloys [1], this shows potential for development in fields such as aerospace and aviation due to the light weightiness of the materials along with their surprising strength, which is comparable to metals.

The process of compounding the material has numerous possibilities, having the ability to look into substitute materials for industries such as biomedical, civil, agricultural etc. While most literature on process of 3D printing discusses the methodology and advantages of novel Plastics and acrylics formed by 3D printing, more recent studies have taken a

direction toward reinforcing materials and industrial usage. This paper discusses the various methods used in 3D printing, how the materials are formed, the reinforcement methods of materials and their effects followed by their advantages, limitations, and uses in various industries.

II. THREE-DIMENSIONAL PRINTING TECHNOLOGIES

3D printing, also known in industries as additive manufacturing (AM) technologies, uses a layer-by-layer process to produce three dimensional objects. Today, several methods of 3D printing are established and further being developed. These include fused deposition modeling (FDM), stereolithography apparatus (SLA), electron beam melting (EBM), laminated object manufacturing (LOM), selective laser sintering (SLS), digital light projection (DLP), etc. [5]. Out of them, the most established techniques will be discussed further in the paper. These techniques are used for composite manufacture depending on desired material, print resolution to be achieved, processing speed, and cost requirements.

A. Fused Deposition Modelling (FDM)

FDM technology also known as fused filament fabrication (FFF) is the most widespread way to rapid production of products utilizing additive manufacturing. Fig. 1 represents the schematic illustration of FDM. FDM consists of depositing a filament of thermoplastic material. The filament from the reel passes through a hot nozzle, with a temperature higher than the melting point of the filament. The material is further extruded out creating a layer of solid material on the build plate. The model is made through depositing a layer contour and further filling the inside with the material in a certain pattern. After each layer, the head moves along the Z axis to initiate the build for the next layer.

This process can help us create complex shapes much more easily when compared to other manufacturing processes. Process parameters such as head movement, feed rate, layer thickness, infill, slicing etc. can be changed to achieve desired products. Post production finishing is required after the 3D printing process is finished. With all the benefits, the technology comes with a few drawbacks. The major one being hygroscopicity, a property of polymer materials that leads to poor print quality. Due to the influence of nozzle temperature, moisture in the material leads to the creation of air bubbles in the model structure. This drawback can be minimized by storing the filament in appropriate conditions. Another flaw is structural inhomogeneity, which refers to uneven structure particle size or insufficient density of the object. Structural inhomogeneity is caused due to selective deposition of the filament on the print bed, temperature differences, or porosity. This can be minimized through certain techniques [6]. Also, stepped layers and support structure removal can cause quality setbacks. Yet, the technology stands strong due to its flexibility in material selection, low cost, user friendly nature and high speed.

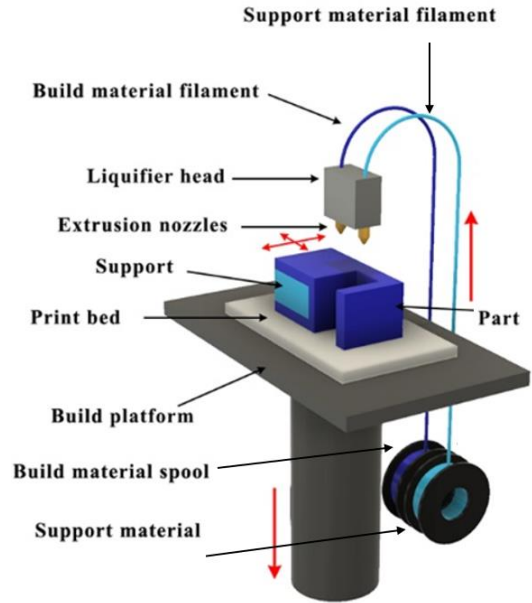


Figure 1. Schematic illustration of Fused Deposition Modelling (FDM)

B. Selective Laser Sintering (SLS)

Selective laser sintering is a technique of 3D printing that uses a powder bed to produce 3D objects. SLS uses a laser for binding the powder particles together. Fig. 2 represents the schematic illustration of SLS. During the process, this laser draws out several patterns onto the surface of the bed. A new layer of powder is distributed by a roller after the first layer is completed. This layer by layer process continues till the final product is ready and can be removed from underneath the production. When compared to FDM, SLS is a one step process as it does not require a prior production of filament and furthermore, offers higher resolution due to the precision of the laser. Commonly used materials are powdered forms of ceramics, plastics, or metal alloys that require high temperature and energy lasers to be sintered [7].

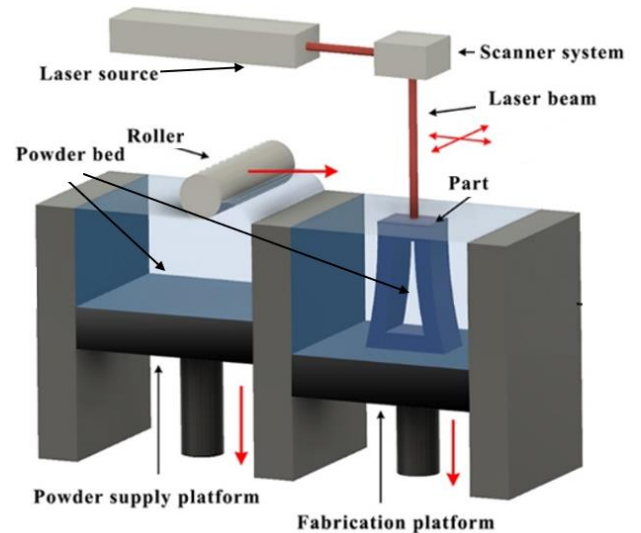


Figure 2. Schematic illustration of Selective Laser Sintering (SLS)

C. Stereolithography Apparatus (SLA)

Stereolithography (SLA), a kind of 3D printing/additive manufacturing and is a Solid Freeform Technique (SFF), it is widely known for its accuracy and high-resolution printing, it uses light to print and require light sensitive materials to as a base, or resin, to print with. Fig. 3 represents the schematic illustration of SLA. The light sensitive material is called photo curable resins and the most common kinds used are acrylic and epoxy resins, the method used in printing is called photo polymerization and the resulting material is said to be made of a photopolymer [3]. The process is conducted using a laser beam usually with Ultraviolet light, which cures the liquid uncured resin on the places specified by the computer program or numerical code, this process like most additive manufacturing processes takes place layer by layer until a whole 3D image is formed. After the desired product is finished the next step would be to clean off the excess liquid resin and wash off the residue, the obtained structure is here after called the “green part” [8], the green part is then taken for post-curing, as after the 3D printing the properties of the material are still not stable, post-curing is done to improve the mechanical properties and strength of the structure [8].

The main parameters to look for in Stereolithography are the power of the light source and the scanning speed of the laser or exposure time in some cases. The accuracy of the SLA makes it a huge boon in the fields of biomedical engineering and the clinical field, forming implants and other surgical devices with sharp accuracy as specified by the CAD or patient Scans, the concerning factors or drawbacks come from the extremely high price which makes it unsuitable for use in the common industry, and the use of UV which may lead to health effects due to the formation of free radicals. Uncured resins may also cause cytotoxicity which is a topic of concern.

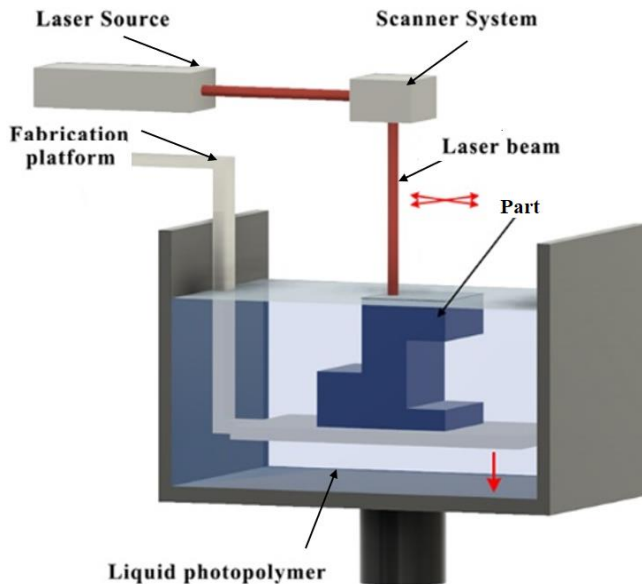


Figure 3. Schematic illustration of Stereolithography Apparatus (SLA)

III. REINFORCEMENT METHODOLOGIES FOR 3D PRINTING

3D printing industries are majorly polymer material based due to their low costs, machining flexibility and light weight. This technology helps achieve manufacture of objects with complex geometry but still could lack in mechanical strength and performance depending on the application. This is why there are very few functional 3D printed parts being used but rather just conceptual prototypes. Recent advancements in metal 3D printing could be a solution to this but due to its high costs and complexity, it may not be an ideal one. To tackle this, combining different materials to achieve desired properties such as mechanical strength, electrical conductivity, thermal stability etc.; proves to be a promising solution. Recent studies in the field of composite manufacture for conventional 3D printing techniques have shown strong results. Methodology involved in producing these special materials to be compatible for 3D printing are discussed in detail below. These can be categorized by reinforcement by particle, nanomaterials or fibers.

A. Particle Reinforced Polymer

Due to advancements in desktop 3D printing, the low costs encourage average users to develop custom filament materials. A cost-effective technique like particle reinforcement in polymers is widely used to achieve desired improved properties. This reinforcement process can be carried out to be in powder, liquid or extruded filament form. Existing data show enhancements in properties such tensile strength, thermal conductivity, thermal degradation temperature, flexural strength, magnetic behavior etc. This shines light on the industrial applications that can be achieved through this process. Majorly metals or nanoparticles are used to improve polymer properties used for 3D printing.

1) Metal/Polymer Composite

There are two major steps in preparation of desired filament material:

a) Material Preparation; Prior to the extrusion process, the particles and polymer are treated together. In the preparation of ABS-Al and ABS-ZrB₂, virgin ABS pellets and reinforcement particles were mixed in a mechanical alloying machine to provide good penetration of the reinforcement particles into ABS pellets [9]. Similarly, in the case of ABS-Cu and ABS-Fe, copper and iron powders were mixed with the thermoplastic ABS. The size of copper and iron particles were comparatively smaller than that of ABS pellets [10]. A novel study conducted using recycled Sm-Co powder, recovered from industrial swarms, loaded in PLA required special preparation technique. PLA was first dissolved in a solvent to obtain a viscous solution, assisted by mechanical homogenization. The recycled Sm-Co powder then was dispersed in this solution by homogenizing. The solvent was evaporated to ensure that the reinforced particles remained dispersed in the binder. Further to perform extrusion, the bonded material was cut into pellets [11]. To obtain pure filaments, polymers and reinforcing particles can be dehydrated to eliminate moisture. Wood flour and PLA required dehydration for about four hours at 103 °C prior to obtaining WF/PLA composite [5]. All studies conducted use varying weight percentages of reinforcement

particles as each composition provides unique results. It is thus important to know this data and its effects on properties of the final product to achieve any desired application.

b) Extrusion; Extrusion of the mixture of the polymer pellets and reinforcement particles takes place in a single/twin screw extruder which operates at very high processing temperatures. The heated material is pushed through a nozzle to extrude the filament to a desired diameter. In the extrusion of Al and ZrB₂ with ABS, die swelling is phenomenon was prevented by setting the drawn rotor speed to an optimum value which had been determined by trial and error process. This helps maintain the consistency of the filament [9].

2) Nanocomposite

Creation of functional composites is possible through mixing nanomaterials into polymers used for 3D printing. Carbon nanotube, graphene nanoplatelets, graphene oxide, montmorillonite etc. are commonly used to provide unique properties especially in the enhancement of mechanical strength of 3D printed parts. The method of preparation is very similar to that of particle reinforced polymer composite.

a) Material Preparation; Special techniques are used to prepare the reinforcing nanomaterial and the polymer. In the case of graphene oxide mixed with ABS, two methods can be applied, either solution mixing or dry mixing. Solution mixing involves dissolving ABS and dispersing GO in DMF. Both are mixed together to obtain ABS/GO composite after solvent evaporation takes place [3]. Similar technique is used in mixing carbon nanotubes with ABS. This is done to avoid inhomogeneity and phase separation [12]. Second method would be dry mixing, involving ABS being mixed mechanically with GO powder [3]. Dry mixing technique is used in preparation of ABS/OMMT nanocomposite as well [13]. Graphene platelets and ABS mixture is prepared with unique technique like compounding. Reinforcement material was melt compounded with ABS and further granulated. This material was then sent for compression molding where it was hot pressed at a high temperature and pressure of 3.9 MPa applied for 10 min at a cooling rate of 20°C to obtain square plaques. Further specimens were built through extrusion process [14].

b) Extrusion; After the material is prepared, it is sent for extrusion process for achieving a 3D printing filament of the desired material. The process is the same as mentioned earlier. ABS/OMMT nanocomposite required a two-step process where first, ABS pellets and OMMT powder undergo a twin-screw extruder and further, this composite undergoes another extrusion through a single screw extruder along with pure ABS pellets. This is done in order to offset the influence on properties caused by mixing and extrusion [13].

B. Fiber Reinforced Polymer

Fiber reinforcement can be very effective in improving material properties. The process of preparation is quite unique and is a bit harder to execute when compared to particle reinforcement. Most common fibers used are CFR, Nylon, Glass, Kevlar etc. Reinforcing polymers and plastics with carbon is a common method to increase the strength and

flexural strength of the material, the novel method of reinforcing carbon into polymers is continuous addition of carbon during 3D printing (FDM), while in many cases Compression Molding is also used, but the method was termed inefficient and not cost-effective [1]. It is prepared by co-extruding carbon fiber from carbon fiber filaments and Polylactic acids (PLA) into an extrusion mold the process is done via a nozzle and heater, commonly called a liquefier, the extruder causes great pressure which allows the molten plastics adhere onto the Carbon Fiber filament [2], [4]. The main effecting factors constraining the reinforcement would be the Hatch spacing, layer thickness, and liquefier temperature [4]. Limiting factors in FDM method include the limit of wt% of carbon fiber, study show CF wt% of above 40% causes clogging the nozzle and while nozzle's can be made wider the feed rate as well as Hatch spacing, and thickness are all duly affected [2], [4]. Additionally, the bonding between Carbon fiber and PLA was discovered to be limited and superficial which didn't allow the maximum potential of the carbon fiber reinforcement [15]

A modified version of the previously discussed CFR PLR, this method was tested to counter the limitation of the weak bonding interface with PLA, the modification occurs before the extrusion where PLA particles are added into Methylene Dichloride where, it is semi dissolved using Magnetic stirring, the solution is then filtered and carbon fiber filaments added to the filtrate, which is then sheared and then emulsified [15]. The result was shown to have less voids and higher bonding area and stronger interfacial strengths as compared to novel CFR PLA. Unlike Carbon reinforced PLA, Glass and Kevlar Fiber reinforcements can be set in Concentric as well as Isotropic settings where the spaces are filled with nylon where concentric is where the fiber is in spirals and isotropic are connected parallel lines, the extrusion methods are similar to that of CFR PLA, due to the limitations of carbon it cannot be set into isotropic shapes possibly due to its high stiffness. While pure nylon samples have high strain values and ductile fractures the fiber reinforcements create a more brittle fracture with strain at 0.1 [1].

IV. MAJOR ADVANTAGES OF REINFORCEMENTS

A. Mechanical Properties

For particle reinforced polymers, copper and iron particles reinforced with ABS bring about a decrease in tensile strength and strain. This is because, the tensile strength of ABS specimen made by FDM processes consists majorly of compressive strength. Thus, this metal/polymer composite could be highly useful in structures requiring resistance to compressive force [10]. An innovative 5 wt. % wood-flour incorporation with polymer show improved compressive strength, just as the mixture of iron particles and ABS [5]. ZrB₂ reinforcement provides ABS with 8.7% improvement in flexural strength and 26% in deflection properties. Relevant engineering applications can utilize these improvements, especially the increase in properties like tensile strength and young's modulus by addition of 2 wt. % graphene oxide into ABS [9]. Nanocomposites like addition of graphene nanoplatelets have shown improvement in tensile modulus of

ABS. 4 wt. % of xGnP increased elastic modulus by 30% and further improves stability under long last loads [14].

Addition of 15% multiwall carbon nanotubes to ABS has shown improvements in mechanical properties. The specimens show 25.6% increase in tensile strength and further 17% increase in flexural strength [16]. When it comes to fiber reinforced polymers, it was discovered that under the certain parameters in temperature, hatch spacing and thickness, and optimal amount of carbon fiber can produce a flexural strength of 335 MPa and flexural modulus of 30 GPa [2]. While Carbon fiber Reinforced PLA has 185% higher and 11% higher tensile strength and flexural strength respectively as compared to PLA, modified CFR PLA was discovered to have 13.8% higher and 164% higher tensile strength and flexural strength respectively, as compared to ordinary carbon fiber reinforcement [15]. The Isotropic formation showed to have an elasticity modulus 20% and 21% higher than the concentric formation for Glass fiber and Kevlar fiber respectively, as well as a 6% and 9% greater increase in tensile strength. While concentric pattern was found to be better in flexibility having a 8.8% and 44% advantage over isotropic formations in flexural modulus. The overall results showed the fiber reinforces nylon to have strength's higher than even some aluminum alloys, having the potential to replace them in certain aspects within the industry [1].

TABLE I. PROPERTY COMPARISON BETWEEN THREE MATERIALS [15]

Properties	PLA	Carbon Fiber Reinforced PLA	Modified Carbon Fiber Reinforced PLA
Tensile strength (MPa)	28	80	91
Flexure strength (MPa)	53	59	156
Storage modulus (GPa)	1.22	0.72	3.25
Loss tangent	2.32	1.52	1.32
Tg (°C)	63.6	65.2	66.8

B. Thermal Properties

A significant 41% increase in thermal conductivity with 50 wt. % copper powder caused by a drop in the coefficient of thermal expansion, provides opportunity to produce electromagnetic products with no distortion in the final product [10]. Further, ABS/OMMT nanocomposite with 5 wt. % OMMT provides thermal stability to ABS. This is confirmed by the author through results which showed the onset of thermal decomposition of the composite, which shifted towards higher temperature range than that of neat ABS [13]. As higher amounts of carbon fiber wt% is added the thermal stability in increased, while there is a limit on the amount of carbon fiber added, modified CFR PLA shows a 2.4% higher Tg value as compared to normal CFR PLA and a 5% higher value when compared to PLA [1].

C. Electrical Properties

A significant decrease in surface resistance also opened doors for electronic applications. FDM process increases electrical conductivity of the 3D printed fibers. Significant reduction of the resistivity of the composite depends on orientation of nanoparticles during extrusion. A 6 wt. % CNT mixed with ABS appeared to be best for electro-conductive applications as explained through its avoidance of thermal degradation during prolonged voltage application [17].

D. Magnetic Properties

A recent study explored the improvement in magnetic properties by utilizing recycled Sm-Co powder recovered from industries. This novel composite is obtained by dissolving PLA polymer into a solvent and further, dispersing Sm-Co powder into this solution. Further on extruding, with 20 vol. % Sm-Co powder, significant improvement in magnetic properties were observed, due to minimized particle to particle interaction and particle rotation. Good uniformity of magnet filament was also demonstrated [11].

V. SUMMARY AND FUTURE OUTLOOK

In this study, the most popular types of 3D printing were reviewed, and the limitations of the materials used to print analyzed, in particular the study analyzes the possible reinforcements which can be done to the materials to enhance mechanical properties to be able to compare to, or exceed, industrial grade materials such as aluminum alloys, and replace them. The main kinds of composites reviewed were fiber reinforced composites, particle reinforced composites, and nanocomposites. Their advantages lay in the fact they complement the corresponding plastic/polymer they are used with such as nanocomposites in SLA due to the ease of curing of nanocomposites. The reinforcement tended to have a positive effect on the materials improving tensile and flexural properties, some materials even reaching the values used in industrial grade. Even then, several limitations still hinder the usage of 3D printing technology for heavy industrial purposes. Although the materials seemed to have enough strength the discerning factor lay within the microstructure, where many voids were found, which occurred during reinforcement due to poor bonding, these voids cause issues such as premature fracture, which in certain uses could be disastrous, thus the lack of trust within industries which translates into less products using 3D printing.

Material selection for any desired application is key for a good engineering design. Feasible 3D printing techniques currently provide limited options in terms of materials and there is a strong need to diversify them. More application-based research is required to understand reinforcements, ideal compositions, effect on properties etc. Recycling materials require exploration as sustainable 3D printing can reduce costs and environmental impact. Furthermore, this could add on to reducing plastic disposal due to waste produced during 3D printing in the form of failed parts or excess supports. There have been deviations seen in the homogeneity of the parts printed which can affect consistency and performance during large scale manufacturing. Tools to bring about uniformity in

properties can be investigated. In the discussed studies, FDM has been widely used to perform tests on polymer composite materials, due to higher costs of production involved with other techniques. More techniques can be explored specific to composite printing that would be feasible and provide better quality with respect to precision of build and surface finish. Post processing activities can also play an important role in enhancement of properties of printed parts [4]. Even with few setbacks, research in the field of 3D printing of polymer composites is still expanding showing great progress in the past few years, newer fields of research attributing to material selection, and performance related to 3D printing are upcoming. This paper offers a concise starting point of view from which researchers may use this data to expand on their own research.

REFERENCES

- [1] A. N. Dickson, J. N. Barry, K. A. McDonnell, and D. P. Dowling, "Fabrication of continuous carbon, glass and Kevlar fibre reinforced polymer composites using additive manufacturing," *Addit. Manuf.*, vol. 16, 146–152, 2017.
- [2] X. Tian, T. Liu, C. Yang, Q. Wang, and D. Li, "Interface and performance of 3D printed continuous carbon fiber reinforced PLA composites," *Compos. Part A Appl. Sci. Manuf.*, vol. 88, 198–205, 2016.
- [3] C. Aumnate, A. Pongwisuthiruchte, P. Pattananuwat, and P. Potiyaraj, "Fabrication of ABS / Graphene Oxide Composite Filament for Fused Filament Fabrication (FFF) 3D Printing," *Advances in Materials Science and Engineering*, Vol 2018, Article ID 2830437, 9, 2018.
- [4] X. Wang, M. Jiang, Z. Zhou, J. Gou, and D. Hui, "3D printing of polymer matrix composites: A review and prospective," *Compos. Part B Eng.*, vol. 110, 442–458, 2017.
- [5] Y. Tao, H. Wang, Z. Li, P. Li, and S. Q. Shi, "Development and application of wood flour-filled polylactic acid composite filament for 3d printing," *Materials (Basel)*, vol. 10, no. 4, 1–6, 2017.
- [6] K. Bryll, E. Piesowicz, P. Szymański, W. Ślaczka, and M. Pijanowski, "Polymer Composite Manufacturing by FDM 3D Printing Technology," vol. 2006, 0–6, 2018.
- [7] F. Fina, A. Goyanes, S. Gaisford, and A. W. Basit, "Selective laser sintering (SLS) 3D printing of medicines," *Int. J. Pharm.*, vol. 529, no. 1–2, 285–293, 2017.
- [8] F. P. W. Melchels, J. Feijen, and D. W. Grijpma, "A review on stereolithography and its applications in biomedical engineering," *Biomaterials*, vol. 31, no. 24, 6121–6130, 2010.
- [9] E. Cantı and M. Aydın, "Effects of micro particle reinforcement on mechanical properties of 3D printed parts," *Rapid Prototyp. J.*, vol. 24, no. 1, 171–176, 2018.
- [10] S. Hwang, E. I. Reyes, K. sik Moon, R. C. Rumpf, and N. S. Kim, "Thermo-mechanical Characterization of Metal/Polymer Composite Filaments and Printing Parameter Study for Fused Deposition Modeling in the 3D Printing Process," *J. Electron. Mater.*, vol. 44, no. 3, 771–777, 2015.
- [11] H. A. Khazdozian, J. S. Manzano, K. Gandha, I. I. Slowing, and I. C. Nlebedim, "Recycled Sm-Co bonded magnet filaments for 3D printing of magnets," *AIP Adv.*, vol. 8, no. 5, 2018.
- [12] K. Gnanasekaran, T. Heijmans, S. Van Bennekom, H. Woldhuis, and S. Wijnia, "3D printing of CNT- and graphene-based conductive polymer nanocomposites by FDM SL," vol. 9, pp., 21–28, 2017.
- [13] Z. Weng, J. Wang, T. Senthil, and L. Wu, "Mechanical and thermal properties of ABS/montmorillonite nanocomposites for fused deposition modeling 3D printing," *Mater. Des.*, vol. 102, 276–283, 2016.
- [14] S. Dul, L. Fambri, and A. Pegoretti, "Composites: Part A Fused deposition modelling with ABS – graphene nanocomposites," vol. 85, 181–191, 2016.
- [15] N. Li, Y. Li, and S. Liu, "Journal of Materials Processing Technology Rapid prototyping of continuous carbon fiber reinforced polylactic acid composites by 3D printing," *J. Mater. Process. Tech.*, vol. 238, 218–225, 2016.
- [16] E. Reddy Cholleti and I. Gibson, "ABS Nano Composite Materials in Additive Manufacturing," *IOP Conf. Ser.: Mater. Sci. Eng.* 455 012038, 2018.
- [17] S. Dul, "Filaments Production and Fused Deposition Modelling of ABS/Carbon Nanotubes Composites," *Nanomaterials*, vol. 8, no. 1, 49, 2018.

ACCURATE ESTIMATION OF PROCESS PARAMETERS IN FUSED DEPOSITION MODELING

Hossein Gohari, Ahmad Barari, Hossam Kishawy

Faculty of Engineering and Applied Science, University of Ontario Institute of Technology (UOIT)
Oshawa, Ontario, Canada

hossein.goahribahabadi@uoit.ca , ahmad.barari@uoit.ca , hossam.kishawy@uoit.ca

Abstract— Additive manufacturing (AM) has brought the flexibility to manufacture products with complexities in geometry, material, and structure. However, the accuracy and repeatability of AM processes require improvement to be used in precise industrial applications. In this paper, a new model is proposed that predicts shapes of strands in fused deposition modeling (FDM) process. In addition to calculate the layer thicknesses, the model can estimate layer widths which is critical especially for the surface integrity of the final part. To verify the model, numerical simulations and experimental tests have been done on five different layer thicknesses. The FDM process is simulated in computational fluid dynamics (CFD) to estimate the shape of the strand considering the process parameters including the gap size (layer thickness), the speed of substrate and the speed of substrate with respect to the nozzle. Results show that the developed model has better compatibility with the process of deposition specifically when the model is implemented to generate machine instructions for an adaptive slicing system.

Keywords—component; Additive manufacturing, Computational Fluid Dynamics (CFD), Adaptive slicing, Fused Deposition Modeling (FDM)

I. INTRODUCTION

Manufacturing complex geometries is highly demanded for the current industrial and medical applications due to the progresses in new design tools such as topology optimization which results in highly complex surfaces. AM processes have significant flexibilities regarding the geometric and material complexities in product development [1]–[3]. However, there are still some concerning issues in employing these processes in manufacturing systems. One of the most concerning issues which demands for the immediate attention of researchers is the surface quality of final products [3]–[6]. Modelling [7], controlling [3], [8], and correction [9] of surface roughness have been the focus of many researchers recently. In addition to the limitations in AM processes, their flexibilities mentioned above also needs to be considered in design of AM products. AM processes can easily fabricate complex and organic shapes which are topologically optimized for the desired functionality in design [10], [11]. Therefore, research on design topology

optimization considering the flexibilities and constraints imposed by various AM processes is necessary.

In AM processes, parts are created using the approximation of the part from two-dimensional (2D) layers. The accuracy of the final part is dependent on the accuracy of the stepped approximation and the accuracy of the deposition process of the 2D layers. Different approaches have been introduced to slice the part to increase the dimensional accuracy of the final part with respect to the manufacturing limitations [4], [5], [12]. However, the accuracy in creation of each layer needs more investigation.

The cross-section shapes of a strand is important because boundaries of strands are the elements that create the surfaces of the final part. The most common shape that has been considered for the cross-section of a strand is a rectangle with two semi-circles at its right and left sides. This shape is created when the distance between the nozzle and the substrate is less than the nozzle diameter. When the distance between the nozzle and substrate is bigger than nozzle diameter the shape of strand approaches to a cylinder with a radius equal to nozzle diameter. Cylindrical strand can be appeared in the process of deposition when the gap distance is bigger than the nozzle diameter and the velocity of the substrate is equal or bigger than the average velocity of the flow inside the nozzle.

The main approach to estimate the setting parameters of the deposition process is to use empirical models because of the difficulties in numerical simulation of the extrusion process in terms of estimating the mechanical properties, fluid dynamics of the flow and temperature-induced deformation during solidification. The empirical models are mostly developed for AM processes using uniform layer thicknesses. The uniform thickness approach is also adopted because of the difficulties that the variable layer thicknesses approach has in terms of changing the setup for applying, curing or sintering material [8], [13]. However, based on the current developments, the number of machines that have the ability to build a part using variable layer thicknesses has been increased. Adaptive selection of layer thicknesses significantly reduces the chordal error and the number of layers. Therefore, a model that can predict the setting parameters for the deposition process of the variable layer thicknesses are significantly required.

The other important information that should be extracted from the printed surface is the amount of deviations from the original surface. Different methods have been introduced to evaluate the deviations. Cusp height [14] and cusp volume [15] and surface roughness evaluations algorithms are the methods to calculate the deviations. The accuracy of the algorithms depends on the accuracy in estimating the cross-sections.

Comminal et al. [16] developed a numerical model to simulate the process of deposition in extrusion-based additive manufacturing. They considered two parameters to control the deposition process which are: 1) the gap distance between the nozzle and substrate; 2) the ratio between the average velocity of the flow inside the nozzle and the velocity of the substrate. They also found that the printing forces of the molten flow on the substrate have negative linear relationship with the velocity ratio. In the next step, they investigated the accuracy of their numerical model with the results of experiments [17]. However, the numerical result can be directly used in the process of the machine code generation.

In this paper, a model has been introduced which estimates the process parameters of deposition in FDM while the layer thicknesses and layer widths can be controlled. This model is highly valuable specifically for the AM processes using variable layer thickness. The main parameters that are considered for the simulations are: 1) gap distance between nozzle and substrate 2) speed of substrate with respect to the nozzle 3) the average speed of the flow inside the nozzle. To verify the model, numerical and experimental simulations have been implemented for the FDM process.

II. PROBLEM DEFINITION

Fused deposition modeling (FDM) also called fused filament fabrication (FFF) is one of the most accessible and easy to use AM machines that can be used to manufacture complex geometries. Part creation in this method is based on putting material on predefined positions extracted from a sliced CAD model. The space between the layers are filled by the molten material. The accuracy of the final part is dependent on the accuracy of the process of slicing and the accuracy of the deposition of material. To increase the accuracy and controllability of the deposition, a model is required to estimate the shape of the cross-section of a strand based on the gap size and the speed of flow inside the nozzle and the speed of nozzle on the path.

G-code is the common programming language to generate a path code for an AM machine. To generate a straight strand following G-code line can be used:

G01 X () Y () E ()

Considering the relative coordinate system, the values for X, Y are defined based on the length of the strand and the value of E is the length of the filament that needs to be fed to fill the gap for the determined length. Fig. 1 shows a schematic for the deposition process of a strand in FDM. The ideal cross-section of a strand for the FDM process is considered to be a rectangle with two semi-circles at its both left and right sides, Fig. 2. For the process printing based on this model, a value E which is the length of the filament needs to be calculated. This length fills a

path with the length of L. Therefore, the E value can be calculated from the equivalency of the volume of the fed filament and the volume of the strand.

$$\frac{\pi E D^2}{4} = \left(t d + \frac{\pi t^2}{4} \right) L \quad (1)$$

Then:

$$E = \left(t d + \frac{\pi t^2}{4} \right) \frac{4L}{\pi D^2} \quad (2)$$

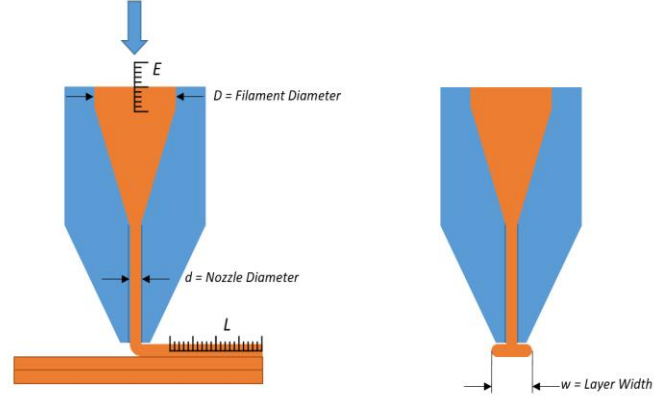


Figure 1. Path-planning parameters for the deposition of a strand in the process of FDM

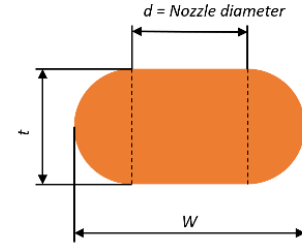


Figure 2. Basic cross-section model for nozzle path generation

III. NUMERICAL SIMULATION

Fluid simulation is based on finding the velocity field for fluid defined in a finite region. A model is defined to simulate the fluid flow of the extruded material from inside the nozzle to the steady state on the substrate using a CFD method. The fluid flow is governed by the dominant laws for the fluid dynamics of a highly viscous material which is the conservation of mass and momentum equation. The conservation of mass, the continuity equation, is defined as the equality between the rate of mass flux out of an element and the rate of reduction of mass inside the element. The conservation of momentum, the Cauchy equation, can be expressed as the equality of the rate of the change of momentum of an element and the net forces on the element. By considering the fluid as incompressible, $\rho = \text{constant}$, and the constitutive relations, an incompressible fluid under isothermal, creeping, steady flow conditions can be governed based on the following equation.

$$\rho \left(\frac{\partial \vec{v}}{\partial t} + \vec{v} \cdot \nabla \vec{v} \right) = -\nabla p + \nabla \tau \quad (3)$$

Where:

$\frac{\partial \vec{v}}{\partial t} = 0$ is the temporal acceleration which considered zero, $\vec{v} \cdot \nabla \vec{v}$ is the speed convection, $-\nabla p = \rho \vec{g}$ and $\nabla \tau = \mu \nabla^2 \vec{v}$ are the volumetric forces. p is the local pressure and g is the gravitational body force. The accelerations are neglected and the flow considered laminar.

The simulation is implemented in the software ANSYS® R19.2 based on the volume fraction model. Hexahedron (hexa) elements defined to mesh the flow simulation region. The element sizes are calculated through an adaptive method and the step time considered as $ST = 0.008$ s. Fig. 3 shows the geometry of the flow simulation region and the mesh defined for the region. We investigated the effect of gap distance on the shape of the boundary of a strand. A constant volume rate is considered to fill the gap while there is moving plane under the gap.

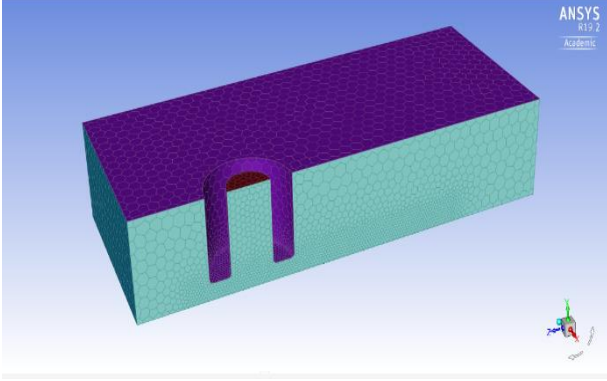


Figure 3. Flow simulation domain inspired from [16]

IV. PROPOSED MODEL

The part creation in FDM is based on the spatial movements of the nozzle and one-dimensional movement of the molten material from inside of the nozzle. As discussed, the speed of the filament is evaluated based on the amount of the material that needs to be deposited for the gap between the slices. Preliminary results showed that the boundary of a strand always is a circle with diameter equal to the layer thickness (Fig. 5). This phenomena can be described by the law of minimal surfaces for fluid flows. Based on this observation and the law of the equivalency of volumes before and after the deposition following model is introduced:

$$E = \frac{\left(tb + \frac{\pi t^2}{4} \right)}{\frac{\pi D^2}{4}} L \quad (4)$$

In the proposed model instead of considering the horizontal edge of the rectangle equal to the diameter of nozzle (d), a new parameter (b) is defined that can be calculated from the

equivalency of the volumes. A schematic simulation for different strand cross-sections is shown in Fig. 5.

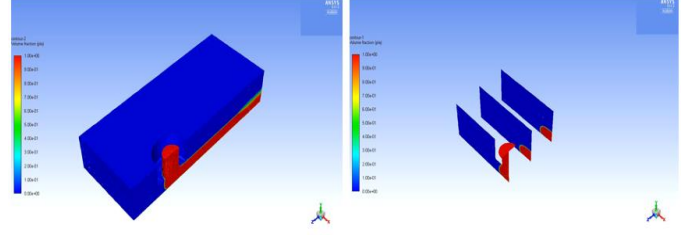


Figure 4. Numerical results of a FDM process when the gap size is $t=0.225$ mm

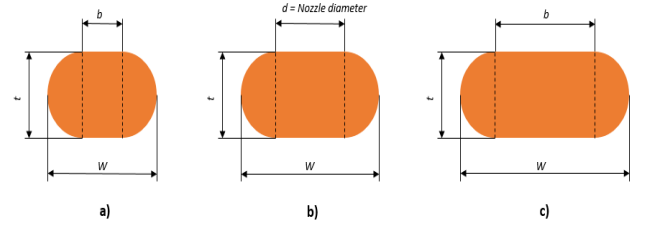


Figure 5. Different shapes of the ideal cross-section defined by the proposed model; a) the fed material is less the ideal cross-section, b) the fed material is equal to the ideal model, c) when the fed material is more than the ideal model

V. NUMERICAL AND EXPERIMENTAL RESULTS FOR THE PROPOSED MODEL

To investigate the accuracy of the numerical results, five different gap sizes have been considered for the experimental tests. These gap sizes cover the common layer thicknesses that can be used for part creation using a FDM machine. The relation between the length path (L) and the length filament that should be fed to fill the path (E) can be derived based on the equivalency of the volume before and after deposition. Therefore, $E = (\text{area of the cross-section of a strand} / \text{area of the cross-section of filament}) \times L$. The ratio between the speed on the path and the average speed of the flow of the filament is the same as the ratio between E and L . For all tests the speed of substrate with respect to the nozzle considered as $F_{path} = 400$ mm/min. Based on the explained relations and the proposed model, the average speed of flow at the nozzle tip can be calculated as:

$$F_{nozzle} = \frac{\left(tb + \frac{\pi t^2}{4} \right)}{\frac{\pi d^2}{4}} F_{path} \quad (5)$$

Using the derived equations, the same parameters value considered for both numerical and experimental simulations. Fig. 6 shows the results of both approaches to construct strand cross-sections. The error between the numerical and experimental simulations are about 8% in average and in rare cases is as high as 14%. This observation shows the capability of the model to predict the process parameters. It is worth mentioning that this discrepancy is not necessarily due to

uncertainties in the developed model. There are many sources of uncertainties in the manufacturing process, the used material, and the inspection process in the experimental study that potentially cause discrepancy. Considering the distribution of the discrepancies in all the observations, it can be concluded that the overall the estimation error using the developed model is less than 10%.

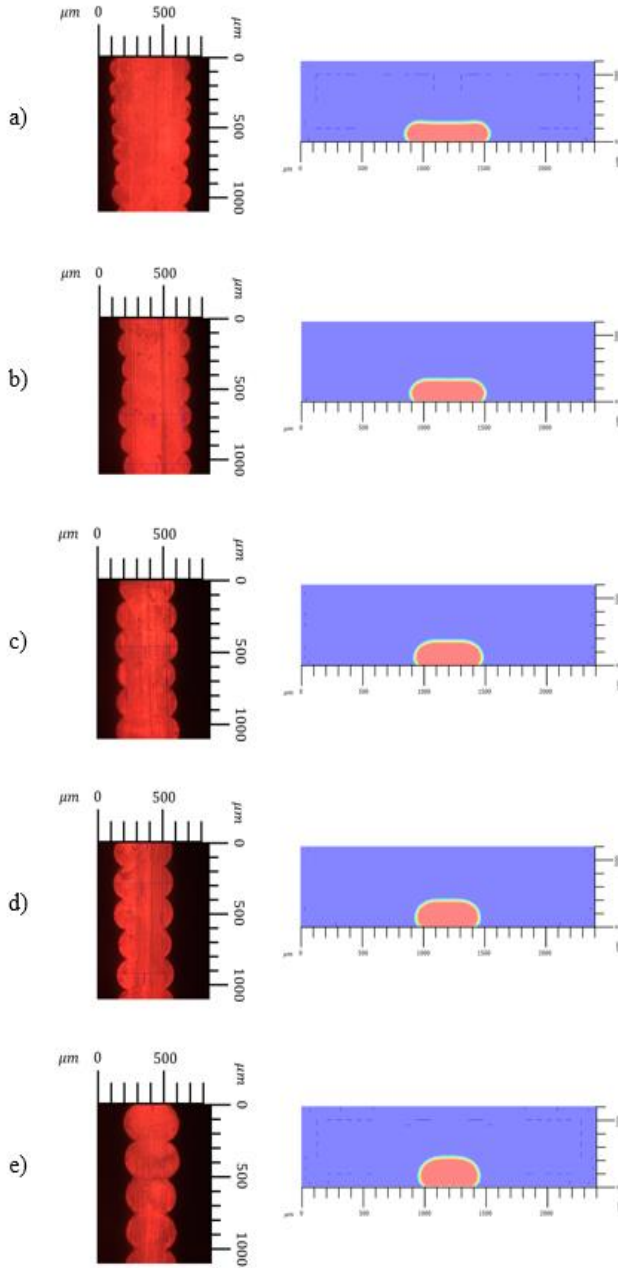


Figure 6. Experimental and numerical results of strand cross-sections in FDM for layer thicknesses of: a) 0.15 mm – Err-ave=14%; b) 0.175 mm – Err-ave=9%; c) 0.2 mm – Err-ave=10%; d) 0.225 mm – Err-ave=6%; e) 0.25 mm – Err-ave=5%;

VI. CONCLUSION

In layer-based AM processes, built time and surface quality are highly dependent on the layer thicknesses. Adaptive layer thickness approaches are cost and time efficient specifically for complex surfaces. In order to implement the variable layer thickness approach without jeopardizing the dimensional accuracy and surface quality of the final part, a model is needed that continuously determines the process parameters for strands with different widths and thicknesses. In this paper a model is developed based on experimental study and computational fluid dynamics to continuously estimate strands widths and thicknesses. For the verification of the results, numerical and experimental approaches are systematically conducted to find the strand cross-sections based the process parameters derived from the proposed model. The results showed that the model can predict the process parameters with less than 10% discrepancy from with the experimental results.

ACKNOWLEDGMENT

The research support provided by the Natural Science and Engineering Research Council of Canada (NSERC) is greatly appreciated.

REFERENCES

- [1] S. Jamiolahmadi and A. Barari, "Surface Topography of Additive Manufacturing Parts Using a Finite Difference Approach," *J. Manuf. Sci. Eng.*, vol. 136, no. 6, p. 061009, Oct. 2014.
- [2] A. Lalehpour and A. Barari, "A more accurate analytical formulation of surface roughness in layer-based additive manufacturing to enhance the product's precision," *Int. J. Adv. Manuf. Technol.*, vol. 96, no. 9–12, pp. 3793–3804, Jun. 2018.
- [3] S. Sikder, A. Barari, and H. A. Kishawy, "Global adaptive slicing of NURBS based sculptured surface for minimum texture error in rapid prototyping," *Rapid Prototyp. J.*, vol. 21, no. 6, pp. 649–661, Oct. 2015.
- [4] H. Gohari, A. Barari, and H. Kishawy, "An efficient methodology for slicing NURBS surfaces using multi-step methods," *Int. J. Adv. Manuf. Technol.*, vol. 95, no. 9–12, pp. 3111–3125, Apr. 2018.
- [5] H. Gohari, A. Barari, and H. Kishawy, "Using Multistep Methods in Slicing 2 ½ Dimensional Parametric Surfaces for Additive Manufacturing Applications," *IFAC-PapersOnLine*, vol. 49, no. 31, pp. 67–72, Jan. 2016.
- [6] A. Barari, H. A. Kishawy, F. Kaji, and M. A. Elbestawi, "On the surface quality of additive manufactured parts," *Int. J. Adv. Manuf. Technol.*, vol. 89, no. 5–8, pp. 1969–1974, Mar. 2017.
- [7] R. Paul and S. Anand, "A new Steiner patch based file format for Additive Manufacturing processes," *Comput. Des.*, vol. 63, pp. 86–100, Jun. 2015.
- [8] M. T. Hayasi, B. Asiabanpour, M. T. Hayasi, and B. Asiabanpour, "A new adaptive slicing approach for the fully dense freeform fabrication (FDFF) process," *J. Intell. Manuf.*, vol. 24, pp. 683–694,

- 2013.
- [9] A. Lalehpour and A. Barari, "Post processing for Fused Deposition Modeling Parts with Acetone Vapour Bath," *IFAC-PapersOnLine*, vol. 49, no. 31, pp. 42–48, Jan. 2016.
- [10] D. Bender and A. Barari, "Convergence Control For Topology Optimization," in *Progress in Canadian Mechanical Engineering*, 2018.
- [11] D. Jankovics, H. Gohari, M. Tayefeh, and A. Barari, "Developing Topology Optimization with Additive Manufacturing Constraints in ANSYS®," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1359–1364, Jan. 2018.
- [12] S. Sikder, A. Barari, and H. A. Kishawy, "Effect of Adaptive Slicing on Surface Integrity in Additive Manufacturing," in *Volume 1A: 34th Computers and Information in Engineering Conference*, 2014, p. V01AT02A052.
- [13] A. Dolenc and I. Mäkelä, "Slicing procedures for layered manufacturing techniques," *Comput. Des.*, vol. 26, no. 2, pp. 119–126, Feb. 1994.
- [14] A. Dolenc, I. M.-C.-A. Design, and undefined 1994, "Slicing procedures for layered manufacturing techniques," *Elsevier*.
- [15] C. Kumar and A. Roy Choudhury, "Volume deviation in direct slicing," *Rapid Prototyp. J.*, vol. 11, no. 3, pp. 174–184, Jul. 2005.
- [16] R. Comminal, M. P. Serdeczny, D. B. Pedersen, and J. Spangenberg, "Numerical modeling of the strand deposition flow in extrusion-based additive manufacturing," *Addit. Manuf.*, 2018.
- [17] M. P. Serdeczny, R. Comminal, D. B. Pedersen, and J. Spangenberg, "Experimental validation of a numerical model for the strand shape in material extrusion additive manufacturing," *Addit. Manuf.*, 2018.

Effect of heat treatment holding time on the microstructure and mechanical properties of Inconel 718 superalloy fabricated by selective laser melting

Eslam Fayed and Mamoun Medraj

*Mechanical, Industrial and Aerospace Engineering
Department*

*Concordia University
Montreal, QC H3G 2W1, Canada*

{e_fayed & mmedraj}@encs.concordia.ca

Davood Shahriari, Mohammad Jahazi and Vladimir
Brailovski

*Department of Mechanical Engineering
École de Technologie Supérieure
Montreal, QC H3C 1K3, Canada*

{davood.shahriari, mohammad.jahazi &
vladimir.brailovski}@etsmtl.ca

Abstract— Improving the performance of parts manufactured using the selective laser melting (SLM) process has been receiving increasing interest especially in the aerospace industry. Homogenizing the as-printed microstructure and mitigating manufacturing defects are among the research current focal points to achieve the required performance. In the present work, Inconel 718 (IN718) samples were fabricated using SLM, and the effect of homogenization and solution heat treatments on the microstructure and precipitates using different dwell times were investigated. The dwell time of the homogenization heat treatment at 1080°C was selected to be 1h and 7h, while for solution heat treatment at 980°C was 15 and 60 minutes. Scanning Electron Microscope (SEM), Energy Dispersive Spectrometer (EDS) and X-Ray Diffractometer (XRD) were employed to characterize the as-printed and heat treated samples. The as-printed condition exhibited a fine dendritic microstructure with a high concentration of refractory elements (Nb and Ti) segregated along the grain boundaries. Increasing the homogenization dwell time resulted in more back diffusion of these segregated elements. In addition, complete recrystallization followed by grain growth was observed after 7h homogenization treatment. However, coarse precipitates and higher carbides content along the grain boundaries and within the grains were observed after longer homogenization (7h). While in case of solution heat treatment, the amount of δ -phase was proportional to the dwell time. The presence of δ -phase at the grain boundaries has some beneficial effects such as pinning of grain growth and enhancing the mechanical properties.

Keywords-Additive manufacturing; selective laser melting; Nickel-based superalloy; microstructure; mechanical properties

I. INTRODUCTION

Over the past four decades, IN718 superalloy experienced outspread evolution in different industrial applications such as aerospace and energy industries [1]. This is due to its outstanding properties such as high strength, high corrosion resistance at both low and elevated temperatures up to 650°C, as well as a very limited expansion coefficient in the

temperature range of 0°C – 100°C [2, 3]. Hence, the average usage of this alloy in modern turbine engine can reach up to 50% by weight [4]. Solid solution strengthening and precipitation hardening are the principle strengthening mechanisms for IN718 superalloys [5]. It is a multiphase superalloy that forms face-centered-cubic (FCC) γ' (Ni₃ (Al, Ti)) and body-centered-tetragonal (BCT) γ'' (Ni₃Nb) phases, which are the primary strengthening phases. In addition, orthorhombic δ -Ni₃Nb, Laves and NbC phases, which are known to degrade tensile strength, creep and fatigue properties, form due to the segregation of some refractory elements such as Nb and Ti [6]. The microstructure and mechanical properties of IN718 superalloy are sensitive to the types, contents, distribution, and size of these precipitations [7]. Inconel 718 is also well known for its good weldability due to its low aluminum and titanium contents which is the main reason rendering it suitable to be processed via selective laser melting (SLM) process [8].

SLM process is one of the additive manufacturing techniques that utilize a layer-by-layer fabrication method to produce solid components using a high-intensity laser beam as a heat source. SLM has gained significant attention in the processing of advanced engineering materials such as aluminide, titanium alloys and nickel-based alloys [2]. Therefore, SLM process has been widely used to fabricate critical metallic parts in various industries such as aerospace, automotive and medical applications [9]. This is due to its great potential to fabricate complex shape components and hollow parts without the need for specific tooling. Besides, the SLM process offers other advantages such as a reduction in manufacturing time, high material utilization and low production cost compared to other conventional subtractive manufacturing methods [10-12]. Moreover, dense metallic components with a high relative density that can reach up to 99.7% at the optimum laser process parameters are fabricated using SLM [13]. However, there are some drawbacks that should be overcome to fully realize and maximize the SLM process potential. Residual stresses, anisotropy and elemental segregation are considered the main drawbacks of SLM which

is attributed to the high thermal gradient and the high cooling rate during solidification process [2, 14-17]. In order to achieve the required microstructure and mechanical properties, post-processing heat treatment are required to obtain homogenous microstructure and defects free components [6].

Numerous studies have been conducted to investigate the effect of heat treatments on the microstructure and mechanical properties of the SLM IN718 components, however, the optimum post heat treatment window has not been established yet. Previous recent reports [1, 18, 19] showed that the homogenization heat treatment at 1080°C for one hour was not enough to completely dissolve the segregated elements Nb and Ti, which resulted in the formation of brittle intermetallic Laves phase. Laves phase is considered one of the most detrimental phases on the mechanical properties of IN718 [20]. They [1, 18] also reported that the solution heat treatment at 980°C for 60 minutes resulted in precipitation of more δ -phase with a large size consuming more Nb element, which is the principle for the precipitation of the strengthening phase γ'' . Therefore, the degradation of the mechanical properties was observed. For instance, Deng et al. [1] investigated the effect of different heat treatments such as homogenization heat treatment followed by aging HA (1080°C/1h/WC + aging), and homogenization followed by solution and aging HSA (1080°C/1h/WC + 980°C/1h/WC + aging), on the microstructure and the mechanical properties of IN718 samples fabricated by SLM. Their [1] results show that the HA and HSA treated microstructure still had Laves phase with the same size and amount which resulted in microvoids during tensile test and consequently deterioration of the mechanical properties. In addition, larger size and quantity of δ -phase were observed in HSA condition, while in the case of HA, δ -phase did not precipitate due to the lower solvus temperature of δ -phase (1020°C). Therefore, lower ductility of the HSA condition in relation to HA was observed due to the formation of larger voids rather than other microvoids due to the presence of δ -phase.

Similarly, Dongyun et al. [19] applied two heat treatments, HSA (1080°C/1.5 h/AC + 980°C/1h/ AC + 720 °C/8 h/FC at 55°C/h to 620°C, 8 h/AC) and HA (1080°C/1.5 h/AC+ 720 °C/8 h/FC at 55°C/h to 620°C/8 h/AC), to analyze the effect of δ -phase on the mechanical behavior of IN718 processed by SLM. Their [19] results reveal that after HSA treatment, the dendritic microstructure transformed to recrystallized grains in addition to fine δ -phase and MC carbides along the grain boundaries. They [19] confirmed that δ -phase does not form in the matrix after HA treatment. Thus, the HA treated samples exhibited improved mechanical behavior (5.8% increase in UTS, 12.2% increase in YS, and 5.9% increase in HRC) in relation to the HAS treated samples. The authors [19] attributed this trend to the fact that the Nb is the principle element

necessary for the formation of both γ'' strengthening phase and δ -phase (Ni_3Nb). Therefore, at a certain concentration of Nb in the matrix, δ -phase formed at the expense of γ'' after applying HSA heat treatment which reduced the mechanical properties [19].

However, there are some beneficial effects of δ -phase in the matrix such as pinning the dislocations and inhibiting the grain growth when the amount and size of the δ -phase particles are controlled [1, 19]. Deng [18] reported that the presence of approximately 4% δ -phase at the grain boundaries could effectively inhibit the grain growth at service condition and during thermal exposure. Hence in order to yield a good combination of high strength and ductility in SLM IN718, they [1] recommended prolonging the homogenization treatment at 1080°C to initially dissolve the Laves phase and homogenize the microstructure. In order to achieve this, they also recommended reducing the duration of the solution treatment to precipitate lower amount of small size δ -phase at the grain boundaries.

In this work, the effect of homogenization and solution treatments with different holding times on the microstructure and precipitates of Inconel 718 superalloy produced by SLM have been studied. Two heat treatment conditions were selected at the extreme times of both homogenization and solution treatments. The aim is generally to investigate the effect of heat treatment time on the phase formation, Laves, carbides and δ -phases of SLM IN718. The microstructure, compositional analysis and phase identification were studied using Scanning Electron Microscope (SEM), Energy Dispersive Spectrometer (EDS) and X-Ray Diffractometer (XRD), respectively.

II. EXPERIMENTAL PROCEDURE AND MATERIALS

Gas atomized IN718 powder, supplied by EOS-GmbH (Krailling, Germany), with an average particle size less 30 μm was used as the raw material to fabricate the test samples via the SLM process. The chemical composition of the gas atomized IN718 powder is given in TABLE I. IN718 test samples were fabricated using an EOS M280 400W (EOS, Krailling, Germany) Yb: YAG fiber laser system with PSW3.6 and Parameter Set IN718_Performance 1.0 (40 μm layer thickness, 1000 mm/s scanning speed, 100 μm laser beam diameter and 67° hatch angle). The test samples were printed horizontally (perpendicular to the building direction) under a nitrogen gas environment. The building platform was pre-heated to 80°C to mitigate the thermal gradient along the building direction and reduce the thermal stresses in the produced parts.

Two heat treatment conditions of homogenization and solution treatment using different holding times were selected based on some recommendations of recent research [1, 18, 19]

TABLE I. CHEMICAL COMPOSITION OF IN718 POWDER [9].

Element	Ni	Cr	Nb	Mo	Ti	Al	Co	C	Si	Mn	Fe
wt. %	52.75	20.20	4.46	2.70	1.12	0.45	-	-	-	-	Balance

on SLM IN718 to obtain a combination of high strength and high ductility. Therefore, homogenization heat treatment was performed at 1080°C for 1 and 7h dwell time, followed by air cooling in order to dissolve more of the segregated elements and Laves phase. Also, solution heat-treatment was carried out at 980°C with dwell time of 60 and 15 minutes, followed by air cooling, to precipitates δ -phase with a lower amount and smaller size. Thus, two samples were used—one of them was homogenized at 1080°C for 1h and followed by solution treatment at 980°C for 60 minutes which is designated as HS1. While, the other sample was homogenized for 7h followed by solution treatment at 980°C for 15 minutes which is designated as HS2. The two heat treatment conditions are as listed in TABLE II. All heat treatment conditions were performed in Carbolite Thermal Cycling Rig (Carbolite TCR, England) using a set of K-type thermocouples for monitoring the temperature.

Both as-printed and heat treated IN718 samples were cut into perpendicular and parallel cross-sections with respect to the building orientation using a slow cutter (Buehler, USA) with a mineral oil bath to prevent the heat generation. For the metallographic analysis, the cut samples were mounted using hot epoxy resin. Mounted samples were mechanically ground gradually from 320 to 1200 grit using SiC papers and polished down to 0.25 μm using alcohol-based alumina suspension. To reveal the precipitated phases and dendritic microstructure, the polished samples were etched using cotton swabbing with 10 ml hydrochloric acid + 1.5 ml 30% hydrogen peroxide etchant [1].

Microstructure evaluations for the IN718 as-printed and heat treated samples were carried out using a Scanning Electron Microscope, SEM, (HITACHI S-3400N, Tokyo, Japan) equipped with an energy dispersive X-ray spectrometer,

TABLE II. HEAT TREATMENT CONDITIONS OF SLM IN718

Designation	Homogenization Heat Treatment	Solution Heat Treatment
As-built	None	None
HS1	1080 °C for 1 h/AC	980 °C for 1 h/AC
HS2	1080 °C for 7 h/AC	980 °C for 15 min/AC

*AC: Air cooling

EDS. It is worth mentioning that all SEM images presented here exhibit the microstructure of the vertical cross-section of the IN718 test samples (parallel to the building direction). The chemical analysis of all precipitated phases was carried out using semi-quantitative EDS. Phase analysis of SLM IN718 as-printed and heat treated samples were performed using XRD, (PANAnalytical X'pert Pro X-ray diffractometer, Almelo, The Netherlands) with a $\text{CuK}\alpha$ radiation at 45 kV and 40 mA.

III. RESULTS AND DISCUSSION

A. Microstructure evaluation in as-printed condition

SEM micrographs of the as-printed SLM IN718 sample sliced parallel to the building direction are shown in Fig. 1(a-b). The microstructure of the as-printed sample is composed of columnar grains oriented parallel to the building direction, which could be attributed to the high thermal gradient along the building direction and the direction of the heat dissipation during the solidification process [9, 19]. Each of these grains contains a number of columnar/cellular sub-grain features which are labeled as 1 and 2 in Figure 1b and these features are consistent with the literature [1, 6]. In some literature, the grains are often referred to as dendrites and the sub-grain features as a secondary dendrite [6].

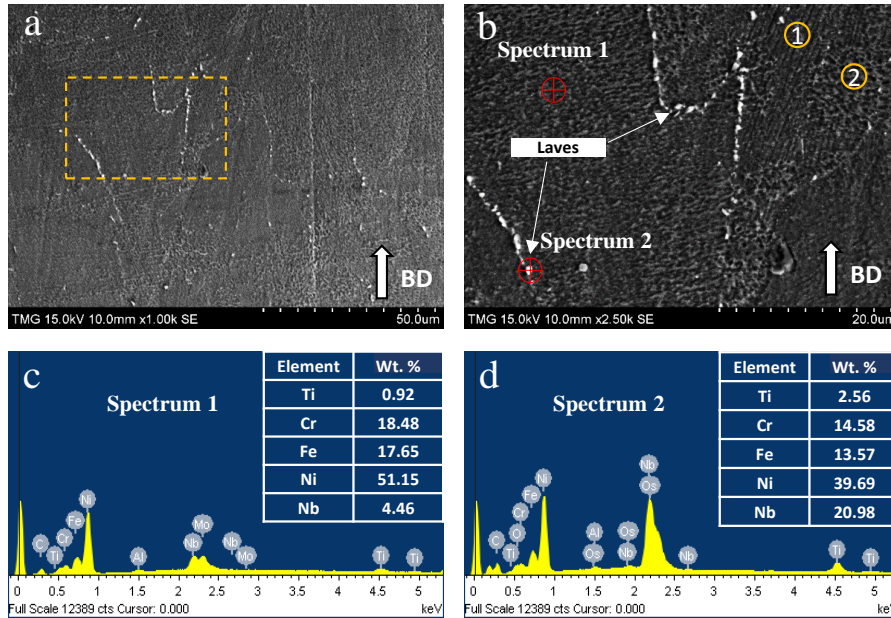


Figure 1. (a, b) SEM micrographs of the as-printed SLM IN718 in the vertical cross section with respect to the building direction (indicated with an arrow in the right low corner); (c, d) EDS spot analysis of spectrum 1 and spectrum 2. The circulated numbers indicate the cellular and columnar features.

Moreover, the grain boundaries were decorated with fine irregular precipitates with the size of about a few hundred nm. According to the EDS analysis of these precipitates, a high concentration of refractory elements such as Nb and Ti was observed as a result of elemental segregation during the solidification process. This concentration reached up to 20.98 wt. % Nb and 2.56 wt. % Ti, and depleted in Ni, Fe and Cr with respect to the nominal composition of the substrate matrix as shown in Fig. 1(c-d). During the solidification process of IN718, segregation of Nb element in the interdendritic region occurred due to its partitioning behavior, and the solidification sequence was well established as: Liquid $\rightarrow \gamma \rightarrow \gamma + \text{NbC} \rightarrow \gamma + \text{NbC} + \text{Laves}$ [1]. In addition, the size of these precipitates was very fine in relation to conventional fabrication methods. This could be attributed to the higher cooling rate of the SLM process that can reach up to 10^4K/s [1]. The majority of these precipitated phases along the intergranular boundaries were identified as a Laves phase based on the work of [6]. They [6] reported that Laves phase formation requires at least 10 wt. % Nb, while δ -phase requires at least 6-8 wt. % Nb. Therefore, the formation of brittle intermetallic Laves phases consume most of Nb and deplete it from the matrix which is confirmed by the EDS analysis. Thus, further heat treatments are required to eliminate the precipitated phase and microsegregation.

B. Microstructure evaluation of heat treated samples

- “Homogenization 1h/Solution 60 min” IN718 alloy

Homogenization of the as-printed samples at 1080°C for 1h followed by solution treatment at 980°C for 60 minutes (HS1) resulted in the initiation of recrystallization and partial dissolution of the microsegregated elements. However, HS1 treated microstructure appeared to maintain some of its initial as-printed features such as columnar grain morphology. This is probably attributed to that the homogenization holding time was not enough to induce complete recrystallization and dissolution of segregated elements. As shown in Fig. 2(a-b), the microstructure of the HS1 condition composed of a mixture of columnar and equiaxed grains which is the core cause for anisotropic properties. The presence of equiaxed grains gave an indication about the initiation of recrystallization which was driven by the residual stresses induced in the IN718 samples during the printing process as a result of the complex thermal cycle of SLM [14, 20]. Moreover, the continuous elemental microsegregation along the grain boundaries was broken into many irregular shape precipitates. According to EDS analysis of these precipitates, partial back diffusion of the segregated Nb and Ti elements was observed in relation to the as-printed samples. The concentrations of these segregated elements in the precipitated particles of HS1 was up to 18.78 wt. % Nb and 2.22 wt. % Ti as shown in Fig. 2(c-d). These precipitates were identified as a Laves phase. In addition, by following the Nb and Ti concentrations in the substrate matrix of as-printed and HS1 samples, higher concentrations were quantified for HS1 condition in relation to the as-printed sample which confirms that more back diffusion of these segregated elements into the matrix took place. However, more homogenization holding

time is required to completely dissolve these segregations which is in agreement with the recommendations of other authors [1, 18, 19].

Furthermore, needle-like δ -phase was clearly observed along the grain boundaries of the HS1 microstructure which is in agreement with other reports [1, 19], and as can be seen in Fig. 2b. The presence of δ -phase at the grain boundaries has some beneficial effects such as pinning of grain growth and enhancing the mechanical properties [18]. However, larger amount of δ -phase consumes more Nb element which is an important element for the precipitation of γ'' strengthening phase. Therefore, δ -phase precipitation occurs at the expense of γ'' precipitation which could result in diminishing the mechanical properties of SLM IN718 superalloy.

- “Homogenization 7h/Solution 15 min” IN718 alloy

The micrograph and the chemical composition analysis of the HS2 condition are shown in Fig. 3. More coarse equiaxed grains were observed in relation to the as-printed, HS1, condition that suggests a near recrystallization completion followed by grain growth. This indicates that the prolonged treatment provided sufficient activation energy at 1080°C for 7h homogenization to break the intermetallic bonds and improve the diffusion which finally resulted in noticeable recrystallization and grain growth in relation to other conditions with shorter holding time [6]. In addition, coarsened precipitates both within the grains and along the grain boundaries were observed compared to HS1 condition. According to the EDS compositional analysis, many precipitates exhibited enrichment in Nb and Ti to be 6.7-7.25 wt. % Nb and 1.13-1.55 wt. % Ti compared to the nominal composition of the matrix as shown in TABLE I. However, the concentration of Nb and Ti in these precipitates was lower compared to the precipitates of HS1 heat treatment condition. This could be attributed to more dissolution of the segregated element with increasing homogenization time, resulting in releasing a considerable amount of Nb and Ti into the matrix. In addition, other particles enriched in C were frequently seen and their size is larger than those particles observed in the matrix of the as-printed and HS1 conditions as shown in Fig. 3. These particles have not been positively identified by EDS but they are likely to be carbides. Compared to the as-printed conditions, this heat treatment resulted in much smaller δ -phase content and precipitate size.

C. Phase identification of as-printed and heat treated samples

The analysis of the XRD patterns of the as-printed and heat treated samples are shown in Fig. 4. It was applied to confirm the results obtained from the microstructure evaluation. The effect of two heat treatment conditions on the presence of MC carbides and δ -phases can be inferred from their diffraction peaks. By focusing on the δ -phase content at the different holding time of solution treatment 60 and 15 minutes (HS1, HS2), it is observed that δ -phase content increases with solution treatment time see Fig. 4b ($2\theta = 42^\circ$). Based on the microstructure results, solution heat treatment at 980°C for less

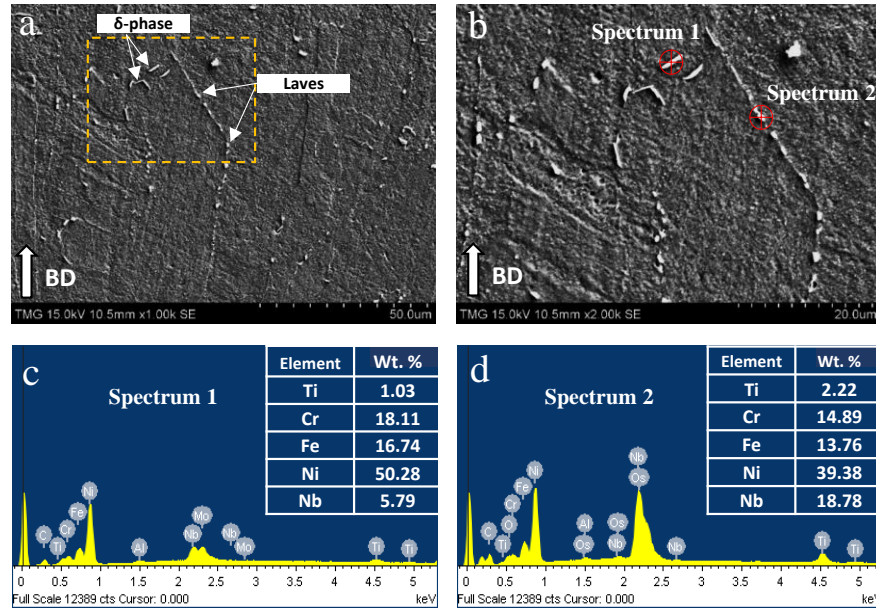


Figure 2. (a-b) SEM micrographs of the HS1 treated SLM IN718; (c-d) EDS spot analysis of spectrum 1 and spectrum 2.

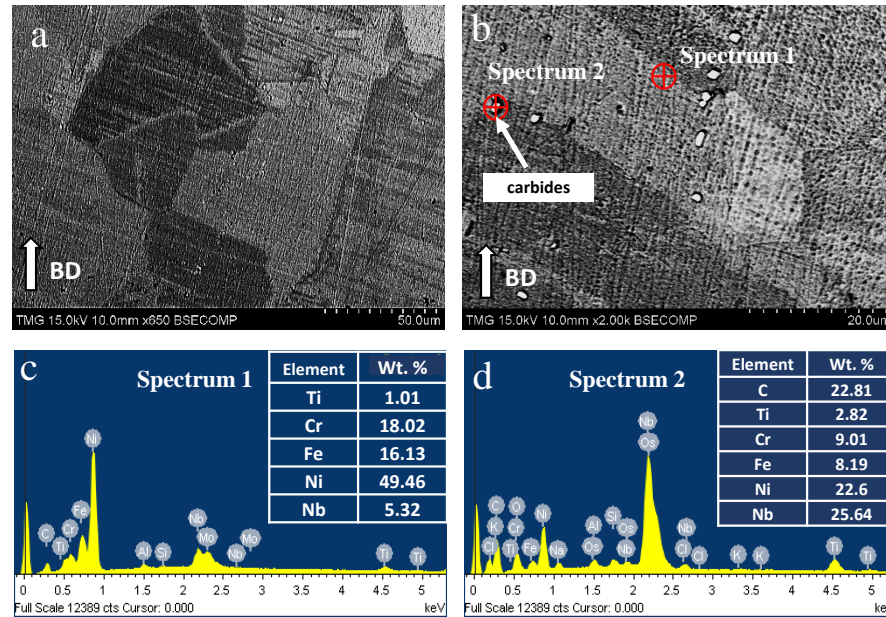


Figure 3. (a-b) SEM micrographs of the HS2 treated SLM IN718; (c-d) EDS spot analysis of spectrum 1 and spectrum 2.

than 60 minutes resulted in less δ -phases which was confirmed by XRD results. While in case of MC carbides, homogenization and solution treatments for longer time results in more carbides content as shown in Fig. 4b ($2\theta = 36^\circ$) which is also in agreement with the microstructure results.

IV. CONCLUSION

The effect of the post-processing heat treatment (homogenization and solution treatment) using different holding times on the microstructure and mechanical properties

of as-printed IN718 fabricated using SLM were presented. The as-printed microstructure exhibited columnar grains oriented parallel to the building direction with fine irregular precipitates decorated the grain boundaries. These precipitates were enriched with Nb and Ti refractory elements which are the primary strengthening elements of IN718. The homogenization treatment at 1080°C for 1h and 7h had a significant impact on the microstructure and precipitates of the SLM IN718. Complete recrystallization followed by grain growth achieved after 7h homogenization heat treatment. In addition, more dissolution of segregated Nb and Ti from the precipitates

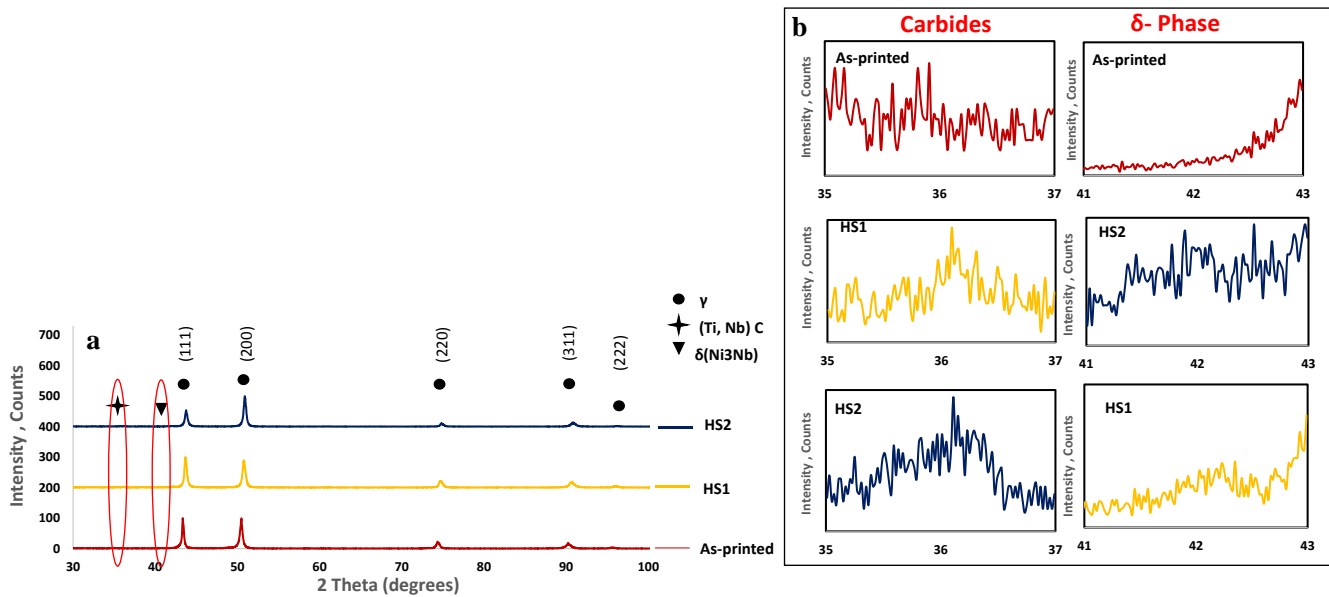


Figure 4. XRD patterns for IN718 superalloy after different heat treatments: (a) total rang 2θ, (b) focus on 2θ= 42° and 2θ=36°.

occurred by increasing the homogenization holding time. However, considerable coarsening of precipitates and more carbides formation were observed with longer homogenization treatment. The amount of δ -phase was proportional to the dwell time of the solution heat treatment. Therefore, solution heat treatment at 980°C for less than 60 minutes resulted in significant reduction in the δ -phase content and size. The results obtained from microstructure characterization are consistent with XRD analysis of as-printed and heat treated conditions.

ACKNOWLEDGMENT

The first author gratefully acknowledges the scholarship and financial support from the Egyptian Military. The authors wish to thank Mazen Samara for his support during using the SEM and XRD analysis at Concordia University.

REFERENCES

[1] D. Deng, R. L. Peng, H. Brodin and J. Moverare, "Microstructure and mechanical properties of Inconel 718 produced by selective laser melting: Sample orientation dependence and effects of post heat treatments," *Materials Science and Engineering: A*, vol. 713, pp. 294-306, 2018.

[2] X. Wang, X. Gong and K. Chou, "Review on powder-bed laser additive manufacturing of Inconel 718 parts," *Proc. Inst. Mech. Eng. Pt. B: J. Eng. Manuf.*, vol. 231, (11), pp. 1890-1903, 2017.

[3] Q. Jia and D. Gu, "Selective laser melting additive manufacturing of Inconel 718 superalloy parts: Densification, microstructure and properties," *J. Alloys Compounds*, vol. 585, pp. 713-721, 2014.

[4] T. M. Pollock and S. Tin, "Nickel-based superalloys for advanced turbine engines: chemistry, microstructure and properties," *J. Propul. Power*, vol. 22, (2), pp. 361-374, 2006.

[5] M. D. Sangid, T. A. Book, D. Naragani, J. Rotella, P. Ravi, P. Kenesei, J. Park, H. Sharma, J. Almer and X. Xiao, "Role of heat treatment and build orientation in the microstructure sensitive deformation characteristics of IN718 produced via SLM additive manufacturing," *Additive Manufacturing*, vol. 22, pp. 479-496, 2018.

[6] W. M. Tucho, P. Cuvillier, A. Sjolyst-Kverneland and V. Hansen, "Microstructure and hardness studies of Inconel 718 manufactured by selective laser melting before and after solution heat treatment," *Materials Science and Engineering: A*, vol. 689, pp. 220-232, 2017.

[7] X. Li, J. Shi, C. Wang, G. Cao, A. Russell, Z. Zhou, C. Li and G. Chen, "Effect of heat treatment on microstructure evolution of Inconel 718 alloy fabricated by selective laser melting," *J. Alloys Compounds*, 2018.

[8] D. Zhang, W. Niu, X. Cao and Z. Liu, "Effect of standard heat treatment on the microstructure and mechanical properties of selective laser melting manufactured Inconel 718 superalloy," *Materials Science and Engineering: A*, vol. 644, pp. 32-40, 2015.

[9] A. Mostafa, I. Picazo Rubio, V. Brailovski, M. Jahazi and M. Medraj, "Structure, texture and phases in 3D printed IN718 alloy subjected to homogenization and HIP treatments," *Metals*, vol. 7, (6), pp. 196, 2017.

- [10] S. Bremen, W. Meiners and A. Diatlov, "Selective laser melting: a manufacturing technology for the future?" *Laser Technik Journal*, vol. 9, (2), pp. 33-38, 2012.
- [11] J. Strößner, M. Terock and U. Glatzel, "Mechanical and microstructural investigation of nickel-based superalloy IN718 manufactured by selective laser melting (SLM)," *Advanced Engineering Materials*, vol. 17, (8), pp. 1099-1105, 2015.
- [12] T. Trosch, J. Strößner, R. Völkl and U. Glatzel, "Microstructure and mechanical properties of selective laser melted Inconel 718 compared to forging and casting," *Mater Lett*, vol. 164, pp. 428-431, 2016.
- [13] E. Chlebus, K. Gruber, B. Kuźnicka, J. Kurzac and T. Kurzynowski, "Effect of heat treatment on the microstructure and mechanical properties of Inconel 718 processed by selective laser melting," *Materials Science and Engineering: A*, vol. 639, pp. 647-655, 2015.
- [14] T. DebRoy, H. Wei, J. Zuback, T. Mukherjee, J. Elmer, J. Milewski, A. M. Beese, A. Wilson-Heid, A. De and W. Zhang, "Additive manufacturing of metallic components—process, structure and properties," *Progress in Materials Science*, vol. 92, pp. 112-224, 2018.
- [15] C. Li, Z. Liu, X. Fang and Y. Guo, "Residual stress in metal additive manufacturing," *Procedia CIRP*, vol. 71, pp. 348-353, 2018.
- [16] F. Zhang, L. E. Levine, A. J. Allen, M. R. Stoudt, G. Lindwall, E. A. Lass, M. E. Williams, Y. Idell and C. E. Campbell, "Effect of heat treatment on the microstructural evolution of a nickel-based superalloy additive-manufactured by laser powder bed fusion," *Acta Materialia*, vol. 152, pp. 200-214, 2018.
- [17] B. E. Carroll, T. A. Palmer and A. M. Beese, "Anisotropic tensile behavior of Ti–6Al–4V components fabricated with directed energy deposition additive manufacturing," *Acta Materialia*, vol. 87, pp. 309-320, 2015.
- [18] D. Deng, "Additively Manufactured Inconel 718: Microstructures and Mechanical Properties," vol. 1798, 2018.
- [19] D. Zhang, Z. Feng, C. Wang, W. Wang, Z. Liu and W. Niu, "Comparison of microstructures and mechanical properties of Inconel 718 alloy processed by selective laser melting and casting," *Materials Science and Engineering: A*, vol. 724, pp. 357-367, 2018.
- [20] B. Song, X. Zhao, S. Li, C. Han, Q. Wei, S. Wen, J. Liu and Y. Shi, "Differences in microstructure and properties between selective laser melting and traditional manufacturing for fabrication of metal parts: A review," *Frontiers of Mechanical Engineering*, vol. 10, (2), pp. 111-125, 2015.

Multi-Element Slicing Approach for Intelligent Additive Manufacturing Preprocessing

Dylan Bender and Ahmad Barari

Department of Automotive, Mechanical, and Manufacturing Engineering
The University of Ontario Institute of Technology
Oshawa, Ontario, Canada
Dylan.Bender@uoit.ca, Ahmad.Barari@uoit.ca

Abstract—Additive manufacturing traditionally starts with an STL file (Standard Tessellation Language) to represent a 3D model for 3D printing. The problem with an STL file is that it is an approximation of the true model which leads to deviations in the geometry of the part. This effect is amplified in the case of topology optimization where the model starts as finite element representation of a structure which itself is an approximation of the true surface which is typically then converted into an STL file for additive manufacturing which introduces compounding deviations of the structure's true geometry all before the manufacturing process itself even introduces its own geometric errors. Slicing of the finite element model directly is introduced to minimize potential deviations in the final manufactured part which also serves to aid in coupling manufacturing and design for the ongoing work of intelligent manufacturing systems.

Keywords—*additive manufacturing; intelligent manufacturing; slicing; finite elements; topology optimization*

I. INTRODUCTION

Additive manufacturing, sometimes referred to as 3D printing, is a cost-effective way to produce plastics, metals, ceramics, and even composites in small batch sizes with manufacturing costs almost completely decoupled from the geometric complexity of the part to be printed [1, 2]. These properties of additive manufacturing synergize with structural topology optimization to easily produce normally unmanufacturable complex shapes and geometries such as internal pathways, cavities and structures [3, 14], which are often resulting features of topology optimization. Although additive manufacturing provides many advantages over traditional manufacturing techniques, it is not without its own drawbacks, best outlined in the context of topology optimization in [4] as well as deviations from the true geometry from manufacturing artifacts such as the “staircase effect” and others [5-12]. The primary purpose of this paper is to eliminate geometric deviations in the manufactured part by slicing the finite element model of a topology optimized structure. Slicing refers to the algorithms responsible for producing the 2D cross-sections of a 3D model that are to be sent to the 3D printer to

manufacture the structure in a layer—by-layer manor. Typically, in the context of topology optimization, the finite element representation of the structure is converted into an STL file by means of a marching cubes algorithm [13] to produce a geometric approximation of the structure which inherently introduces deviations and simplifications of the true geometric shape.

Topology optimization and AM are getting very close recently due to flexibility of AM in fabrication of parts with high complexities in geometry and material [14], and high capability of topology optimization in designing the most efficient use of material in product development process. Therefore, it needs immediate attention of researchers to work on the requirements for establishing proper interfaces between these two processes. The preprocessing stage in AM needs to be compatible with the output of TO process. The typical STL format is not the most suitable way to establish communication between TO and AM-preprocessing. The typical output format from the TO process includes a combination of various solid elements that are defined in finite element software. However, typical AM preprocessing only recognizes the surface representation provided by triangular facets. This paper provides a generic solution to allow direct communication of the TO process and AM preprocessing by direct slicing of multi-elements.

Communication between TO and AM preprocessing can be even conducted iteratively. Traditionally, the TO process and AM preprocessing are carried out in sequential order which is not in line with the design principles of agile design. However, in more advanced applications, it is even desirable to conduct these two processes simultaneously, they need to communicate iteratively and in a closed loop structure. [15-17] For a proper bridge between the engineering simulation, that is the structural topology optimization and the AM preprocessor, a level of information should be maintained that exists in a full solid element representation of the part all while removing predefined rigid interfaces in order to enable a flexible design. In essence, a transfer of the solid elements to STL is basically a reduction in the quality of the associated part as well as being against the design principles of agile design outlined in [16], this should be avoided if and when possible.

This paper presents an algorithm for slicing the 3D model of a structure that is to be additively manufactured and is already described in a finite element representation such the presented use case for topology optimization. Traditionally, a tessellated surface representation of the model is sent to the 3D printing software for slicing which not only adds additional computational time for manufacturing topology optimized structures, it also introduces losses in geometric information and is considered a rigid interface to the desired flexible, agile and intelligent design and manufacturing process that is demanded by industry 4.0.

II. BACKGROUND

Commercially available additive manufacturing systems today's mainly worked based on stacking uniform-thickness solid layers which are created by a material phase change from liquid to solid, material powder sintering, material depositing, or cutting sheets of material. Despite the physics of the process, in almost all of them the solid object is created by virtue of the combination of many 2 ½ D layers [7]. As a result, these technologies are also referred as Layer-based manufacturing.

The instruction for AM machine to create each one these 2 ½ D layers needs to be generated from the solid model of the product developed in the Computer Aided-Design (CAD) system. The process of creating information of a 2 ½ D layer from a solid CAD geometry is referred as "Slicing" [6]. The information of the designed product in the CAD system can be stored using various advanced computational geometry methods including polynomial representation of the curves and surfaces, B-rep, CSG-Tree, etc. In order to make the AM-preprocessing slicing a uniform process applicable for all various CAD systems, the STL format was introduced which in fact presents all the surfaces in a CAD solid model by a combination of triangular facets.

Slicing of STL triangular surfaces is a simple algorithm. However, slicing is still a challenging task in many AM – preprocessing systems for the following reasons. (i) When the number of triangles increases in a STL file, the slicing process computationally becomes very expensive [8]. (ii) In addition, several errors including degenerated surfaces, manifold structures, voids, and tilted normal vectors in generated STL file from various commercial CAD systems have been reported frequently which also cause challenges in the slicing process. (iii) Representation of a complex surface by triangular tessellation introduce a chord error. The chord error is the difference between the ideal surface and the approximation using the triangular element. This limits the issue of accuracy from the CAD system to AM-preprocessing. In order to avoid this error, many researchers have been consider "Direct Slicing" of the CAD model recently [6 - 9].

The slicing process also has direct effects on the accuracy and surface quality of the final AM product. Stacking of the 2 ½ D layers introduced by slicing will result in cusps on the profile of the fabricated surfaces which results undesired surface roughness [1, 6, 9-11]. "Adaptive Slicing" is an approach to decide for the optimum thickness of individual layers based on the local characteristics of the geometric features on the product [6, 9].

III. METHODOLOGY

In this section, the developed slicing procedure is described. The procedure includes six computational tasks as follows:

A. Transfer of multi-elements structure from finite element system to AM preprocessor

Topology optimization (TO) is the most general branch of structural optimization, the other main branches are size, and shape optimization. Many methods of topology optimization have been implemented for solving structural design problems but this paper will focus on an evolutionary method known as Solid Isotropic Material with Penalization (SIMP) [18]. The working principles of SIMP are simple. First, the structural design problem is defined by a design domain (as a 3D CAD model) and by its boundary conditions such as loads and constraints. The design domain is then meshed into a finite element representation of the model, where the FEA solver then analyzes the stress and strain conditions of the design. Elements of the design domain that do not contribute significantly to the stiffness of the overall structure are made to be less dense and conversely, the elements which contribute greatly to the overall stiffness of the structure are made to be denser while maintaining a predefined weight from iteration to iteration. Changes to the design (finite element densities) are slightly modified and the new structure is analyzed by FEA. This iterative process continues typically for a number of iterations until an optimal design has been reached. The fact that the SIMP topology optimization algorithm utilizes a finite element mesh means that its designs can be directly sliced using the techniques described in this paper. Topology optimized solution for a cantilever beam, its loading condition, and the corresponding design domain are presented in fig. 1 as an example.

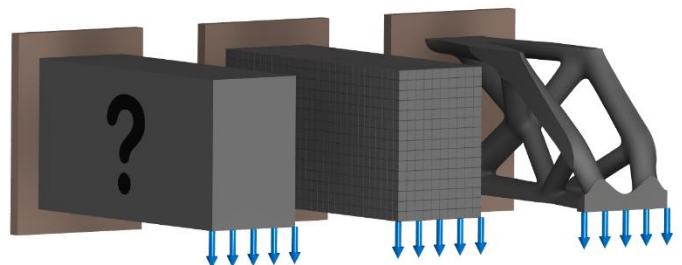


Figure 1. Topology optimization of a cantilever beam under bending condition

Topology optimization process uses FEA as its solving engine. A 3D model is decomposed into a finite element mesh for the purpose of finite element analysis utilizing numerical solving techniques to analyze the mechanical behavior of a structure under loading conditions. Designs may consist of anywhere from thousands to millions of finite elements in any combination of the element types. The user typically has control over the element sizes, types, and mesh refinement strategies in order to adequately approximate the true to life structure for their tailored needs. Topology optimization utilizes the finite element representation of a structure to carry

out the evolutionary structural optimization process, the utilization of the finite mesh for TO will be further described in the next section.

Finite elements are not necessarily constrained to be any particular size, shape, or even complexity but the majority of commercially available meshing software will decompose a 3D model into a subset of the four fundamental solid elements, known as a cuboid, a tetrahedral, a pyramid and prism like shapes for the sake of simplicity. These four elements and various relevant characteristics are outlined below in table 1.

TABLE I. THE FOUR FUNDAMENTAL FINITE ELEMENT TYPES

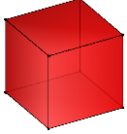
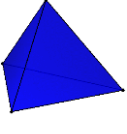
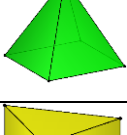
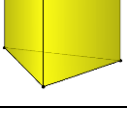
	Visual	# of Faces	# of Nodes	# of Edges
<i>Cuboid</i>		6	8	12
<i>Tetrahedral</i>		4	4	6
<i>Pyramid</i>		5	5	8
<i>Wedge</i>		5	6	9

Figure 2 presents an example of a 3D finite element model with a combination of various elements that are transferred to AM preprocessing system through the data transfer process explained above. This model consists of 5222 elements and 14010 nodes and is composed of all four typical element types described in Table 1. This mesh will be used in the slicing process that will be explained in the following steps.

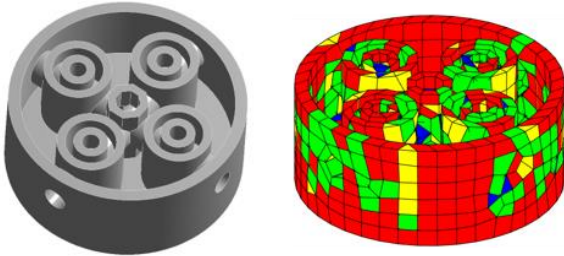


Figure 2. 3D model and sample finite element mesh used for slicing testing

B. Element Filtration

The second of the six processes is to determine which of the elements will be passed to the remainder of the slicing algorithm since passing all of the elements would be inefficient. Once a slice height has been given from either the user or perhaps from an adaptive slicing algorithm, it is easy to identify the elements which intersect the slice plane. First, the max and min nodes of each element with respect to the normal vector of the plane are identified. If the slice plane lies between the min and max nodes of an element, it will be passed to the next step for slicing. A sample slice is shown on the sample mesh in Fig. 3 with the elements which lie on the slice plane highlighted in red. The total number of filtered elements in this case study has been 684. This is only about 13% of the total number of elements in the sample mesh.

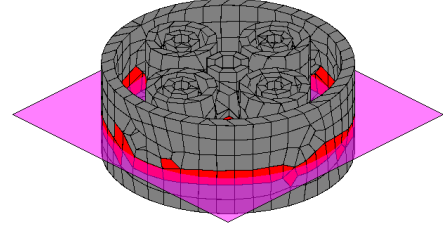


Figure 3. A visualization of the elements that lie on the slice plane

C. Intersection of Elements with Plane

In this task, the filtered set of elements is processed to find all of the facets that intersect with the slicing plane. The intersection of each one of the facets with the slicing plane creates a candidate line segment to form the printing contours. The output of this process is a set of all of the line segments resulting by this intersection analysis. Figure 4 presents the filtered set of elements (a) and the full set of all the resulting intersection line segments. There are 2448 line segments found to intersect the slice plane in the case.

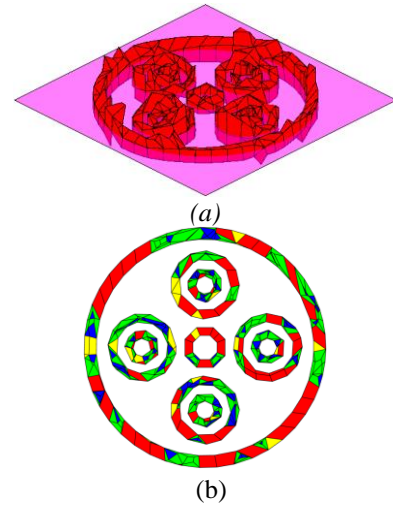


Figure 4. Finding the candidate line segments of the contours using the filtered set of elements a) Filtered set of elements b) Candidate line segments of contours

D. Eliminating Internal and Redundant Contour Edges

Only a subset of all of the line segments generated in the previous step are part of an internal or external contour. In order to recognize these particular line segments the following technique is used. A line segment can be part of an internal or external contour if and only if it only belongs to one element. Figure 5 presents all of non-redundant line segments which will be used to form the polygonal contours of the previous example. The total number line segments after removing the redundant edges is 316 which is approximately 12% of all of the identified line segments.

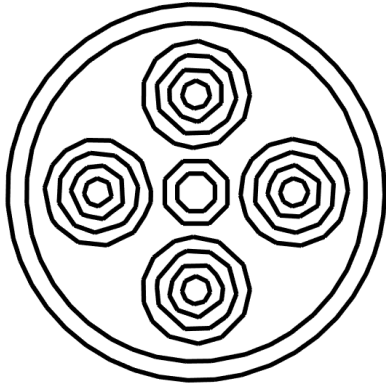


Figure 5. All of the detected line segments which will form contours for the sliced layer

E. Ordering Contour Edges into Polygons

The lists of ordered edges from the set of non-redundant edges are created in this step. Each of these lists forms a closed polygon, which in fact will be an internal or an external contour defining the sliced layer. There are 20 individual polygons recognized in this example.

F. Determine the Direction of the Polygons (Solid or Void)

A ray-tracing approach is used to identify the internal and external contours as two classes. The combination all of the classified contours provide all of the required detailed information of the sliced layer. Figure 6 presents the two different classes of contours detected for the sliced layer in the previous example. The internal contours are specified by green, and the external contours are shown by the red colour for visualization. Ten of the polygons of the 20 are recognized as external contours and the other 10 have been identified as internal contours.

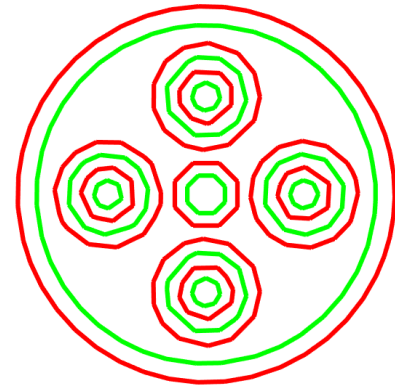


Figure 6. The contours of the slice layer identified as internal (green) or external (red)

The resulting classified contours are used directly to generate instructions required for the additive manufacturing machine. As an example, Fig. 7 presents the digital image required to print a layer of the model corresponding to the previous example using an Digital Light Processing (DLP) Stereolithography machine.

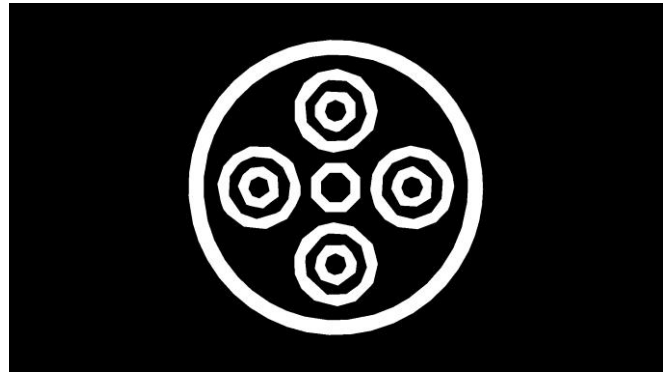


Figure 7. Printing instruction for DLP Stereolithography machine

The described procedure is highly reliable and flexible for any finite element represented geometry. A variety of experiments have been conducted and the results always were satisfactory. Figure 8 presents the intermediate results in a TO-AM preprocessing loop. The topology optimiation in this case study receives some feedbacks from the calculated layer to decide for its next iteration.

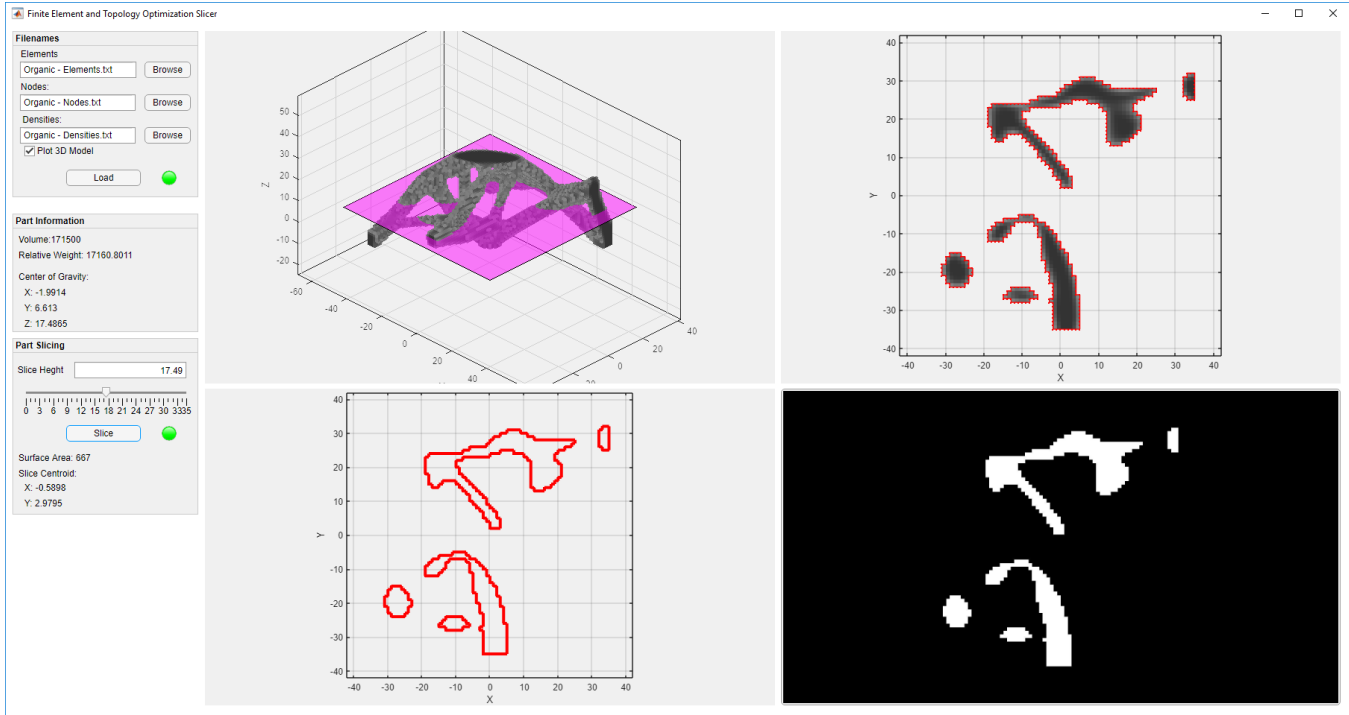


Figure 8. An iteration of the loop between topology optimization and AM preprocessing

IV. CONCLUSION

This paper presents a procedure for direct communication of topology optimization process and additive manufacturing preprocessing computation tasks. The information resulting by AM preprocessing can be used by TO to modify the process. This will allow consideration of the manufacturing requirements during the design process and it is an important element in design for additive manufacturing. The developed procedure have been fully implemented and validated by conducting a variety of complex case studies. It has been robust and efficient by eliminating the redundant step of converting the finite element mesh into a tessellated surface. This will result in removing the predefined rigid interfaces between the computational processes which hinders a flexible, and agile design.

REFERENCES

- [1] Barari, A., Kishawy, H.A., Kaji, F. and Elbestawi, M.A., 2017. On the surface quality of additive manufactured parts. *The International Journal of Advanced Manufacturing Technology*, 89(5-8), pp.1969-1974.
- [2] Atzeni, E., Iuliano, L., Marchiandi, G., Minetola, P., Salmi, A., Bassoli, E., Denti, L. and Gatto, A., 2014. Additive manufacturing as a cost-effective way to produce metal parts. In *High Value Manufacturing: Advanced Research in Virtual and Rapid Prototyping-Proceedings of the 6th International Conference on Advanced Research and Rapid Prototyping, VR@ P* (Vol. 2013, pp. 3-8).
- [3] Aidibe, A., Tahan, A. and Brailovski, V., 2016. Metrological investigation of a selective laser melting additive manufacturing system: A case study. *IFAC-PapersOnLine*, 49(31), pp.25-29.
- [4] Liu, J., Gaynor, A.T., Chen, S., Kang, Z., Suresh, K., Takezawa, A., Li, L., Kato, J., Tang, J., Wang, C.C. and Cheng, L., 2018. Current and future trends in topology optimization for additive manufacturing. *Structural and Multidisciplinary Optimization*, pp.1-27.
- [5] Lalehpour, A. and Barari, A., 2016. Post processing for fused deposition modeling parts with acetone vapour bath. *IFAC-PapersOnLine*, 49(31), pp.42-48.
- [6] Sikder, S., Barari, A. and Kishawy, H.A., 2014, August. Effect of adaptive slicing on surface integrity in additive manufacturing. In *ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. V01AT02A052-V01AT02A052). American Society of Mechanical Engineers.
- [7] Gohari, H., Barari, A. and Kishawy, H., 2016. Using Multistep Methods in Slicing 2 1/2 Dimensional Parametric Surfaces for Additive Manufacturing Applications. *IFAC-PapersOnLine*, 49(31), pp.67-72.
- [8] Gohari, H., Barari, A. and Kishawy, H., 2018. An efficient methodology for slicing NURBS surfaces using multi-step methods. *The International Journal of Advanced Manufacturing Technology*, 95(9-12), pp.3111-3125.
- [9] Sikder, S., Barari, A. and Kishawy, H.A., 2015. Global adaptive slicing of NURBS based sculptured surface for minimum texture error in rapid prototyping. *Rapid prototyping journal*, 21(6), pp.649-661.
- [10] Jamiolahmadi, S. and Barari, A., 2014. Surface topography of additive manufacturing parts using a finite difference approach. *Journal of Manufacturing Science and Engineering*, 136(6), p.061009.
- [11] Kaji, F. and Barari, A., 2015. Evaluation of the surface roughness of additive manufacturing parts based on the modelling of cusp geometry. *IFAC-PapersOnLine*, 48(3), pp.658-663.
- [12] Lalehpour, A. and Barari, A., 2018. A more accurate analytical formulation of surface roughness in layer-based additive manufacturing to enhance the product's precision. *The International Journal of Advanced Manufacturing Technology*, 96(9-12), pp.3793-3804.
- [13] Lorensen, W.E. and Cline, H.E., 1987, August. Marching cubes: A high resolution 3D surface construction algorithm. In *ACM siggraph computer graphics* (Vol. 21, No. 4, pp. 163-169). ACM.
- [14] Marchesi, T.R., Lahuerta, R.D., Silva, E.C., Tsuzuki, M.S., Martins, T.C., Barari, A. and Wood, I., 2015. Topologically optimized diesel

- engine support manufactured with additive manufacturing. *IFAC-PapersOnLine*, 48(3), pp.2333-2338.
- [15] Bender, D. and Barari, A., 2018. Convergence Control For Topology Optimization.
- [16] Barari, A. and Pop-Iliev, R., 2009. Reducing rigidity by implementing closed-loop engineering in adaptable design and manufacturing systems. *Journal of Manufacturing Systems*, 28(2-3), pp.47-54.
- [17] Jankovics, D., Gohari, H. and Barari, A., 2018. Constrained Topology Optimization For Additive Manufacturing Of Structural Components In Ansys®.
- [18] Huang, X. and Xie, M., 2010. *Evolutionary topology optimization of continuum structures: methods and applications*. John Wiley & Sons.

Advanced Energy Systems

Exergy Analysis of Renewable Resources for Energy Production in Remote Communities

Mudit Nijhawan

Department of Mechanical, Automotive & Materials
Engineering
University of Windsor
Windsor, Canada
Nijhawam@uwindsor.ca

Dr. Ofelia A. Jianu

Department of Mechanical, Automotive & Materials
Engineering
University of Windsor
Windsor, Canada
Ofelia.Jianu@uwindsor.ca

ABSTRACT

Remote communities in the North of Ontario survive in isolation as their proximity to the southern industrial sector of the province limits their accessibility to the major grid. Within Ontario, there are approximately 30 remote communities with a combined population of 15,000 people. The lack of grid connection has led to antiquated methods of power generation which pollute the environment and deplete the planet of its natural resources. Aside from the primary means of electricity generation being through the use of diesel generators, generation infrastructure is deteriorating due to age and poor maintenance practices. The stagnation of the power supply and the growth of the community has led to communities facing load restrictions. These challenges may be resolved by introducing clean energy alternatives and providing a fuel blend option. The primary energy sources investigated in this research are solar, wind, and hydrogen. To assess the viability of these energy production methods in Northern communities, an exergy analysis will be employed as it utilizes both the first and second law of thermodynamics to determine systems' efficiency and performance in the surroundings. Exergy efficiency provides a more accurate method of assessing the resources by taking into consideration their physical, kinetic and chemical properties. Local weather patterns will be used to determine the viability of using wind turbines, solar panels and/or hydrogen fuel cells in a remote community. Weather data will be obtained from local weather stations using the government of Canada's weather repository. Based on the exergy efficiencies of the individual energy conversion methods the optimal fuel blend will be provided. This method can be used in conjunction with other analysis methods such as cost analysis and environmental assessments to develop a holistic approach that predicts the optimal fuel blend for remote communities.

Thermodynamic Modeling and Optimization of Ejector Refrigeration Systems Using Alternative Working Fluids

Aggrey Mwesigye, Amir Kiamari, Seth B. Dworkin
Mechanical and Industrial Engineering Department
Ryerson University
Toronto, Canada
Corresponding Author: seth.dworkin@ryerson.ca

ABSTRACT

Heating, ventilation, air conditioning and refrigeration (HVAC&R) systems account for a significant portion of world energy use. They therefore contribute a significant portion of global CO₂ emissions, exacerbating climate change. In addition, conventional HVAC&R systems based on the vapor compression cycle make use of refrigerants with high global warming potentials. As such, there are increased efforts to develop systems that are less reliant on fossil fuels and that use environmentally benign working fluids. The ejector refrigeration system is one such system with the potential to be incorporated in HVAC&R. It is simple, has no moving parts, thus uses less power and requires limited maintenance. However, the coefficients of performance (COPs) of this technology are still very low and most studies have considered performance with refrigerants that are harmful to the environment. Moreover, current models for determining the performance of ejector refrigeration systems give large errors in comparison with available experimental data and consider mostly the performance in the critical mode of operation with, fewer studies presenting results on performance in the subcritical mode of operation.

In this study, a novel and precise numerical model that predicts the performance of an ejector refrigeration system in both the critical and subcritical modes of operation when working with dry and isentropic refrigerants was developed. Furthermore, a chart presenting the relationship between the COP and the generator temperature for given evaporator temperatures, condensing temperatures and area ratios is suggested for use in the design of these systems. The performance with environmentally benign working fluids: R600, R1233zd(E), R1234ze(Z) and RE245fa2 was compared with the performance of conventional working fluids: R141b and R245fa. Engineering Equation Solver (EES) was used for the present study, with evaporator temperatures between 0°C and 6°C, condensing temperatures between 20°C and 40°C, ejector area ratios between 4.45 and 12.98, and generator temperatures between 70°C and 110°C were used.

Results showed that R600, R1234ze(Z) and R1233zd(E) gave better performance than R141b and R245fa, the widely used and studied working fluids in ejector refrigeration systems. Results further showed that for each refrigerant, there is a generator temperature that gives optimal performance. The COPs at the optimal performance point were in the range 0.35 – 0.60 for R600, 0.14 to 0.30 for R1233zd(E), 0.20 – 0.45 for R1234ze(Z) compared to 0.15 - 0.29 for R141b when the condensing temperature is 35°C with an evaporator temperature of 8°C.

Keywords- *coefficient of performance; critical mode of operation; ejector refrigeration system; HVAC&R; optimal performance*

Numerical Modeling of a Thermal Energy Storage System for Ground Source Heat Pump Applications

Hiep V. Nguyen, Aggrey Mwesigye, Pedram Atefrad, Reza
Daneshazarian, Seth B. Dworkin
Mechanical and Industrial Engineering Department
Ryerson University
Toronto, Canada
Corresponding Author: seth.dworkin@ryerson.ca

David Salt
Capture Technologies
McMaster Innovation Park
Hamilton, Canada

ABSTRACT

Ground source heat pump systems have the potential to reduce energy consumption in buildings. They rely on the relatively stable ground temperatures to supply heating with reasonably low power consumption when ambient temperatures drop below freezing. However, in cases where the heating and cooling loads of a ground source heat pump system differ significantly the result is a ground thermal imbalance. With the thermal equilibrium destroyed the ground temperature drifts leading to a gradual loss in performance and in some cases a system failure. With this, several research efforts are directed towards improving performance and alleviating thermal balance issues affecting ground source heat pump systems. For improving thermal imbalance, ground thermal energy storage is becoming increasingly appealing. Moreover, the use of high thermal conductivity material to transfer energy to and from the borehole is seen as an effective way of further improving performance.

In this study, a numerical model of the performance of a thermal storage medium for ground source heat pump applications is presented. The finite element model is developed and validated for a thermal storage tank of 1 m diameter and 1.5 m height, equipped with temperature sensors at different radial positions and depths. The tank includes a steel pipe of 15 cm diameter to emulate a borehole in which a 2.54-cm diameter u-shaped pipe is located. To study the energy transfer rate from the borehole to the surrounding material in the tank, different combinations of materials were considered in the borehole and in the volume surrounding the borehole. In the first part of the study, the use of high thermal conductivity bentonite-sand-water mixtures was considered to determine the temperature distribution and rates of heat transfer.

In the second part of the study, the potential for energy storage using phase change material was investigated. Nine pipes, 5.08 cm diameter and 122 cm long were filled with a paraffin-based phase change material with a latent heat of 190 kJ/kg and a melting temperature of 24°C and placed 5.08 cm away from the steel pipe that emulates a borehole. The bentonite-sand mixture giving high thermal conductivities shows better performance than the bentonite-only configuration. The temperature distribution in the tank improves as the thermal conductivity improved. A reduction in the tank temperature difference of about 34% is achieved by using bentonite-sand mixtures in both the borehole and the tank respectively than with the bentonite-and-sand only mixture. Conversely, the heat transfer performance increases by about 90% with the use of the bentonite-sand mixture. Moreover, the potential for energy storage with phase change material has been demonstrated. During the charging phase and depending on the amount of phase change material used, energy is effectively stored and released during the discharge phase.

Keywords- *bentonite-sand mixture; ground source heat pump; phase change material; thermal energy storage, thermal imbalance*

Economic analysis of off-grid hybrid solar-hydrogen energy systems with energy storage

Amin Ghobeity

School of Mechanical and Electrical Engineering
Sheridan College

7899 McLaughlin Rd., Brampton, L6Y5H9, ON

amin.ghobeity@sheridancollege.ca

ABSTRACT

Numerical simulation and economics analysis of three types off-grid hybrid energy systems involving a solar energy system and a hydrogen fuel cell is presented. Most of the small-scale, off-grid hybrid energy systems considered in the literature are either photovoltaics (PV) with electricity storage (batteries), or (PV) hybridized with a diesel generator. If an emission free off-grid power system is desired, then the choice is primarily PV with batteries. In this presentation, other than PV with batteries two other different types of hybrid clean energy systems are considered. Electrolyzers and hydrogen fuel cells can be hybridized with PV to store energy and augment PV when needed (e.g., at nights and cloudy days). The capability to store energy at a relatively large scale in the form of hydrogen provides significant operational flexibilities for an emission-free islanded micro-grid. In this study, economic viability of three different systems are discussed: a) solar photovoltaics (PV) and electricity storage using large scale batteries [PV+Battery], b) solar PV with hydrogen polymer exchange membrane fuel cell (PV+FC), hydrogen generator (H2GEN) and hydrogen storage (high pressure tanks) [PV+FC], and c) Solar PV used with a hybrid energy storage involving both electricity storage and hydrogen storage [PV+FC+Battery]. Hourly electricity demand, energy production, energy storage charging/discharging, and various energy storage options are considered in the yearly simulation of the systems. Optimization of the design and operation is achieved considering a number of optimization variables.

The location selected for the yearly simulation of the proposed systems is Brampton, ON. The hourly electricity load of a typical residential home 5.75 kW average demand, 10 kW peak and consumption of 137.9 kWh/day, scaled to 8.33 kW average demand, 14.5 kW peak and consumption of 200 kWh/day, is used for the yearly simulation of the systems considered. A time step of 1 hour is considered.

The optimization variables considered are: PV=[50,60,...,200kW]; DC/AD Inverter=[10,12.5,15,...,25 kW]; FC=[0.5,7.5,...,25 kW]; H2GEN=[0,10,15,...,70 kW]; Battery (67.2 kWh capacity)=[0,1,2,...,6], and H2TANK=[0,10,15,20,...,50]. The costs assumed in the simulation are PV=2,000 \$/kW installed with negligible operating cost; FC= 2000 \$/kW installed; H2GEN= 400 \$/kW; Battery= 47,500 \$/battery (67.2 kWh battery), and H2TANK= 600 \$/kg of H2 stored.

The levelized cost of electricity (LCOE) produced using proposed systems are found using HOMER, a high fidelity micro-grid simulation software package. The results of the optimization case studies are:

- System (a): [70 kW PV+ 4 x 67.2 kWh Battery], with capital cost of approximately \$337,500 yielding a LCOE for this system is **0.327 \$/kWh**
- System (b): [90 kW PV+ 12.5 kW FC + 30 kg H2 Storage + 70 kW H2GEN], with capital cost of \$259,750 yielding a LCOE for this system is **0.295 \$/kWh**
- System (c): [80 kW PV+ 1 x 67.2 kWh Battery + 5 kW FC + 15 kg H2 Storage + 45 H2GEN], with capital cost of 252,000 yielding a LCOE for this system is **0.272 \$/kWh**

The results show that using a highly hybridized system such as PV+FC+Battery, which is compromised of two technologies for power generation (PV and FC) and two technologies for energy storage (Hydrogen and batteries), provides significant operational flexibilities. From an economical point of view this results in the adequacy of smaller equipment to meet the electricity demand and consequently lower capital and operating costs. This ultimately yield a lower LCOE. Such configuration may be deemed as impractical for small scale off-grid power system requirements (e.g., a remote off-grid farm), nevertheless one can conclude that larger and more complex systems can benefit from such configuration. For the smaller systems, requiring less complexities, the study shows that the use of hydrogen storage results in lower energy costs, considering yearly simulation of the system.

INVESTIGATING n-HEPTANE SPRAY CHARACTERISTICS INSIDE THE IGNITION QUALITY TESTER (IQT™) CHAMBER THROUGH TEMPERATURE VARIATION MEASUREMENTS

Osama Hmood and Edgar Matida

Department of Mechanical and Aerospace Engineering
Carleton University
Ottawa-Ontario, Canada
e-mail: osama.hmood@carleton.ca

Dave Gardiner, Luc Menard, Omar Ramadan, Aaron

Wilcox, and Gary Webster
Advanced Engine Technology Ltd.
Ottawa-Ontario, Canada
e-mail: gary@aet.ca

ABSTRACT

The IQT™ (Ignition Quality Tester) is a constant-volume combustion chamber apparatus used to measure diesel-fuel ratings (i.e., the derived Cetane number). In normal IQT™ operation, the time between the start of the fuel spray and its auto-ignition is measured using information from the chamber-pressure variation in time. In the current work, n-heptane (C_7H_{16}) is used as fuel and temperature measurements inside the IQT™ chamber are investigated along with the pressure variation. Fine gauge type K thermocouples were used to record the temperature variations at 46 pre-selected locations inside the IQT™ for the period prior to the n-heptane injection up to the fuel-air mixture combustion. The temperature was recorded with intervals of 0.1 ms by using a high frequency data acquisition system. Each run consisted of 20 injections and the mean of the last 15 injections was presented in the results. Wide range of initial (chamber set-point) temperature (530 – 590 °C) was covered in this study at three values of initial pressure (10, 15, and 21 bar). The temperature measurement results indicate that the spray velocity decreases gradually while propagating in the chamber. In addition, the spray droplets do not impinge on the side walls and seem to evaporate in the main part of the chamber before reaching the downstream walls.

Numerical Modelling of Helical Steel Piles as In-Ground Heat Exchangers for Ground-Source Heat Pumps

Sarah R. Nicholson, Dr. Aggrey Mwesigye, & Dr. Seth B. Dworkin
Mechanical and Industrial Engineering
Ryerson University
Toronto, Canada

ABSTRACT

In climates with significant seasonal temperature differences, Ground-Source Heat Pumps (GSHPs) can improve a building's efficiency by taking advantage of mild year-round ground temperatures. Cold climates with hot, humid summers, are thermodynamically suitable for GSHPs - but are limited by the high costs and large footprint of borehole fields. Canada's building energy use can be made more sustainable by addressing feasibility issues of existing energy efficient systems. By combining existing structural or foundation equipment with a GSHP, there is a potential to have shared cost and equipment for dual benefit.

To explore this potential, the authors and an industry partner are investigating the use of helical steel piles with integrated internal flow systems as in-ground heat exchangers. These piles can provide dual function (delivering structural support for a building as well as its heating and cooling loads) within one cost of installation. The potential reduction in helical pile GSHP retrofit and installation costs means that efficient heat pumps may be installed for small scale residential uses - creating the possibility of more energy efficient buildings in a wider range of sizes and locations. As these piles are a common structural engineering tool and traditionally exist as anchors for buildings, they are relatively inexpensive components, have well-known installation procedures, and a small footprint. The helical steel piles also provide increased heat exchanger fluid volume, which can act as thermal storage.

A fully transient, time-accurate finite element analysis model was developed utilizing computational fluid dynamics (CFD) across various pile geometries to optimize the heat transfer between the heat pump's working fluid and the soil. In addition, the temperature distribution within the soil across years of use was studied to provide critical understanding of the wider impacts of this new heat exchanger design. In parallel, an experimental testing site is being developed at a commercial building to gather experimental data on the flow, space heating/cooling capacities, and temperatures resulting from a fully functional helical steel pile GSHP system). This experimental data will be used to validate and improve the initial computer models and simulations.

Ultimately, this research aims to quantify the potential improvements to performance, costs, and GHG emission reductions of the piles, and assess their potential for deeper market penetration. Initial detailed transient finite element modelling results show some fundamental advantages to the helical installation - which can increase the surface contact of the soil to improve heat transfer. This research characterizes heat transfer and system performance of the piles in seasonal climates and directs new borehole field configurations by quantifying ground temperature responses and determining optimal spacing for various loads (with a GSHP). The approach will seek to develop a fundamental understanding of the performance, economics, and operation of these helical piles in a variety of configurations, to provide guidance for implementation in the market.

A Data-Driven Thermal Fault Detection Method in Lithium-Ion Batteries

Olaoluwa Joseph Ojo

Department of Automotive,
Mechanical and Manufacturing
Engineering,
University of Ontario Institute of
Technology,
Oshawa, Canada

Xianke Lin

Department of Automotive,
Mechanical and Manufacturing
Engineering,
University of Ontario Institute of
Technology,
Oshawa, Canada

Haoxiang Lang

Department of Automotive,
Mechanical and Manufacturing
Engineering,
University of Ontario Institute of
Technology,
Oshawa, Canada

ABSTRACT

The modern realization that climate change has unwanted consequences is causing an increasing demand for sustainable product alternatives. The automotive industry is an example of where this zeal for change exists. Electric vehicles (EVs) being the solution for the automotive industry is only viable because of the electric storage systems available such as hydrogen fuel cells, ultracapacitors, and batteries. Lithium-ion batteries are the most popular in this regard and have seen great contributions in research. Lithium-ion (Li-ion) batteries, however, suffer from a short lifespan when compared to the alternatives. This short lifespan can be attributed to degradation or the diminishing state of health of the battery over time. Knowledge of the internal dynamics as well as safety efforts in management systems is important to understand and mitigate this issue.

Thermal fault detection is an area that is important to the safety of Li-ion batteries but has not been widely investigated. The majority of published literature on lithium-ion fault-based predictions focus on faults relating to the electrochemical properties in li-ion batteries, overcharging and over-discharging, as well as sensor and actuators in li-ion battery applications. The latest approach in thermal fault detection uses a partial differential equation (PDE) model-based method to diagnose thermal faults. This method is very resource-intensive due to the highly complicated models used. The models require parameter estimations and an expert-level understanding on the working of the models. Other methods include using one or two-state lumped parameter thermal models to diagnose thermal faults. These methods also require parameter identification.

The proposed technique in this abstract is a data-driven approach that uses a trained neural network with data typically found in the battery management system (BMS) such as current, voltage, SOC, ambient and surface temperatures to detect thermal faults in the batteries. This method does not involve any physics-based or lumped parameter models, situational assumptions, or the need to identify model parameters. Accurate detection is accomplished by training the neural network based on experimental measurements from which it learns.

This method is attractive due to the ease of implementation for non-expert adopters, versatility with different batteries, and low cost on computing resources compared to the other methods.

UTILIZING GEOTHERMAL LOOPING FOR THERMAL REGULATION OF AN AQUACULTURE POND

Mitchell H. Kuska

Mechanical & Materials Engineering
Western University
London, Canada
mkuska@uwo.ca

Christopher T. DeGroot

Mechanical & Materials Engineering
Western University
London, Canada
cdegroo5@uwo.ca

Kamran Siddiqui

Mechanical & Materials Engineering
Western University
London, Canada
ksiddiqui@uwo.ca

ABSTRACT

The demand for high quality sources of dietary protein is increasing due to the ever-increasing global population. With fish being a very well-rounded source of animal protein, increasing the consumption of fish could give a growing population the protein that it needs to survive. However, decades of overfishing in natural habitats and poor management of wild fish stocks have jeopardized the sustenance of natural fish populations. Therefore, reliance on wild fish capture to meet the dietary protein requirement in years to come, may bring these aquatic species to an extinction level.

One promising alternative to natural fish capture is aquaculture. Aquaculture has been steadily increasing globally in recent decades and nearly rivals captured fish as the dominant method of global fish production [1]. With aquaculture being a sustainable source of high-quality food, it is important that it continues to grow to become the primary source of global fish supply. However, the economic competitiveness of aquaculture is affected by fluctuating water temperatures, which lead to lower fish yield through disease-driven fish kills and non-ideal growing conditions [1]. Literature review shows that very little research has been conducted to address this issue. Active heating and cooling methods, such as the use of natural gas or electricity, are not feasible due to their high cost and environmental downfalls. Therefore, an alternative, ideally passive heating and cooling, method is required.

Only a handful of published works have touched on this subject, with the majority focusing on utilizing direct injection of hot geothermal ground water to thermally regulate aquaculture pond water [2]. While this method is feasible from a thermal regulation aspect, there are some concerns over its environmental and ecological impacts, such as the depletion or salinification of aquifer water. Additionally, this option is restricted in regions that have active geothermally heated aquifers or groundwater. One feasible option that can be widely implemented is the use of the ground as a heat sink or source to thermally regulate the aquafarm via a closed loop geothermal heat exchanger system. The seasonal outdoor temperature variations influence the ground temperature only within a few meters from the surface. Typically, at depths of 10 m and more, ground temperature remains almost constant throughout the year. For example, in Southern Ontario, the ground maintains a constant temperature of 9-10 °C below 10 m depths. This temperature lies in the middle between hot and cold temperatures typically experienced in summer and winter seasons, respectively.

The objective of this study is to investigate the feasibility of thermally regulating aquaculture raceway water with geothermal borehole heat exchangers. An energy model of an aquaculture raceway was developed, taking into account all energy exchange processes that would influence the temperature of the water (e.g. solar heat flux, evaporation, convection, etc.). The model was validated against data taken from an aquaculture raceway site in Cambridge, Ontario. Next, an energy model of a set of geothermal borehole heat exchangers was developed and validated against experimental data found in the literature for a single-borehole. This quasi-three-dimensional model accounts for aspects such as thermal interference between the inlet and outlet u-pipes, u-pipe placement in the borehole, and the interference between multiple boreholes in proximity to each other. Literature shows these aspects are not typically included in borehole heat exchanger models making this a potentially more accurate model to real-world conditions [3]. Detailed results describing the number and configuration of boreholes and the heat transfer capacity of the borehole heat exchanger system to thermally regulate the aquaculture raceway will be presented and discussed at the conference.

- [1] J. Gelegenis, P. Dalabakis and A. Ilias, "Heating of a fish wintering pond using low-temperature geothermal fluids, Porto Lagos, Greece," *Geothermics*, vol. 35, pp. 87-103, 2006.
- [2] FAO, "The State of World Fisheries and Aquaculture 2016," Food and Agriculture Organization of the United Nations, Rome, 2016.
- [3] H. Zeng, N. Diao and Z. Fang, "Heat transfer analysis of boreholes in vertical ground heat exchangers," *International Journal of Heat and Mass Transfer*, vol. 46, pp. 4467-4481, 2003.

Efficient non-contact steam generation and superheating using sunlight

Thomas A. Cooper

Department of Mechanical Engineering
 York University
 Toronto, Canada
 tcooper@yorku.ca

ABSTRACT

Providing planet with a sustainable source of fresh drinking water is one of the grand challenges facing our society. Solar-driven distillation technologies are a promising emerging technology offering a simple, scalable, and low-cost solution to this problem. may provide the answer to this problem. A solar still is a device which absorbs sunlight and transfers this heat to the water thus driving evaporation. When water evaporates, salt and other contaminants are left behind, resulting in a stream of pure water in the vapour phase. This vapour can be condensed and collected to provide a source of clean drinking water. Within the last 5 years [1]–[3], there has been a concerted research effort to develop low-cost floating structures which can be deployed on bodies of seawater to passively generate steam and clean water.

All previous methods to generate steam using solar energy have relied on heat conduction from a solar absorber material *in contact* with the water. We recently demonstrated [4], for the first time, evaporation driven by a solar absorber *not in contact* with the water (see Figure 1). In this contactless configuration, energy is transferred to the water via thermal radiation, which is absorbed in a sub-100 μm layer near the water surface. Because the absorber is not in contact with the water, problems of fouling and clogging are entirely avoided. Moreover, the absorber is no longer pinned to the boiling point, and can be used to superheat the generated vapour. Under a solar flux of 1000 W/m^2 , we achieved an efficiency of 27% including vapour collection, and steam temperatures in excess of 130°C , thus marking the first demonstration of contactless solar evaporation and generation of superheated steam under one sun illumination. In addition to providing clean drinking water, the elevated temperatures achieved by this system may provides the basis for sanitization, cooking, cleaning and process heating technologies. In this presentation, recent results toward scaling up the process and methods to increasing system efficiency through innovative thermal design are presented.

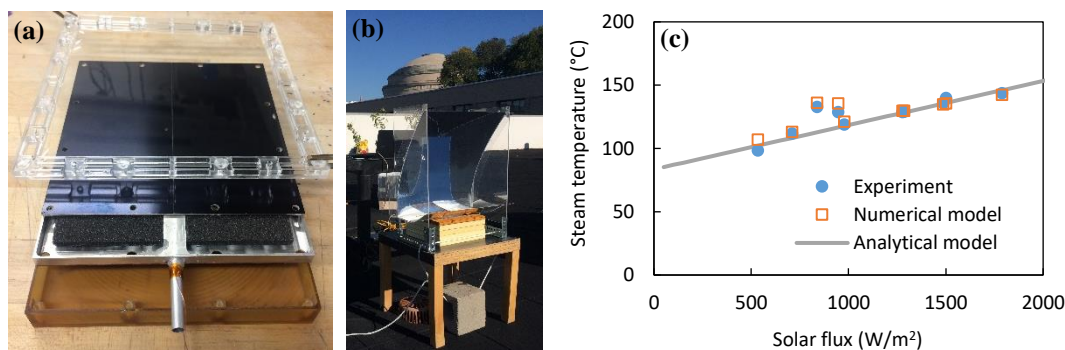


Figure 1. (a) Exploded view of contactless solar evaporation structure (CSES) prototype; (b) photograph of CSES prototype operating on the MIT rooftop; (c) achieved steam temperature versus solar flux.

References

- [1] H. Ghasemi, G. Ni, A. M. Marconnet, J. Loomis, S. Yerci, N. Miljkovic, and G. Chen, “Solar steam generation by heat localization,” *Nat. Commun.*, vol. 5, p. 4449, 2014.
- [2] G. Ni, G. Li, S. V. Boriskina, H. Li, W. Yang, T. Zhang, and G. Chen, “Steam generation under one sun enabled by a floating structure with thermal concentration,” *Nat. Energy*, vol. 1, no. August, p. 16126, 2016.
- [3] G. Ni, S. H. Zandavi, S. M. Javid, S. V. Boriskina, T. A. Cooper, and G. Chen, “A salt-rejecting floating solar still for low-cost desalination,” *Energy Environ. Sci.*, vol. 11, no. 6, pp. 1510–1519, 2018.
- [4] T. A. Cooper, S. H. Zandavi, G. W. Ni, Y. Tsurimaki, S. V Boriskina, and G. Chen, “Contactless steam generation and superheating under one sun illumination,” *Nat. Commun.*, vol. 9, p. 5086, 2018.

SIZING, MODELING AND ANALYSIS OF A SOLAR SEASONAL ENERGY STORAGE SYSTEM FOR SPACE HEATING IN NEWFOUNDLAND

Rabbani Rasha
Faculty of Engineering and
Applied Science
Memorial University
St. John's, Canada
rrasha@mun.ca

Md. Habibur Rahaman
Faculty of Engineering and
Applied Science
Memorial University
St. John's, Canada
mhrahaman@mun.ca

M. Tariq Iqbal
Faculty of Engineering and
Applied Science
Memorial University
St. John's, Canada
tariq@mun.ca

Abstract—this paper presents thermal modeling of a house and design of a solar space heating system with seasonal storage to meet the annual heating demand for residential application. Load estimation is a very important part of designing an energy system of any house. In addition, sizing of various components of the thermal system is a complex procedure requiring demand and local climate data to get the desired solar fraction and seems one of the biggest challenges of this research work. In this paper, thermal modeling is done by BE opt software and a simple calculation approach is used to find out the components design parameters of this system. A simulation software Polysun was used to get the optimal energy output of the proposed system. Design, simulation results and detailed analysis of this research work are included in this paper.

Keywords- Thermal modeling; BE opt; solar water heating; seasonal energy storage; Polysun simulation software

I. INTRODUCTION

The global energy demand, as well as the use of conventional energy production, has been increasing rapidly that can affect the environment in the form of global warming, wildfires and so on. Reducing the greenhouse gas (GHG) emission and increasing the use of renewable energy are the major challenges nowadays that can lead better environment. According to Natural Resources Canada (2015), in the residential sector, space heating consumed around 62% of total energy whereas consumption of space heating in Canada is about 55% in commercial use. It is noticeable that space heating consumed the highest energy in both cases. Also, in Newfoundland, the consumption of space heating energy higher than other secondary energy and possess approximately 70.9% with GHG emissions about 91.8% [1] [2]. Researchers and energy companies of the world have been emphasized to find a solution that can use to produce and store thermal energy from solar energy during summer season and to be used during heating period winter as it is free energy but not readily available. To overcome these flaws, seasonal solar thermal energy systems with longer storage period can be an effective way for colder countries throughout the world for feasible utilization of the

abundant resource of solar energy. System performance is one of the important factors for the end users and relies on various system configurations like solar collector, low-temperature storage water tank, auxiliary heater, heat pump and the like [3]. Extensive researches have been going on to get the best and feasible options to store thermal energy seasonally that can mitigate the mismatch between supply and demand. Out of them, sensible heat energy storage system is one of the suitable choices because of cost-effectiveness, longer lifespan and easy to maintain. C.N. Antoniadis and G. Martinopoulos [4] designed a solar thermal system with seasonal thermal energy storage and simulated using TRNSYS software for a single-family detached house in Thessaloniki, Greece. They found almost 52.3 % space heating demand covered by the STES system but was able to fulfill the hot water load over the whole year. A seasonal solar thermal energy storage system using TRNSYS simulation software made for a student housing project at Virginia Commonwealth University followed the American Society of Heating and Air-conditioning Engineers (ASHRAE) specifications. Due to the availability and cost, sand was used to the storage medium. Simulations were done for five years to make the sand based system steady state and concluded that an effective and efficient SSTES system could meet up to 91% heating energy demand of the building [5]. A solar assisted heat pump system's performance was simulated by Lie et al. [6] with the help of TRNSYS software for covering the space heating and domestic hot water demand with seasonal storage. Water used as a storage medium of their system as well as an air-to-water heat pump unit, a water-to-water heat pump unit with other necessary components used for making the simulation model. After completing one-year simulation period and comparing the findings to conventional space heating system, it was noticeable that domestic hot water solar fraction and energy saving ratio was around 68.1% and 52% respectively on a monthly basis.

On the other hand, the annual seasonal storage tank's energy storage efficiency was about 64% that was for space heating. A loss free thermal energy storage system was built by the researchers in Germany using small size Fraunhofer's zeolite can store up the heat to four times more than the water for

indefinite periods [7]. A water tank based system was designed and simulated by Lund [8] for central solar heating plants with seasonal storage (CSHPSS) and concluded that about 35% to 60% of solar fractions for this type of system was increased using stratified storage tank rather than the fully mixed tank. Wills et al. [9] designed and simulated a solar thermal system with seasonal storage as a part of C-RISE project at Carleton University. A domestic hot water tank and buried concrete tank were used to provide hot water and space heating demand for the selected house. A co-simulation tool was used to do the different parametric and sensitivity analysis and found that this system at C-RISE house met about 89.2 % space heating and domestic hot water demand. A study showed that renewable energy has been contributing to meet around 17% of primary energy demand in Canada that leads them to lessen GHG emissions by 30% underneath 2005 levels by 2030 [10].

From the literature search, it found that renewable energy based system and projects should be implemented in the residential sector to reduce the greenhouse gas emissions and provide a better and livable environment. The best option to utilize this solar energy is residential and industrial use through the proper design, sizing, and control to make this system economically viable and more practical. If it happens, then the proposed solar space heating system will be useful to supply thermal energy in the houses of Canadian's.

In this research work, section 2 will present the thermal modeling and analysis. Section 3 will show the proposed system and methodologies. Sizing of proposed system components and result with discussion will be discussed respectively and will end with a conclusion and references.

II. THERMAL MODELING AND ANALYSIS OF A HOUSE

To design a solar system the total amount of space heating load per year is required. A thermal model of the house was developed using BEopt (Building Engineering Optimization) software to get that load. The house has two floors, contains four rooms with two washrooms on the first floor and four rooms with a half washroom on the ground floor. Overall dimension (length: 45 feet and width: 30 feet) gave as an input at Geometry screen also the correct values for all outside walls, windows, doors, roof, orientation, location, all major electrical appliances, and their types gave as inputs to BEopt. The geometry and output screen of this software is shown in figure 1 and figure 2.

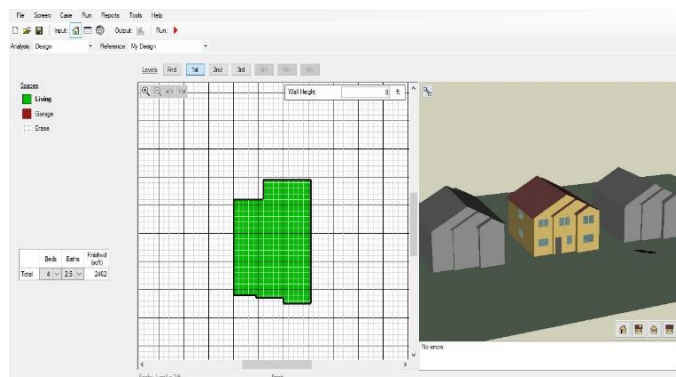


Fig. 1. Geometry screen of a designed house (input)

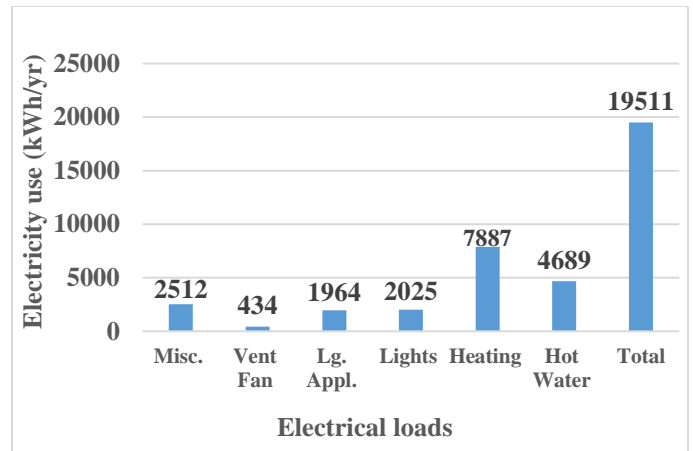


Fig. 2. Electricity consumption of designed house for one year

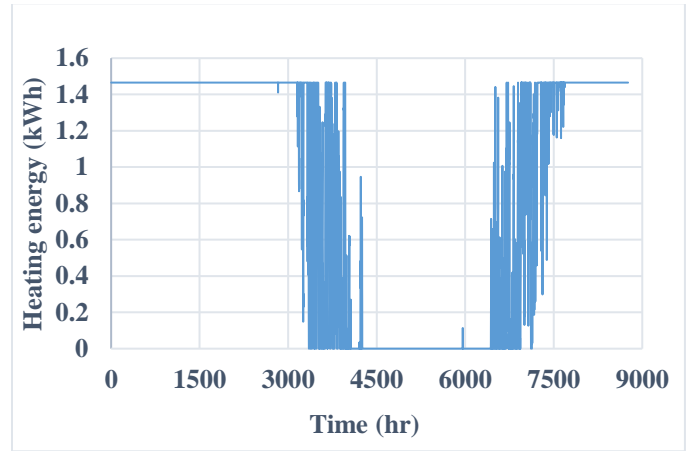


Fig. 3. Hourly space heating consumption profile from BE opt data viewer

From BEopt, the hourly space heating consumption data was found and it shows the average hourly space heating demand around 0.90 kWh. Figure 3 above represents the data. To get output result accurately and precisely, some selected input parameters regarding the house location, properties of making house materials and all major home appliances were used as an input of thermal modeling and analysis software.

After analyzing the result, it found that total electrical energy consumption of the designed house is 19511 kWh/yr. Among these, the hot water consumption 4689 kWh/yr., Space heating 7887 kWh/yr., Lights 2025 kWh/yr., Lg. Appl. 1964 kWh/yr., and Miscellaneous 2512 kWh/yr. and Vent. Pump 433 kWh/yr. Therefore, our main concern is to design a solar heating system that can replace the electric space heating energy consumption 7887 kWh/yr. and reduce the overall electricity bill with seasonal energy storage. A list of these properties is given below in table 1.

TABLE I. SELECTED INPUT PARAMETERS OF A HOUSE FOR OPTIONS SCREEN:

House parameters	Particulars	Descriptions
Building	Orientation	Northeast
Wall	Wood Stud	R-15 fiberglass batt 2*4, 16 in o.c
Ceilings/ Roofs	Unfinished attic	Ceiling R-38 fiberglass, vented
Windows & Doors	Window areas	A new input of window areas has inserted with clear, double, non-metal, air
Space conditioning	Electric baseboard	100 % efficiency
	Ducts	7.5 % leakage, uninsulated
Airflow	Air leakage	4 ACH50
	Mechanical ventilation	2010, HRV, 70%
Water Heating	Water heater	Electric standard
	Distribution	Uninsulated, HomeRun, PEX
Lighting	Lighting	34% CFL hardwired, 34 % CFL plugin
Appliances and fixtures	Refrigerator	Top freezer, EF= 21.9
	Cooking range	Electric
	Dishwasher	290 rated kWh, 80% usage
	Clothes washer	Energy star cold only
	Clothes dryer	Electric

III. PROPOSED SYSTEM AND METHODOLOGIES

According to our literature search and the geographical feature, one of the suitable option for our selected area in St. John's, NL, Canada, latitude (47.56 °N) and longitude (52.71 °W), is active closed loop solar heating system that can produce our required space heating energy with seasonal storage. Outside temperature of this area drop below zero, easy to maintain, more reliable, longer lifespan and cost-effective due to use of water storage tank are the primary reasons to choose this type of system. The proposed solar space heating system (SSHS) is as follows:

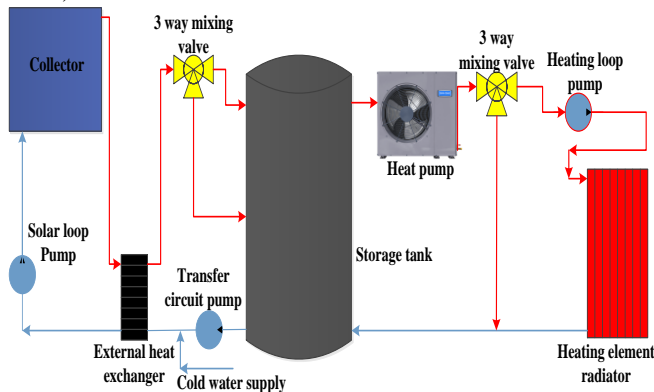


Fig. 4. Proposed space heating system with components

The proposed system consists of a solar collector, storage tank, external heat exchanger, pump, heat pump, three-way mixing valve, heating element radiator with necessary control devices. In this proposed work, all control strategies neglected. The working principle of this system is similar to active closed-loop water heating system where solar energy transferred directly to working fluid and circulates this working fluid from the collector up to the storage tank. An external heat exchanger transfers heat from working fluid to storage tank water before back it again with the help of a pump. Here, the water tank worked as a seasonal storage tank for space heating purpose during the heating periods. A heat pump is used to keep and supply the desired temperature of the system throughout the year. Thus, the hot water is passed through the radiator heating element of the building and finally stored in the seasonal storage tank. Various components design with detailed parameters according to our research criteria discussed in the next section.

IV. SIZING AND SELECTION OF PROPOSED SYSTEM COMPONENTS

A. Solar collector

System performance and the efficiency depend on the perfect sizing of the components of any systems. In our system, the main components are mainly the solar collector and storage tank. To design the solar collector size perfectly, different equations were used from the literature we reviewed that can meet the space heating demand of the house in the heating period. The calculation method for getting the solar collector area is as follows:

$$\text{Collector area, } a1 = Q_{\text{demand}} / Q_{\text{solar1}} \quad (1)$$

Here, Q_{demand} = Monthly energy demand of the house from BE opt and Q_{solar1} = Monthly solar radiation of the selected area from RET Screen (3.06 kWh/m²/day). The initial collector area, $a1$ has calculated by using (1) and using this collector area, the new Q_{solar2} has calculated from (2) and found out the minimum excess energy from (3)

$$Q_{\text{solar2}} = a1 * Q_{\text{solar1}} \quad (2)$$

$$Q_{\text{excess}} = Q_{\text{solar2}} - Q_{\text{demand}} \quad (3)$$

After that, some assumptions have been taken from literature like collector efficiency $\eta=50\%$, standby loss of insulated storage tank=0.05 kWh/hr. and circulation loss through insulated pipes=5% and using below equation Q_{solar3} was found.

$$Q_{\text{solar3}} = \eta * Q_{\text{solar2}} - \eta * Q_{\text{solar2}} * 0.05 - 0.05 \quad (4)$$

$$\text{Collector area, } a2 = Q_{\text{demand}} / Q_{\text{solar3}} \quad (5)$$

Finally, using (5) got the new collector area, which one is larger than the previous one. However, following the same procedure our desired solar collector area with other parameters like final collector area, $a=16 \text{ m}^2$, excess energy amount: 0.150 kWh were found. For our proposed system, flat plate collector has chosen and selected collector model: Honeycomb collector HC1-A, manufacturer: TIGI Ltd., test standard: North America, data source: SRCC. In addition, a total number of collectors and arrays were used 9 and 8 respectively as well as the collector tilt angle and orientation were 47° and 0°. This type of collectors are available in the market with affordable cost.

B. Storage water tank

Based on our proposed system, a storage tank commonly referred to seasonal storage tank was designed without using any sizing software. Storage tank plays a vital role to hold the hot water up to the desired temperature. The amount of solar energy available in the summer month of selected area St. John's, NL, Canada used to determine the tank volume. May to September is considered the summer months in this research work.

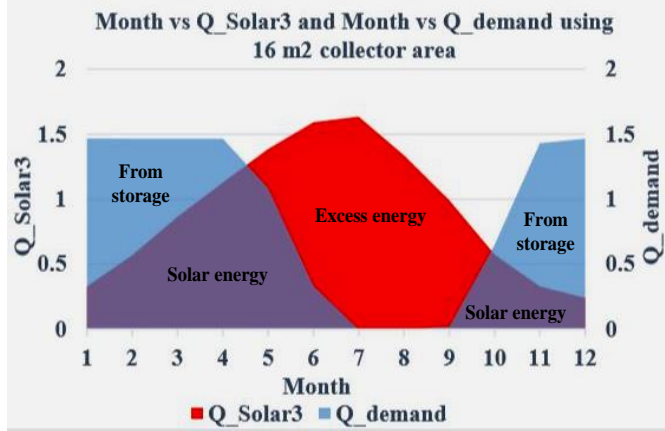


Fig. 5. Month vs Q_{Solar3} and Month vs Q_{demand} for storage tank design

The above figure shows the demand of space heating in summer months is less than the colder months and the difference between available solar energy and energy demand from summer months considered as the excess energy demonstrated by red color. So, the total amount of excess electrical energy, $E=4100$ kWh from this summer months, need to be stored in the storage tank for winter use. Several equations have been used to find out the storage water tank diameter and height, and these are as follows:

$$\text{Excess energy, } E = m * C_p * \Delta T \quad (6)$$

$$\text{Mass of water, } m = E / (C_p * \Delta T) \quad (7)$$

$$\text{Volume of water tank, } V = m / \rho \quad (8)$$

$$\text{Volume, } V = (\pi * D^2 * H) / 4 \quad (9)$$

Here, $D=H$ =Height and respective diameter of the tank (m), ΔT = temperature difference (75°C), C_p = specific heat of water ($4.2 \text{ KJ/Kg}^\circ\text{C}$) and ρ =water density (1000 Kg/m^3). After the calculation, it figured out the following values of water storage tank such as mass of water, $m= 47000$ liters, volume of water tank, $V= 47 \text{ m}^3$ and $D=H= 3.91$ m. The assumed storage tank height to diameter ratio, $H/D: 1$ and several important parameters were taken into account in order to design the storage tank more accurately and perfectly from default value of Polysun software such as material of the tank: enameled steel, insulation material: flexible polyurethane foam with thickness 101 mm. Besides, the thickness of the top and base of the tank were 100 mm and 76 mm respectively.

It is noticeable that the volume of the designed water tank is larger because if we take one-day solar irradiance that will not provide the required energy to get the desired hot water. Therefore, back up stored hot water was inevitable for space heating purpose during winter. This type of storage water tanks

is available in the market that can incorporate our proposed solar system design.

C. Pump

Pump commonly referred to as circulator plays an important role in circulating the hot water in the system. In the proposed system, there are three types of loop available. These are the solar loop pump, transfer circuit pump, and heating loop pump. Several parameters should take into account under consideration like flow rate, electricity consumption, pressure drop, pump speed, etc. In this proposed system, variable speed pump used. Some simple equations can be used to get the power consumption of these pumps. According to the International energy agency-solar heating and cooling program Task 26 [11], below equation can find the power consumption of the solar collector loop pump,

$$\text{Power1} = [44.6 * \exp(0.0181 * A_c)] \quad (10)$$

Also, for the other loops, the power consumption can be found by below equation,

$$\text{Power2} = [78.3 * \exp(0.0156 * A_c)] - \text{power1} \quad (11)$$

Using the above equation, it is found that the power requirement of a solar loop pump is 59.54 W and other two loop pump is around 40.91 W each respectively.

D. Heat pump

In this research work, a heat pump is used to transfer thermal energy from the heat source. It mainly absorbed the heat from the colder space like air and released it to warmer water. It followed the refrigeration cycle when the use of it is heating. Though the installing cost is higher than the electric heater, it is more efficient and can provide three to four times more energy depend upon the weather condition [12]. A coefficient of performance, seasonal coefficient of performance and seasonal performance factor are the important parameters for selecting the most efficient heat pump in typical space heating application. For this research, the selected model was Belaria 8 KW Air to the water heat pump. The coefficient of performance (COP) and the performance factor of this heat pump are 3.5 and 2.96 respectively with a variable flow rate. Simulation result with necessary graphs discussed in the result section.

E. External heat exchanger

A heat exchanger is a device, which can be used to transfer thermal energy (enthalpy) between two or more fluids, such as between cold water and hot water. There are many types of heat exchangers available in the market. Different heat exchangers have different flow pattern, and these divided into the parallel flow, counter flow, multi-pass and cross-flow and the like. In this proposed system, an external heat exchanger was used with the heat transfer capacity 5000W/K and the number of heat exchanger plates 20.

F. Working fluid

Working fluid is one of the important features where freezing condition can occur like colder countries. Therefore, the selection of suitable working fluid depends on many factors. Working fluid will be more efficient when it extracts the maximum amount of thermal energy from the collector field.

Propylene mixture used as a working fluid with a fluid concentration around 33.3 % in this proposed system.

G. Radiator

A radiator is one type of heat exchanger that can be used to transfer thermal energy from one medium to another for space heating. It mainly follows the convection heat process where hot water circulates through the pipes with extended surface area or often called fins [13]. In this proposed system, 2.5 m² area based radiator was used with 1000 W power per heating element under standard condition. Moreover, the selected nominal inlet and return temperature of the radiator were 45°C and 35°C respectively.

V. SIMULATION RESULT AND DISCUSSION

After sizing the major components of the proposed system, the full system was simulated for one year using professional Polysun design software, from Switzerland. Version 11 was used of this simulation software to find out various components output as well as system overall performance along with various conditions. In addition, it took different parameters as an assumption to link with individual components, which were suitable and more convenient to get an accurate and appropriate result. Fig. 6 demonstrated the overall production and demand of electrical energy for a given year of this selected house. From this figure, it can be said that proposed system met the demand of space heating completely and the overall production of electrical energy is higher than the demand not only in the winter months but also for the whole year. The heating period was considered from November to April and the total demand for these months is 6337 kWh, and the production from this system is 6932 kWh. On the other hand, the total annual demand of this house is 7887 kWh, and the total production is 12751 kWh including the losses of system components and the loss to an indoor room.

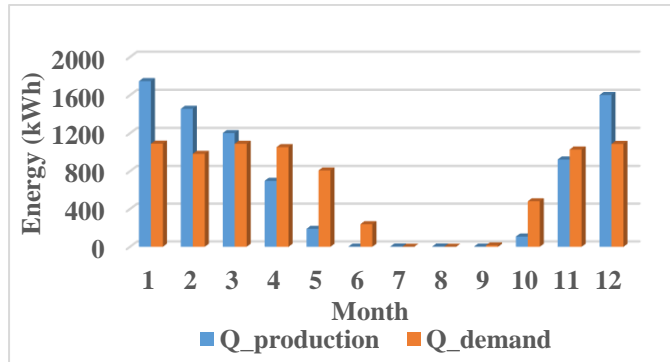


Fig. 6. Annual demand and production of the proposed system

Fig.7 demonstrated the various layer temperature of the storage tank. Storage tank divided by ten layers, the top layer possessed the higher temperature comparing with the bottom layer. It was visible that in summer months all three layers of temperature were more than winter months. The heating setpoint temperature was 17 °C. From below figure, it is concluded that this setting temperature reached in storage tank through the whole year with the average top layer temperature was around 43.9 °C and the bottom layer 33.2 °C.

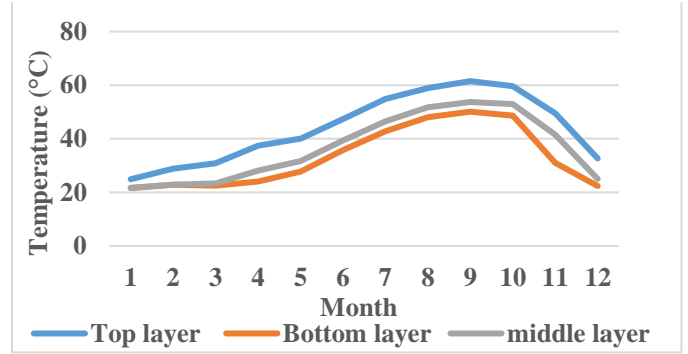


Fig. 7. Different layer temperature of the storage tank

Fig. 8 demonstrated the daily top layer temperature of the storage tank. The highest top layer temperature possessed 61.5 °C in September and lower 24.9 °C in extreme colder month January and January 1 was set as the reference day for hourly analysis in this research work. In the case of daily profile, the temperature started to rise after 8 AM and reached highest between 1 PM to 4 PM. Finally, this value ended with almost constant configuration. Either summer or winter, hourly temperature difference was lower and almost similar.

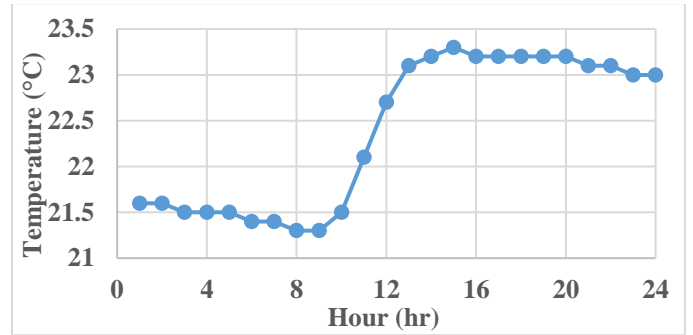


Fig. 8. Daily top layer temperature of the storage tank

Fig. 9 discussed the operation and mean temperature of the solar collector field in constant flow rate. Operation temperature was started at 28 °C and ended above 25 °C similarly, the mean temperature was started and ended with 5 °C. Though the temperature range is different for both but the temperature pattern was almost similar and in summer the operation and mean temperature was higher because of less heat loss and more irradiation on collector field.

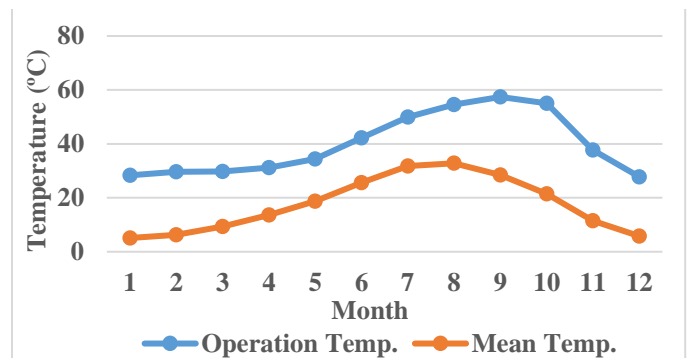


Fig. 9. Operation and mean temperature of the solar collector

Solar irradiation on the collector and the collector field yield were discussed by below fig. 10. Form the graph it is noteworthy that solar irradiation on the collector field was higher than the collector field production. Due to the losses of collector area through the year, the rate of useful solar collector field yield became like this.

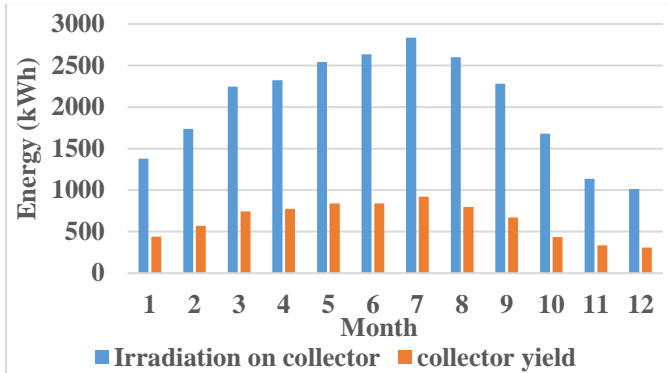


Fig. 10. Irradiation on collector and collector field yield of the proposed system

Fig. 11 showed the heat pump electricity consumption concerning the total production for one year. A heat pump was used to provide additional energy to the system in only winter months. An air source heat pump was used that took air from outside to increase hot water temperature of the storage tank up to the desired level for space heating. The electricity production was three times more than the consumption. The coefficient of the performance (COP) of the heat pump was 3.5. Also, the production and consumption of it were 4812 kWh and 1623 kWh respectively. So, the temperature of the heat pump's thermal energy was higher in colder months and firmly constant in summer months.

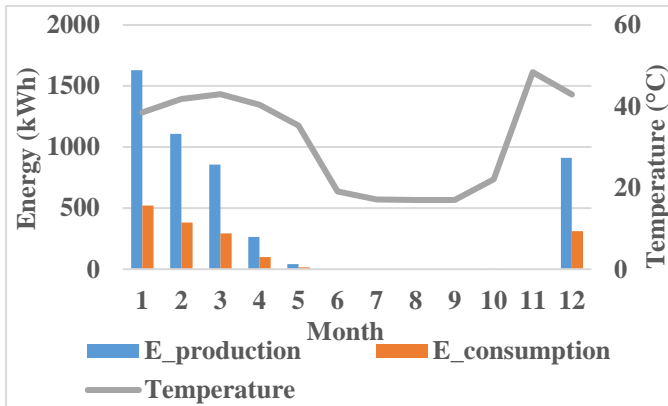


Fig.11. Electricity consumption and production with temperature of heat pump

The fig.12 described the temperature profile of the heating element radiator. Actual inlet and the outlet temperature of the radiator showed a similar character. Inlet temperature was comparatively higher than the outlet temperature. Both were started from the temperature range of 20°C to 25 °C and ended with the similar range. After the heating periods, the graphs started to rise because in summer there was extra thermal energy as well the demand for space heating was less with a constant flow rate.

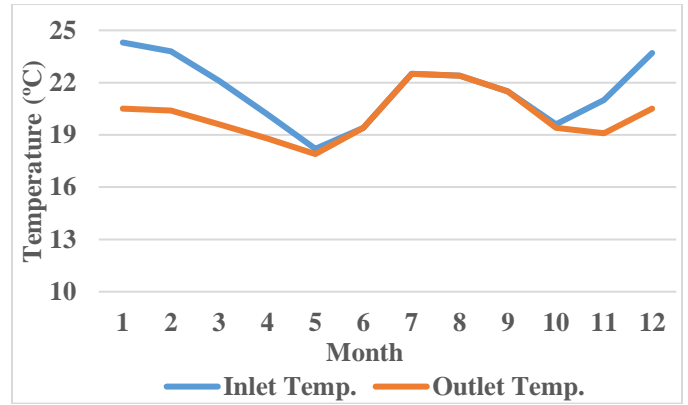


Fig.12. Inlet and outlet temperature profile of heating element radiator

Fig. 13 described the external heat exchanger supply and demand temperature with its transfer efficiency. From the result, it found that the temperature difference between supply side and demand side was very less though in summer months both the temperature were higher than the colder months. Besides, the transfer efficiency of heat was constant through the winter months, and this value was decreased by only 1 % in summer months.

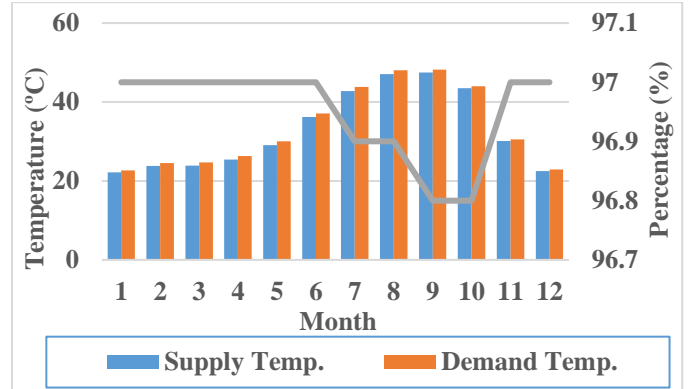


Fig.13. Temperature and transfer efficiency of the external heat exchanger

Fig. 14 discussed the daily pump solar loop temperature. In proposed system, three different pump loops were used, mentioned earlier. After analyzing the daily loop temperature of these, it was obvious that pump solar loop and transfer circuit loop have similar temperature pattern, means after 9 AM the temperature started to rise, reached highest at noon, and then started to decrease the value because solar collector got the highest solar radiation on that time. On the other hand, pump heating loop showed slightly different profile with less temperature difference. But in case of yearly loop temperature, pump solar loop and transfer circuit loop possessed the highest temperature in summer and the pump heating loop achieved the highest temperature in winter.

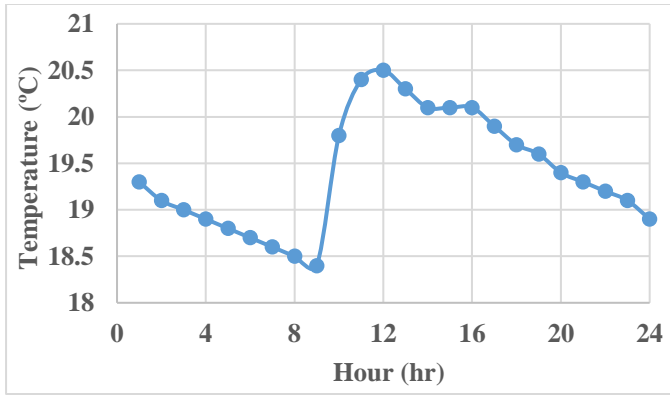


Fig.14. Daily pump solar loop temperature

Fig. 15 demonstrated the supplied heating energy and the room temperature of the building for the designed system. The room temperature was above and near about 17 °C during the heating periods as it was the setting temperature for the building, but in summer months it went higher and reached at 22.5°C because of comparatively lower space heating consumption mentioned earlier section. Besides, the supplied heating energy was sufficient to meet the space heating demand.

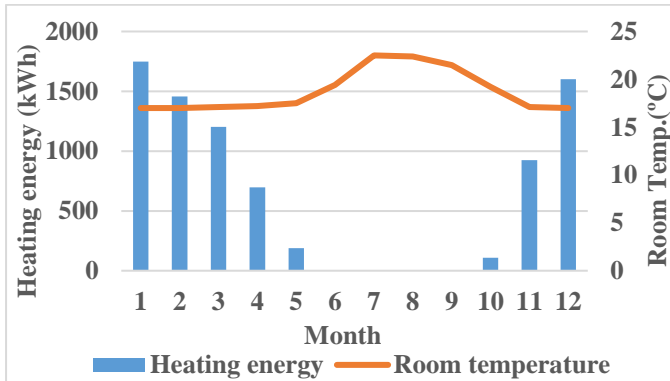


Fig.15. Supplied heating energy and room temperature of the building

Fig. 16 described the overall solar fraction and irradiation fraction of the proposed system. From the result, it can be said that the overall solar fraction was 61.4 %, but in the summer months, the percentage was 100. On the other hand, the overall solar irradiation on the collector was 93.8% but during the summer months, it was 100%. Due to the various loss, lower efficient collector and more space heating demand in winter were behind of this.

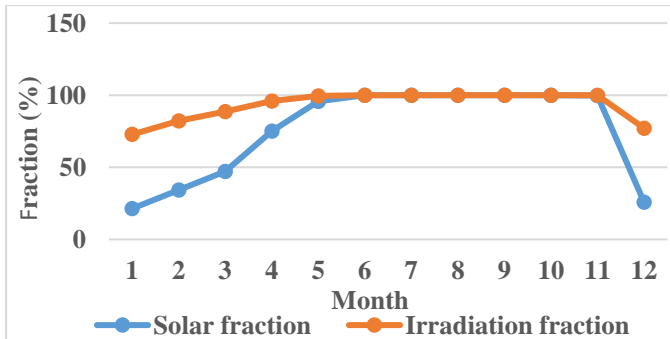


Fig.16. Solar and irradiation fraction on the collector of the proposed system

Different losses from the system were described by fig. 17. Four major losses through the system were discussed in this research work. These are a tank, building, ventilation, and infiltration. Out of them, building was the primary source of loss. In our next phase of research, minimizing the overall losses of the system will be prime target.

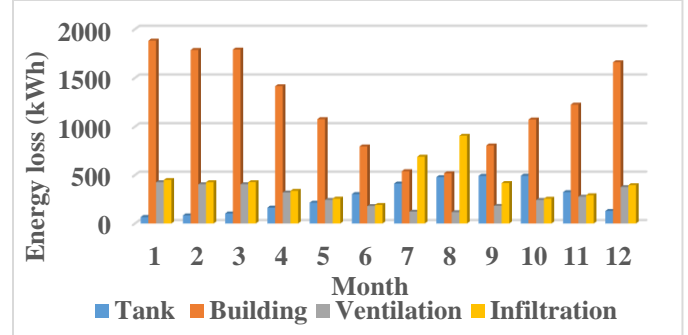


Fig.17. Different losses throughout the proposed system

VI. CONCLUSION

In this research work, a solar space heating system was designed and simulated with professional Polysun software to fulfill the annual space heating demand with seasonal energy storage of a single-family house. Simulation result depends on the actual sizing of the system components and the other value given as an input of the simulation environment. After demonstrating the output, it can be said that the proposed system was more practical, profitable and efficient in the colder countries where the seasonal difference is more. It concluded that the proposed system should consist of 16 m² flat plate collector area, 47 m³ proper insulated storage tank with height and diameter 3.91 m and 8 kW heat pump. Then this system will meet the space heating demand in heating periods of the selected house. A heat pump was used to boost up the backup heat to this system for heating months that discussed earlier. This research work can be further modified by controlling the individual components and the cost analysis of the full system to make it more affordable and reliable.

ACKNOWLEDGMENT

The authors would like to thank Memorial University for supporting this research. Also, the authors would like to thank NSERC for supporting the research as well.

REFERENCES

- [1] N. R. Canada, "Heating equipment for residential use," 05-Dec-2013. [Online]. Available: <https://www.nrcan.gc.ca/energy/products/categories/heating/13740>. [Accessed: 02-Jan-2019].
- [2] N. R. C. Government of Canada, "Residential Sector Newfoundland Table 2: Secondary Energy Use and GHG Emissions by End-Use," 30-Nov-2018. [Online]. Available: <http://oe.nrcan.gc.ca/corporate/statistics/neud/dpa/showTable.cfm?type=CP§or=res&juris=nf&rn=2&page=0>. [Accessed: 02-Jan-2019].
- [3] S. K. Shah, L. Aye, and B. Rismanchi, "Seasonal thermal energy storage system for cold climate zones: A review of

- recent developments,” *Renew. Sustain. Energy Rev.*, vol. 97, pp. 38–49, Dec. 2018.
- [4] C. N. Antoniadis and G. Martinopoulos, “Simulation of Solar Thermal Systems with Seasonal Storage Operation for Residential Scale Applications,” *Procedia Environ. Sci.*, vol. 38, pp. 405–412, Jan. 2017.
 - [5] L. T. Terziotti, M. L. Sweet, and J. T. McLeskey, “Modeling seasonal solar thermal energy storage in a large urban residential building using TRNSYS 16,” *Energy Build.*, vol. 45, pp. 28–31, Feb. 2012.
 - [6] H. Li, L. Sun, and Y. Zhang, “Performance investigation of a combined solar thermal heat pump heating system,” *Appl. Therm. Eng.*, vol. 71, no. 1, pp. 460–468, Oct. 2014.
 - [7] “Zeolite thermal storage retains heat indefinitely, absorbs four times more heat than water - ExtremeTech.” [Online]. Available: <https://www.extremetech.com/extreme/130523-zeolite-thermal-storage-retains-heat-indefinitely-absorbs-four-times-more-heat-than-water>. [Accessed: 02-Jan-2019].
 - [8] P. D. Lund, “Effect of storage thermal behavior in seasonal storage solar heating systems,” *Sol. Energy*, vol. 40, no. 3, pp. 249–258, Jan. 1988.
 - [9] A. D. Wills, “Design and co-simulation of a seasonal solar thermal system for a Canadian single-family detached house,” 2014.
 - [10] N. R. Canada, “Renewable energy facts,” 06-Oct-2017. [Online]. Available: <https://www.nrcan.gc.ca/energy/facts/renewable-energy/20069>. [Accessed: 02-Jan-2019].
 - [11] W. Weiss, *Solar Heating Systems for Houses : A Design Handbook for Solar Combisystems*. Routledge, 2003.
 - [12] “Heat pump,” *Wikipedia*. 15-Dec-2018.
 - [13] “Radiator (heating),” *Wikipedia*. 15-Oct-2018.

Natural Gas Engine Driven Heat Pump System – A Case Study of an Office Building

Farshad Kimiaghalam, Altamash Baig, Rakesh Kumar, JuanLi (Bess) Ma, Alan S. Fung
Department of Mechanical and Industrial Engineering
Ryerson University
Toronto, Canada

Abstract—An eQUEST model was developed to conduct a study of a natural gas engine driven heat pump (GEHP) for an office building in Woodstock, Ontario. Prior to the installation of the GEHP, the heating and cooling demands of the office building were provided by rooftop units (RTUs), comprising of natural gas heater and electric chiller. Energy consumption for both GEHP and RTUs were measured and then used in the regression analysis. The developed eQUEST models were also validated and calibrated with the regression analysis result with respect to the ASHRAE Guideline 14-2002. The eventual models were then applied to investigate the potential annual energy consumption, greenhouse gas (GHG) emission and energy cost savings by using the GEHP in Woodstock, Ontario.

Keywords: *GEHP, Energy Saving, GHG Saving*

I. INTRODUCTION

Heat pumps (HPs) are cyclic devices that transfer heat from low temperature medium to high temperature one. In winter time, heat pumps absorb heat from the outdoor environment and transfer it to the building that requires heating whereas in the summer time they reject heat from the building to the outdoor environment. They usually operate on vapor-compression cycle or absorption-compression cycle. Depending on the thermal energy sources, heat pumps can be classified as air, water or ground source heat pumps. Based on driving power, they can be categorized as electric driven heat pumps (EHPs), air source heat pump (ASHP), ground source HP, geothermal energy HP, solar assisted HPs or gas engine driven heat pumps [1-3]. Furthermore, air-source heat pumps (ASHPs) can be classified into electric heat pump, gas engine-driven heat pump (GEHP), and gas absorption heat pump (GAHP).

II. LITERATURE REVIEW

A. GEHP System

Gas engine-driven heat pump (GEHP) is a relatively new energy-efficient heating and air conditioning equipment that utilizes natural gas as the fuel in the gas engine that drives compressor coupled to it [4-7]. The GEHP system is

predominantly composed of vapor compression heat pump system that is driven by a gas fueled internal combustion engine, the cold and heat source system, the data acquisition and controlling system. Among the heat pump systems, recently GEHPs have been given more emphasis as a preferable choice in the heating and air-conditioning system [7]. This is due to their notable advantages such as capability to recover about 80% waste heat from internal combustion engine [8], higher primary energy ratio [5], elimination of losses caused by electricity production and transportation [9] and easy modulation of the compressor speed (by adjusting the gas supply) to meet the different load demands [10, 11].

B. Performance Studies of GEHP

Many researchers focused on the performance characteristics and performance improvement of the gas engine driven heat pump system [12-15]. Elgandy et al. [13] developed an experimental test facility to evaluate the performance of GEHP system with heat recovery for heating and cooling application. They investigated the effects of engine speed, ambient air temperature, evaporator water flowrate and the evaporator water inlet temperature on the performance characteristics of the GEHP. The studied performance characteristics were cooling capacity, heating capacity and primary energy ratio (PER). The result of their study showed that increasing evaporator water inlet temperature from 12.2°C to 23°C, cooling capacity, heat recovery and PER increased by 18%, 31% and 22%, respectively.

C. Modelling and Simulation of GEHP

Many studies have been conducted on modeling and simulation of GEHP. The first modeling study on GEHP was conducted by MacArthur et al. [16]. They performed a dynamic model of vapor compression HP with a detailed mathematical treatment of condenser, evaporator and accumulator. A lumped-parameter model was developed for the compressor, expansion device and natural-gas-fueled internal combustion. A model of an GEHP, comprising of an engine and a heat pump model together was established by Rusk et al. [17].

D. Control Strategy of GEHPs

Some studies focused on the control strategies for the GEHP since improving control strategy (optimum control) increases efficiency of GEHP [8] and decreases consumption of energy. Li et al. [18] proposed a cascade fuzzy control algorithm as a control strategy for the GEHP system. It controls the return water temperature and engine speed via its outer loop and inner loop respectively. In their study, they proved that as compared to the performance of the common control for the heat pump system, PI (proportional and integral) control strategy, the cascade fuzzy control provides better performance, smaller overshoot temperature and reduced reaction time. Yang et al. [19] developed an intelligent control scheme for GEHP system to investigate the dynamic characteristics of the system in the heating operation. In their study, a model that consists of the main components of the GEHP was made to test the effect of the intelligent control scheme. The result of the study showed that the model was very successful in analyzing the effects of the control system. Furthermore, the steady state accuracy of the intelligent control scheme was higher than that of the fuzzy controller.

III. CASE STUDY: GEHP FOR OFFICE BUILDING IN WOODSTOCK, ONTARIO

The GEHP system considered in this case study consists of natural-gas-fueled (ICE engine based) driven heat pump system that delivers thermal energy to the office building, Woodstock, Ontario. It is similar to variable refrigerant flow (VRF) system. Some VRF systems have an option for heat recovery to utilize heat rejected from cooling zones to heating zones. However, the GEHP considered in this case study has heat recovery from engine to assist heat transfer from the outdoor evaporator coil to enhance the heating performance. Figure 1 shows the GEHP system considered in this case study. The studied system can provide either heating or cooling, depending on season. The GEHP system has two basic unit configurations: *outdoor unit* and *indoor unit* configuration.



FIGURE 1. NATURAL GAS FUELED INTERNAL COMBUSTION ENGINE DRIVEN HEAT PUMP

The GEHP system considered in this case study is a multi-split variable refrigerant flow (VRF) system in which the refrigerant is used as the heating or cooling medium. The gas engine drives the compressors utilizing natural gas as the fuel. The outdoor unit consists of two compressors (one with variable speed), a heat exchanger and a four-way valve.

Some VRF systems have inverter driven variable speed compressors that enable them to have a wide capacity modulation with high part-load efficiency [20]. For such system, by changing the frequency of the inverter, the outdoor unit varies its capacity by varying the discharged refrigerant mass flowrate so as to meet the desired heating or cooling loads of the various zones. In this way, the fluctuating heating or cooling load requirements of different zones are met by variable refrigerant flow system. The GEHP considered in this case study has overall cooling and heating capacity of 56kW and 63kW, respectively as indicated in Table 1. Heating and cooling COPs are 1.5 and 1.3, respectively.

TABLE 1. SUMMARY OF THE OUTDOOR UNIT SPECIFICATION

Rated Cooling Capacity	56kW
Rated Heating Capacity	63kW
Fuel Type	Natural Gas
Power Supply	Single Phase
Power Supply Frequency	AC 200V
Fuel Consumption Rate (HHV)	For Cooling: 43.3kW, For Heating: 41kW
Rated Power Consumption	For Cooling: 0.98kW, For Heating: 0.91kW
Refrigerant	R410A
Design Pressure	H4/L2.2 MPa
Air Tight Test Pressure	H4/L2.2 MPa

The indoor unit of the VRF system under consideration consists of heat exchangers, expansion valves, temperature sensors and fans. Many indoor units are connected to a single outdoor unit. In our GEHP system, 14 indoor units are connected to one outdoor unit. The indoor units of the system under consideration have different configurations and capacities as shown in the Table 2.

TABLE 2. INDOOR UNITS CONFIGURATION AND CAPACITIES

Configuration	Quantity	Cooling Capacity of Each Unit (Tons)	Total Cooling Capacity (kW)	Heating Capacity (kW)
Ceiling Suspended	1	1	3.52	3.6
Wall Mount	4	0.6	8.45	$4 \times 2.2 = 8.8$
2'x2' 4-Way Cassette	2	1	44	$9 \times 5.6 = 50.4$
2'x2' 4-Way Cassette	7	1.5	44	$9 \times 5.6 = 50.4$
Total	14	15.9	56	62.8

IV. BUILDING DESCRIPTION AND INFORMATION

The building under consideration is an office building with an estimated area of 5413ft². The building has divisions that includes general office, three meeting rooms, waiting room, lunch room, data and voice room, guest room, equipment and storage room and vestibule room as indicated in Figure 2.

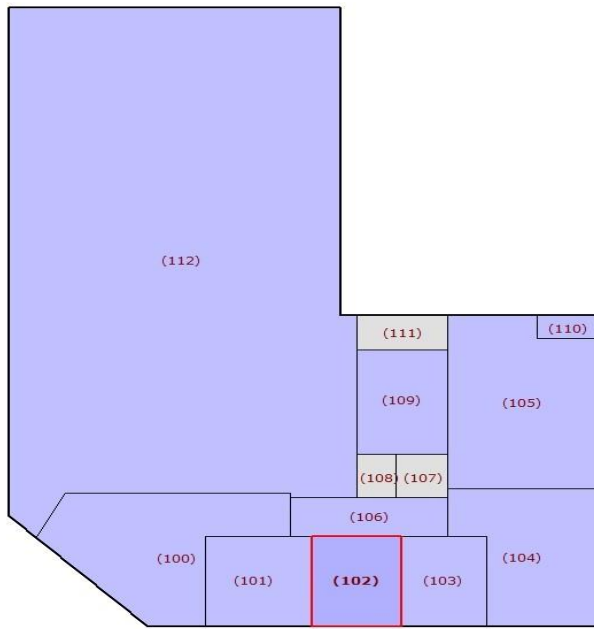


FIGURE 2. TOP VIEW AND DIVISIONS OF THE OFFICE BUILDING

V. OBJECTIVE AND METHODOLOGY

One of the objectives of the case study is to make a base building model for the roof top units (RTUs) as well as for the GEHP using eQUEST software. In addition, it is also aimed to calibrate the eQUEST model for both systems so that the baseline design consumption of gas and electricity for both GEHP and RTUs can be compared with regression analysis, which was obtained from measured data.

In this study, the eQUEST v3.65 software has been used in combination with the weather data obtained from the nearby weather station close to the Woodstock office and has been converted to bin file using CanMeteo software [21] as eQUEST only accept bin files. In addition, the input data such as the building data (parameters and dimensions), GEHP specifications, RTUs specifications, temperature set points, schedules, etc., were utilized as inputs to the models in such a way that the required outputs are generated by eQUEST. The main inputs used in the simulation were: actual weather data, cooling and heating capacities and GEHP's manufacturer product specification.

Each of the zones indicated in Figure 2 was modelled as a separate zone in the eQUEST simulations. The model is developed for both cooling and heating seasons during the summer and winter.

VI. MEASURING AIR LEAKAGE (INFILTRATION)

It is important to identify and measure the amount of infiltration in a building since the infiltration can have a significant impact on indoor air quality, building energy consumption (heating and cooling loads in buildings). From the standard methods used to measure the infiltration through a building envelope, the fan pressurization method was used in this study. This method evaluates the air leakage

characteristics from the airflow rate at given indoor-outdoor static pressure differences based on mechanical pressurization or de-pressurization of a building. To conduct the blower door test, two blower door fans were placed in the main entry door to depressurize the building so that the air from the inside of the building is sucked out causing the outside air to be drawn into the building through the leakage points.

The normal procedure was followed and described as: HVAC equipment were switched off before the test and all external doors and windows were closed. The test fans were switched on and the flow through them increased till the pressure differential between indoor and outdoor reaches 50Pa. The total air flow through the fan and the building pressure differential (inside to outside) were recorded. The fan speed was then slowly reduced in steps with fan flow and pressure difference data were recorded at each step.

The blower door infiltration test was performed based on the above procedure to determine the air change per hour (ACH) of the space, excluding the warehouse. Based on 5 data points, the building leakage curve found that 4311 cfm was the total airflow leakage at differential pressure of 50Pa which was equivalent to 6.63 ACH for the given building volume.

Figure 3 shows the blower infiltration test components used in the measurement of the office building air-leakage. It consists of a variable-speed fan, a pressure gauge to measure the pressure differences between inside and outside the building, and an airflow manometer and hoses for measuring airflow.



FIGURE 3. BLOWER DOOR TEST

Table 3 provides the summary of the main input data and important parameters used as input to the eQUEST for the modeling of both GEHP and RTUs. The specific input data used for the simulation of RTU is represented in Table 4. Table 5 indicates the specific input data used for the eQUEST simulation of GEHP.

TABLE 3. INPUT DATA USED IN SIMULATION OF GEHP AND RTU

Input and Design Model Characteristics			
General			
Location	Woodstock, Ontario		
Building Area	5413 ft ²		
Infiltration (Shell Tightness)			
Perimeter Zones (ACH) @ 50Pa	6.63		
Building Construction			
	Roof Surface	Above Grade Walls	
Construction	2 Inch Concrete	4 Inch Concrete	
Exterior Finish	Roof, Built-up	Brick	
Exterior Insulation	2 Inch Polystyrene (R-8)	2 Inch Polystyrene (R-8)	
Interior Insulation	---	R-8 Metal Furring	
Ground Floor			
Exposure	Earth Contact		
Construction	4 Inch Concrete		
Interior Finish	Vinyl Tile		
Activity Areas Allocation			
	Design Maximum Occupancy (ft ² /person)	Design Ventilation (CFM/person)	Actual Number of Person
Office	150	20	35
Dining Room	22.5	20	10
Meeting Room	22.5	20	11
Waiting Room	150	15	1
Storage	450	67	1

TABLE 4. INPUT DATA USED FOR THE SIMULATION OF RTU

Thermostat		
	Occupied	Unoccupied
Cooling Set Point (°F)	76	82
Heating Set Point (°F)	75	70
Minimum Design Flow (CFM/ft²)	0.5	
Cooling		
Overall Size (Tons)	17.5	
Condenser Type	Air-Cooled	
Efficiency (EER)	10.1	
Heating		
Size (kBtu/h)	323	
Efficiency	0.6	
Economizer		
Type	Dry bulb Temperature	
High Limit (°F)	65	
Low Limit (°F)	-2	
Ventilation		
Ventilation Fans (BHP)	1.2	

TABLE 5. INPUT DATA USED FOR THE SIMULATION OF THE GEHP

HVAC System		
Cooling Source	DX Coils	
Heating Source	DX Coils (Heat Pump)	
Heat Pump Source	Air	
System Type	Packaged Terminal Heat Pump	
Supply Fan	0.619 kW (0.83 BHP)	
Overall Cooling size	56 kW (15.9 Tons)	
Overall heating size	63 kW (215 kBtu/h)	
Cooling COP (EER)	1.3 (4.4)	
Heating COP	1.4	
Design Temperature and Air Flows		
	Indoor	Supply
Cooling Design Temperature (°F)	75	55
Heating Design Temperature (°F)	75	100
Minimum Design Flow (CFM/ft²)	0.5	

VII. SIMULATION VALIDATION

The simulated energy consumption was compared with the regression analysis obtained from the actual energy consumption. The data which were provided to the team was not complete for a full year and for some months, the data were not fully complete for the whole month. Moreover, the data were for every other month which means RTU gas and electricity consumption were gathered for a month, then the building had been switched over to the GEHP for the next month. Therefore, regression analysis was done based on actual data. However, only matching between the simulation result and regression is not enough. The accuracy of the comparison needs to be determined. For this, the ASHRAE Guideline 14-2002 [21], a building simulation criterion, has been adopted in order to verify the accuracy the eQUEST simulation. Two statistical indices, the Normalized Mean Bias Error (NMBE) and Coefficient of Variation of the Root Mean Square Error (CVRMSE) were determined. The NMBE and CVRMSE were calculated by using Equations (1) and (2):

$$NMBE = \frac{1}{\bar{y}} \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \times 100 \quad (1)$$

$$CVRMSE = \frac{1}{\bar{y}} \left[\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \right]^{1/2} \times 100 \quad (2)$$

where

y_i is measured or actual energy consumption

\hat{y}_i simulated energy consumption

\bar{y} mean of the actual energy consumption

n the number of data used in the calibration.

These ASHRAE criteria can be applied for the data on monthly or hourly basis. For monthly data, the range of acceptance is such that the absolute value of NMBE is less than 5% and that of CVRMSE is less than 15% [22] as indicated in Table 6.

TABLE 6. RANGE OF ACCEPTANCE OF CALIBRATION CRITERIA FROM ASHRAE GUIDELINE 14-2002, FOR NMBE AND CVRMSE [22]

	NMBE (%)	CVRMSE (%)
Monthly	±5	±15
Hourly	±10	±30

RTUs and GEHP calibration results are shown in Table 7. It has been calculated that NMBE and CVRMSE for the GEHP gas consumption was 4.02% and 13.93% respectively. For the GEHP electricity consumption, the calculated NMBE was 3.91% while CVRMSE was 13.53%. For the Rooftop Unit (RTU), the calculated NMBE and CVRMSE value for the gas consumption were -1.38% and 4.79% respectively. For the electricity consumption, the value obtained for NMBE was -3.43%, while that of CVRMSE was 11.89%. It can be easily seen that the NMBE and CVRMSE for both the gas and electricity consumption of GEHP and RTU are within the range of the ASHRAE calibration criteria with the calibrated monthly plug loads in Table 7. It should be mentioned that the regression analysis was performed based on the measured data from the same year that both RTU and GEHP models were developed.

Since both GEHP and RTUs simulation results are in the range of acceptance of calibration criteria from ASHRAE Guideline 14-2002, these models are developed for other major

cities in Ontario to compare how GEHP and RTU perform in different regions.

TABLE 7. GEHP AND RTU CALIBRATION RESULTS OBTAINED FROM ASHRAE GUIDELINE 14-2002

	NMBE (%)	CVRMSE (%)
Monthly GEHP Gas Consumption	4.02	13.93
Monthly GEHP Electricity Consumption	3.91	13.53
Monthly RTU Gas Consumption	-1.38	4.79
Monthly RTU Electricity Consumption	-3.43	11.89

VIII. SIMULATION RESULTS AND DISCUSSION

In this section, results obtained by running eQUEST simulation software for both GEHP and RTUs are discussed. Comparison of the simulation results and regression analysis are presented in graphs. The regression analysis obtained from measured data for GEHP and RTUs systems are presented in Table 8 and Table 9, respectively.

TABLE 8. MONTHLY GAS AND ELECTRICITY CONSUMPTION OF GEHP OBTAINED FROM REGRESSION ANALYSIS FOR WOODSTOCK

Months	Regression Analysis Data		
	Gas Consumption (m ³)	Electricity Consumption (kWh)	Total Energy Consumption (MJ)
January	1,520	1,773	64,148
February	1,274	1,659	54,389
March	1,009	1,716	44,523
April	817	1,544	36,607
May	417	1,178	20,088
June	299	878	14,524
July	415	878	18,932
August	445	1,089	20,832
September	370	913	17,348
October	428	1,144	20,384
November	688	1,430	31,294
December	961	1,773	42,904
Total	8,643	15,975	385,978

TABLE 9. MONTHLY GAS AND ELECTRICITY CONSUMPTION OF RTU OBTAINED FROM REGRESSION ANALYSIS FOR WOODSTOCK

Months	Regression Analysis Data		
	Gas Consumption (m ³)	Electricity Consumption (kWh)	Total Energy Consumption (MJ)
January	1,795	1,153	72,367
February	1,540	1,079	62,410
March	1,302	1,116	53,498
April	1,087	1,004	44,924
May	-	2,151	7,743
June	-	2,486	8,949
July	-	3,501	12,603
August	-	3,728	13,420
September	-	3,095	11,142
October	638	744	26,924
November	941	930	39,109
December	1,269	1,153	52,377
Total	8,572	22,140	405,474

Figure 4 shows the monthly variation of the gas consumption of the roof top units (RTUs) (obtained from the eQUEST simulation) and regression analysis. It can be seen that the gas consumption increases after summer months towards fall and winter months. In order to provide the heating supply, the RTU has the furnace that runs on natural gas. Since from May to September is the cooling season (no space heating is needed), the RTUs are only utilized to provide cooling load, hence the RTUs furnace was off during this period. As a result, there was no gas consumed by the RTUs during the cooling season. Due to this reason, for both regression analysis and eQUEST model, the consumption of the gas is nil for the summer months.

From Figure 4, it can also be noticed that the maximum difference between the model and regression analysis is 23.53% and this was seen in October. However, for the other months the difference was less than 10%. The difference of the total annual gas consumption between the regression analysis and the model was 1.38%.

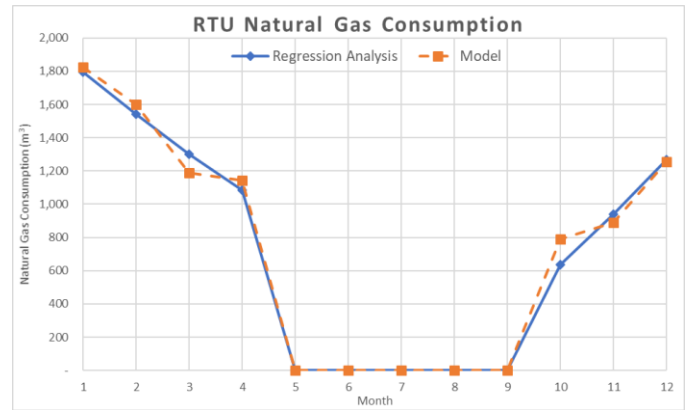


FIGURE 4. COMPARISON OF SIMULATION AND REGRESSION ANALYSIS OF MONTHLY VARIATION OF RTU GAS CONSUMPTION FOR WOODSTOCK

The roof top unit monthly electricity consumption was indicated in Figure 5. It can be seen from Figure 5 that the overall trend of the electrical consumption estimated by the eQUEST simulation matches with regression analysis. The total annual difference between eQUEST and regression analysis is only 3.43%.

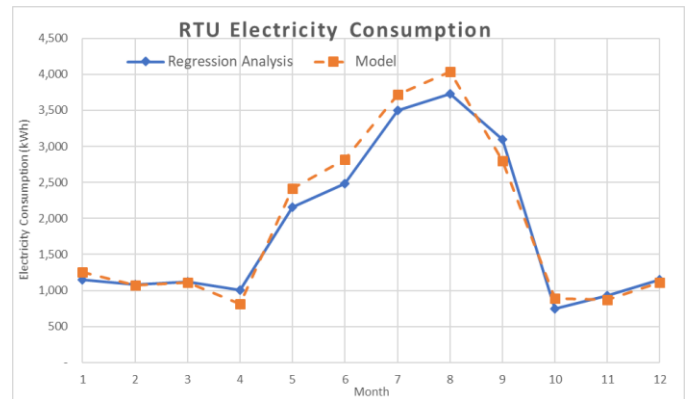


FIGURE 5. COMPARISON OF SIMULATION AND REGRESSION ANALYSIS OF MONTHLY VARIATION OF RTU ELECTRICITY CONSUMPTION FOR WOODSTOCK

Figure 6 describes the variation of monthly gas consumption utilized by GEHP. This Figure also compares the results of the modeling and regression analysis. The simulation result obtained from eQUEST modeling is very close to the regression analysis. The percentage difference of the annual consumption between the two analyses is around 4.02%.

Figure 7 presents the monthly electrical energy consumed by the GEHP. It can be seen that the overall electrical consumption trend of the model is similar to that of regression analysis. The percentage of difference of the total electrical energy consumption between the two analyses is 3.91%.

The monthly gas and electricity consumption for GEHP and RTUs are shown in Table 10 and Table 11, respectively, for Woodstock building. It could be seen that gas, electricity and total energy consumption of GEHP are less than RTUs consumption.

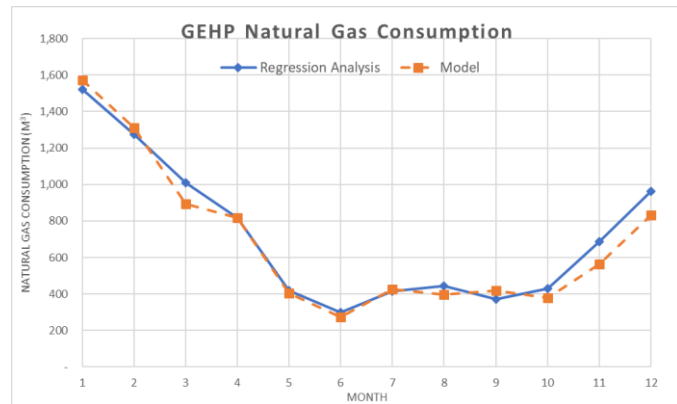


FIGURE 6. COMPARISON OF SIMULATION AND REGRESSION ANALYSIS OF MONTHLY VARIATION OF GEHP NATURAL GAS CONSUMPTION FOR WOODSTOCK

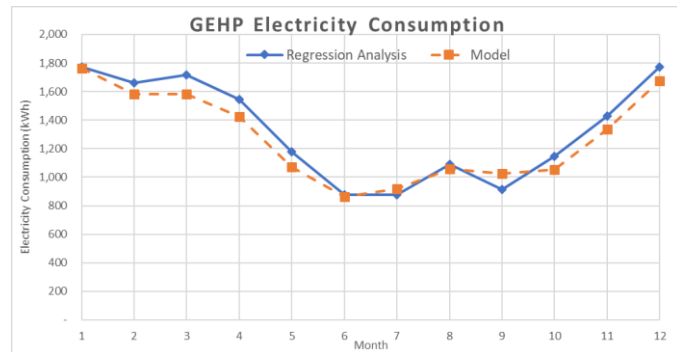


FIGURE 7. COMPARISON OF SIMULATION AND REGRESSION ANALYSIS OF MONTHLY VARIATION OF GEHP ELECTRICITY CONSUMPTION FOR WOODSTOCK

TABLE 10. SIMULATED MONTHLY GAS AND ELECTRICITY CONSUMPTION OF GEHP FOR WOODSTOCK

eQUEST Simulation Results for Woodstock			
Months	GEHP		
	Gas Consumption (m ³)	Electricity Consumption (kWh)	Total Energy Consumption (MJ)
January	1,574	1,764	64,956
February	1,312	1,584	54,581
March	895	1,580	39,037

April	819	1,425	35,630
May	405	1,072	18,946
June	272	864	13,238
July	425	920	19,148
August	399	1,056	18,645
September	418	1,024	19,258
October	380	1,054	17,942
November	564	1,334	25,818
December	832	1,674	37,001
Total	8,295	15,351	364,199

TABLE 11. SIMULATED MONTHLY GAS AND ELECTRICITY CONSUMPTION OF RTU FOR WOODSTOCK

eQUEST Simulation Results for Woodstock			
Months	RTU		
	Gas Consumption (m ³)	Electricity Consumption (kWh)	Total Energy Consumption (MJ)
January	1,826	1,250	73,908
February	1,598	1,070	64,594
March	1,188	1,110	49,149
April	1,145	810	46,432
May	-	2,410	8,676
June	-	2,820	10,152
July	-	3,720	13,392
August	-	4,040	14,544
September	-	2,800	10,080
October	788	890	33,155
November	888	870	36,873
December	1,257	1,110	51,765
Total	8,691	22,900	412,720

Table 12 shows the annual natural gas, electricity and total energy consumption saving by applying GEHP instead of RTUs in Woodstock. The annual natural gas and electricity consumption savings are 395 m³ and 7,549 kWh, respectively. Annual energy consumption saving is 48,520 MJ.

Emission factors are the rate at which a pollutant is released into the atmosphere due to a process activity. Emission factors are expressed in kilograms (kg) or metric tonnes (t) of GHG emissions per unit of consumption activity [23]. Typically, the factors for a given category of activity for instance, building energy or fleet fuel consumption – are expressed in common units to enable comparison across different fuel types, travel modes, etc. Usually, emission factors are specified by individual gas. In some cases, a cumulative factor for multiple gases is provided in kilograms (kg) or tonnes (t) of CO₂ equivalent (CO₂eq) emissions. Overall, CO₂eq is the standard unit for measuring and comparing emissions across GHGs of varying potency in the atmosphere.

For later calculation, the GHG emission intensity of 29 gCO₂eq/kWh has been taken as the GHG emission factor for the electric energy consumption [24]. The GHG emission factor for the natural gas was taken as 1,863g CO₂eq/m³ (or 49.01 kgCO₂/GJ) as reported in the guide line version 2017 for the Greenhouse Gas Emissions reporting on Section 4(1) of Ontario Regulation 452/09 [25]. Therefore, the annual GHG emission saving by utilizing GEHP instead of RTUs in Woodstock is 955.15 kg of CO₂eq.

Marginal cost of electricity is the operating cost required to produce each kWh of electric energy. The marginal cost of

electricity generation varies depending on the supply and demand for electricity, rising and falling by season, day and hour. The marginal cost of natural gas is based on the consumption and it varies in different region. With respect to marginal costs in Woodstock, the annual energy cost saving by implementing GEHP instead of RTUs in Woodstock is \$ 1,200 per year, as shown in Table 12.

TABLE 12. ANNUAL GAS, ELECTRICITY, ENERGY CONSUMPTION, GHG EMISSION AND COST SAVING FOR USING GEHP IN WOODSTOCK

Annual Natural Gas Consumption Saving (m3)	395
Annual Electricity Consumption Saving (kWh)	7,549
Annual Energy Consumption Saving (MJ)	48,520
Annual GHG Emission Saving (CO ₂ eq (kg))	955.15
Annual Cost Saving (\$)	1,200

IX. CONCLUSION

In this investigation, a case study of energy modeling for the GEHP that delivers heating and cooling energy to the office building located in Woodstock, Ontario, was conducted.

Blower door infiltration test was performed (by depressurizing the building) to determine the amount of infiltration of the building. The test result of the building leakage curve revealed that 4311 cfm was the total airflow leakage at differential pressure of 50Pa which was equivalent to 6.63 ACH for the Woodstock office building volume.

Additionally, comparison was made between the eQUEST simulated energy consumption and regression analysis, which was obtained from measured data. The accuracy of the eQUEST simulation was verified by using building simulation criterion, ASHRAE Guideline 14-2002. The comparison between the simulated and regression analysis of both the gas and electricity consumption of GEHP and RTU are within the range of the ASHRAE calibration criteria.

The results of the simulation for Woodstock office building showed that the percentage of the difference of the total annual GEHP consumption of natural gas and electricity between the simulation and regression analysis were 4.02% and 3.91%, respectively. On the other hand, the percentage of the total difference for the RTU annual natural gas and electricity consumption between the simulation and regression analysis were 1.38% and 3.43%, respectively.

Furthermore, annual natural gas, electricity and total energy consumption saving by applying GEHP instead of RTU for Woodstock are represented. The annual energy consumption saving is 48,520 MJ which shows the potential energy saving of this region. The annual cost and GHG emission savings are \$ 1,200 per year and 955.15 kg of CO₂eq, respectively.

ACKNOWLEDGMENT

This work was made possible through the financial support of Union Gas Ltd. and Enbridge Gas Distribution. The authors are grateful to NSERC, Ontario Center of Excellence (OCE) and TTCI and their staff.

The authors would like to thank Ms. Anna Wang and Mr. Esa Kerme who provided great aid throughout completing this project.

REFERENCES

- [1] Wongsuwan W, Kumar S, Neveu P, Meunier F. A review of chemical heat pump technology and applications. *Appl Thermal Eng.* 21(2001) 1489–519.
- [2] Hepbasli A, Ozgener L. Development of geothermal energy utilization in Turkey: a review. *Renewable Sustainable Energy Rev.* 8(2004)433–60.
- [3] Ozgener O, Hepbasli A. A review on the energy and exergy analysis of solar assisted heat pump systems. *Renewable Sustainable Energy Rev.* (2007)482-96.
- [4] S. Sanaye, M.A. Meybodi, M. Chahartaghi, Modeling and economic analysis of gas engine heat pumps for residential and commercial buildings in various climate regions of Iran, *Energy and Buildings* .42 (2010) 1129 –1138.
- [5] Huanwei Liu, Qiushu Zhou, Haibo Zhao. Experimental study on cooling performance and energy saving of gas engine-driven heat pump system with evaporative condenser. *Energy Conversion and Management* 123 (2016) 200–208.
- [6] Behnaz Rezaie, Marc A. Rosen. District heating and cooling: Review of technology and potential enhancements. *Applied Energy*, 2012 (93)2-10.
- [7] Y.G. Chen, Z. Yang, Theoretical simulation and experimental research on the system of air source energy independence driven by internal-combustion engine, *Energy and Buildings*. 43(2011) 1351-1358.
- [8] Hepbasli A, Erbay Z, Icier F, Colak N, Hancioglu E. A review of gas engine driven heat pumps (GEHPs) for residential and industrial applications. *Renew Sustain Energy Rev*; 13(2009)85-99.
- [9] J.Brenn.P.Soltic,C.Bach, Comparison of natural gas driven heat pumps and electrically driven heat pumps with conventional systems for building heating purposes, *Energy and Buildings* 42(2010)904-908.
- [10] M.D. d'Accadia, Survey on GHP technology, *Proceedings of the ASME Advanced Energy Systems Division 1* (1998) 313–323.
- [11] D.D. Colosimo, Introduction to engine-driven heat pumps-Concepts, approach, and economics, *ASHRAE Transactions* 93 (Pt. 2) (1987) 987-996.
- [12] Rajeev Kamal, Arun Kumar Narasimhan, Chatura Wickramaratne, Abhinav Bhardwaj, D. Yogi Goswami , Elias K. Stefanakos , Herbert A. Ingle. Field performance of gas-engine driven heat pumps in a commercial building. *International journal of refrigeration* 68 (2016) 15-27.
- [13] E. Elgendy, J. Schmidt,A. Khalil, M. Fatouh. Performance of a gas engine heat pump (GEHP) using R410A for heating and cooling applications. *Energy* 35 (2010) 4941- 4948.
- [14] E. Elgendy, J. Schmidt. Experimental study of gas engine driven air to water heat pump in cooling mode. *Energy* 35 (2010) 2461-2467.
- [15] Feng-Guo Liu, Zhong-Yun Tian, Fu-Jiang Dong, Chao Yan, Rui Zhang, Ai-Bin Yan. Experimental study on the performance of a gas engine heat pump for heating and domestic hot water. *Energy and Buildings* 152 (2017) 273-278.
- [16] MacArthur JW, Grald EW. Unsteady compressible two-phase flow model for predicting cyclic heat pump performance and a comparison with experimental data. *International Journal of Refrigeration* 12 (1989) 29–41.
- [17] Rusk RP, Van Gerpen JN, Nelson RM, Pate MB. Development and use of a mathematical model of an engine-driven heat pump. *ASHRAE Trans.* Pt. 2 (1990)282-90.
- [18] Li S, Zhang W, Zhang R, Lv D, Huang Z. Cascade fuzzy control for gas engine driven heat pump. *Energy Conversion and Management* 46(2005)1757-66.
- [19] Yang Z, Haibo Z, Zheng F. Modeling and dynamic control simulation of unitary gas engine heat pump. *Energy Conversion and Management* 48 (2007)3146-53.

- [20] R.Afify, Designing VRF systems, ASHRAE Journal 50(2008)52-55.
- [21] <https://www.nrcan.gc.ca/energy/software-tools/19908>
- [22] ASHRAE, ANSI/ASHRAE Guideline 14-2002: Measurement of Energy and Demand Savings, ASHRAE, Atlanta, 2002.
- [23] [https:// weatherstats.ca/metrics/hdd.html](https://weatherstats.ca/metrics/hdd.html)
- [24] Ministry of Environment. 2016. “2016/17 B.C. BEST PRACTICES METHODOLOGY FOR QUANTIFYING GREENHOUSE GAS EMISSIONS.” <http://www2.gov.bc.ca/assets/gov/environment/climatechange/cng/methodology/2016-17-pso-methodology.pdf>.
- [25] [www.ieso.ca /- /media/files/ IESO/ Document- Library /engage/ cf/ CF 20170427 collaboration. pdf?la=en](http://www.ieso.ca/-/media/files/IESO/Document-Library/engage/cf/CF20170427_collaboration.pdf?la=en)

Optimal Design of Li-Ion Batteries for Health-Aware Fast Charging

Xianke Lin

Department of Automotive, Mechanical and
 Manufacturing Engineering,
 University of Ontario Institute of Technology,
 Oshawa, Canada

Youngki Kim

Department of Mechanical Engineering,
 University of Michigan Dearborn
 Michigan, USA

Abstract—Fast charging is urgently needed for the wide acceptance of electric vehicles. Most of the existing charging methods are developed based on the understanding and observations of battery dynamics and do not achieve the best charging performance in terms of charging speed and degradation. In this study, a health-aware fast charging protocol is proposed, which achieves a fast charging speed and avoids lithium plating during fast charging. Besides the health-aware fast charging protocol, several key design parameters are selected, and the design optimization is carried out with the goal of improving the charging performance. The results show that the design optimization framework can significantly improve charging performance while the energy and power density requirements are satisfied.

Keywords—fast charging; Lithium-ion batteries; optimal design

I. INTRODUCTION

Lithium-ion batteries play an increasingly important role in transportation electrification and renewable energy systems [1,2]. However, there are still several issues facing lithium-ion batteries [3]. One important issue is the long charging time [4]. Compared to the conventional gasoline vehicles which take only a few minutes to refuel, the recharge process for the electric vehicles with big battery packs last for several hours. It is thus urgently needed to develop a systematic approach to improve the charging performance.

There are many charging protocols developed in the literature, such as constant-current / constant-voltage (CC/CV) [5], multi-stage charging [7], neural networks [8], pulse current charging [6], and fuzzy logic [9]. However, most of them are developed based on basic understanding and observations of battery dynamics and do not achieve the best charging performance in terms of charging speed and degradation. Following our previous work on charging protocol optimization [10], a health-aware fast charging protocol is proposed, which achieves a fast charging speed and avoids the main degradation mechanism, lithium plating [11], during fast charging.

In order to achieve the best charging performance, the battery design needs to be optimized as well. An optimal design problem is formulated based on a physics-based battery model charged

by the health-aware fast charging protocol. The design parameters are optimized to further improve charging performance.

The rest of this paper is structured as follows. In section II, the physics-based pseudo-2D model is presented. In section III, the health-aware fast charging protocol is presented in details, and its charging performance is examined. In section IV, the battery design optimization problem is formulated, and the design parameters are optimized using the gradient descent approach to further improve the charging performance. In section V, the optimized battery design is discussed in details, followed by the conclusions in Section VI.

II. PHYSICS-BASED BATTERY MODEL

In this section, the details of the physics-based pseudo-2D battery model [12] are described. A graphite/ LiFePO₄ cell is selected in this study. The battery model consists of two levels: cell level, and particle level. A schematic of the battery model is shown in Figure 1.

In the particles of both anode electrode and cathode electrode, the mass conservation of lithium-ion is governed by the diffusion equation in a spherical coordinate system [12], as follows,

$$\frac{\partial c_1}{\partial t} = \frac{D_1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial c_1}{\partial r} \right) = D_1 \left(\frac{\partial^2 c_1}{\partial r^2} + \frac{2}{r} \frac{\partial c_1}{\partial r} \right) \quad (1)$$

$$\frac{\partial c_1}{\partial r} \Big|_{r=0} = 0, D_1 \frac{\partial c_1}{\partial r} \Big|_{r=R_s} = -\frac{j_{loc}}{a_s F} \quad (2)$$

where c_1 is the lithium ion concentration, D_1 is the lithium ion diffusivity, a_s is the active surface area per electrode unit volume, j_{loc} is the intercalation/deintercalation current density, R_s is the particle radius.

In the cell level, the mass transport of lithium ions is governed by Newman's porous composite electrode theory [12]. The cell-level model consists of three domains: anode, separator, and cathode. The mass transportation process is governed by the following equation [12],

$$\frac{\partial(\varepsilon_2 c_2)}{\partial t} = \frac{\partial}{\partial x} \left(D_2^{eff} \frac{\partial c_2}{\partial x} \right) + \frac{1-t_+^0}{F} j_{loc}, \quad (3)$$

$$\left. \frac{\partial c_2}{\partial x} \right|_{x=0} = \left. \frac{\partial c_2}{\partial x} \right|_{x=L} = 0, \quad (4)$$

where c_2 is the electrolyte concentration, ε_2 is the electrolyte volume fraction, D_2^{eff} is the effective diffusivity of electrolyte, t_+^0 is the transference number.

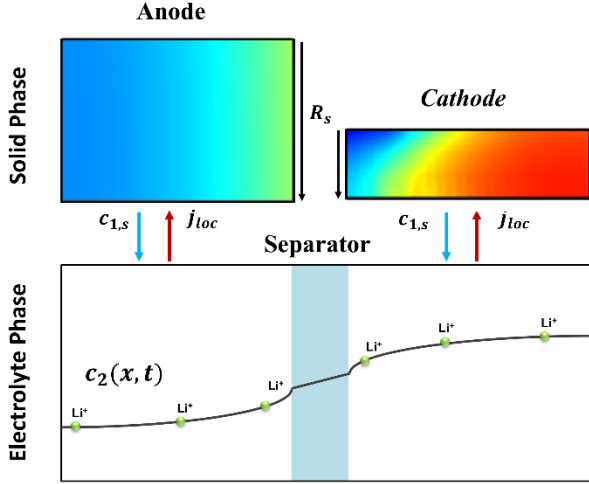


Figure 1. Physics-based pseudo-2D battery model

Besides the mass conservation in both cell level and particle level models, the governing equations for the electric potential distributions in both the electrolyte phase and solid phase are needed. In the electrolyte, the potential distribution is described by the following equation [12],

$$\frac{\partial}{\partial x} \left(k_2^{eff} \frac{\partial \phi_2}{\partial x} \right) - \frac{\partial}{\partial x} \left(k_{2D}^{eff} \frac{\partial \ln c_2}{\partial x} \right) + j_{loc} = 0, \quad (5)$$

$$\left. \frac{\partial \phi_2}{\partial x} \right|_{x=0} = \left. \frac{\partial \phi_2}{\partial x} \right|_{x=L} = 0. \quad (6)$$

where ϕ_2 is the electrolyte electric potential, k_2^{eff} is the effective conductivity in the electrolyte.

The potential distribution in the solid phase is governed by the following equation,

$$\frac{\partial}{\partial x} \left(k_1^{eff} \frac{\partial \phi_1}{\partial x} \right) - j_{loc} = 0, \quad (7)$$

$$\begin{aligned} -k_1^{eff} \frac{\partial \phi_1}{\partial x} \Big|_{x=0} &= k_1^{eff} \frac{\partial \phi_1}{\partial x} \Big|_{x=L} = \frac{I}{A}, \frac{\partial \phi_1}{\partial x} \Big|_{x=L-} \\ &= \frac{\partial \phi_1}{\partial x} \Big|_{x=L-L_{sep}} = 0, \end{aligned} \quad (8)$$

where k_1^{eff} is the effective conductivity.

The potential difference between the solid phase and the electrolyte phase provides a driving force for lithium ion intercalation/deintercalation. The rate of lithium ion intercalation/deintercalation is governed by the Butler-Volmer equation as follows,

$$j_{loc} = a_s i_0 \left\{ \exp \left[\frac{\alpha_a F}{RT} (\eta - R_{SEI} j_{loc}) \right] - \exp \left[-\frac{\alpha_c F}{RT} (\eta - R_{SEI} j_{loc}) \right] \right\}, \quad (9)$$

$$\eta = \phi_1 - \phi_2 - U, \quad (10)$$

where R_{SEI} is the resistance of the SEI layer, the overpotential η is the driving force for lithium intercalation/deintercalation, i_0 is the exchange current density, U is the open-circuit potential.

Table I. Battery Model Parameters

Parameter	Anode	Sep.	Cathode	Description
L (um)	35	20	60	Electrode thickness
R_s (um)	2		0.072	Particle radius
ε_1	0.7503		0.5080	Electrode volume fraction
ε_3	0.07		0.172	Filler volume fraction
ε_2	0.1797	1	0.3200	Electrolyte volume fraction
k_1 (S/m)	100		3.8	Solid phase conductivity
$c_{1,max}$ (mol m ⁻³)	26390		22806	Max. solid phase conc.
D_1 (m ² s ⁻¹)	20×10^{-14}		4×10^{-18}	Solid phase Li diffusivity
c_2 (mol m ⁻³)	1000	1000	1000	Initial electrolyte conc.
t_+^0	0.363	0.363	0.363	Li Transference number
D_2 (m ² s ⁻¹)	1×10^{-10}	1×10^{-10}	1×10^{-10}	Electrolyte Li diffusivity
f_-^+	1	1	1	Electrolyte activity coefficient
p	1.5	1.5	1.5	Bruggeman porosity exponent
α_a, α_c	0.5		0.5	Charge transfers coefficient
k_0 (m/s)	0.4×10^{-11}		0.4×10^{-11}	Reaction rate coefficient
R_{SEI} (Ω m ²)	0.022			SEI film resistance
a_s (m ⁻¹)	$\frac{3\varepsilon_1}{R_s}$		$\frac{3\varepsilon_1}{R_s}$	Active surface area per unit electrode volume
U_n (V) ^c	Interp_neg($c_{1,s}/c_{1,max}$)			Negative electrode Equilibrium potential
U_p (V) ^c			Interp_pos($c_{1,s}/c_{1,max}$)	Positive electrode Equilibrium potential
k_1^{eff} (S/m)	$k_1^{eff} = k_1 \varepsilon_1^p$		$k_1^{eff} = k_1 \varepsilon_1^p$	Effective conductivity of solid phase
i_0 (Am ⁻²)	$F k_0 c_2^{\alpha_a} (c_{1,max} - c_{1,s})^{\alpha_a} c_{1,s}^{\alpha_c}$			Exchange current density
k_{2D}^{eff} (A/m)	$k_{2D}^{eff} = \frac{2RTk_2}{F} (t_+^0 - 1) (1 + \frac{d \ln f_-^+}{d \ln c_2})$			Effective electrolyte phase Li diffusion conductivity
k_2^{eff} (S/m)	$k_2^{eff} = k_2 (\varepsilon_2)^p$			Effective electrolyte ionic conductivity
k_2 (S/m) ^c	Interp_k2(c_2)			Electrolyte ionic conductivity

^cInterpolation curves in the reference [13] [14]

The physics-based battery model is validated against experimental data as shown in Figure 2. The Root Mean Square Error (RMSE) of voltage is about 21.3 mV. The battery model parameters are listed in Table I. More details are available in the reference [12,13,15].

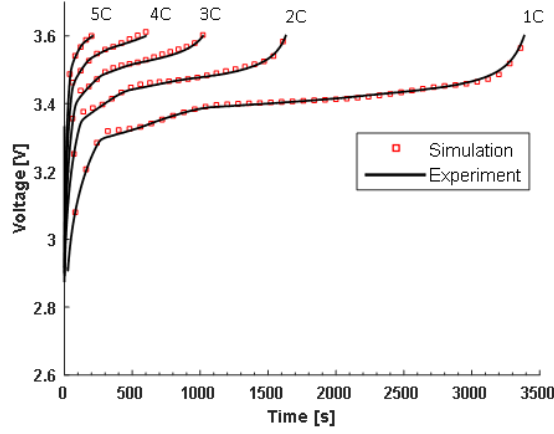


Figure 2. Comparison of voltage between model prediction and experimental data

III. HEALTH-AWARE FAST CHARGING PROTOCOL

Following our previous work on optimal charging strategy [10], a health-aware fast charging protocol is developed by considering the lithium plating during fast charging. As shown in Figure 3, this health-aware fast charging protocol applies a maximum charging current throttled by four limiting factors in order to prevent electrolyte from depletion, active materials from depletion or saturation, anode potential from going below zero, battery voltage from exceeding cut-off voltage. By preventing anode potential from going below zero, lithium plating is avoided.

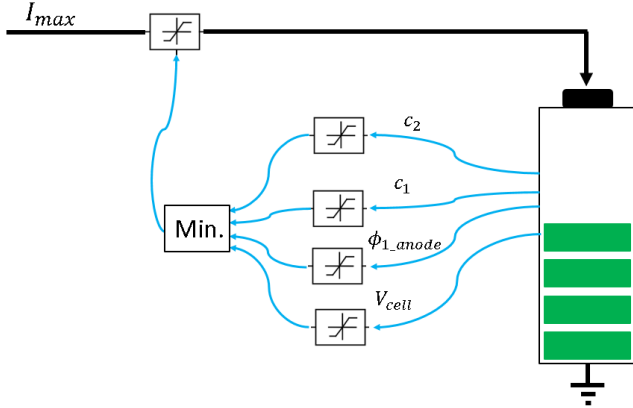


Figure 3. Health-aware fast charging protocol

The four limiting factors can also be expressed in the following equation,

$$c_2 \geq c_{min}, \quad (11)$$

$$c_{1,max} \geq c_1 \geq 0, \quad (12)$$

$$\phi_{1_anode} - \phi_2 > 0, \quad (13)$$

$$V_{cell} \leq 3.6V, \quad (14)$$

When the limiting variable is away from its upper and lower bounds, the corresponding limiter outputs 1. When the limiting variable approaches the upper or lower bounds, the limiter's output also approaches 0, which throttles down or even cuts off the charging current. Therefore, the charging current is throttled by the above four limiting factors in real-time until the charging process finishes.

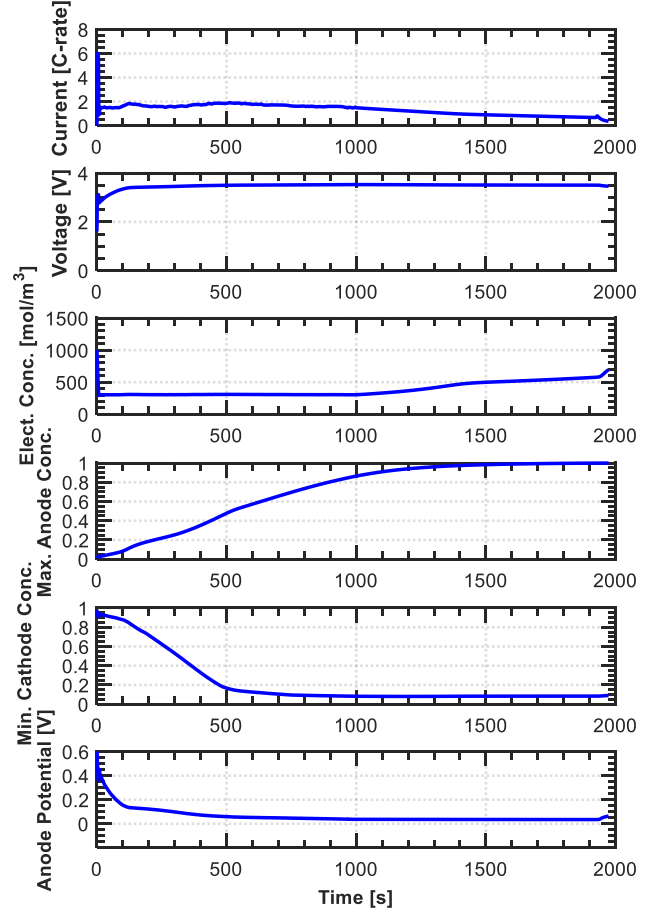


Figure 4. Health-aware fast charging process

The health-aware fast charging protocol is applied to the physics-based model, and the results are examined. As shown in Figure 4, at the beginning of the charging process, the current is maximum. When the electrolyte concentration decreases quickly, the current is throttled down to prevent electrolyte from depletion. Here, we set the minimum electrolyte concentration to be 200 mol/m³. As the battery SOC increases, the anode potential decreases and approaches 0. The charging current is still throttled to prevent the anode potential from going below zero. And near the end of the charging process, the anode concentration is close to maximum, and cathode concentration is close to empty, the throttling is to prevent the anode saturation and cathode depletion. This charging process takes about 32 mins, and the lithium plating is avoided.

IV. BATTERY DESIGN OPTIMIZATION

The design parameters have a significant impact on fast charging performance. In several studies [16–18], battery design

parameters are optimized to achieve better energy density and power density. However, energy density and power density optimization do not guarantee a good charging performance.

In this work, the key design parameters will be optimized to improve the charging performance while the battery energy density and power density requirements are satisfied. An optimization problem is formulated as follows [19],

$$\text{minimize} \quad f(x), \quad f: \mathbb{R}^n \rightarrow \mathbb{R} \quad (1)$$

$$\text{subject to} \quad \begin{cases} x_{lower} \leq x \leq x_{upper} \\ g_i(x) \leq 0, \quad i = 1, \dots, m \\ h_i(x) = 0, \quad i = 1, \dots, p \end{cases} \quad (2)$$

where $f(x)$ is the cost function which needs to be minimized by adjusting the design variables, x_{lower} is the lower bound for design variables, x_{upper} is the upper bound for design variables, $g_i(x)$ is the inequality conditions that need to be satisfied. $h_i(x)$ is the equality condition that needs to be satisfied. The above optimization problem is solved by using a gradient-descent scheme [19].

Table II. Key Design Parameters

Parameter	Description	Lower bound	Upper bound
R_{s+} (um)	Cathode particle radius	0.03	5
L_+ (um)	Cathode electrode thickness	20	200
ϵ_{2+}	Cathode porosity	0.1	0.8
k_{1+} (S/m)	Cathode conductivity	1	10
R_{s-} (um)	Anode particle radius	0.5	30
L_- (um)	Anode electrode thickness	20	200
ϵ_{2-}	Anode porosity	0.1	0.8

Seven main design parameters are selected for battery design optimization. The design parameters and their bounds are listed in Table II.

The experimentally validated battery model is used to test the battery power density and energy density. The energy density is calculated by discharging the battery from full state to the cut-off voltage 2.7V for 3600s. The discharging current is adjusted so that the discharge lasts for 3600s. The power density is calculated by discharging the battery to the cut-off voltage 2.2V for 10s [18]. The equations for calculating the energy density and power density are shown below,

$$E_{cell} = \frac{1}{M_{cell}} \int_0^{3600} V_{cell}(t) Idt \quad (3)$$

$$P_{cell} = \frac{1}{M_{cell}} \int_0^{10} V_{cell}(t) Idt \quad (4)$$

$$M_{cell} = M_{cc-} + M_- + M_{sep} + M_+ + M_{cc+} \quad (5)$$

where E_{cell} is the energy density, M_{cell} is the mass of the battery cell, P_{cell} is the power density, $V_{cell}(t)$ is the cell voltage, I is the current.

In this study, the validated battery model can provide 2390 W/kg of power density, and 98 Wh/kg of energy density.

The charging performance optimization is to minimize the charging time while satisfying the energy density and power density requirements as shown below,

$$\text{minimize} \quad f(x) = t_{charge} \quad (6)$$

$$\text{subject to} \quad \begin{cases} x_{lower} \leq x \leq x_{upper} \\ g_1(x) = P_{req} - P_{cell} \leq 0 \\ g_2(x) = E_{req} - E_{cell} \leq 0 \end{cases} \quad (7)$$

where P_{req} is the power density requirement, E_{req} is the energy density requirement. The energy density requirement and the power density are both lower slightly than the calculated values based on the battery model so that the optimization algorithm has more freedom to search for better parameters to improve charging performance.

V. BATTERY DESIGN OPTIMIZATION RESULTS

In this section, the optimal design problem is solved by using the gradient descent method. The optimization results are examined and discussed in detail.

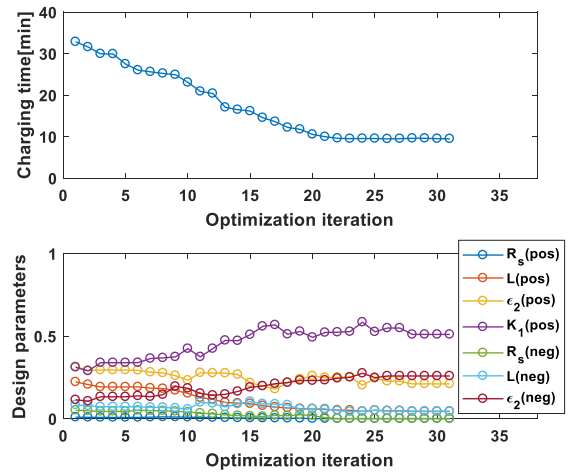


Figure 5. Charging performance optimization

The power density requirement varies from different applications. In this work, we choose 2000 W/kg which is in the range of the reported values [20,21]. The energy density optimization is set to be 90 Wh/kg. The optimization algorithm is used to improve charging performance while fulfilling the power density and energy density requirements. The battery parameters from Table I are used as the starting point.

As shown in Figure 5, after many iterations, the optimization converges. The charging time is reduced from 1923s, about 32 mins, to 568s, about 9.5 mins. A 70% reduction of charging time is achieved by design optimization. The design parameters before and after charging performance optimization are listed in Table III. From the comparison, it is found the most significant changes are the particle size and electrode thickness. Both particle sizes for anode and cathode are reduced, which facilitates the diffusion process. Electrode thickness for both

electrodes is reduced as well because thinner electrodes reduce the diffusion distance for lithium ions in the electrolyte and therefore speeds up the diffusion process.

Table III. Optimized Parameter Values for Charging Performance

Design parameters	Optimized for energy density	Optimized for charging performance
Cathode particle radius (um)	0.072	0.03
Cathode thickness (um)	60	27.8
Cathode porosity	0.32	0.2472
Cathode conductivity (S/m)	3.8	5.5898
Anode particle radius (um)	2	0.5
Anode thickness (um)	35	28.1
Anode porosity	0.179	0.2807
Power density (W/kg)	2390	2000
Energy density (Wh/kg)	98	90
Charging time (s)	1923	568

VI. CONCLUSION

In this paper, a health-aware fast charging protocol is proposed, which achieves a fast charging speed and avoids lithium plating during fast charging. Furthermore, several key design parameters are selected, and design optimization is carried out with the goal of improving the charging performance. The energy density and power density requirements are applied to the optimization problem as inequal constraints. After the charging performance optimization, the charging time is reduced from 1923s, about 32 mins, to 568s, about 9.5 mins. A 70% reduction of charging time is achieved by design optimization. This study demonstrates that the design optimization framework can significantly improve charging performance while fulfilling the energy and power density requirements.

ACKNOWLEDGMENT

This work was supported by the NSERC Discovery Program (RGPIN-2018-05471).

REFERENCES

- [1] A.A. Pesarani, T. Markel, H.S. Tataria, D. Howell, Battery Requirements for Plug-in Hybrid Electric Vehicles--analysis and Rationale, 2009.
- [2] J. Axsen, A. Burke, K.S. Kurani (2008).
- [3] J.-M. Tarascon, M. Armand, in: Materials For Sustainable Energy: A Collection of Peer-Reviewed Research and Review Articles from Nature Publishing Group, World Scientific, 2011, pp. 171–179.
- [4] X.-H. Sun, T. Yamamoto, T. Morikawa, Transportation Research Part D: Transport and Environment 37 (2015) 97–107.
- [5] S.S. Zhang, K. Xu, T.R. Jow, Journal of power sources 160 (2006) 1349–1354.
- [6] B.K. Purushothaman, U. Landau, Journal of The Electrochemical Society 153 (2006) A533-A542.
- [7] T. Ikeya, N. Sawada, J.-i. Murakami, K. Kobayashi, M. Hattori, N. Murotani, S. Ujiie, K. Kajiyama, H. Nasu, H. Narisoko, Journal of power sources 105 (2002) 6–12.
- [8] Z. Ullah, B. Burford, S. Dillip, IEEE Aerospace and Electronic Systems Magazine 11 (1996) 26–34.
- [9] H. Surmann, IEEE Transactions on Industrial Electronics 43 (1996) 541–548.
- [10] X. Lin, X. Hao, Z. Liu, W. Jia, 17735502 (2018).
- [11] S.S. Zhang, Journal of power sources 161 (2006) 1385–1391.
- [12] M. Doyle, T.F. Fuller, J. Newman, Journal of The Electrochemical Society 140 (1993) 1526–1533.
- [13] X. Lin, J. Park, L. Liu, Y. Lee, A.M. Sastry, W. Lu, Journal of The Electrochemical Society 160 (2013) A1701-A1710. <https://doi.org/10.1149/2.040310jes>.
- [14] M. Safari, C. Delacourt, Journal of The Electrochemical Society 158 (2011) A562. <https://doi.org/10.1149/1.3567007>.
- [15] X. Lin, X. Hao, Z. Liu, W. Jia, Journal of power sources 400 (2018) 305–316. <https://doi.org/10.1016/j.jpowsour.2018.08.030>.
- [16] X. Lin, W. Lu, Journal of The Electrochemical Society 165 (2018) A3380-A3388. <https://doi.org/10.1149/2.0741814jes>.
- [17] S. Golmon, K. Maute, M.L. Dunn, Int. J. Numer. Meth. Engng 92 (2012) 475–494. <https://doi.org/10.1002/nme.4347>.
- [18] N. Xue, W. Du, T.A. Greszler, W. Shyy, J.R.R.A. Martins, Applied Energy 115 (2014) 591–602. <https://doi.org/10.1016/j.apenergy.2013.10.044>.
- [19] R. Baldick, Applied optimization: formulation and algorithms for engineering systems, Cambridge University Press, 2006.
- [20] W.F. Howard, R.M. Spotnitz, Journal of power sources 165 (2007) 887–891.
- [21] T. Horiba, K. Hironaka, T. Matsumura, T. Kai, M. Koseki, Y. Muranaka, Journal of power sources 97 (2001) 719–721.

Through-plane permeability of a successively strained gas diffusion layer in polymer electrolyte membrane fuel cells

James Park, Jian Zhao, Xianguo Li

20/20 Laboratory for Fuel Cell and Green Energy RD&D, Department of Mechanical and Mechatronics Engineering
University of Waterloo
Waterloo, Canada
xianguo.li@uwaterloo.ca

Abstract— Gas diffusion layer (GDL) is often strained in an actual polymer electrolyte membrane (PEM) fuel cell as a result of assembly force to minimize ohmic losses due to electrical contact resistance. The compression of GDLs will reduce the porosity, and hence deteriorate the mass transport of reactant gases. However, accurately measuring the mass transport properties under compression conditions remains challenging. In this study, a commercial GDL sample (TGP-H-120 carbon fibre with nominal 40 wt% polytetrafluoroethylene treatment) was pre-stressed which resulted in residual strain, the morphology was characterized by an optical microscope, and the permeability was measured by a custom-engineered apparatus based on Darcy's law. With the stress ranging from 0 to 30.3 MPa, well above typical cell assembly conditions, a linear relation between the stress and residual strain (up to 0.240) is observed. As the sample is continuously compressed, the porosity is reduced from 62.8% to 51.1%, resulting in a decrease in permeability from $4.3 \times 10^{-12} \text{ m}^2$ to $1.1 \times 10^{-12} \text{ m}^2$. The experimental results are consistent with theory and the data reported in literature. These characterizations can lead to enhanced numerical models accounting for the compression of the GDL and the validation of predictive models for permeability.

Keywords—Polymer electrolyte membrane fuel cell; gas diffusion layer; compression; porosity; permeability

I. INTRODUCTION

Polymer electrolyte membrane (PEM) fuel cell has been recognized as a promising power source for vehicular applications due to its low emissions, high energy conversion efficiency, and low vibration and noise level [1]–[3]. A typical PEM fuel cell is composed of a solid membrane, two catalyst layers, two gas diffusion layers (GDLs), two flow channels, and two bipolar plates [4], [5]. During the cell operation, hydrogen is usually supplied at the anode channel and transported by the GDL toward the catalyst layer, where hydrogen is catalyzed into protons and electrons. The electrons are conducted back through the GDL and bipolar plate and routed towards the

cathode. The protons, on the other hand, are transported through the membrane towards the cathode, where they are combined with electrons and oxygen, generating water. The water may be present in a liquid form, and will be expelled by the GDL toward the cathode flow channels above a certain accumulation threshold. Therefore, the GDL is one of the key components in PEM fuel cells to support the mechanically weak catalyst layer, to transport reactants from the channel to reaction sites, to conduct electrons and heat, and to manage water.

A typical GDL is composed of two layers – one microporous layer (MPL) and one carbon fibre paper [6], [7]. The MPL is made of carbon particles and polytetrafluoroethylene (PTFE) agent [7] for an improved contact resistance and water management. The carbon fibre paper is often PTFE treated to enhance its hydrophobicity [8]. During the actual cell assembly, an optimal assembly force is often applied to the fuel cell stack by the torque on the bolts [9]. After a long cell operation, an obvious thickness change in GDLs can be observed due to the clamping stress created by the assembly force [10], [11]. In addition to preventing gas leakage, the clamping stresses, optimally between 1.0 and 2.0 MPa [12], decrease the electrical contact resistance between the GDL and bipolar plates [13], a considerable source of ohmic losses. However, the reduction of electrical contact resistance is at the expense of the GDL porosity and mass transport resistances [13]–[15].

Many experimental and numerical studies indicate that the mass transport of the reactant gases through GDLs is dominated by convection due to its relatively large pore sizes [16], [17]. In numerical studies, the permeability of the GDL may be assumed constant [10], [18], or taken from manufacturer's specifications [19], or estimated by empirical models that may correlate porosity [20] and, to a lesser extent, incorporate tortuosity, fibre diameter, and pore surface area in the calculation [14]. These structural parameters can be significantly affected by compression behaviours. Unfortunately, limited experimental studies have been conducted into the correlation between compression and permeability. While most studies have been an investigation into the effect of the in-plane permeability [21],

[22], through-plane permeability studies of compressed samples have encountered the same issue: the compressive stress is acting in the same axis as the flow of gas. In situ, fine pressure probes have been inserted into active fuel cells [23]. These can be intrusive and disruptive to the surrounding porous structure. Ex situ, porous plastic plugs [24] have been used in an attempt to overcome the technical challenge of applying a stress to a sample and have fluid flow in the same axis. However, the plugs alter the up- and down-stream flow characteristics from actual flow channels, introducing measurement uncertainties to the permeability calculation. Tamayol et al. [8] measured the permeability through compressed raw GDL samples. Samples, with varying initial thicknesses, were compressed to different thicknesses. It was assumed that initial bulk properties of each sample met the manufacturer's specifications. These experimental studies may either lose the appropriate comparison for the actual GDL compression in a cell or lead to inaccurate results.

The objective of this study is therefore to investigate the permeability of GDLs under compression conditions. In the present study, commercial GDLs (TGP-H-120) are compressed with various thickness changes, the surface morphology of different compressed GDLs is characterized by an optical microscope, and the permeability is measured using a custom-engineered apparatus established based on Darcy's law. These experimental data will be useful for predictive permeability models and to enhance numerical investigations into fuel cell performance with clamping stresses.

II. EXPERIMENTAL METHODOLOGY

A. Sample Preparation

Toray® TGP-H-120, which consists of carbon fibres and PTFE with a nominal 40 wt% ($46.6\% \pm 9\%$ based on the sample bulk density and the manufacturer's specification for untreated carbon paper). The samples are prepared in a disk-like shape with a diameter of 3.8 cm, thickness of 370.4 μm (Fowler® 0-1"0-25mm Electronic IP54 Disc Micrometer), and mass of 356.4 ± 0.1 mg (Sartorius® Praxium 224-1s).

Samples were compressed (Carver® Auto Series NE) under loads of 4.3, 13.0, 21.6, and 30.3 MPa, for 150 s at an ambient room temperature of 19 °C. Samples were measured in their residual strained condition. GDLs under large cyclic loads observe initial hysteresis and a residual strain stabilizes with time [25]. It should be noted that the samples with a residual strain may not be microstructurally comparable to that of a sample being compressed in situ. Taking in the popular assumption that the highly compressible pore volume is the primary altered variable in a compressed sample [21, 22], resulting in a reduced porosity. As well, by virtue of the overall importance of porosity in predicting permeability [26, 27], the porosity of a compressed *in-situ* sample and a sample with residual strain, both with the same thickness, would be the same. To take this further, it is plausible that other factors are required to reliably predict permeability, such as pore size distribution, mean path length, and tortuosity, that have not been considered in many of the models. Through-plane permeability studies often

neglect the impact that these variables [8] may play in the prediction of the permeability as a means to compare to such models. Given the randomized overlapping carbon fibres in the carbon paper sample and the distribution and size of the PTFE agglomerates, these otherwise neglected variables are more like the baseline condition in the residual strained sample than a different sample with the same thickness and porosity. While the individual variables may hold a negligible effect, combined their impact can be confounded. The dissimilarity with the baseline condition of these variables would be smaller, thus closer to negligible, than another sample with similar bulk properties to the compressed sample.

B. Optical microscopy

The morphology of the GDL sample was characterized by an optical microscope (ZEISS® Axio Zoom.V16). The fibre orientation, diameter, cracks, and locations are examined carefully after compression.

C. Porosity

The porosity was measured as a function of GDL compressed thickness. It is assumed that the GDL fibres are incompressible and all bulk volume reduction arises from the reduction of pore volume. It is also assumed that the sample only deforms in the direction of compression. The porosity of the compressed sample, ε_c , can be determined

$$\varepsilon_c = 1 - \frac{z_0(1 - \varepsilon_0)}{z_c} \quad (1)$$

where ε_0 is the original porosity (0.6281 ± 0.0035 [28]) and z is the sample thickness [m].

D. Permeability

The permeability was measured by a custom-engineered apparatus based on Darcy's law similar to that in [17]. The apparatus consists of two gas chambers with a total interior diameter of 3.8 cm and length of 42.5 cm. A thermocouple (OMEGA®) and pressure sensor (OMEGA® PX419-26BV) are installed in each chamber to measure the temperature and pressure of the gas, respectively. A flow meter (Omega® FMA 5500 Mass Flow Controller) at the inlet controls the mass flow rates of the gas. During the measurement, dry oxygen gas (99.993%) is fed into the top chamber, forced to pass through the samples, and vented from the bottom chamber. By measuring the pressure difference under particular flow rates, the permeability of the GDL can be evaluated. The current configuration is suitable for the catalyzed electrodes with the permeability in the order of 10^{-14} m^2 [17]; however, for carbon papers with higher permeability, the pressure difference is tiny, which will result in large measurement uncertainty. To overcome this challenge, two samples were stacked together [24] and the samples were sandwiched with three thin (20 μm) non-porous films with a 2.3 cm opening. The pressure drop from the films, without a sample, were used to adjust the pressure difference from the sample based on resistor theory in series [24].

The gas permeability of the porous samples is determined by using Darcy's law. The pore Reynolds number for air in PEM fuel cell electrodes is in the order of 10^{-4} [24] and the current experimental setup had a maximum value of 1.5×10^{-5} . This allows for the omission of the inertial effect found in the Forchheimer-Darcy Law [29].

The general form of Darcy's law is,

$$v = -\frac{k}{\mu} \nabla p \quad (2)$$

where v is the superficial velocity [m s^{-1}], k is the permeability [m^2], μ is the gas viscosity [Pa s], and p is the pressure [Pa]. Assuming ideal gas properties, Darcy's law can be simplified [30] to:

$$k = \frac{2\mu zRT\dot{m}}{AM(p_{top}^2 - p_{bot}^2)} \quad (3)$$

where R is the universal gas constant [$\text{J mol}^{-1} \text{K}^{-1}$], T is the temperature [K], \dot{m} is the mass flux [kg s^{-1}], A is the cross-sectional area of the samples [m^2], M is the gas molecular weight [kg mol^{-1}], and top and bot refer to the top and bottom chambers, respectively.

III. RESULTS

A. Compression on Sample

Table 1 summarizes the physical properties of the uncompressed and compressed TGP-H-120 samples. The maximum residual strain can be as much as 0.240, which is the typical range of thickness change during actual cell operation [31]. In line with Gigos et al. [25], with increasing peak stresses, the residual strain remains approximately constant. It is observed that this change is independent of stress loading time. The peak stress and residual strain indicate a linear relationship, as seen in Fig. 1. This linearity shows good agreements with the studies in [15], in which a GDL with 5% PTFE treatment is strained to 0.189. As the samples are continuously compressed, the pore volume of the samples is reduced, resulting in a decrease in porosity from 62.8% to 51.1 %.

Following the assumption of incompressible fibres and compressible pore volumes, the changes in thickness (from 740.8 to 562.9 μm) is a result of the re-orientation of the random fibre structure into the pore volumes and some damage-type changes. Damage-type changes include crushing and breakage, examples seen in Fig. 2. The PTFE treatment, found randomly throughout the sample, seems to act as a flexible support for the fibres. Localized regions of the sample with PTFE were less prone to changes from stress. It has been reported [25] that the PTFE treatment would affect the strain rate, and it is expected that samples with less PTFE would experience greater damage-type changes and residual strain in response to the same peak stress loads.

It can also be noted that, while the stress range is beyond that of a typical operating fuel cell, the physical changes, at least

TABLE I. PHYSICAL PROPERTIES OF THE COMPRESSED TGP-H-120 SAMPLE.

Peak Stress [MPa]	Thickness [μm]	Residual Strain	Porosity
0	740.8	0.0	62.8%
4.3	713.0	0.038	61.4%
13.0	667.9	0.098	58.8%
21.6	624.3	0.157	55.9%
30.3	562.9	0.240	51.1%

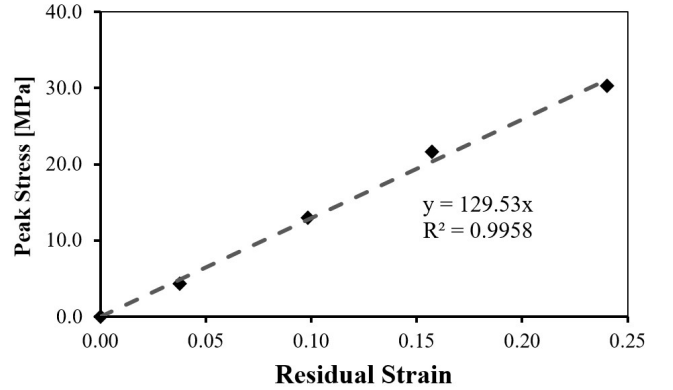


Figure 1. Peak stress – residual strain plot of 40 wt% PTFE TGP-H-120 sample.

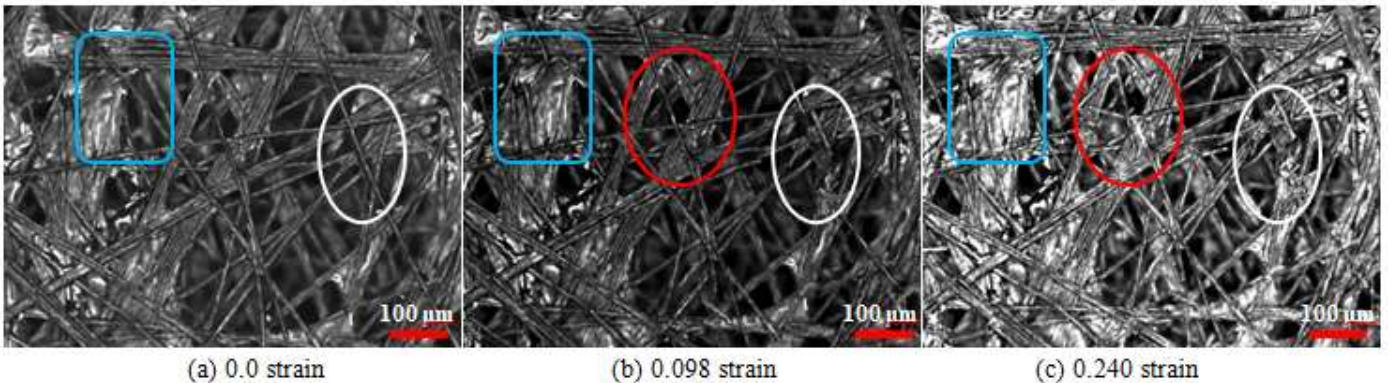


Figure 2. Surface morphology of a successively strained sample of TGP-H-120 with 40 wt% PTFE. Signs of breakage and crushing are visible.

on the surface, were minimal. This would be in support that the compression method and load did not significantly damage or alter the overall structure of the GDL sample within the range of clamping stress resulted from the actual cell assembly, which is expected during normal operating loads.

B. Compression on Permeability

Fig. 3 presents the relationship between permeability and strain of the same sample. The permeability is found to decrease from $4.3 \times 10^{-12} \text{ m}^2$ (unstrained) to $1.1 \times 10^{-12} \text{ m}^2$ (strained to 0.240). These values are of the same order of magnitude as a similar study with samples without any PTFE treatment [8]. Fig. 4 presents the relations between permeability and porosity of the same TGP-H-120 sample that has been successively compressed. As expected, with decreasing porosity, the permeability of the sample is reduced. Major damage-type changes to the GDL would be characterized by an increase in permeability by introducing large flow paths through the sample, however, this was not observed and further supports that the tested stress range did not overly damage the samples.

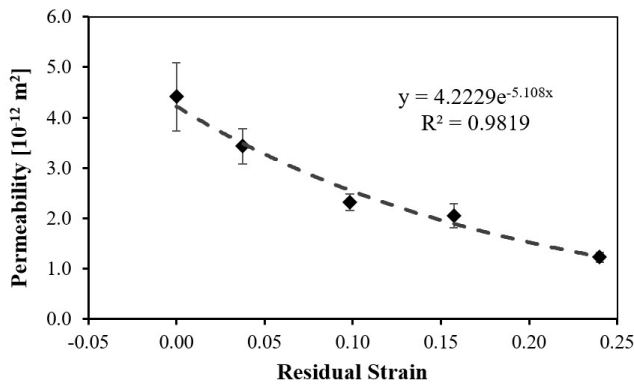


Figure 3. Relation between permeability and residual strain of the same TGP-H-120 sample that has been successively compressed. Error bars represent the repeatability of at least five tests within 95% confidence interval.

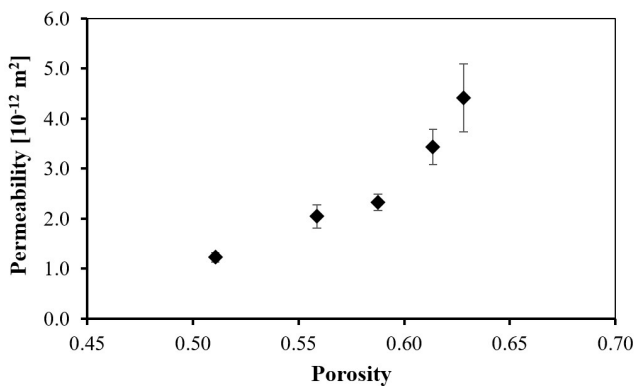


Figure 4. Relation between permeability and porosity of the same TGP-H-120 sample that has been successively compressed. Error bars represent the repeatability of at least five tests within 95% confidential interval.

IV. CONCLUSIONS

In this study, the permeability of an increasingly strained gas diffusion layer (GDL) for polymer electrolyte membrane (PEM) fuel cells is measured. In the nominal 40 wt% PTFE-treated Toray® TGP-H-120 carbon paper, a linear relationship is observed between the peak nominal stress and the residual strain. As the GDL is strained from 0 to 0.240, the porosity is observed to decrease from 62.8% to 51.1%. Consistent with the porosity changes, the permeability of the uncompressed GDL is $4.3 \times 10^{-12} \text{ m}^2$, though when strained to 0.240, the permeability is observed to be reduced by almost 80% to $1.1 \times 10^{-12} \text{ m}^2$. The reduction in porosity is responsible for the decreased permeability without any indication of excessive physical changes beyond a decreased thickness. Inspection of changes in surface morphology by microscopy confirmed limited breakage and crushing of fibres while most of the structure remained unchanged. The experimental data on permeability of strained GDLs at various degrees would be beneficial to the understanding of mass transport phenomena in compressed GDLs in an operating PEM fuel cell.

ACKNOWLEDGEMENT

This work received financial support from Canadian Urban Transit Research and Innovation Consortium (CUTRIC Project Number: 160028) and the Natural Sciences and Engineering Research Council of Canada (NSERC) via a Discovery Grant.

REFERENCES

- [1] L. Cindrella *et al.*, "Gas diffusion layer for proton exchange membrane fuel cells-A review," *J. Power Sources*, vol. 194, no. 1, pp. 146–160, 2009.
- [2] J.-H. Wee, "Applications of proton exchange membrane fuel cell systems," *Renew. Sustain. Energy Rev.*, vol. 11, pp. 1720–1738, 2007.
- [3] H. Zhao and A. F. Burke, "Optimization of fuel cell system operating conditions for fuel cell vehicles," *J. Power Sources*, vol. 186, no. 2, pp. 408–416, 2009.
- [4] J. Zhao, A. Ozden, S. Shahgaldi, I. E. Alaefour, X. Li, and F. Hamdullahpur, "Effect of Pt loading and catalyst type on the pore structure of porous electrodes in polymer electrolyte membrane (PEM) fuel cells," *Energy*, vol. 150, pp. 69–76, 2018.
- [5] J. Zhao, S. Shahgaldi, I. Alaefour, S. Yang, and X. Li, "Pore structure and effective diffusion coefficient of catalyzed electrodes in polymer electrolyte membrane fuel cells," *Int. J. Hydrogen Energy*, vol. 43, no. 7, pp. 3776–3785, 2018.
- [6] J. Zhao, S. Shahgaldi, A. Ozden, I. E. Alaefour, X. Li, and F. Hamdullahpur, "Geometric pore surface area and fractal dimension of catalyzed electrodes in polymer electrolyte membrane fuel cells," *Int. J. energy Res.*, pp. 1–9, 2018.
- [7] A. Ozden, S. Shahgaldi, J. Zhao, X. Li, and F. Hamdullahpur, "Assessment of graphene as an alternative microporous layer material for proton exchange membrane fuel cells," *Fuel*, vol. 215, pp. 726–734, 2018.
- [8] A. Tamayol, F. Mcgregor, and M. Bahrani, "Single phase through-plane permeability of carbon paper gas diffusion layers," *J. Power Sources*, vol. 204, pp. 94–99, 2011.
- [9] I. Gatto, F. Urbani, G. Giaccoppo, O. Barbera, and E. Passalacqua, "Influence of the bolt torque on PEFC performance with different gasket materials," *Int. J. Hydrogen Energy*, vol. 36, no. 20, pp. 13043–13050, 2011.
- [10] Y. Zhou, G. Lin, A. J. Shih, and S. J. Hu, "Assembly pressure and membrane swelling in PEM fuel cells," *J. Power Sources*, vol. 192, no. 2, pp. 544–551, 2009.
- [11] I. Taymaz and M. Benli, "Numerical study of assembly pressure effect on the performance of proton exchange membrane fuel cell," *Energy*, vol. 35, no. 5, pp. 2134–2140, 2010.

- [12] P. K. Sow, S. Prass, P. Kalisvaart, and W. Mérida, "Deconvolution of electrical contact and bulk resistance of gas diffusion layers for fuel cell applications," *Int. J. Hydrogen Energy*, vol. 40, no. 6, pp. 2850–2861, 2015.
- [13] A. El-Kharouf and R. Steinberger-Wilckens, "The Effect of Clamping Pressure on Gas Diffusion Layer Performance in Polymer Electrolyte Fuel Cells," *Fuel Cells*, vol. 15, no. 6, pp. 802–812, 2015.
- [14] P. H. Chi, S. H. Chan, F. B. Weng, A. Su, P. C. Sui, and N. Djilali, "On the effects of non-uniform property distribution due to compression in the gas diffusion layer of a PEMFC," *Int. J. Hydrogen Energy*, vol. 35, no. 7, pp. 2936–2948, 2010.
- [15] T. J. Mason, J. Millichamp, T. P. Neville, A. El-Kharouf, B. G. Pollet, and D. J. L. Brett, "Effect of clamping pressure on ohmic resistance and compression of gas diffusion layers for polymer electrolyte fuel cells," *J. Power Sources*, vol. 219, pp. 52–59, 2012.
- [16] J. Zhao, "Catalyst Layers in Polymer Electrolyte Membrane Fuel Cells: Formation, Characterization and Performance," University of Waterloo, 2019.
- [17] J. Zhao, S. Shahgaldi, I. Alaefour, Q. Xu, and X. Li, "Gas permeability of catalyzed electrodes in polymer electrolyte membrane fuel cells," *Appl. Energy*, vol. 209, pp. 203–210, 2018.
- [18] O. M. Orogbemi, D. B. Ingham, M. S. Ismail, K. J. Hughes, L. Ma, and M. Pourkashanian, "On the gas permeability of the microporous layer used in polymer electrolyte fuel cells," *J. Energy Inst.*, vol. 91, no. 6, pp. 894–901, 2018.
- [19] A. Tamayol, F. McGregor, and M. Bahrami, "Single phase through-plane permeability of carbon paper gas diffusion layers," *J. Power Sources*, vol. 204, pp. 94–99, 2012.
- [20] M. Hossain, S. Z. Islam, and P. Pollard, "Numerical study of the effect of effective diffusivity and permeability of the gas diffusion layer on fuel cell performance," *Proc. Inst. Mech. Eng. Part A J. Power Energy*, vol. 226, no. 7, pp. 907–921, 2012.
- [21] J. T. Gostick, M. W. Fowler, M. D. Pritzker, M. A. Ioannidis, and L. M. Behra, "In-plane and through-plane gas permeability of carbon fiber electrode backing layers," *J. Power Sources*, vol. 162, no. 1, pp. 228–238, 2006.
- [22] J. Ihonen, M. Mikkola, and G. Lindbergh, "Flooding of Gas Diffusion Backing in PEFCs," *J. Electrochem. Soc.*, vol. 151, no. 8, p. A1152, 2004.
- [23] W. K. Epting and S. Litster, "Microscale measurements of oxygen concentration across the thickness of diffusion media in operating polymer electrolyte fuel cells," *J. Power Sources*, vol. 306, pp. 674–684, 2016.
- [24] I. S. Hussaini and C. Y. Wang, "Measurement of relative permeability of fuel cell diffusion media," *J. Power Sources*, vol. 195, no. 12, pp. 3830–3840, 2010.
- [25] P. A. Gigos, Y. Faydi, and Y. Meyer, "Mechanical characterization and analytical modeling of gas diffusion layers under cyclic compression," *Int. J. Hydrogen Energy*, vol. 40, no. 17, pp. 5958–5965, 2015.
- [26] M. M. Tomadakis, "Viscous Permeability of Random Fiber Structures: Comparison of Electrical and Diffusional Estimates with Experimental and Analytical Results," *J. Compos. Mater.*, vol. 39, no. 2, pp. 163–188, 2005.
- [27] F. J. Valdes-Parada, J. A. Ochoa-Tapia, and J. Alvarez-Ramirez, "Validity of the permeability Carman-Kozeny equation: A volume averaging approach," *Phys. A Stat. Mech. its Appl.*, vol. 388, no. 6, pp. 789–798, 2009.
- [28] Y. Shen, "Mechanical Degradation of Membrane Electrode Assemblies in Proton Exchange Membrane Fuel Cells," *Ph.D. Diss. MME, Univ. Waterloo*, p. 205, 2017.
- [29] H. Taira and H. Liu, "In-situ measurements of GDL effective permeability and under-land cross-flow in a PEM fuel cell," *Int. J. Hydrogen Energy*, vol. 37, no. 18, pp. 13725–13730, 2012.
- [30] J. Geertsma, "Estimating the Coefficient of Inertial Resistance in Fluid Flow Through Porous Media," *Soc. Pet. Eng. J.*, vol. 14, no. 05, pp. 445–450, 1974.
- [31] Y. Zhou *et al.*, "Enhancement of Pt and Pt-alloy fuel cell catalyst activity and durability via nitrogen-modified carbon supports," *Energy Environ. Sci.*, vol. 3, no. 10, pp. 1437–1446, 2010.

MODELING COUPLE LAYOUT OF A THERMOELECTRIC GENERATOR

Xi Wang

Turbulence & Energy Lab
 University of Windsor
 Windsor, Ontario, Canada

Paul Henshaw

Turbulence & Energy Lab
 University of Windsor
 Windsor, Ontario, Canada

David S-K. Ting

Turbulence & Energy Lab
 University of Windsor
 Windsor, Ontario, Canada

Abstract—A solid-state device, named thermoelectric generator (TEG), can be impelled by heat flow to generate power. The device has been applied in heat recovery, with sources ranging from hot liquids to exhaust gases. In this paper, a TEG module involving 199 pairs of couples (each couple is actually a pair of PN junction) was modeled by ANSYS Workbench. After being validated using experimental results from the literature, the model was used to analyze the effect of couple layouts on TEG performance. Four configurations of couple layouts were considered. It was indicated that these couple layouts influence on the TEG's open circuit voltage and current; however, the effects on its output power and efficiency are limited.

Keywords- *heat recovery technology; thermoelectric effect; TEG; ANSYS Workbench*

I. INTRODUCTION

A thermoelectric generator (TEG) is a semiconductor-based energy converter. A TEG system typically consists of a thermoelectric module, sandwiched between a heat source and a radiator [1]. The working principle of the TEG is the thermoelectric effect, involving both Seebeck and Peltier effects, *i.e.*, utilizing temperature difference to produce electricity [2-4]. As such, TEGs are considered as an effective method for recovering waste heat energy [1].

TEGs are compact and reliable, and thus, have been applied in many industrial domains [1]. Many automobile companies, such as Ford and BMW, are investigating TEG devices to recover the waste heat. A device with 72 TEG modules of 40 mm by 40 mm produced 75 W of electricity when the vehicle speed reached 80 km/h [5]. In addition, a famous Japanese watch maker, Seiko, developed a watch powered by TEG, which can create 300 mV open circuit voltage from only a 1.5°C temperature gradient [5].

However, an outstanding challenge of TEG is its low efficiency (less than 5%) [5]. In an effort to boost its efficiency, current research mainly focuses on two promising aspects. Firstly, thermoelectric materials play a dominant role in TEG performance. With the development of material technology, especially the application of nano-materials and superlattice materials, the efficiency has reached 7% for commercial materials, and 20% for the latest research materials [5-7].

Secondly, in order to improve the TEG's efficiency, some researchers have tried to increase the performance of the radiator to better dissipate heat to the cold temperature sink. Therefore, many advanced thermal dissipation technologies have been applied to TEG systems, especially, heat pipes and micro-channels [8,9].

Although the properties of semiconductor materials have a decisive effect on TEG efficiency, one should not lose sight of the important influence of couple layouts of a TEG on its performance [10]. In this paper, a TEG model was created through ANSYS Workbench and validated using the experimental results of Hsu *et al.* [11] for a TEG with 199 pairs of couples. Upon validation, it was used to model TEG performance with varying couple layouts. Thereby, four kinds of couple layouts were considered and electrical output characteristics were calculated, including voltage, output power and efficiency.

II. METHODOLOGY

A TEG module normally consists of three main parts: the semiconductor, the electrical material and the insulating material [12,13]. Each couple consists of a pair of semiconductors (P material and N material), connected by an electrically conductive material, such as copper or aluminum. Finally, the module is packaged on a substrate, which is an electrically insulating material, such as Al₂O₃ ceramic.

Based on the working principle of a TEG, it exhibits thermoelectric coupling. Therefore, the Seebeck effect and Peltier effect functions are considered as important coupling equations. The Seebeck effect means that the temperature gradient can create voltage between two kinds of semiconductors. The voltage is [5]:

$$\begin{aligned}\Delta V &= V_P - V_N = (S_P + |S_N|) \Delta T \\ &= S (T_h - T_c)\end{aligned}\tag{1}$$

where ΔV is the voltage; V_P and V_N are the electric potentials of the P semiconductor and N semiconductor, respectively; S_P and S_N are the Seebeck coefficients of the P semiconductor and N semiconductor, respectively; S is the relative Seebeck coefficient; T_h and T_c are the temperatures of the hot side and cold sides, respectively.

The Peltier effect means that the electrical current can create a cooling or heating effect at a junction of semiconductors, depending on the direction of the current flow. The Peltier heat transfer rate is [13]:

$$\dot{Q}_P = \Pi I \quad (2)$$

where \dot{Q}_P is the Peltier heat transfer rate; I is the current flow; Π is the relative Peltier coefficient.

Due to the Kelvin relation, the Peltier coefficient is:

$$\Pi = ST \quad (3)$$

in which T is the temperature of the junction.

Therefore,

$$\dot{Q}_P = STI \quad (4)$$

A. Electrical Characteristics.

A TEG module can be considered as a kind of battery, which creates voltage under a temperature gradient. Therefore, the TEG will produce electric power when connected with a load [14].

$$I = \frac{\Delta V}{r_i + r_L} = \frac{S(T_h - T_c)}{r_i + r_L} \quad (5)$$

$$P = I^2 r_L = \frac{S^2 (T_h - T_c)^2 r_L}{(r_i + r_L)^2} \quad (6)$$

$$r_i = r_P + r_N = \frac{\rho_P L_P}{A_P} + \frac{\rho_N L_N}{A_N} \quad (7)$$

in which r_i and r_L are the internal resistance and load resistance; r_P , ρ_P , L_P and A_P are the resistance, resistivity, length and cross-sectional area of the P semiconductor, respectively; r_N , ρ_N , L_N and A_N are the resistance, resistivity, length and cross-sectional area of the N semiconductor, respectively.

When $r_L = r_i$, the output power reaches a maximum:

$$P_{max} = \frac{S^2 (T_h - T_c)^2}{4r_i} \quad (8)$$

where P_{max} is the maximum output power.

B. Thermal Transfer Process.

Figure 1 depicts the thermal transfer processes in a TEG system. The thermal energy flows through a TEG system to produce electrical power. However, according to the second

law of thermodynamics [14], it is impossible to use all the thermal energy for work. Therefore, some input energy will be lost through Fourier heat and Peltier heat, *etc.*

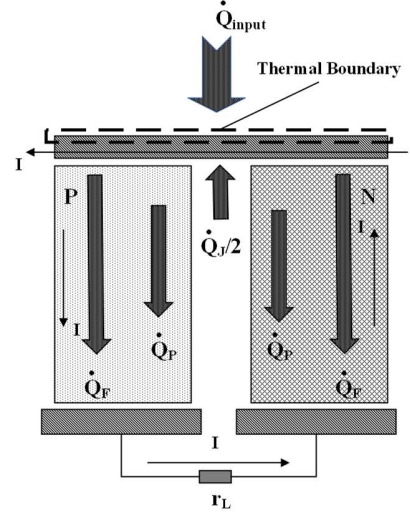


Figure 1. Schematic diagram of a TEG heat transfer process

As shown in Figure 1, there are two kinds of energy absorbed by the hot side thermal boundary: input heat and Joule heat. This boundary absorbs only half of the Joule heat, for the other half is absorbed by the cold side. Meanwhile, there are two kinds of energy coming out of this boundary, which are Fourier heat and Peltier heat. Thus, according to Figure 1 [15], the heat input rate is:

$$\dot{Q}_{input} = \dot{Q}_F + \dot{Q}_P - \frac{1}{2}\dot{Q}_J \quad (9)$$

The Fourier heat transfer rate is:

$$\begin{aligned} \dot{Q}_F &= \frac{k_P \Delta T A_P}{L_P} + \frac{k_N \Delta T A_N}{L_N} = \left(\frac{k_P A_P}{L_P} + \frac{k_N A_N}{L_N} \right) \Delta T \\ &= k(T_h - T_c) \end{aligned} \quad (10)$$

The Peltier heat transfer rate is:

$$\dot{Q}_P = ST_h I \quad (11)$$

And the Joule heat transfer rate is:

$$\dot{Q}_J = I^2 r_i \quad (12)$$

in which ΔT is the temperature gradient; \dot{Q}_{input} , \dot{Q}_F , \dot{Q}_P and \dot{Q}_J are the input, Fourier, Peltier and Joule heat transfer rates, respectively; k , k_P and k_N are the relative thermal conductivity,

and thermal conductivities of the P semiconductor and N semiconductor, respectively.

C. Efficiency.

As is known, the efficiency of a TEG is the ratio of output electric power and input heat [15].

$$\eta = \frac{P}{\dot{Q}_{input}} = \frac{\frac{S^2(T_h - T_c)^2 r_L}{(r_i + r_L)^2}}{\dot{Q}_F + \dot{Q}_P - \frac{1}{2}\dot{Q}_J}$$

$$= \frac{\frac{S^2(T_h - T_c)^2 r_L}{(r_i + r_L)^2}}{k(T_h - T_c) + ST_h \frac{S(T_h - T_c)}{r_i + r_L} - \frac{1}{2} \left[\frac{S(T_h - T_c)}{r_i + r_L} \right]^2 r} \quad (13)$$

In order to simplify Equation (13), some parameters are introduced:

$$Z = \frac{S^2}{kr_i} \quad (14)$$

$$m = \frac{r_L}{r_i} \quad (15)$$

Finally, the efficiency of a TEG is:

$$\eta = \frac{T_h - T_c}{T_h} \frac{\frac{m}{m+1}}{\left(1 + \frac{m+1}{ZT_h} - \frac{1}{2} \frac{T_h - T_c}{T_h} \frac{1}{m+1}\right)} \quad (16)$$

where Z is the thermoelectric figure of merit; m is the ratio of the load resistance to the internal resistance.

III. TEG MODEL ESTABLISHMENT AND VALIDATION

This section describes a TEG modelled by ANSYS Workbench. Meanwhile, an experimental result reported by Cheng-Ting Hsu's group was used to validate the TEG model.

A. Establishment and Simplification of TEG Model.

Based on Cheng-Ting Hsu's experiment, a TEG model with 199 pairs of couples was established, as shown in Figure 2 (a). The initial dimension of the TEG was 60 mm × 60 mm × 2.91 mm. Inside, the initial dimension of the individual semiconductors and electric conductors were 2 mm × 2 mm × 0.64 mm and 4.5 mm × 2 mm × 0.5 mm.

The TEG model consisted of 199 pairs of couples, being considered as 199 batteries in series. Hence, the total voltage is the sum of each couple's voltage. In this way, the model was simplified as 2 pairs of couples. The simplified model can be seen in Figure 2 (b). Table 1 provides some information about material properties.

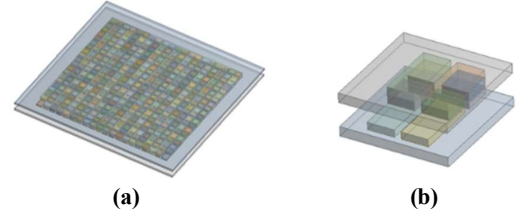


Figure 2. TEG model geometrical structure, TEG with 199 couples (a), TEG with 2 couples (b)

Table 1. Material properties in the TEG model [11,16]

Parameter	P semiconductor	N semiconductor	copper	ceramic
Seebeck coefficient (VK ⁻¹)	Constant 2.15×10 ⁻⁴ Variable ± (22224+930.6×T- 0.9905×T ²) ×10 ⁻⁹	Constant -2.12×10 ⁻⁴	—	—
Thermal conductivity (Wm ⁻¹ K ⁻¹)	Constant 1.373 Variable (62605-277.7×T+0.4131×T ²) ×10 ⁻⁴	Constant 1.456	385	22
Resistivity (Ωm)	Constant 1.04×10 ⁻⁵ Variable (5112+163.4×T+0.6279×T ²) ×10 ⁻¹⁰	Constant 1.04×10 ⁻⁵	3.2×10 ⁻⁸	1×10 ¹²
Contact area (m ²)	4×10 ⁻⁶	4×10 ⁻⁶	9×10 ⁻⁶	9×10 ⁻⁶
Thickness (m)	6.4×10 ⁻⁴	6.4×10 ⁻⁴	5×10 ⁻⁴	6.35×10 ⁻⁴

The Seebeck coefficient of the P semiconductor is positive and that of the N is negative.

During the process of modelling studies, it was assumed that the contact thermal resistance, electric resistance and heat loss could be ignored. Meanwhile, the first thermal boundary condition was set for both cold and hot sides. In addition, the electric potential of a surface with 0 mV was set as a reference. The boundary conditions were shown in Figure 3.

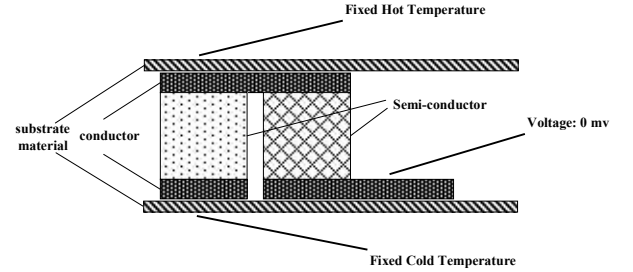


Figure 3. Schematic cross-section of a PN couple showing boundary conditions

B. Model Validation.

The validation of the TEG model was done before using it in further studies. In this paper, experimental data from Hsu's group was used to check the simulation result [11]. Therefore, the temperature of the cold side was fixed at 300 °C, and the temperature of the hot side was varied from 305 °C to 340 °C under these conditions. Then, the open circuit voltage from the model was compared with the Hsu group results.

In this work, three parameters used in the model could be constant or variable (Table 1). According to the results (Figure 4), there is a similar trend of open circuit voltage with temperature difference in the three cases: the voltage value had a linear increase with the temperature gradient growth. Compared with the results of the experiment, the average error of the simulation result with constant properties (25.6%) was

higher than when using variable properties (16.7%), when the temperature difference was 40 °C. In order to make the TEG performance more meaningful, it was necessary to increase the temperature difference in further studies. Using a linear relationship between the experimental open circuit voltage and the temperature difference, and running the model at a temperature difference of 300 °C, it was predicted that the average error of the simulation result with variable properties was 19.1%, which was also lower than that for constant properties (25.5%). Hence, the TEG model with variable properties was used for further studies.

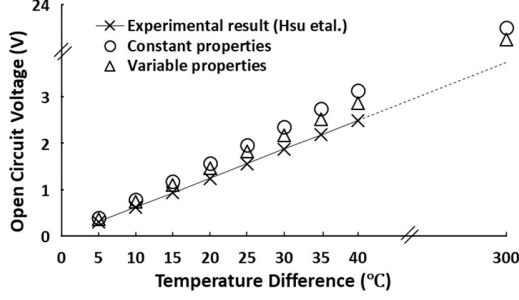


Figure 4. Comparison of simulations with experimental results

IV. THE EFFECTS OF LAYOUT ON THE TEG PERFORMANCE

The thermoelectric properties of TEGs, such as internal resistance, may vary across couple layouts. It is imagined that couple layouts have important influences on TEG's performance. In this section, four kinds of couple layouts were considered, shown in Figure 5. For each couple layout, the substrate area was 8 mm × 8 mm and the total couple area was 16 mm². During this set of simulations, the temperature of the cold side was fixed at 30°C, and the temperature of the hot side was varied from 100°C to 300°C. Moreover, the value of m was 1, namely; the internal resistance equaled the load resistance.

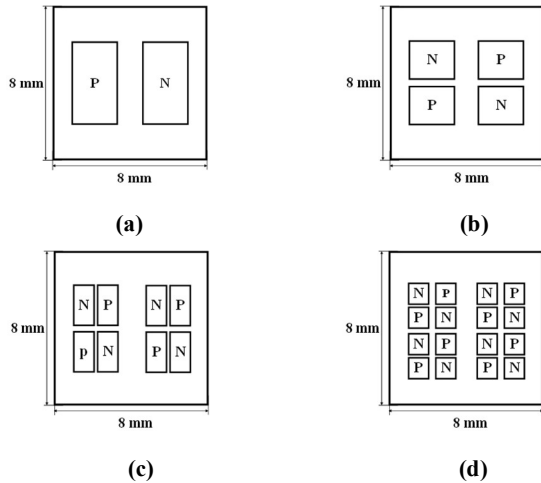


Figure 5. Schematic of four kinds of couple layouts, (a) 1-couple layout, (b) 2-couples layout, (c) 4-couples layout and (d) 8-couples layout

According to Figure 6, as for a specific layout, the output power experienced an increasing trend with the growth of hot side temperature (T_h). The reason is that a higher temperature on the hot side means a larger temperature difference. At each specific T_h , there was a slight increase in output power with different layouts. When the temperature of the hot side was 300°C, the output power of the 8-couples layout was highest in the graph with 0.90 W, while the smallest was for the 1-couple layout, which is 0.84 W. As for the 2-couples layout and 4-couples layout, the output power were 0.86 W and 0.87 W, respectively.

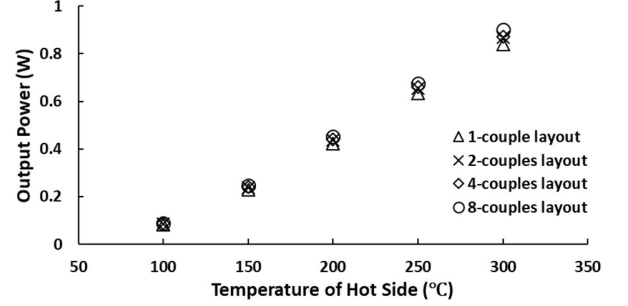


Figure 6. Output power at different hot temperature and layouts

In addition, some similar trends are found in Figure 7. The most obvious difference happened in the change of the efficiency with different T_h for a specific layout. The effect of T_h on the TEG efficiency is not a linear relationship. Although the output power of the TEG module increases with temperature of the hot side, the increase in T_h results in a higher input thermal energy (denominator of Equation 13).

Based on Equation (7), there is a higher internal resistance for each couple with a smaller across-sectional area, such as 8-couples layout. However, all the couples are in series, so the output power is sum of each couple's power. In this way, although a high internal resistance may weaken the individual couple's TEG performance, it can be compensated by the quantity of couples. This is the main reason why influences of layout on the TEG output power and efficiency are inconspicuous.

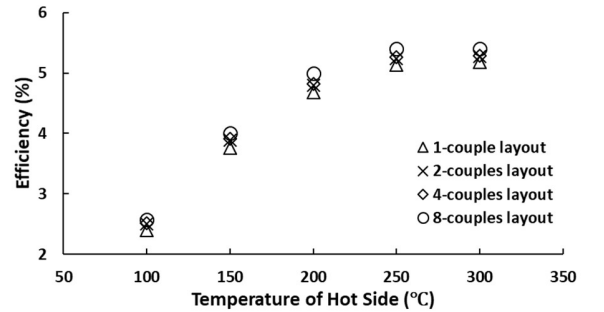


Figure 7. Efficiency at different hot temperature and layouts

As for a specific layout, a higher T_h led to a higher open circuit voltage and current, as shown in Figure 8 and 9. However, at a specific T_h , there was a totally different

tendency in the open circuit voltage and current among layouts. Because all couples were in series, the open circuit voltage was the total of each couple voltage. Thus, the TEG with more couples (like 8-couples layout) produced a higher open circuit voltage under the same temperature condition. Inversely, the total current equaled each couple's current as they were in series. The couple in 1-couple layout had the smallest internal resistance resulting from having the biggest cross-sectional area of all couples. Therefore, the current of 1-couple layout was the highest.

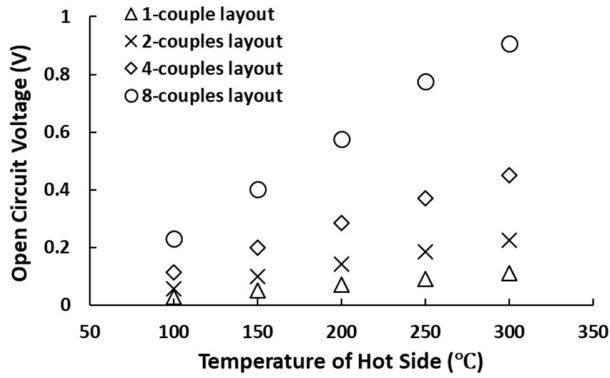


Figure 8 Open circuit voltage at different hot temperature and layouts

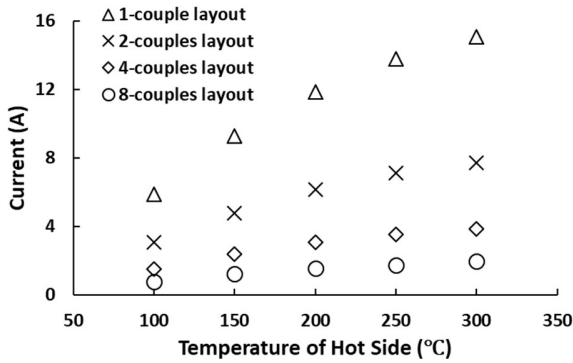


Figure 9. Current at different hot temperature and layouts

V. CONCLUSIONS

In this paper, a TEG with constant and variable material properties was simulated with ANSYS-Workbench. The base-case simulation results were validated by experimental results. The results with variable properties were much closer to the experimental data than those with constant properties. Moreover, the modelling study indicates that the variation of the output power and efficiency is inconspicuous for different layouts. However, layouts have inverse influences on open circuit voltage and current as couples are in series: the 1-couple layout and 8-couples layout had the maximum current and voltage respectively.

In future work, the assumptions invoked in this study can be relaxed one at a time. For example, the effects of heat loss and welding resistance can be modeled, instead of neglecting them, as was done here.

ACKNOWLEDGEMENTS

This work is made possible by funding from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

- [1] M. Chen, J. L. Gao and Z. D. Kang et al., 2012, "Design Methodology of Large-Scale Thermoelectric Generation: A Hierarchical Modeling Approach," *Journal of Thermal Science and Engineering Applications*, 4 (4), pp. 041003-1–041003-9.
- [2] S. Kim, S. Park and S. Kim et al., 2011, "A Thermoelectric Generator Using Engine Coolant for Light-Duty Internal Combustion Engine-Powered Vehicles," *Journal of Electronic Materials*, 40 (5), pp. 812–816.
- [3] C. T. Hsu, G. Y. Huang and H. S. Chu et al., 2011, "An Effective Seebeck Coefficient Obtained by Experimental Results of a Thermoelectric Generator Module," *Applied Energy*, 88 (12), pp. 5173–5279.
- [4] A. R. M. Siddique, S. Mahmud and B. V. Heyst, 2017, "A Review of the State of the Science on Wearable Thermoelectric Power Generators (TEGs) and Their Existing Challenges," *Renewable and Sustainable Energy Reviews*, 73, pp. 730–744.
- [5] O. H. Ando Junior, A. L. O. Maran and N. C. Henao, 2018, "A Review of the Development and Applications of Thermoelectric Microgenerators for Energy Harvesting," *Renewable and Sustainable Energy Reviews*, 91, pp. 376–393.
- [6] W. S. Liu, Q. Jie and H. S. Kim et al., 2015, "Current Progress and Future Challenges in Thermoelectric Power Generation: From Materials to Devices," *ScienceDirect*, 87, pp. 357–376.
- [7] N. Farahi, S. Prabhudev and M. Bugnet et al., 2016, "Effect of Silicon Carbide Nanoparticles on the Grain Boundary Segregation and Thermoelectric Properties of Bismuth Doped $\text{Mg}_2\text{Si}_{0.7}\text{Ge}_{0.3}$," *J. Electron. Mater.*, 45 (12), pp. 6052–6058.
- [8] P. J. Kim, S. H. Rhi and K. B. Lee et al., 2013, "Heat-Pipe-Associated Localized Thermoelectric Power Generation System," *Journal of Electronic Materials*, 43 (6), pp. 1613–1619.
- [9] D. J. Yang and H. M. Yin, 2011, "Energy Conversion Efficiency of a Novel Hybrid Solar System for Photovoltaic, Thermoelectric, and Heat Utilization," *IEEE Transactions on Energy Conversion*, 26 (2), pp. 662–670.
- [10] T. Z. Ming, W. Yang and X. M. Huang et al., 2017, "Analytical and Numerical Investigation on a New Compact Thermoelectric Generator," *Energy Conversion and Management*, 132, pp. 261–271.
- [11] C. T. Hsu, G. Y. Huang and H. S. Chu et al., 2010, "Experiments and Simulations on Low-Temperature Waste Heat Harvesting System by Thermoelectric Power Generators," *Applied Energy*, 88 (4), pp. 1291–1297.
- [12] A. Rezaia and L. A. Rosendahl, 2012, "Thermal Effect of Ceramic Substrate on Heat Distribution in Thermoelectric Generators," *Journal of Electronic Materials*, 41 (6), pp. 1343–1347.
- [13] T. H. Kwan, X. F. Wu and Q. H. Yao, 2018, "Thermoelectric Device Multi-Objective Optimization Using a Simultaneous TEG and TEC Characterization," *Energy Conversion and Management*, 168, pp. 85–97.
- [14] Y. A. Çengel and M. A. Boles, 2008, "Thermodynamics: An Engineering Approach, Sixth Edition," McGraw-Hill, New York.
- [15] S. Twaha, J. Zhun and Y. Yan et al., 2015, "A Comprehensive Review of Thermoelectric Technology: Materials, Applications, Modelling and Performance Improvement," *Renewable and Sustainable Energy Reviews*, 65, pp. 698–726.
- [16] G. Fraisse, J. Ramousse and D. Sgorlon et al., 2012, "Comparison of Different Modeling Approaches for Thermoelectric Elements," *Energy Conversion and Management*, 65, pp. 351–356.

Detailed Characterization of Novel Radiant Floor Heating and Cooling Systems

S. Shukla, R. Daneshazarian, A. Mwesigye, SB. Dworkin
 Mechanical and Industrial Engineering Department
 Ryerson University
 Toronto, Canada

J. Swift
 Grip Metal Manufacturing
 NUCAP Industries Inc.
 Scarborough, Canada

Abstract—Traditional radiant floor heating systems consist of a heat source- hot water pipes or electrical heating wires, embedded in concrete and other building materials. Radiant floor cooling systems typically use water close to room temperature. Both systems have the potential to offer benefits to homeowners as compared to the widely adopted forced air and hot water radiator systems. They exhibit quiet operation, minimal airborne dust production, and furniture space savings. Also, they optimize thermal comfort and result in energy savings of 15-30%. However, these floor heating systems are more expensive than conventional alternatives, and there is still potential for improved heat transfer efficiency with novel design improvements. Floor cooling systems are also limited by factors such as acceptable floor temperature, dew point temperature, and responsiveness of the floor to fluctuations in thermal load. Recently, a retrofittable radiant panel configuration has been proposed by the industry partner as a component in floor heating systems, comprising a metal plate with small spikes that can be pressed into cement board or wood. The spikes can serve two purposes; they can bind materials together reducing manufacturing costs and improve heat transfer. The behavior of this configuration was simulated for different materials for the metal plate, spike dimensions, and varying spacing between the spikes. Two scenarios for this configuration were considered: radiant floor heating in an office environment and radiant floor cooling in a basketball court. The optimal configuration comprises an aluminum sheet with an enhanced surface consisting of a series of spikes, increasing surface area and penetrating surrounding material, with spike length of 0.09 in (2.3 mm). It yields approximately 60% energy reduction and heat transfer enhancement compared to a no-spike metal plate configuration. The outcome of this research is fundamental understanding leading to an optimized design of a cost-effective and energy-efficient floor-heating and floor-cooling system that can be easily installed in new or retrofitted buildings. By significantly improving radiant floor heating systems, more and more traditional natural gas heating can be offset, leading to reduced greenhouse gas emissions.

Keywords- *Metal Plate with Spikes; Radiant Floor Heating and Cooling; Energy Efficiency; Thermal Comfort; Computer Simulation; Economic Optimization*

I. INTRODUCTION

A. Background

1) *Energy Use in Buildings:* According to the International Energy Agency (IEA), building sector is responsible for 36% of total energy consumption and this contributes to 40% of total CO₂ emissions. Over the next 40 years, the building sector will grow by nearly 230 billion square meters which is equivalent to adding the floor area of Japan's landmass to the planet every year until 2060 [3]. Hence, there has been a spark of interest globally to reduce the energy consumption of commercial & industrial buildings and residential housing. The combination of strategies such as improving efficiency of heating, ventilation, and air conditioning (HVAC) equipment and reducing the thermal demand of the house by improving envelope conditions, better control, and introducing renewable energy technologies could yield significant reduction of greenhouse gas emissions (GHG). The potential for GHG emission, CO₂ primarily, is presented in Fig.1.

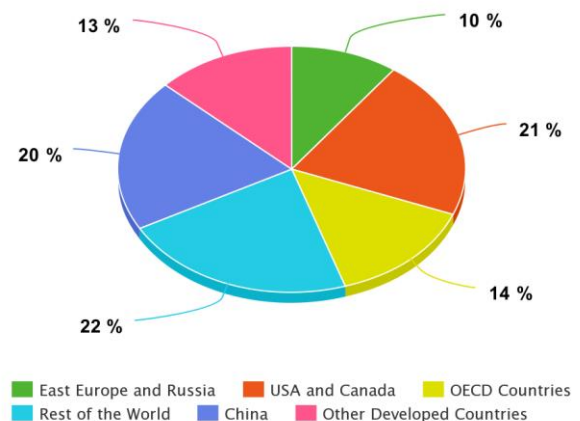


Figure 1. Potential CO₂ reductions from the building sector [3]

According to a survey of commercial and institutional energy use, commercial and institutional buildings in Canada represent 26% of the total energy consumption [4]. As

presented in Figure 2, 65% of the energy consumed in the buildings sector is dedicated to space heating and thus the highest GHG intensity- greenhouse gas emissions per unit floor area (m²)- is seen for space heating equipment as well.

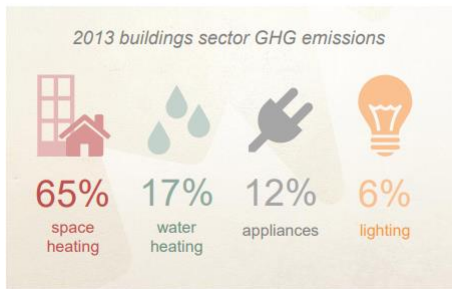


Figure 2. Breakdown of building sector GHG emissions [5]

At 41.6% of the total energy consumption of buildings in Canada, Ontario's building sector consumes the highest amount of energy and this is because buildings, commercial and institutional, account for 35.6% of buildings in Canada [6]. According to the National Energy Board (NEB), the GHG emission factor for electricity and natural gas in Ontario is 40 g CO₂ per kWh electricity generation and 1860 g CO₂ per m³ of natural gas burned [7]. By significantly improving radiant floor heating systems, more and more traditional natural gas heating can be offset, leading to reduced GHG emissions. Since buildings are the biggest consumer of energy in Canada, it is important that space heating systems used in buildings operate in a more economical and sustainable way. This research is in line with Canada's Climate Change Action Plan (CCCAP) and the short-term goal of reducing GHG emissions by 15% below 1990 levels by the year 2020 [8].

2) *Radiant Floor Heating and Cooling Systems:* The Romans pioneered floor-heating systems; evident from ancient sites known as "hypocausts", the floors were heated by directing exhaust gases from wood fires under raised floors [9]. In early 20th century, when building envelopes were not as airtight as they are now, radiant floor systems were not adopted as much in cold climate countries such as Canada. To compensate for harsh cold climatic conditions, floors had to be heated to uncomfortably high temperatures and thus there was little to no effort made in navigating nuances of space heating systems. However, since the energy crisis of the 1970s, there has been increased focus on implementing buildings code that foster 'energy-efficient' building envelopes such that those pertaining to minimizing transmission and ventilation heat losses [9].

Radiant floor heating and cooling systems have the potential to offer more benefits to homeowners than the widely adopted forced-air convection system. State-of-the-art systems comprise hot water pipes embedded in concrete underneath wood, with a floor surface temperature between 19°C and 29°C [7–9]. They are integral to the sustainable energy market as they function well with geothermal, or efficient heat pump

systems. They exhibit quiet operation, minimal airborne dust production, furniture space savings, and simultaneously optimize comfort and energy savings [10,11].

Ideal heating curve starts at the lowest temperature at the head and increases towards the feet [13]. Radiant floor systems follow the same pattern. This demonstrates enhancement in thermal comfort conditions and it is shown in Fig. 3.

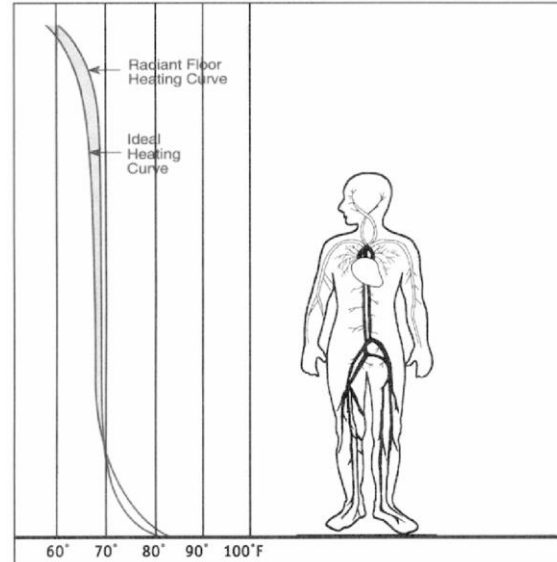


Figure 3. Ideal heating curve and radiant floor heating curve [13]

Radiant floor cooling systems emerged later than floor heating systems, having first been introduced in dry climates of Europe in the mid-20th century. They were first employed in the Copenhagen Opera House where outdoor conditions are relatively dry and cool and the floor cooling system's sole purpose is to remove solar heat gain by circulating water at 15-18°C [14].

The main challenge of radiant floor systems, however, is that they are expensive to construct and install. Furthermore, due to the insulating characteristics of many floor materials (e.g. wood), there is potential to improve energy transfer with advanced system design. Also, they cannot be easily retrofitted and the tendency of concrete to hold heat makes them less responsive to fluctuations in thermal loads [15].

Since their introduction, a number of studies have considered different variations of radiant floor heating and cooling systems. The next section details a review of related studies.

B. Literature Review

Different variations of radiant floor heating systems can be found in the literature. Izquierdo and Agustin-Camacho considered micro photovoltaic-heat pump systems in conjunction with radiant floor heating systems [16]. Zhou and He assessed thermal performance of a radiant floor heating

system with different heat storage materials [17]. Most of well-researched radiant floor heating systems fall in the following categories: heavy- pipes laid out in the concrete layer of the construction or lightweight- pipes placed in aluminum foil [18]. Also, numerous studies have been performed to demonstrate thermal comfort achieved by radiant floor systems; they show that the temperature distribution is uniform and maintains thermal comfort conditions prescribed by the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) [16–18].

Radiant floor cooling systems have been an economically viable option for large space buildings such as airports, convention centers, atria, and entrance halls [2]. This is primarily because, as identified through experimental analyses of multiple studies, the indoor thermal environment in these spaces possess high-intensity solar radiation and high temperature internal wall surfaces [20, 21]. For large space buildings whose envelopes are mainly composed of glass facades and skylights, floor cooling is an effective means of removing sensible heat because of direct absorption of solar radiation and long wave radiant heat exchange with a building envelope [2]. Research conducted at Lawrence Berkeley National Laboratory concluded that radiant panel cooling used less than 5% of the otherwise necessary fan energy to remove a given amount of indoor sensible heat [24]. Also, on average, it has the potential of saving 30% of the overall cooling energy in applications across a range of representative climates in North America [22, 23].

There has been little to no research performed on retrofitted radiant floor panel configurations that optimize thermal comfort and save energy, all while operating within 19°C to 29°C floor surface temperature ranges and minimizing the risk of condensation [9]. This study is novel and aims to accomplish an economically viable and energy-efficient radiant floor configuration that maintains thermal comfort standards. It incorporates metal sheeting [26] with small strong spikes (0.76 mm to 2.41 mm long), that can both decrease construction costs through mechanical adhesion, and increase conductivity through the floor. The physical model of this proposed design is discussed in the following section.

II. PHYSICAL MODEL

A schematic two-dimensional representation of the proposed radiant panel is depicted in Fig. 4. The heat source can be an electrical heating element or hot/cold water through a network of pipes.

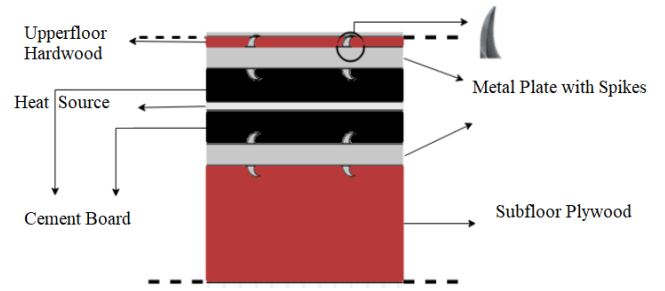


Figure 4. Radiant panel layout

A three-dimensional representation of the panel layout and spike variations for the metal plate are shown in Fig. 5 and Fig. 6, respectively.

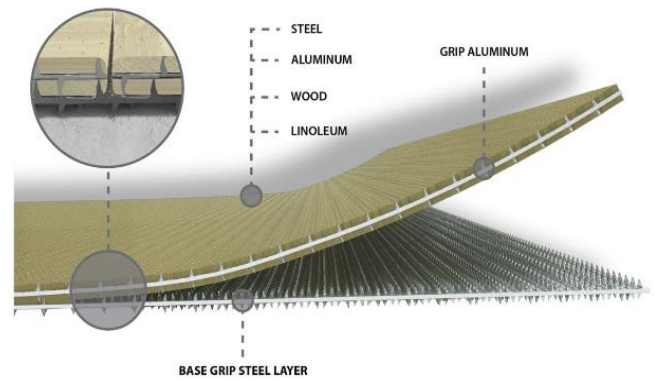


Figure 5. Three-dimensional panel layout [26]

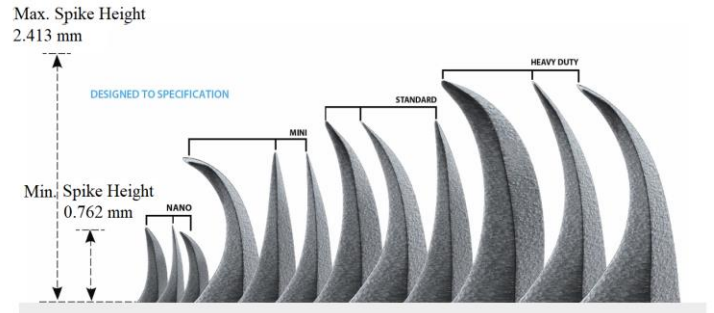


Figure 6. Spike variations for metal plate [26]

For the metal plate with spikes, two material alloys were considered: aluminum alloy 5052 H38 and carbon steel 1010 full hardness (H85-95).

Thermophysical properties of each layer in the panel layout presented in Fig. 4 are shown in Table I. Here, the panel layout and the materials used were the same for the radiant floor heating and the radiant floor cooling applications.

Table I. Thermophysical properties of each layer in radiant panel [27]

Layer	Material	Density (kg/m ³)	Thermal Conductivity (W/m.K)	Heat Capacity (J/ kg. K)
Extended Surface	Aluminum 5052 H38	2680	138	880
	Carbon Steel 1010	7870	49.8	450
Upper Floor	Hardwood	720	0.16	1255
Subfloor	Plywood	540	0.1154	1210
Concrete Board	Cement	1920	0.58	1006

III. METHODOLOGY

The investigation of the enhanced heat transfer surface started with creating a computational heat transfer model in COMSOL [27]. Detailed conductive heat transfer within the enhanced floor panel, along with convective and radiative heat transfer between the panel and a conditioned interior space, were characterized numerically. The model was used systematically to parametrically quantify surface configurations (varying spike spacing and length), as there were multiple versions available for integration. Newly proposed design modifications were then assessed by performing simulations and comparing the predicted power consumption needed to maintain room temperature. The radiant floor heating system was simulated for an office-building environment. The radiant floor cooling system was simulated for an indoor gymnasium-based basketball court.

A. Computational Domain

For the heat transfer simulation, two orthogonal cross sections have been considered; parallel to the heat source and perpendicular to the heat source. The lengths of the spikes are 2.29 mm (0.09'') and 1.52 mm (0.06'') based on ease of manufacturability. The dimensions of the various layers are given in Fig. 7. Here, the length of the subfloor is the same as spike length. Moreover, the heat source is a hot/cold water pipe.

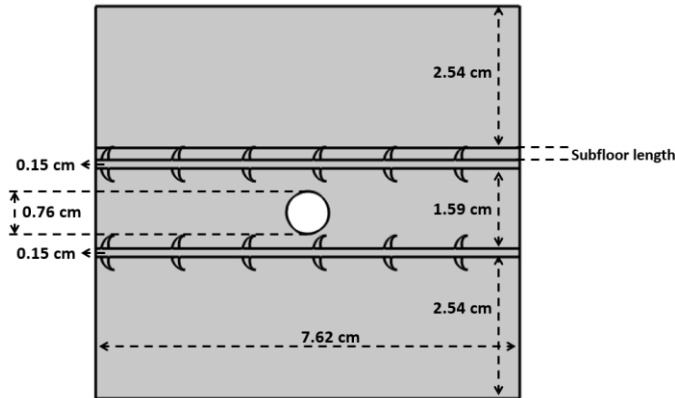


Figure 7. Dimensions of the radiant panel (perpendicular to the heat source)

B. Mesh Independency

The mesh independency analysis is carried out to find the optimum number of nodes for the simulation. As presented in Fig. 8, the optimum number is 27409 which is based on the time to achieve to the set point temperature of 22°C and 24°C for radiant floor heating and radiant floor cooling respectively. All the simulation results presented in subsequent sections will have achieved at least 27409 nodes with the use of a triangular mesh.

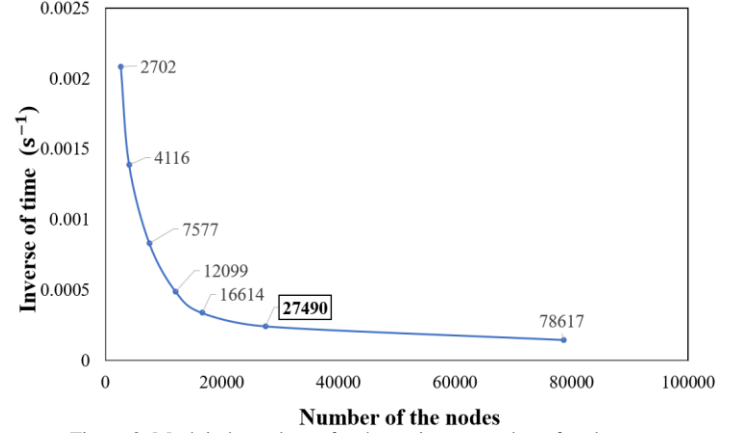


Figure 8. Mesh independency for the optimum number of nodes

C. Boundary Conditions

1) *Radiant Floor Heating*: For the radiant floor heating application, the initial temperature of the entire panel and air layer, topmost layer of thickness 2.54 cm on Fig. 7, were assumed to be 10°C. The left and right side of the panel layout had a periodic condition applied to it and thus both sides of the wall were maintained at the same temperatures. In accordance with ASHRAE Standard 55-1992, 30% relative humidity was considered representing typical indoor environment conditions for a winter season [10].

2) *Radiant Floor Cooling*: In accordance with ASHRAE standard 55-1992, 50% relative humidity was considered representing typical indoor environment conditions for the summer season [10]. An initial condition of 27°C was used representing a typical indoor temperature on a sunny day in the summer. Also, a periodic condition was used for both sides of the model.

D. Solution Procedure

1) *Radiant Floor Heating*: Here, different nuances of the panel layouts: different spike variations (spacing, length and fitting with the upper layer) and material for metal plate, were first modelled. Through a careful review of these simulation results, an optimal configuration was determined and then the effect of having more than two metal plate with spikes were studied.

2) *Radiant Floor Cooling*: For the radiant floor cooling application, two different simulations were run. Similar to the radiant floor heating application, the first of the simulations was run to quantify efficacy of enhanced metal plate with spikes and to understand how spike spacing, and material properties would influence heat distribution. Once optimal panel dimensions and metal plate material were determined, a radiant floor cooling application in basketball court was simulated.

For the radiant floor cooling application in a basketball court, it was important to understand three heat transfer mechanisms of cooling: conduction through the floor, evaporative cooling of sweat from the basketball player, and radiation due to the temperature difference between the floor and the player's body. Using Fanger's scale and considering an active metabolism of a basketball player, a body temperature of 38°C was used [9]. A typical indoor basketball court height of 17' was used, along with a basketball player height of 6' similar to a study in [20]. The model including the human body is shown in Fig. 9. At this stage in the project, only the influence of evaporative cooling has been studied and hence the model was simplified to a sphere at 38°C and 400W/m² metabolic rate [22].

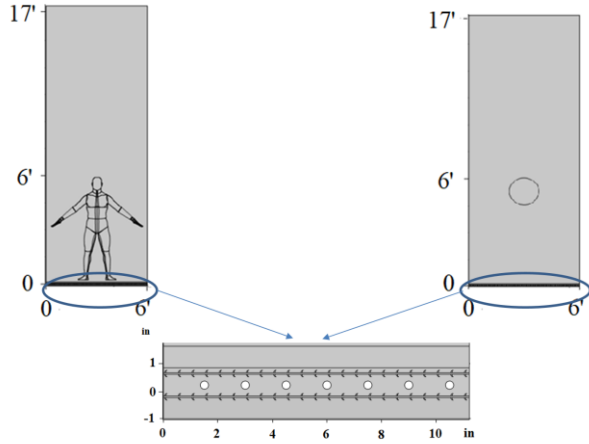


Figure 9. Basketball court radiant floor cooling application: left to right shows the transition to the simplified model and the bottom image shows a detailed view of array of an array of cold water pipes embedded in a radiant panel configuration

IV. RESULTS AND DISCUSSION

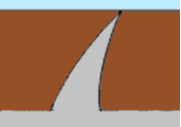


A. Radiant Floor Heating

Initial simulations were performed to observe the impact of spike protrusion, the material of the metal plate with spikes, and spike length. The results obtained for each of these variations are summarized in Table II. Table II shows the amount of time it takes to heat the floor from 10°C, the initial condition, to 22°C, the setpoint temperature for four scenarios; a metal plate that is not enhanced with small spikes, a spike barely touching the upper floor plywood, a spike nailed down to the upper floor plywood, and a spike under the plywood.

Energy reduction shown in Table II was calculated using (1). E_f and E_{nf} denote energy consumption for the metal plate with spikes configuration and the metal plate without spikes — just a flat metal plate configuration.

$$\text{Energy Reduced (\%)} = \frac{E_f - E_{nf}}{E_{nf}} \times 100 \quad (1)$$

Table II. Optimization of metal plate with spikes

Scenario	Plate Material	Spike Length (in)	Setpoint Time (min)	Energy Reduction (%)
Without spikes	-	-	68	-
	Al	0.06	34	50
	C	0.09	39	43
	Al	0.06	30	56
	C	0.09	36	47
	Al	0.06	28	59
	C	0.09	33	51
	Al	0.09	25	63
	C	0.09	31	54
	Al	0.06	46	32
	C	0.09	52	23

It can be seen from Table II that an aluminum metal plate with spikes with 0.09" spike length is the optimal configuration.

For the optimal configuration, the temperature distributions across the air layer were simulated. Using an electrical heating element as heat source, temperature distributions after the air layer has reached the setpoint temperature were plotted as shown in Fig. 10 and Fig. 11, respectively. These results were plotted for a heat source emitting 10 W/ ft², which is the minimum amount of power used for radiant floor heating in industry [14-16].

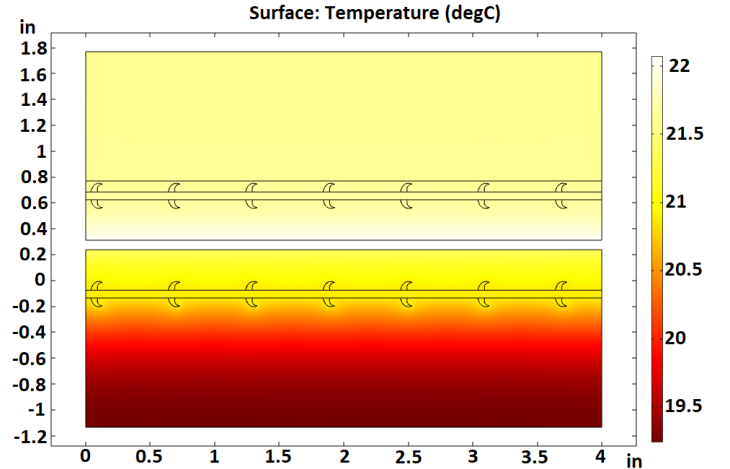


Figure 10. Temperature distribution parallel to the heat source

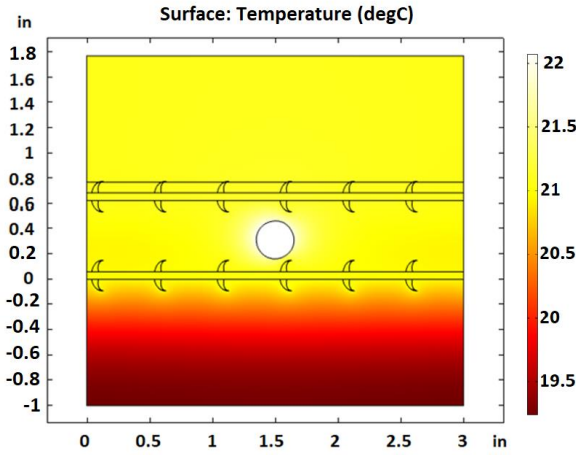


Figure 11. Temperature distribution perpendicular to the heat source

As can be seen from Fig. 10 and Fig. 11, parallel and perpendicular to the heat source, both, yield uniform temperature distributions and reach the 22°C setpoint temperature in under 28 minutes, as opposed to the without-spikes configuration which takes 68 minutes to reach the setpoint temperature.

In addition, the effect of having extra metal plate with spikes layers was observed. Consistent with expectations, the addition of extra layers resulted in extra time needed to reach the setpoint temperature. This can be explained from the fact that having an extra metal plate with spikes sandwiched in between concrete boards adds thermal mass to the system. This effect is illustrated in Fig. 12 and Fig. 13 for two different heat sources; a resistive element, and hot water pipes, respectively.

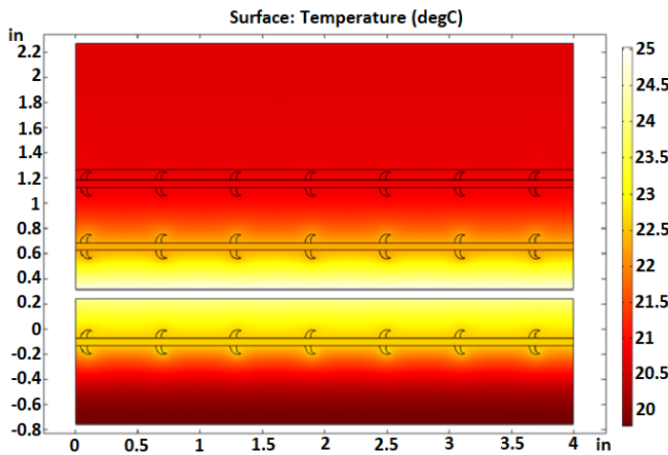


Figure 12. Temperature distribution associated with the addition of an extra plate with spikes (parallel to the heat source)

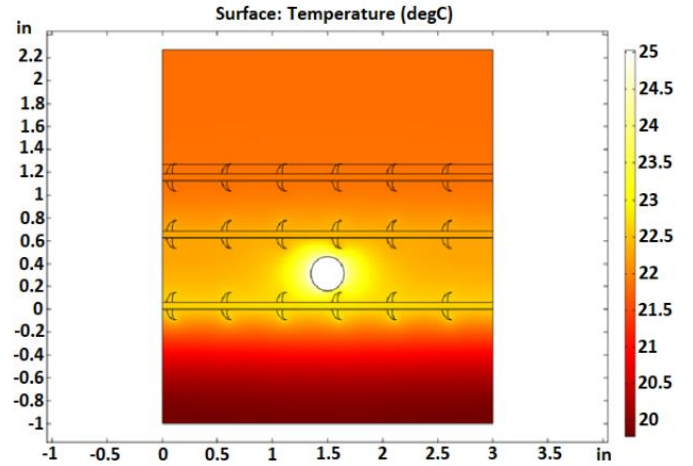


Figure 13. Temperature distribution associated with the addition of an extra plate with spikes (perpendicular to the heat source)

Here, surface temperatures for all configurations were in the ‘comfortable range’ of 19°C to 29°C which is recommended by ASHRAE and some European standards [9-11].

B. Radiant Floor Cooling

For the radiant floor cooling application in the basketball court, only the spike-under-the-surface scenario was considered. This is in line with construction methods of floors of basketball courts [2]. Hence, the optimal configuration had to be iteratively simulated only for different materials for metal plate with spikes. Consistent with a radiant floor heating layout, only two material candidates were considered: aluminum alloy and carbon steel. Since water of 15°C-18°C is to be used to cool the floor, an air to water heat pump was assumed to be used to supply water to the pipes. Table III shows the amount of time it takes to cool the air layer from the 30°C, initial condition to 24°C, cooling setpoint temperature for a coefficient of performance (COP) of the heat pump of 3, and under various cooling load scenarios. Similar to prior results for the radiant floor heating layout, it was determined that using an aluminum plate with small spikes yields the most efficient configuration.

Also, temperature distributions for the air layer were uniform and floor surface temperatures were not lower than the design dew point temperature of 16°C, or lower than 19°C, which are the ASHRAE and European standards’ recommendations for lowest floor surface temperature for radiant floor cooling systems [8-10].

Table III. Optimization of Metal plate with spikes

Cooling Load (W)	Power Consumption of Heat Pump (W)	Cooldown Time- Al ^a (min)	Cooldown Time- C ^a (min)
20	7	45	53
30	10	35	39
50	17	26	28
100	33	17	19

a. Al-Aluminum and C- Carbon Steel

Next, the optimal radiant panel configuration was simulated for an indoor basketball court environment. As mentioned earlier, only evaporative cooling has been accounted for at this stage in the project. A moist air transport module was added and 25 CFM/person flow rate was used, in accordance with ASHRAE Standard 62-1999 [7, 8]. An air velocity of 0.1 m/s was used. As mentioned in [2], this is most commonly used in large space buildings. The initial condition and temperature distribution once the setpoint temperature is reached are shown in Fig. 14.

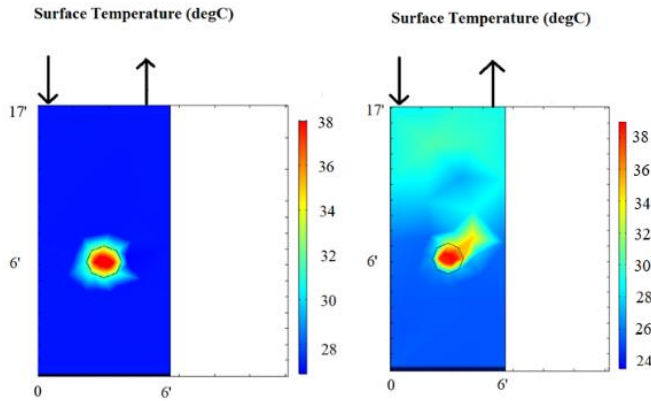


Figure 14. Evaporative cooling through a radiant floor. The image on the left represents the initial condition of a sphere starting at 38°C. The image on the right is a screen capture once the setpoint temperature of 24°C has been reached.

V. FUTURE WORK AND CONCLUSION

This novel study aimed at configuring the optimal radiant panel layout composed of a metal plate with spikes and flooring materials rendering it retrofittable and energy-efficient. Using an aluminum metal plate with 0.09” spikes nailed down to upper floor hardwood yielded an energy reduction of 63% compared to the base case scenario. Preliminary calculations show that such retrofittable systems have potential to save up to 30% energy in space heating and space cooling systems. It was observed that the addition of extra layers of a metal plate with spikes adds extra thermal mass to the system. A potential pathway of interest may also be to consider off-peak electricity during the night to start heating up the floor and carry out economic projections for its use in office buildings. For radiant floor cooling system, the influence of radiation and conduction still needs to be investigated and the work is in progress. The optimal layout does not lead to condensation and is expected to yield greater thermal comfort for basketball players.

ACKNOWLEDGEMENTS

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and Canada Research Chair (CRC) program for providing

financial support. The authors also acknowledge technical expertise provided by Dr. Wey H. Leong from Ryerson University.

REFERENCES

- [1] A. Frank and L. Wright, “Radiant Floor Heating In Theory and Practice,” no. July, pp. 19–24, 2002.
- [2] K. Zhao, X. H. Liu, and Y. Jiang, “Application of radiant floor cooling in large space buildings - A review,” *Renewable and Sustainable Energy Reviews*, 2016.
- [3] International Energy Agency, “Buildings.” [Online]. Available: <https://www.iea.org/buildings/>. [Accessed: 02-Dec-2018].
- [4] National Energy Board, “Provincial and Territorial Energy Profiles-Canada,” 2018.
- [5] S. Stinson, “Energy Efficiency in the Buildings Sector,” 2017.
- [6] Statistics Canada, “Survey of Commercial and Institutional Energy Use, 2014,” 2016.
- [7] National Energy Board, “Canada’s Renewable Power Landscape 2017- Energy Market Analysis,” 2017.
- [8] C. Workman, “Minister of Environmental and Climate Change Marks Canada’s Ratification of the Paris Agreement,” 2016.
- [9] R. D. Woodson, *Radiant Floor Heating*, 2nd ed. McGraw-Hill Education, 2010.
- [10] “International Organization for Standardization. Moderate thermal environments - Moderate thermal environments - Determination of the PMV and PPD index and specification of the conditions for thermal comfort, ISO 7730,” Geneva, 1994.
- [11] “ANSI/ASHRAE Standard 55-1992, Thermal Environmental Conditions for Human Occupancy.”
- [12] “CEN (1994) EN 1264: Floor Heating - Systems and Components. European Committee for Standardization.”
- [13] B. W. Olesen, “Radiant floor heating in theory and practice,” *ASHRAE J.*, vol. 44, no. 7, p. 19, 2002.
- [14] B. Olesen, “Radiant floor cooling systems,” *ASHRAE J.*, vol. 50, no. 9, p. 16, 2008.
- [15] L. Zhang and B. Cao, “Design and Analysis of a Floor Radiant Heating System Based on Energy Substitution Technology,” *Appl. Sci.*, vol. 8, no. 4, 2018.
- [16] M. Izquierdo and P. de Agustín-Camacho, “Solar heating by radiant floor: Experimental results and emission reduction obtained with a micro photovoltaic-heat pump system,” *Appl. Energy*, vol. 147, pp. 297–307, Jun. 2015.
- [17] G. Zhou and J. He, “Thermal performance of a radiant floor heating system with different heat storage materials and heating pipes,” *Appl. Energy*, 2015.
- [18] D. Zhang, N. Cai, and Z. Wang, “Experimental and numerical analysis of lightweight radiant floor heating system,” *Energy Build.*, 2013.

- [19] N. Ghaddar, M. Salam, and K. Ghali, "Steady thermal comfort by radiant heat transfer: The impact of the heater position," *Heat Transfer Engineering*. 2006.
- [20] T. Sakoi, K. Tsuzuki, S. Kato, R. Ooka, D. Song, and S. Zhu, "Thermal comfort, skin temperature distribution, and sensible heat loss distribution in the sitting posture in various asymmetric radiant fields," *Build. Environ.*, 2007.
- [21] Wang, Z., Zhang, H., Arens, E., Lehrer, D., Huizenga, C., Yu, T., Hoffman, S., "Modelling thermal comfort with radiant floors and ceilings. In: Indoor Environmental Quality," 2009.
- [22] B. W. Olesen, "Possibilities and limitations of radiant floor cooling," in *ASHRAE Transactions*, 1997.
- [23] P. Simmonds, S. Holst, S. Reuss, and W. Gaw, "Using radiant cooled floors to condition large spaces and maintain comfort conditions," in *ASHRAE Transactions*, 2000.
- [24] H. E. Feustel and C. Stetiu, "Hydronic radiant cooling - preliminary assessment," *Energy Build.*, 1995.
- [25] C. Stetiu, "Energy and peak power savings potential of radiant cooling systems in US commercial buildings," *Energy Build.*, 1999.
- [26] "Radiant Floor System, NUCAP Industries Inc." .
- [27] V. Gerlich, K. Sulovská, and M. Zálešák, "COMSOL Multiphysics validation as simulation software for heat transfer calculation in buildings: Building simulation software validation," *Meas. J. Int. Meas. Confed.*, 2013.

ENERGY AND EXERGY ANALYSIS OF A COMBINED BRAYTON/BRAYTON POWER CYCLE WITH HUMIDIFICATION

AK Mossi Idrissa

Department of Mechanical and Aerospace Engineering
Royal Military College of Canada
Kingston, Canada
Abdoul.Mossi-Idrissa@rmc.ca

K Goni Boulama

Department of Mechanical and Aerospace Engineering
Royal Military College of Canada
Kingston, Canada
Goni.Boulama@rmc.ca

Abstract— First and second law approaches have been used to analyze the performance of a humidified Brayton/Brayton power cycle. The energy efficiency and exergy destruction rates consistently improved when the combustion temperature was increased. Both performance indicators improved, reached an optimum, and then deteriorated when the topping cycle pressure ratio increased, while their sensitivity to the bottoming cycle pressure ratio depended on the humidification rate used at the bottoming cycle. Upon increasing the mass flowrate of air through the bottoming cycle, the energy efficiency of the power cycle increased linearly, while the irreversibility generation had a non-monotonic variation. In all cases, a higher degree of humidification always resulted in greater first and second law performances.

Combined gas power cycles; Humidification; Energy; Exergy.

I. INTRODUCTION

Improving the performance of energy conversion devices has been the focus of significant research and development work for decades. Brayton cycle based gas power plants have received a particular attention because of their essential role in the world electricity generation infrastructure. Ibrahim et al. [1] identified three main directions for efficiency improvement: increasing the turbine inlet temperature, increasing the efficiencies of individual cycle components, and adding some modifications to the basic cycle. One such modification consists in routing the warm gases leaving the expander so that the residual thermal energy is used to preheat the compressed air before it enters the combustion chamber. In some cases, this was shown to result in a twofold increase of the efficiency [2, 3]. Cooling the working fluid during the compression process [4, 5] or reheating it between successive expansion stages are other modifications proposed by other authors [6], although in some cases the latter can result in a poorer energy efficiency and increased electricity costs compared to the basic cycle [7]. In general, the power plant efficiency was shown to increase, reach a maximum and then decrease when the pressure ratio was increased, while it consistently increased when the combustion temperature was increased [2, 6, 7].

Besides these relatively minor modifications to the basic Brayton cycle, combined power cycles in which the thermal energy of the exhaust stream from the expander is used to drive a bottoming power cycle have been proposed. Brayton/Rankine and Brayton/Brayton combinations are now well-established technologies [8, 9, 10]. Bolland et al. [9] showed that the Brayton/Brayton power cycle is economically viable and allows an increase of the thermal efficiency by more than 10 percentage points compared to the base cycle, while the fuel consumption and specific power output potential are improved in even greater proportions. Other authors proposed other combined cycle configurations, including the cases where the bottoming cycle is an inverse Brayton cycle [11, 12], a Stirling engine [13], or an Ericsson cycle [14].

On the other hand, many studies have demonstrated that humidified gas turbine cycles can have a greater power output potential, at a lower cost than a cycle re-design as discussed in the previous paragraphs [15]. The humidification could be achieved by water spray or steam injection before [16], during [17] or after [18, 19] the compression process, or right at the combustion chamber [8, 20].

Saghafifar and Gadalla [21] compared a standard Brayton/Brayton cycle to a Brayton cycle to which a humidified gas cycle was attached. A sensitivity analysis was conducted for the combustion temperature, bottoming cycle pressure ratio, mass flowrate ratio between the topping and bottoming cycles, and degree of humidification. Under their respective optimized operation conditions, the two power plant arrangements differed by less than 4% in terms of energy efficiency and specific power output. More recently, the present authors investigated a similar combined power plant with different degrees of humidification at the bottoming cycle [22]. Using the first law approach, significant beneficial effects of the humidification were illustrated. To our knowledge, these are the only two papers available to date on humidified Brayton/Brayton power cycles. The present paper proposes to extend the previous study by adopting the second law approach. Amongst other results, the sensitivity of the component based and overall exergy destruction with the humidification ratio will be discussed.

$$m_g = m_a^T + m_f \quad (5)$$

$$m_m = m_a^B + m_w \quad (6)$$

$$m_a^T h_2 + m_f \text{LHV} = m_g h_3 \quad (7)$$

$$W_{\text{TCC}} = m_a^T \quad w_{\text{TCC}} = m_a^T (h_1 - h_2) \quad (1)$$

$$W_{\text{BCC}} = m_a^{\text{B}} w_{\text{BCC}} = m_a^{\text{B}} (h_6 - h_7) \quad (2)$$

$$Q_{\text{in}} = m_{\text{f}} \text{LHV} \quad (8)$$

$$\eta_{\text{HEX}} = (T_8 - T_9) / (T_8 - T_4) \quad (9)$$

$$W_{\text{TCT}} = m_g w_{\text{TCT}} = m_g (h_3 - h_4) \quad (3)$$

$$W_{\text{BCT}} = m_{\text{m}} w_{\text{BCT}} = m_{\text{m}} (h_9 - h_{10}) \quad (4)$$

$$m_{w,\text{sat}} = n_{w,\text{sat}} M_w = m_a^B (M_w/M_a) P_{\text{sat}}(T_8)/(P_7 + P_{\text{sat}}(T_8)) \quad (11)$$

$$W_{\text{pump}} = (m_{\text{w,sat}} / \rho) (P_{11} - P_{12}) / \eta_{\text{pump}} \quad (12)$$

$$\alpha = m_w / m_{w,\text{sat}} \quad (13)$$

In this case, the pumping energy requirement (12) will use the actual water mass flowrate, m_w , instead of the maximum amount corresponding to the saturation conditions.

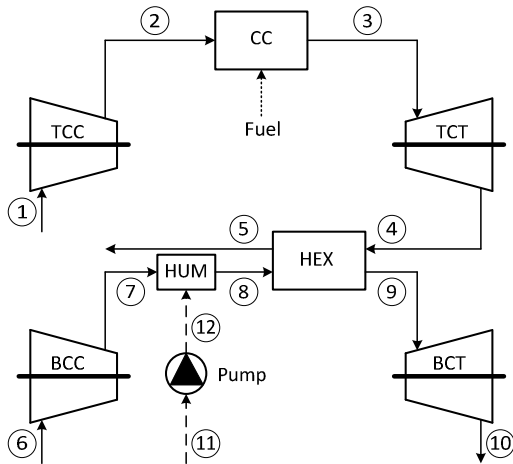


Figure 1. Combined Brayton/Brayton power cycle with humidification

Finally, the first law efficiency of the combined power cycle is defined as the ratio between the net power output of the cycle and the input thermal energy:

$$\eta_{\text{net}} = W_{\text{net}} / Q_{\text{in}} \quad (14)$$

For the purpose of the exergy analysis, the total exergy destruction of a component k is defined as the difference between the exergy of the fuel and that of the product:

$$E_{D,k} = E_{f,k} - E_{p,k} \quad (15)$$

The fuel refers to the main exergy input to a component, while the product is the useful exergy output from the component. When calculating the exergy of individual streams, both the physical and chemical exergies are accounted for:

$$E_i = E_i^{\text{PH}} + E_i^{\text{CH}} \quad (16)$$

The fuel exergy is calculated using its exergy grade, or the exergy to heating value ratio, which is taken from [26]. The chemical exergy e^{CH} of the combustion products and moist air at the exit of the humidifier is calculated in a similar fashion as in [27, 28].

Equations 1 – 16 are integrated in a computer code which calculates the air properties at all points of the cycle, the fuel consumption and water usage at humidifier, the compressor and pump power requirements, the turbine power outputs, and the heat transfer rate at the recovery heat exchanger. These are then used to calculate the overall power generation and first and second law performances of the power cycle.

IV. RESULTS AND DISCUSSION

Before processing the results, the validity of the model and its implementation has been confirmed through a series of tests. First, fluid property calculation subroutines were validated against data published in the IAPWS reports [24, 25] and other thermodynamics textbooks. Simplified versions of the code were also used to simulate the operation of single gas turbine cycles and combined Brayton/Brayton cycles, with and without humidification [22].

In the following, the first and second law performances of the combined power cycle are presented and their sensitivity to variations of select operating conditions is discussed. The base values and variation ranges for these operating parameters are given in Table 1, noting that the exergy destruction data have been calculated for a net power output of 100 MW. The results are conveniently presented so that the effect of the humidification is apparent in all graphs.

Using the base values of the operating parameters, the thermal efficiency of the combined power plant is about 36%, an approximately 12% increase compared to a single gas turbine cycle, while the exergy efficiency is of about 51%.

According to (7), for a given value of the topping cycle pressure ratio, the enthalpy at the exit of the compressor is

TABLE I. VARIATION RANGES FOR OPERATING PARAMETERS

Parameters	Symbol	Base value	Variation range
Topping cycle pressure ratio	TCPR	14	1.5 – 30
Bottoming cycle pressure ratio	BCPR	4	1.5 – 14
Fuel lower heating value	LHV	50,050 kJ/kg	
Turbine inlet temperature	TIT	1400 K	1000 – 1600 K
Turbine & compressor isentropic efficiencies	$\eta_{\text{TCT}}, \eta_{\text{TCC}}, \eta_{\text{BCT}}, \eta_{\text{BCC}}$	0.85	
Bottoming to topping cycle mass flowrate ratio	β	0.5	0 – 1.6
Water pump efficiency	η_{pump}	0.9	
Humidification ratio	α	0.5	0 – 1
Recovery heat exchanger effectiveness	η_{HEX}	0.8	0.2 – 1
Reference pressure	P_{ref}	101.325 kPa	
Reference temperature	T_{ref}	298.15 K	

constant, and an increase of the combustion temperature is obtained by increasing the fuel to air flowrate ratio. The increase of TIT means a higher quality energy at the inlet of the turbine and hence a greater power generation potential. This also implies that more thermal energy is available for driving the bottoming cycle. As a result, the overall power generation potential increases significantly, with Fig. 2 showing that the overall plant energy efficiency initially increases sharply, and then the increase rate gradually decreases. In the absence of humidification, the overall energy efficiency starts at 17.6% when TIT = 1000 K and it is doubled when TIT = 1600 K. On the other hand, when the air in the bottoming cycle is saturated before entering the recovery heat exchanger, $\eta_{\text{net}} = 22.5\%$ when TIT = 1000 K and it reaches $\eta_{\text{net}} = 37.4\%$ when TIT = 1600 K, which reveals a smaller effect of the humidification ratio on the overall energy efficiency when the combustion temperature is high.

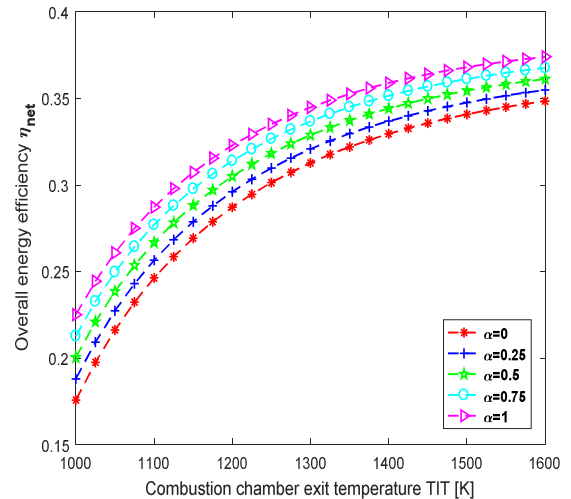


Figure 2. Effect of the combustion temperature on the energy efficiency

Fig. 3 shows that the exergy destruction initially decreases very sharply, and then the decrease rate gradually smoothens out at high TIT values. In total, the exergy destruction is reduced by a factor of 2/3 when TIT is varied between 1000 K and 1600 K. The small effect of the humidification ratio at high TIT is consistent with that observed in Fig. 2.

When the pressure ratio is increased, research on single gas turbine cycles predicts that the power generation initially increases, reaches a maximum value and then decreases [2, 6, 18, 20]. It has been observed in the present study that the overall power generation potential of the combined Brayton/Brayton cycle also initially increases, reaches a maximum value at about TCPR = 7.5, and then gradually decreases when TCPR is further increased, consistently with results in published studies on combined non-humidified [11, 12] and humidified [22] gas cycles. It is noted that according to (7), in order to maintain the combustion chamber exit temperature constant at TIT = 1400 K while TCPR is varied, the fuel to air flowrate ratio has to be reduced. In the meantime, less thermal energy is available to drive the bottoming cycle, and Fig. 4 shows that η_{net} increases, reaches a maximum value at around TCPR = 19.5, and then gradually decreases. It is noted that, because the variation of η_{net} is small over the TCPR = 11 to 30 range, operating the plant at the lower end of this interval is more practical, as this will correspond to reduced mechanical stress on the equipment, i.e. a reduced cost. Finally, Fig. 4 shows that the humidification ratio has an essentially constant effect on the first law efficiency: about 10% improvement between the un-humidified case and the case where the moist air exiting the humidifier is fully saturated, regardless of the operating value of TCPR.

Fig. 5 shows that the total exergy destruction is extremely high when TCPR is small, and it quickly decreases as TCPR increases. At about TCPR = 15, the exergy destruction is minimal, and it starts gradually increasing again as TCPR is further increased. The variation of the exergy efficiency could be easily inferred: a sharp increase, a maximum value at about TCPR = 15, and then a gradual decrease.

Similar trends for the total exergy destruction and overall exergy efficiency were predicted for a Brayton/Brayton cycle without humidification [29]. In fact, Fig. 5 also shows that humidification only marginally affects the exergy destruction, although the second law performance is consistently greater when the humidification ratio is high.

When the bottoming cycle pressure ratio is varied without changing the topping cycle operation, Fig. 6 shows that the overall energy efficiency increases, reaches a maximum, and then decreases. The maximum value reached increases with increasing humidification ratios, and the value of BCPR corresponding to the maximum power output shifts to the right when α increases. For a given value of α , the increase of BCPR also implies an increase of the water flowrate through the bottoming cycle turbine. From a practical point of view, this suggests that when the bottoming cycle is operated with a low pressure ratio, humidification effects are small.

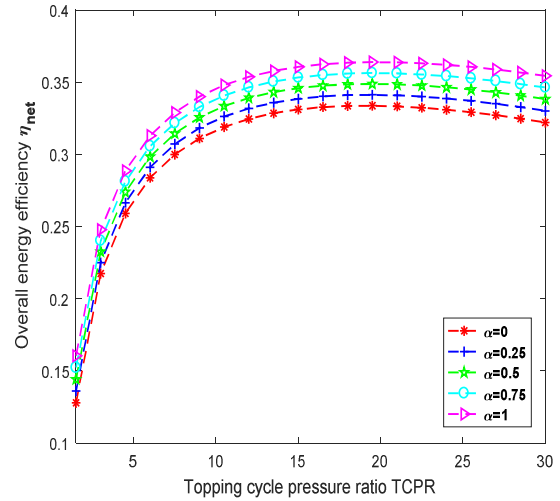


Figure 4. Effect of the topping cycle pressure ratio on the energy efficiency

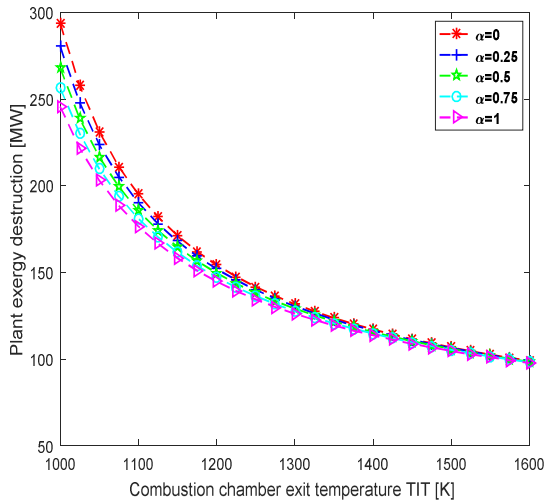


Figure 3. Effect of the combustion temperature on the exergy destruction

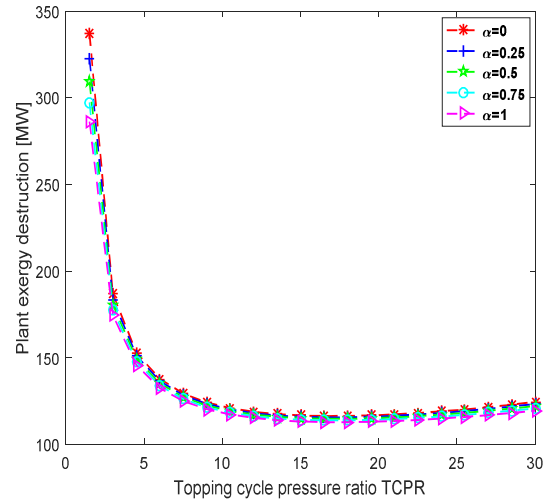


Figure 5. Effect of the topping cycle pressure ratio on the exergy destruction

However, when BCPR is high, humidification has a dramatic effect on the plant performance. In terms of specific power output, when $BCPR = 14$, $w_{net} = 268$ kJ/kg when $\alpha = 0$, while $w_{net} = 368$ kJ/kg when $\alpha = 1$. This could justify efforts to humidify the bottoming cycle to even greater extents than possible with the current plant layout. Because the energy consumption is constant throughout these simulations, the energy efficiency η_{net} has the same variations as the specific work output.

Fig. 7 shows the effects of the bottoming cycle pressure ratio on the total exergy destruction. For the un-humidified cycle, the irreversibility generation initially decreases when BCPR increases, it reaches a minimum value at a BCPR value close to 3, and then it increases in a seemingly linear fashion as BCPR is further increased. The same behaviour is reproduced when α is equal to 0.25 and 0.5, with the optimum BCPR value gradually shifting to the right. For the two higher

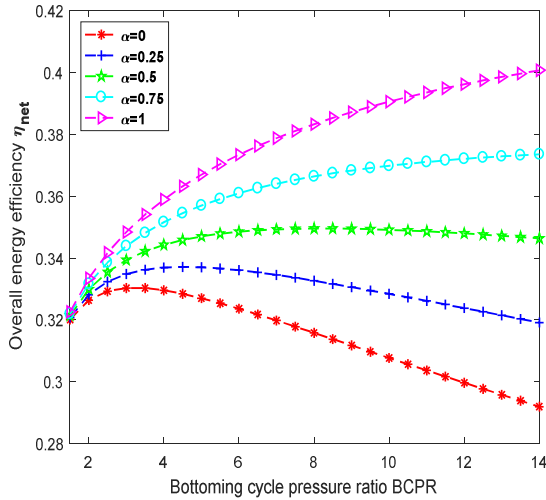


Figure 6. Effect of the bottoming cycle pressure ratio on the energy efficiency

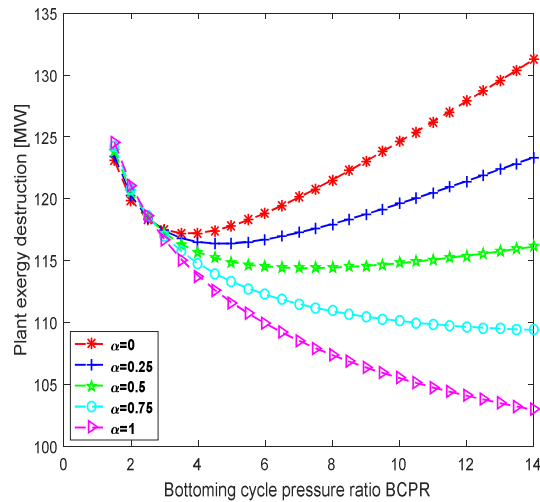


Figure 7. Effect of the bottoming cycle pressure ratio on exergy destruction

humidification ratio cases considered in this study, the cycle's irreversibility generation decreases over the entire BCPR variation range, although the decrease rate is smaller at high BCPR. Fig. 7 also shows that the sensitivity to α is small at low BCPR, but becomes increasingly important as BCPR increases. At the upper end of the investigated BCPR variation range, the un-humidified cycle generates 27% more irreversibility than the cycle with $\alpha = 1$.

When the air mass flowrate ratio β between the bottoming and topping cycles is varied while all other plant operation parameters are maintained constant at their base values listed in Table 1, the power generation by the topping cycle is constant, and a linearly increasing use of the heat available at the exit of the topping cycle turbine is made to generate additional power at the bottoming cycle. Furthermore, when the air leaving the bottoming cycle compressor is humidified before the expansion, an even greater shaft power is generated at the turbine for the same compression cost, and the power production by the entire cycle increases even faster. During these tests, the fuel consumption at the topping cycle is also constant, which means that the energy efficiency of the combined cycle (Fig. 8) varies in the same way as the specific power generation. These variations are consistent with the initial variations of the specific power generation and energy efficiency calculated in [21]. That being said, the variations in Fig. 8 are only valid for the considered combination of operating parameters. Changing the combustion chamber exit temperature from $TIT = 1400$ K to $TIT = 1100$ K, the variation of the overall cycle efficiency with β remains linear, but the slope is negative when the bottoming cycle is not humidified, and it quickly becomes positive and increases when α is increased. Similarly, changing the bottoming cycle pressure ratio from $BCPR = 4$ to $BCPR = 14$, results in a linear decrease of the cycle overall efficiency when β increases for the non-humidified case, and the slope becomes positive and increases as α increases. It is obvious that the combined power cycle should not be operated when it returns an efficiency that is poorer than that of a single gas turbine.

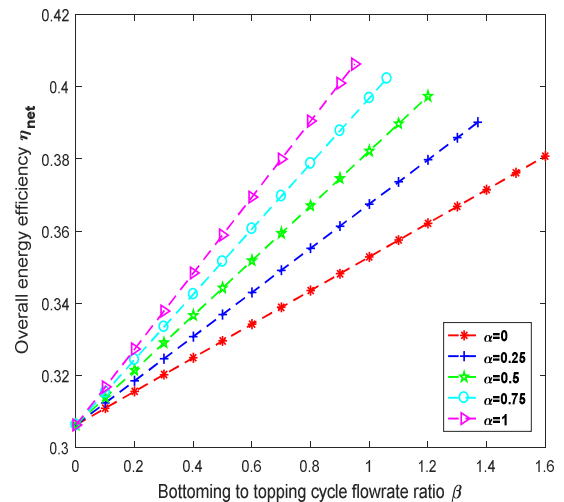


Figure 8. Effect of the mass flowrate ratio on the energy efficiency

The variation of the overall power plant exergy destruction with the bottoming cycle working fluid flowrate is illustrated in Fig. 9. As β increases, it is seen that the exergy destruction first increases, reaches a maximum value, and then it reverts variation, decreasing at a more and more pronounced rate. The initial increase of the exergy destruction is attributable to the irreversibility generated in the recovery heat exchanger and bottoming cycle turbine and compressor; these three components are absent for the $\beta = 0$ case. However, as the bottoming cycle working fluid flowrate increases, and increasing share of the exergy that would be otherwise destroyed by the topping cycle is recovered and converted into useful work by the bottoming cycle. As far as the humidification ratio is concerned, Fig. 9 shows that its effect is small at low β values, which is expected, and becomes increasingly pronounced when β increases. The un-humidified ($\alpha = 0$) cycle is always the one with the largest irreversibility generation, and the second law performance consistently improves when α increases. The value of β corresponding to the worst second law performance shifts to the left as α increases. Finally, it is worth noting that the working fluid mass flowrate at the bottoming cycle can only be varied within a certain range, corresponding to the availability of waste heat rejected by the topping cycle. This explains why the different curves in Figs. 8 and 9 have different end points.

Fig. 10 shows that the total power plant exergy destruction gradually decreases when the humidity ratio increases, dropping from $E_D = 118$ MW when $\alpha = 0$ to $E_D = 115$ MW when $\alpha = 1$. The combustion chamber is the component with the largest exergy destruction, which is unsurprising because combustion is a typical irreversible process due to large temperature and composition non-uniformities. The irreversibility generation at the combustion chamber has a similar variation as the total power plant exergy destruction. The topping cycle turbine and compressor are the next largest irreversibility generation components, and their irreversibility generation seems to be independent of the humidification ratio at the bottoming cycle.

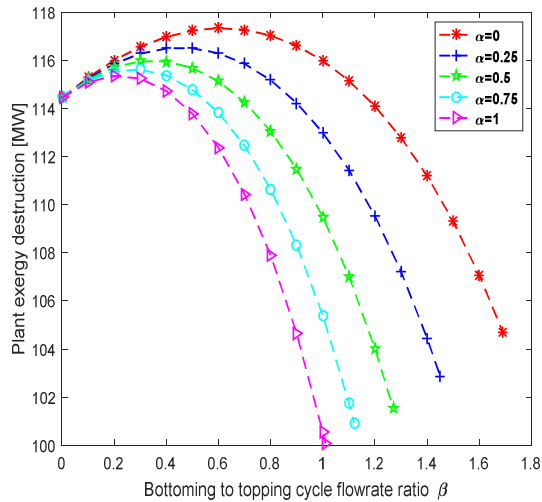


Figure 9. Effect of the mass flowrate ratio on the exergy destruction

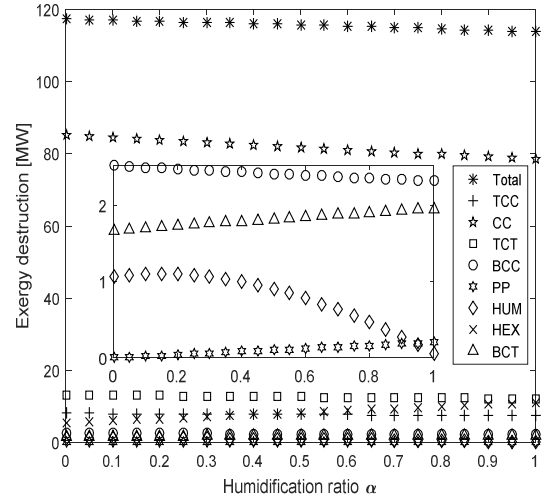


Figure 10. Effect of the humidification ratio on the exergy destruction

The recovery heat exchanger is the fourth least performing component when α is small, but $E_{D,HEX}$ increases in a quasi linear fashion and, at about $\alpha = 0.4$, the recovery heat exchanger is outperformed by the topping cycle compressor and becomes the third least performing component of the cycle. The exergy destruction is of about the same magnitude at the humidifier and bottoming cycle compressor and turbine, but it is differently affected by α . The pump is the best performing component due to the low flowrate of water, high isentropic efficiency and small temperature change.

V. CONCLUSION

A combined Brayton/Brayton power cycle with humidification at the bottoming cycle has been investigated using first and second law approaches. Both the cycle energy efficiency and exergy destruction rates consistently improved when the combustion temperature increased, though the performance gain reduced at high values of the combustion temperature. Both performance indicators used in this paper also improved, reached an optimum value, and then slightly deteriorated when the topping cycle pressure ratio increased. The energy efficiency and exergy destruction rates displayed a different sensitivity to the bottoming cycle pressure ratio depending on the humidification ratio at the bottoming cycle. In fact, for low values of the humidification ratio, the energy efficiency and exergy destruction rates initially improved, hit an optimum, and then decreased when the bottoming cycle pressure ratio was increased, while both performance indicators consistently deteriorated when the bottoming cycle pressure ratio was increased for high values of the humidification ratio. The energy efficiency of the combined power cycle linearly increased when the ratio between the bottoming and topping cycle mass flowrates increased. However, the irreversibility generation was shown to first increase until a maximum is reached, after which it quickly decreased. In all cases, humidifying the bottoming cycle was shown to be an efficient way of improving both the first and second law performances.

ACKNOWLEDGMENT

The financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Defense Academic Research Program (CDARP) is acknowledged.

REFERENCES

- [1] T.K. Ibrahim, F. Basrawi, O.I. Awad, A.N. Abdullah, G. Najafi, R. Mamat, and F.Y. Hagos, "Thermal performance of gas turbine power plant based on exergy," *Applied Thermal Engineering*, vol. 115, pp. 977–985, 2017.
- [2] C. Wu, L. Chen, and F. Sun, "Performance of a regenerative Brayton heat engine," *Energy*, vol. 21, pp. 71–76, 1996.
- [3] M. Goodarzi, "Comparative energy analysis on a new regenerative Brayton cycle," *Energy Conversion and Management*, vol. 120, pp. 25–31, 2016.
- [4] R. Bhargava and A. Peretto, "A unique approach for thermodynamic optimization of intercooled, reheat, and recuperated gas turbine for cogeneration applications," *Journal of Engineering for Gas Turbines and Power*, vol. 124, pp. 881–891, 2002.
- [5] Y. Haseli, "Efficiency of irreversible Brayton cycles at minimum entropy generation," *Applied Mathematical Modelling*, vol. 40, pp. 8366–8376, 2016.
- [6] A. Tripathi, M. Dubey, H. Chandra, and A. Kumar, "Thermodynamics performance analysis of reversible reheat Joule-Brayton cycle with cogeneration system," *International Journal of Energy Engineering*, vol. 6, pp. 1–6, 2016.
- [7] M.K. Sahu and N. Sanjay, "Comparative exergoeconomic analysis of basic and reheat gas turbine with air film blade cooling," *Energy*, vol. 132, pp. 100–170, 2017.
- [8] K. Jesionek, A. Chrzczonowski, P. Ziolkowski, and J. Badur, "Power enhancement of the Brayton cycle by steam utilization," *Archives of Thermodynamics*, vol. 33, pp. 39–50, 2012.
- [9] O. Bolland, M. Forde, and B. Hande, "Air botoming cycle: Use of gas turbine waste heat for power generation," *Journal of Engineering for turbines and power*, vol. 118, pp. 359–366, 1996.
- [10] M. Ghazikhani, I. Khazaei, and E. Abdekhoodaie, "Exergy analysis of gas turbine with air bottoming cycle," *Energy*, vol. 72, pp. 599–607, 2014.
- [11] B. Agnew, A. Anderson, I. Potts, T.H. Frost, and M.A. Alabdoadaim, "Simulation of combined Brayton and inverse Brayton cycles," *Applied Thermal Engineering*, vol. 23, pp. 953–963, 2003.
- [12] W. Zhang, L. Chen, and F. Sun, "Power and efficiency optimization of combined Brayton and inverse Brayton cycles," *Applied Thermal Engineering*, vol. 29, pp. 2885–2894, 2009.
- [13] A. Entezari, A. Manizadeh, and R. Ahmadi, "Energetical, exergetical and economical optimization analysis of combined power generation system of gas turbine and Stirling engine," *Energy Conversion and Management*, vol. 159, pp. 189–203, 2018.
- [14] J. Zheng, F. Sun, L. Chen, and C. Wu, "Exergy analysis for a Braysson cycle," *Exergy*, vol. 1, pp. 41–45, 2001.
- [15] M. Jonsson and J. Yan, "Humidified gas turbines – a review of proposed and implemented cycle," *Energy*, vol. 30, pp. 1013–1078, 2005.
- [16] A. Omar, M. Saghaififar, and M. gadalla, "Thermo-economic analysis of air saturator integration in conventional combined power cycles," *Applied Thermal Engineering*, vol. 107, pp. 1104–1122, 2016.
- [17] K.H. Kim and H.J. Ko, "Exergy analysis of wet-compression gas turbine cycle with recuperator and turbine blade cooling," *Journal of Automation and Control Engineering*, vol. 1, pp. 140–143, 2013.
- [18] P. Jenkins, M. Cerza, and M. Al Said, "Analysis of using the M-cycle regenerative-humidification process on a gas turbine," *Journal of Energy and Power Engineering*, vol. 8, pp. 1824–1837, 2014.
- [19] M. Saghaififar and M. Gadalla, "Analysis of Maisotsenko open gas turbine power cycle with a detailed air saturator model," *Applied Energy*, vol. 149, pp. 338–353, 2015.
- [20] H.K. Kayadelen and Y. Ust, "Performance and environment as objectives in multi-criterion optimization of steam injected gas turbine cycles," *Applied Thermal Engineering*, vol. 71, pp. 184–196, 2014.
- [21] M. Saghaififar and M. Gadalla, "Analysis of Maisotsenko open gas turbine bottoming cycle," *Applied Thermal Engineering*, vol. 82, pp. 351–359, 2015.
- [22] A.K. Mossi Idrissa and K. Goni Boulama, "Investigation of the performance of a combined Brayton/Brayton cycle with humidification," *Energy*, vol. 141, pp. 492–505, 2017.
- [23] B.L. McBride, M.J. Zehe, and S. Gordon, "Coefficients for calculating thermodynamic properties of individual species," NASA Glenn Research Centre, Cleveland, OH, Report no. NASA/TP-2002-211556, 2002.
- [24] The International Association for the Properties of Water and Steam: "Revised Supplementary Release on Saturation Properties of Ordinary Water Substance," St. Petersburg, Russia, 1992.
- [25] The International Association for the Properties of Water and Steam: "Revised Release on the IAPWS Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam," Lucerne, Switzerland, 2007.
- [26] J.P. Szybist, K. Chakravathy, and C.S. Daw, "Analysis of the impact of selected fuel thermochemical properties on internal combustion engine efficiency," *Energy Fuels*, vol. 26, pp. 2798–2810, 2012.
- [27] H. Nami and E. Akrami, "Analysis of gas turbine based hybrid system by utilizing energy, exergy and exergoeconomic methodologies for steam, power and hydrogen production," *Energy Conversion Management*, vol. 143, pp. 326–337, 2017.
- [28] M. Fallah, H. Siyahi, R. Akbarpour Ghiasi, S.M.S. Mahmoudi, M. Yari, and M.A. Rosen, "Comparison of different gas turbine cycles and advanced exergy analysis of the most effective," *Energy*, vol. 116, pp. 701–715, 2016.
- [29] A.K. Mossi Idrissa and K. Goni Boulama, "Advanced exergy analysis of a combined Brayton/Brayton power cycle," *Energy*, vol. 166, pp. 724–737, 2019.
- [30] YA Cengel and MA Boles (2015): *Thermodynamics: An Engineering Approach*. 8th ed. McGraw-Hill, New York.

Design and Modeling of MRF Impact Base Frequency Enhancement for a Piezoelectric Energy Generator

Sylvester Sedem Djokoto

*Dept. of Automotive, Mechanical and Manufacturing
Engineering, Faculty of Engineering and Applied Science
University of Ontario Institute of Technology 2000 Simcoe
Street, North, Oshawa, Canada ON L1H 7K4
sylvester.djokoto@uoit.ca*

Hayford Azangbebil

*Dept. of Automotive, Mechanical and Manufacturing
Engineering, Faculty of Engineering and Applied Science
University of Ontario Institute of Technology 2000 Simcoe
Street, North, Oshawa, Canada ON L1H 7K4
Hayford.Azangbebil@uoit.ca*

Martin Agelin-Chaab

*Dept. of Automotive, Mechanical and Manufacturing
Engineering, Faculty of Engineering and Applied Science
University of Ontario Institute of Technology 2000 Simcoe
Street, North, Oshawa, Canada ON L1H 7K4
martin.agelin-chaab@uoit.ca*

Egidijus Dragašius

*Department of Production Engineering, Kaunas University of
Technology, Studentu 56, 51424,
Kaunas, Lithuania
egidijus.dragasius@ktu.lt*

Vytautas Jūrėnas

*Institute of Mechatronics, Kaunas University of Technology,
Studentu 56, 51424
Kaunas, Lithuania
vytautas.jurenas@ktu.lt*

Abstract— This paper proposed a new concept using magnetorheological fluids (MRF) as an impacting object for frequency tuning as well as power enhancement for a vibrating piezoelectric bimorph cantilever beam. MRF is a smart fluid that changes from a liquid state into semi-solid state within milliseconds when a magnetic field is applied. The paper presented a computed results compared to the FEM results. The FEM analysis was done using Comsol Multiphysics tool. The computed results show that the frequency of the vibrational system was increased by 15 percent when the MRF was activated at 0.1T compared to the free vibrating beam. The Comsol showed 32 percent of an impacting MRF with 0.1T magnetic field application compared to the free vibrating piezoelectric bimorph cantilever beam. The computed power output for the impact MRF with 0.1T gave a maximum power of $282\mu W$ and a matching resistance of $8.9k\Omega$. The FEM results on the hand showed a maximum power output of $95\mu W$ and a matching resistance of $8.8k\Omega$ for the impacting MRF with 0.1T magnetic application

Keywords- frequency tuning; MRF, Piezoelectric energy generator;

I. INTRODUCTION

The generation of power in micro systems have gained considerable research attention. This is due to the development of micro-electromechanical systems (MEMS) and wireless technologies. It is known that batteries are used to power these wireless systems. There have been a lot of studies and research into the development of an alternative, renewable, sustainable and environmentally friendly form of power generator to replace batteries for large networks and small devices. The frequency generated from vibrating piezoelectric objects like cantilever beams offers a good source of energy for conversion into usable electrical power.

Literature form recent and previous research showed that piezoelectric materials are one of the three general vibration-to-electric energy conversion mechanisms. The other two form of electric conversion are electrostatic and electromagnetic transduction [1]. Piezoelectric transducers are mostly used for energy harvesting because the energy density is three times higher compared to electrostatic and electromagnetic transduction [2].

The drawback in generating power from vibrating piezoelectric beam is the necessary frequency needed to match the maximum power generated compared with the excitation frequency from the ambient vibration sources. The change in the

excitation frequency and electric load influences frequency matching [3].

One of technique and methods employed by engineers and researchers in improving the efficiency of the piezoelectric energy harvesters is an impact vibration tuning method. An impact-based frequency has improved the output power in the low frequency applications environment as reported in several publications [4-7]. The impact vibration tuning devices uses different kinds of materials as their impacting objects. A ball moving freely between two dielectric membranes located at a certain distance from each other was proposed by [8]. The proposed system generated electricity when the ball moving due to ambient vibrations impacts on one of the membranes. The research results were useful for selecting the system parameters in achieving an optimal output performance. A potential application of the proposed system for energy harvesting from car engine vibrations. A frequency increased piezoelectric vibration energy harvesting device where the low-frequency periodic impact of a driving beam with a horizontally extended rectangular tip makes two piezoelectric generating beams to vibrate at the same time was proposed by [9].

The proposed technique presented in this article, is a new form of an impact-based frequency enhancement using vibrating bimorph piezoelectric energy generator and introducing a novel typing of impacting object which is an activated magnetorheological fluid (MRF). The MRF changes from fluid state to semi-fluid state when being activated due to its rheological properties. The goal is to therefore design an energy generator with the capabilities of producing up to $500\mu W$ using varying magnetic field application to the MRF.

The system was modeled as Single Degree of Freedom (SDOF) with the vibrating cantilever in first mode of vibration. The model was done by hand computation with Matlab assistance and compared to the FEM analysis using Comsol Multiphysics 5a. FEA tool.

The computed results show that the frequency of the vibrational system was increased by 15 percent when the MRF was activated at 0.1T compared to the free vibrating beam. The Comsol showed 32 percent of an impacting MRF with 0.1T magnetic field application compared to the free vibrating piezoelectric bimorph cantilever beam. The computed power output for the impact MRF with 0.1T gave a maximum power of $282\mu W$ and a matching resistance of $8.9k\Omega$. The FEM results on the hand showed a maximum power output of $95\mu W$ and a matching resistance of $8.8k\Omega$ for the impacting MRF with 0.1T magnetic application.

II. MODELING OF THE DEVICE

A. Device Framework

The proposed piezoelectric bimorph energy generator and mode of operation is shown in Fig.1. The system consists of vibrating piezoelectric bimorph cantilever beam, two permanent magnets mounted on the opposite sides of the beam. The MRF as an impacting element attached to one of the magnets. The permanent magnets are arranged to face each other with opposite poles and a gap d apart in the horizontal x-plane.

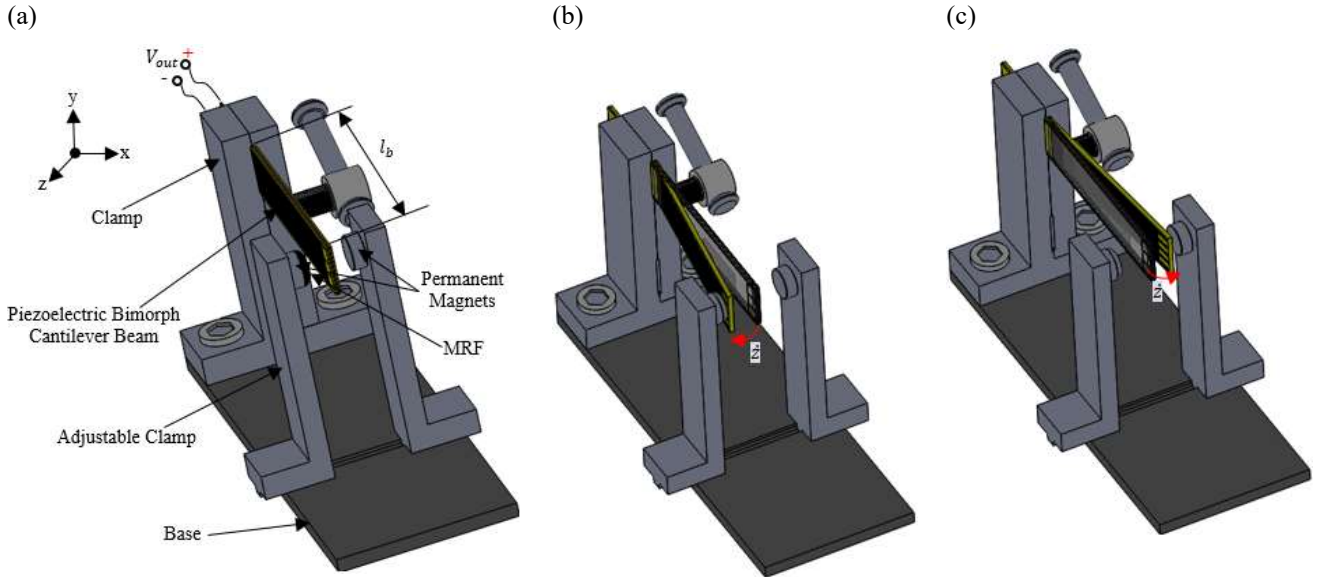
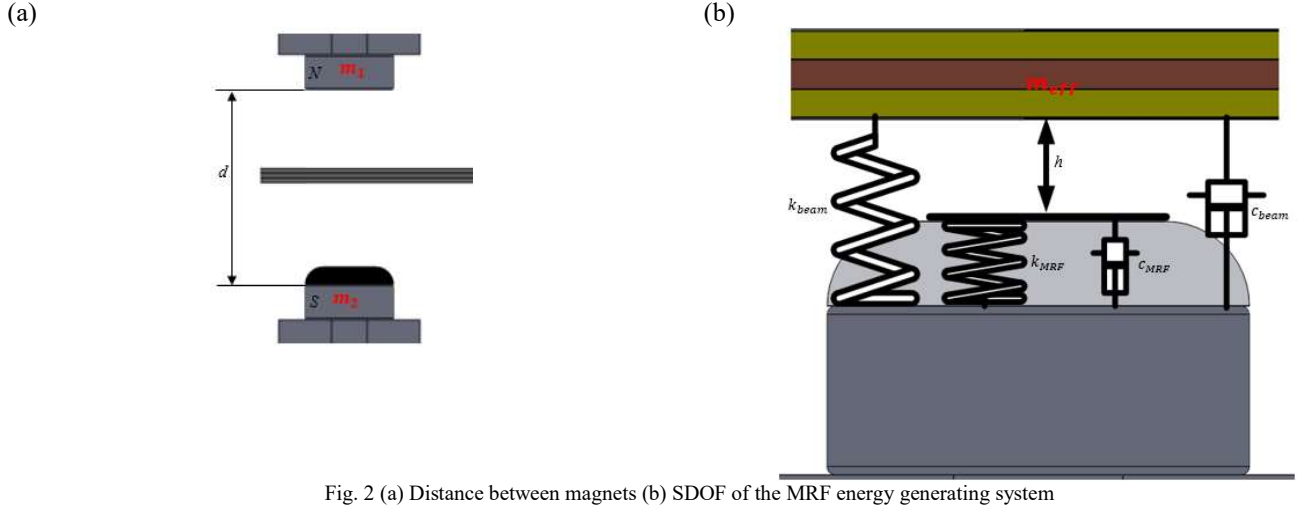


Fig. 1 (a) Proposed MRF energy generator, (b), displacement of excited cantilever impacting the MRF, (c) pull away displacement of the excited cantilever,



A cantilever piezoelectric bimorph cantilever was chosen for this application. This kind of cantilever is a typical piezoelectric device which has many vibration beam applications such as energy harvesting for battery-free remote-control switches, battery-free power supply of sensors in the field of automation technology, control systems and tire pressure monitoring systems and many more. The bimorph cantilever is made of a shim layer sandwiched between two piezoceramic layers. The piezoelectric bimorph cantilever beam used is also known as “Piezo Bending Actuator 427.0085.11Z from Johnson Matthey Piezo Products GmbH, [10].

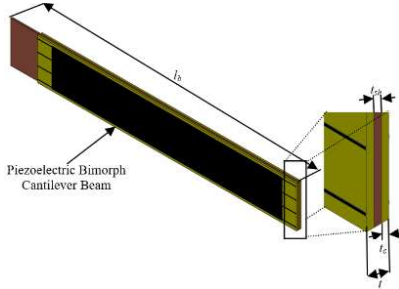


Fig. 3 Bimorph (Piezo Bending Actuators 427.0085.11Z [10])

TABLE 1 PROPERTIES AND DIMENSIONS OF PIEZOELECTRIC BIMORPH

Symbol	Description	Value	Units
l	Total length of beam	49.95 ± 0.05	mm
l_b	Free (vibrating) length	29.5	mm
w	Width of the beam	7.2 ± 0.05	mm
t	Total beam thickness	0.78 ± 0.03	mm
t_p	Thickness of piezoelectric layer	0.25 ± 0.05	Mm
t_{sh}	Shim layer thickness	0.28 ± 0.05	Mm

ρ_p	Piezoelectric layer density	8000	Kg/m ³
ρ_{sh}	Shim layer density	1800	Kg/m ³
ϵ_{33}	Permittivity	30975e-12	nF/m
E_{sh}	Shim layer modulus of elasticity	120e9	N/m ²
E_p	Piezoelectric layer modulus of elasticity	6.3e10	N/m ²
e_{31}	piezoelectric constant	250	

B. Characteristics of MRF

To design a device using MRF, it requires the knowledge and the characterization of the properties of the materials involved. Magnetorheological fluids (MRF) have attracted much research attention in recent times because of their unique rheological characteristics when a magnetic field is exerted on them and also because of the enormous industrial applications. MRF, just like electrorheological fluids (ERF), are part of a special type of fluids termed as smart fluids whose viscosity can be changed significantly by the application of a suitable stimulus. MRF have the special property that their viscosity or resistance to flow can be altered from liquids to semi-solids by the application of a magnetic field [11].

The apparent change in viscosity when a magnetic force is exerted on MRF makes them excellent choices in vibration damping and mechanical shock attenuating in the field of engineering. Magnetorheological fluids, foams, and elastomers comprise a class of smart materials whose rheological properties may be controlled by the application of an external magnetic field [12].

The devices that use MRF are classified, based on the shape of the particles, into two types: i) sphere iron particles and ii) plate iron particles [13,14]. The operation of MRF are classified into three types: i) shear mode, ii) squeeze mode and iii) flow mode as shown in Figs.4 (a), (b) and (c) respectively. One of the most important and potentially applicable modes for designing MRF products is in the squeeze mode. The squeeze mode could be classified into two categories namely; constant area squeeze

mode and constant volume squeeze mode. There has however been little literature on this subject area [15].

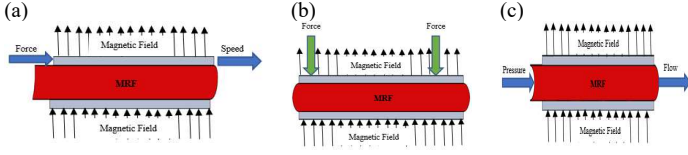


Fig. 4 (a) Shear mode, (b) Squeeze mode and (c) Flow mode [15]

TABLE 2 PHYSICAL PROPERTIES OF FLUID MR140CG LORD CORPORATION [16]

Property	Value
Base Fluid	Oil hydrocarbon
Working temperature	-40 to 130 (°C)
Density	3540 (kg/m ³)
Colour	Dark grey
Viscosity (slope between 800 and 500 Hz at 40°C)	0.280 (± 0.070) Pa·s
Flash point	>150 (°C)
Weight percentage of particles	85.44%
Thermal conductivity at 25 (°C)	0.25–1.06 (W/m, K)

C. Theoretical Modeling

Fig. 2(b) show a single degree of freedom (SDOF) spring mass-damper model to illustrate the linear modeling of the MRF based impact piezoelectric energy generating device. The equations of motion for the system is given ;

$$m_{eff}\ddot{x} + c_{total}\dot{x} + k_{eff}x = f \quad (1)$$

where, m_{eff} is the cantilever beam giving as;

$$m_{eff} = (2\rho_p t_p + \rho_{sh} t_{sh}) \times l \times w \quad (2)$$

c_{total} , is the total damping coefficient which is the sum of the damping of cantilever beam, c_{beam} and damping caused by MRF c_{MRF} , k_{eff} is the effective stiffness which is the sum of the beam stiffness k_{beam} and stiffness caused by the MRF k_{MRF} and f is the external force applied to the cantilever beam. x is the tip displacement of the cantilever beam at a distance h between the MRF and cantilever beam.

The rheological properties of the MRF due to the intensity of magnetic field determines the effective stiffness of the beam at the point of impact. The stiffness of the cantilever beam, k_{beam} is constant during displacement and the vibration behaviour is linear. At the point of impact with the MRF, the stiffness of the vibrating cantilever beam is converted from k_{beam} to k_{MRF} , where $k_{MRF} \gg k_{beam}$. From this stiffness effect, the vibration of the cantilever is transformed from linear oscillations to nonlinear impact oscillations, as a result of the displacement constraints given by the MRF stopper. The stiffness then becomes k_{MRF} ($k_{MRF} > k_{beam}$), making the effective stiffness of the cantilever beam (after impact) change from k_{beam} to $(k_{MRF} + k_{beam})$ and this shows that there is an adhesive or tensile force between the vibrating cantilever and MRF for part of the vibrating cycle until the magnetic particles dispersed. This behaviour gives an increase in the effective stiffness, which in effect changes the frequency response of the cantilever beam and diverges from its normal behaviour

allowing resonance to extend over a wider interval, in the vicinity of its resonant frequency. The kinetic energy of the cantilever beam is transformed to the potential energy of the layer of MRF without an additional loss [17].

The first point to consider in modeling the MRF is the yield shear stress of the fluid as a function of the magnetic field. In this work, the MRF 140-CG [16] was used. The yield stress $\tau_B(H)$ of this type of fluid is given by an experimentally-derived equation which depends on the magnetic field intensity and the particle volume fraction ϕ as [18].

$$\tau_B(H) = C \times 271700 \times \phi^{1.5239} \times \tanh(6.33 \times 10^{-6} H) \quad (3)$$

Where C is a coefficient dependent on the carrier fluid of the MRF and H is the magnetic field intensity.

The change in the viscosity of the MRF is used to generate a viscous force between the two surfaces of both magnets, in relative motion to the vibrating cantilever beam. The equation which relates the magnetic flux density B to the magnetic field intensity H [18] is:

$$B = 1.91 \times \phi^{1.133} \{1 - [\mu_0 e^{-10.97 \mu_0 H}]\} + \mu_0 H \quad (4)$$

The MRF stiffness k_{MRF} and squeeze damping c_{MRF} depends therefore on the force F_{MRF} , the yield stress $\tau_B(H)$. The determination of the $\tau_B(H)$ depends on the distance d between the permanent magnets 1 and 2.

The force generated due the magnetic field intensity and with the fluid F_{MRF} fixed at its free end is given as [19]:

$$F_{MRF} = \frac{4}{3} \frac{\pi r_m^2 \tau_B(H)}{(h+x)} \text{sign}(\dot{x}) \quad (5)$$

Where $\tau_B(H)$ is the yield stress of the MRF giving in Equation (3), r_m is the radius of the magnet, h is the initial gap between the MRF and cantilever beam and x, \dot{x} are the displacement of cantilever beam respectively.

The force due to the MRF is just an approximation and does not consider the hysteresis effect between the magnets while the MRF is either being compressed or separated from the oscillation of the cantilever beam. The magnitude of the magnetic stiffness k_{MRF} is computed from F_{MRF} as shown [19].

$$k_{MRF}(h) = \left| \frac{\delta k_{MRF}}{\delta h} \right| = \left| \frac{\partial k_{MRF}}{\partial h} \right| \quad (6)$$

$$k_{MRF} = \frac{4}{3} \frac{\pi r_m^3 \tau_B(H)}{(h+x)^2} \quad (7)$$

The equivalent stiffness of the cantilever beam can be expressed as [20]:

$$k_{beam} = \frac{3E_p I}{l_b^3} \quad (8)$$

Where I is the moment of inertia. Therefore, the moment of inertia of the reference layer can be expressed in equation 9 as and l_b is the Free (vibrating) length of the beam [21].

$$I = \frac{w}{12} \left(\frac{E_{sh}}{E_p} t_{sh}^3 + 2t_p^3 \right) + \frac{wt_p}{2} (t_{sh} + t_p)^2 \quad (9)$$

Where w is the beam width, E_{sh} is the shim layer modulus of elasticity, E_p is Young's modulus of piezoelectric material, t_{sh} is the shim layer thickness and t_p is the thickness of the ceramics layer of the piezoelectric beam.

The resonant frequency of the generating system is given as:

$$\omega_{sys} = \sqrt{\frac{k_{eff}}{m_{eff}}} \quad (10)$$

where k_{eff} and the effective mass of the cantilever m_{eff}

The damping ratio of the system is defined as the mechanical energy losses due to the fluid squeeze force, [19]. Therefore, the total mechanical damping gave as, $c_{total} = c_{MRF} + c_{beam}$

The damping effect owing to the squeeze force results due to the flow of viscous fluid through the small gaps which can be calculated as [19]:

$$c_{MRF} = \frac{3}{2} \frac{\pi \eta}{(h+x)^3} \frac{4}{m} \quad (11)$$

Where η is the viscosity of the MRF. The damping coefficient of the piezoelectric beam c_{beam} on the other hand can be expressed as [22];

$$c_{beam} = 2 \times \zeta \times m_{eff} \times \omega_{beam} \quad (12)$$

where ζ is the damping ratio the free vibrating beam giving as $\delta/2\pi$. δ is the log decrement corresponding to $\delta = (1/n)(\ln z_1/z_{n+1})$.

D. Modeling of the output power

From the maximum resonant frequency of the vibrating cantilever shown in Equation (10), the piezoelectric material, thus conserve most of the power generated. To generate electrical power, a resistive load is required. The electrical load is modeled as a single resistor. It is described by [23] that the output power reaches its maximum value when the resistance is described by Equation (14). The averaged power generated by a bimorph piezoelectric cantilever is given by,

$$Power = \frac{\omega_{sys}^2 t_{sh}^2 w^2 e_{31}^2 x^2}{4 \left(1 + w l_b \epsilon_{33} x \frac{\omega_{sys} R}{t_p} \right)^2} \times R \quad (14)$$

where ω_{sys} is the resonant frequency of the MRF-based system, t_{sh} is the shim layer thickness, w is the width of the beam, x is the max deflection and t_p is the thickness of the piezoelectric layer. e_{31} is piezoelectric constant. The matching resistance R of the piezoelectric material and is given as:

$$R = \frac{t_p}{w l_b \epsilon_{33} \omega_{sys}} \quad (15)$$

III. RESULTS AND DISCUSSION

The finite element model was performed in Comsol mutiphysics5a FEM package. The FE model was examined by comparing its responses in terms of the four natural frequencies. The purpose of this research only the first mode was considered. Table 3 summarizes the comparisons of the computed results to the FEM at different magnetic density application

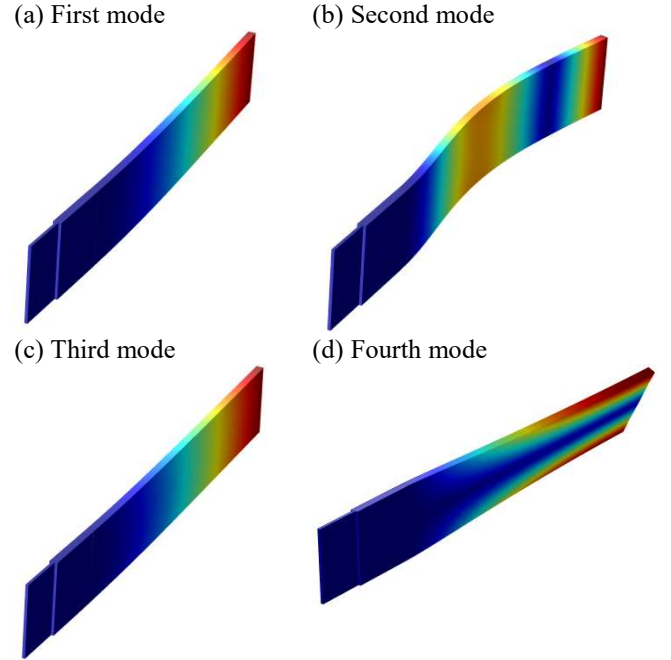


Fig. 5 Deflection modes of the piezoelectric Bimorph beam corresponding to the four natural frequencies.

TABLE 3 COMPARISONS OF COMPUTED FREQUENCY RESPONSE TO FEM

Magnetic Field Condition (T)	Natural Frequency (ω_n) (Hz)		Difference (%)	Peak Amplitude (μm)		Difference (%)
	Computed.	Comsol Model		Computed	Comsol Model	
No field	241.14	242.00	0.35	41	48	17
0.1	241.51	242.78	0.53	17	28	65
0.02	241.49	242.50	0.42	22	26	18
0.03	241.48	242.21	0.30	24	29	21

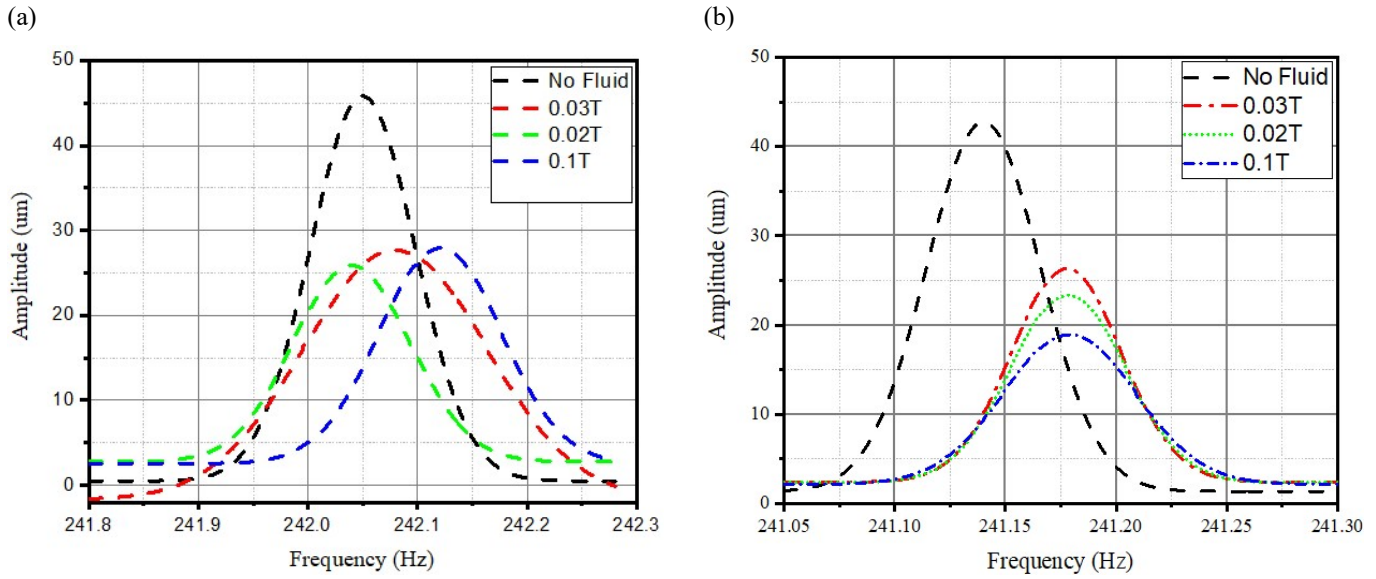


Fig. 6 Frequency response of the piezoelectric bimorph cantilever with MRF stopper (a) Computed and (b) FEM

The results from table 3 demonstrated that frequency could be tuned with the impact MRF activation. A frequency increase is observed when there is an increased in magnetic field application to the MRF. There is a 15 percent increase from a free vibrating cantilever without fluid impact to the highest magnetic field of 0.1T application to MRF by hand computation and a 32 percent increase in FEM. The results also show a decrease in amplitude when the magnetic field is increased that explaining the effects of the rheological properties of the MRF on the structural damping. The Fig. 5 (a) and (b) shows the graph of the computed and FEM results respectively.

Tables 4 and 5 shows the computed and FEM results for the theoretical output power. From equation 10 and the results shown in table 4, it was clear that the ability to tune the frequency of the system has paved the way to model and compute the amount power that can be harness from this system. The computed matching resistance and corresponding

power output was computed using equation 15 and 14 respectively. The maximum power achieved from the FEM study was $95\mu W$ and a matching resistance of $8.8k\Omega$. The computed results on the other hand shows a maximum power output $282\mu W$ and a matching resistance of $8.9k\Omega$. There has also been a 100 percent increased in power from the vibrating cantilever beam without MRF impact to an impact of MRF with 0.1T application, for the FEM analysis. The computed results gave an 88 percent increased in the out power for the MRF impact vibrating with a 0.1T application compared to the free vibrating cantilever beam. The figs 7 (a) and (b) shows the output power of FEM analysis and computed results respectively.

TABLE 4 COMPARISONS OF POWER OUTPUT TO FEM RESPONSE

Magnetic Field Condition (T)	Power Out-power (μW)		Matching Resistance ($k\Omega$)	
	Computed.	Cmsol Model	Computed	Cmsol Model
No field	150	47	8.9	8.9
0.1	282	95	8.9	8.8
0.02	245	82	8.9	8.9
0.03	223	63	8.9	8.9

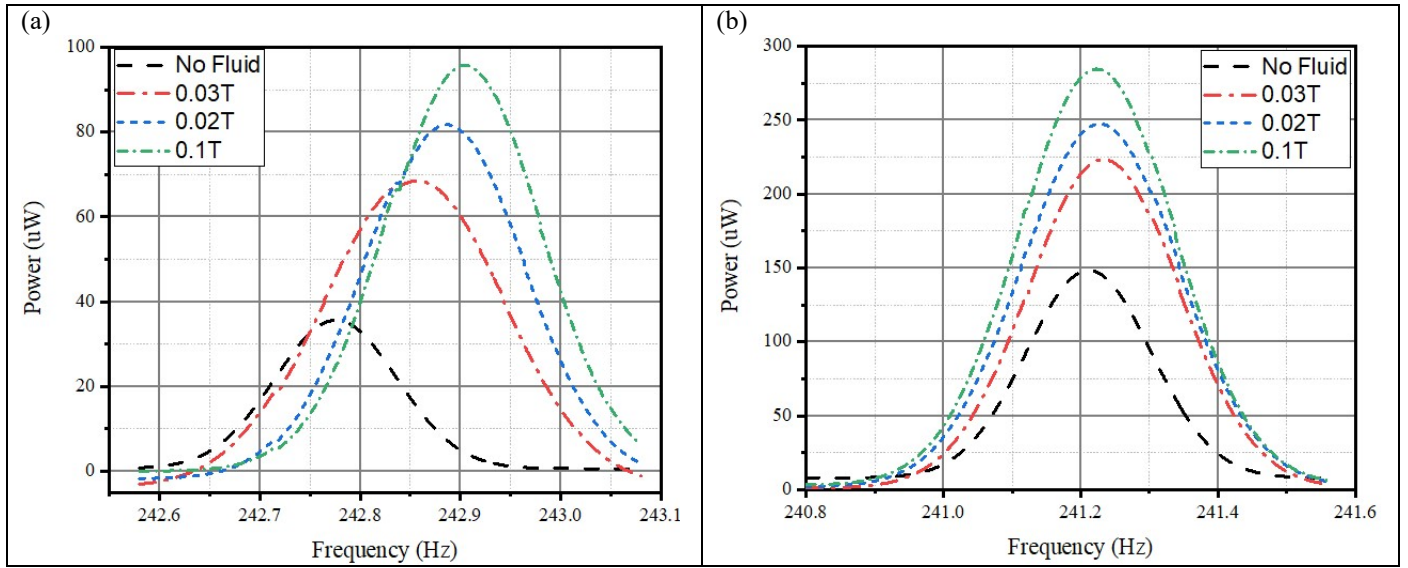


Fig. 7 Output power verses frequency response of the piezoelectric bimorph cantilever with MRF stopper (a) FEM and (b) Computed

IV. CONCLUSIONS

This paper presented an analytical analysis into frequency tuning method using the change in rheological properties of an MRF as magnetic field is applied as an impacting object. The system was computed as Single Degree of Freedom (SDOF) and results compared with FEM in Comsol Multiphysics FEA tool. Two set of results were achieved in the study. Firstly, computing frequency was tuned 15 percent for the 0.1T magnetic field application to the MRF as an impacting object compared to the free vibrating cantilever beam. The FEM analysis results shows that frequency was tuned at a 32 percent of an impacting MRF with 0.1T magnetic field application compared to the free vibrating piezoelectric bimorph cantilever beam.

Secondly, the computed power output for the impact MRF with 0.1T gave a maximum power of $282\mu W$ and a matching resistance of $8.9k\Omega$. The FEM results on the hand showed a maximum power output of $95\mu W$ and a matching resistance of $8.8k\Omega$ for the impacting MRF with 0.1T magnetic application.

These modeled results demonstrated that the influence of MRF increases as the magnetic field increases. MRF changes from liquid to semi-solid state on the application of the magnetic field. By changing the magnetic fields, have the frequency of beam increased during impact on the activated MRF. It was also observed that there was significant amount of power harnessed for the impacting beam.

REFERENCES

- [1] C.B. Williams and R.B. Yates 'Analysis of a micro-electric generator for microsystems'. *Sensors and Actuators*, 1996, A 52: 8-11
- [2] S. Priya 'Advances in energy harvesting using low profile piezoelectric transducers'. *Journal of Electroceramics* 19 2007: 165-182
- [3] E. Bryn, C.A. Kean and P.H. Aiguo, 'Mechanical frequency up-conversion for sub-resonance, low-frequency vibration harvesting' *Journal of Intelligent Material Systems and Structures* 2016, Vol. 27(16) 2145-2159
- [4] M. Umeda, K. Nakamura, and S. Ueha, "Analysis of the transformation of mechanical impact energy to electric energy using piezoelectric vibration," *Jpn. J. Appl. Phys.*, vol. 35, pp. 3267-3273, 1996.
- [5] M. Renaud, P. Fiorini, R. Schaijk, and C. Hoof, "Harvesting energy from the motion of human limbs: the design and analysis of an impact based piezoelectric generator," *Smart Mater. and Struct.*, vol. 18, pp. 1-16, Jan. 2009.
- [6] D. G. Lee, G. P. Carman, D. Murphy, and C. Schulenburg, "Novel micro vibration energy harvesting device using frequency up conversion," *IEEE Transducer '07*, pp. 871-874, 2007.
- [7] L. Gu, and C. Livermore, "Impact-driven, frequency up-converting, coupled vibration energy harvesting device for low frequency operation," *Smart Mater. Struct.*, vol. 20, pp. 10, Mar. 2011.
- [8] Z.H. Lai, G. Thomson, D. Yurchenko, D.V. Val, E. Rodgers 'On energy harvesting from a vibro-impact oscillator with dielectric membranes', *Mechanical Systems and Signal Processing* 107 (2018) 105-121, <https://doi.org/10.1016/j.ymssp.2018.01.025>
- [9] M. A. Halim, S. Khym, J. Y. Park, 'Impact based Frequency Increased Piezoelectric Vibration Energy Harvester for Human Motion Related Environments' *The 8th Annual IEEE International Conference on Nano/Micro Engineered and Molecular Systems*, (2013), 10.1109/NEMS.2013.6559879
- [10] Data Sheet Piezoceramic Trimorph Bending Actuator, Part No. 427.0085.11Z, Johnson Matthey Piezoproducts Bahnhofstraße 43, D-96254 Redwitz
- [11] B. Sapiński, J. Snamina, 'Vibration of a beam with magnetorheological fluid in non-homogenous magnetic field', *Modelowanie Inżynierskie*, 2009 6, 241-248.
- [12] W. H. Li, H. Du, G. Chen, S. H. Yeo, and N. Guo, "Nonlinear viscoelastic properties of MR fluids under large-amplitude-oscillatory-shear," *Rheol. Acta*, vol. 42, no. 3, pp. 280-286, 2003.
- [13] X. H. Liu, Z. M. Fu, X. Y. Yao and F. F. Li 'Performance of magnetorheological fluids flowing through metal foams Measurement' *Science Review* 2011, 11, 144-8
- [14] K. Shah, D. X. Phu and S. B. Choi 'Rheological properties of bi-dispersed magnetorheological fluids based on plate-like iron particles with application to a small-sized damper' *J. Appl. Phys.* 2014 115 203907
- [15] A. S. Mazlan, N.B. Ekrem and A. G. Olabi 'The performance of magnetorheological fluid in squeeze mode'. *Smart Mater Struct* 2007 16: 1678-1682.

- [16] L. Corporation, MRF-140CG, Magneto-Rheological Fluids, [http://www.lord.com/products-and-solutions/magnetorheological-\(mr\)/product.xml/1646](http://www.lord.com/products-and-solutions/magnetorheological-(mr)/product.xml/1646). (2014)
- [17] J. D. Carlson, , Magnetorheological fluids, in Smart Materials, New York, NY, USA: CRC Press, (2008)
- [18] M.F. Narimani, G.N Golnaraghi,. Jazar, ‘Frequency response of a piecewise linear vibration isolator’, J. Vib. Control (2004) 10 (12) 1775–1794.
- [19] S. Kaluvan, K. Shah and S-B. Choi, “A new resonant based measurement method for squeeze mode yield stress of magnetorheological fluids” Smart Mater. Struct (2014), 23 115017 (8pp). doi:10.1088/0964-1726/23/11/115017
- [20] H. Kim, Y. Tadesse and S. Priya ‘Piezoelectric Energy Harvesting’. In: Priya S and Inman D (ed.) Energy Harvesting Technologies. New York: Springer, 2009 41-77.
- [21] F. P. Beer and E. R. Johnston Jr ‘Mechanics of Materials’, 2nd Edition 1992. New York: McGraw-Hill
- [22] B. Richter ‘Modellbasierter Entwurf resonant betriebener, piezoelektrischer’ Biegeschwinger in Energy Harvesting Generatoren. Aachen: Shaker Verlag 2010
- [23] F. Lu, H. P. Lee and S. P. Lim, “Modeling and analysis of micro piezoelectric power generators for micro-electromechanical systems applications.” Smart Mater, Struct. 2004 13 57-

Advanced Manufacturing

LASER POLISHING AND STRUCTURING OF TOOLING AND FUNCTIONAL SURFACES

Evgueni V. Bordatchev

Automotive and Surface Transportation Research Center
 National Research Council of Canada
 London, Canada
 evgueni.bordatchev@nrc-cnrc.gc.ca

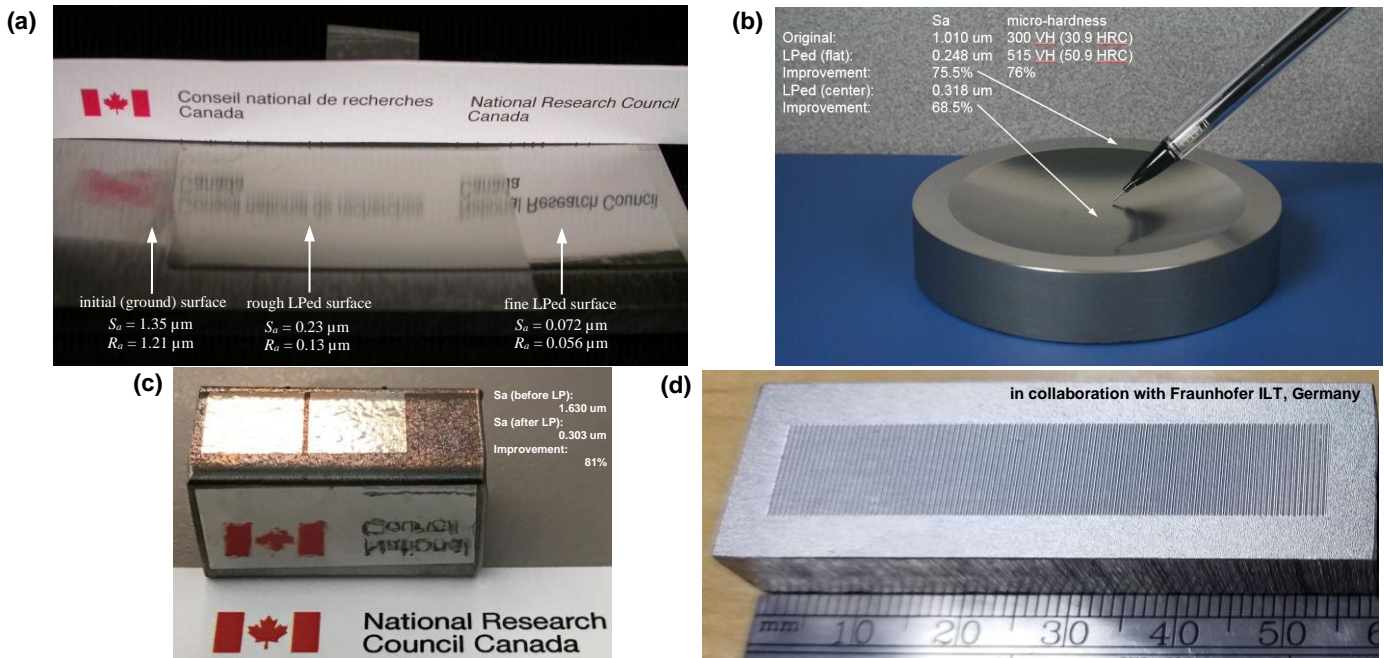
O. Remus Tutunea-Fatan

Mechanical and Materials Engineering
 Western University
 London, Canada
 rtutunea@eng.uwo.ca

Since its inception, laser polishing (LP) technology has been receiving an increasing attention as a plausible alternative to the conventional polishing techniques. The main driver behind the development of LP technology resides in the fact that >40% of the tooling cost is associated with high-cost and time-consuming manual polishing. By contrast, LP can significantly reduce these costs by the high level of automation and the precision provided through its coupling with CNC technology. Considering its potential applications in automotive, aerospace and biomedical industries, the Automotive Portfolio of the National Research Council (NRC), Canada has been actively engaged in the development of LP technology. Building on this activity, the main objective of the present report is to introduce some of the achievements and developments of LP technology at NRC over the past four years.

This presentation will focus on detail description of the laser polishing process, its advantages and disadvantages with respect to the conventical abrasive polishing techniques, and examples of LP process technical implementations along with examples of LPed parts. Then common understanding the process physics, process classification and its variants, material and surface characterization, and modeling capability will be presented. In addition, effect of most critical process parameters, laser type and characteristics, laser path trajectory, and process planning methodologies on achieved surface quality and physical-mechanical characteristics, e.g. gloss, micro-hardness, metallographic structures, corrosion resistance and others.

Special attention will be placed on technical applications of the LP process in manufacturing tooling, molds and dies, medical implants, additive manufactured parts, optics, and others. The presentation will conclude with techno-economic analysis of the LP implementation, an outlook on the future of the technology, and technical and knowledge gaps that still need to be filled.



Photographs of laser polished samples and their surface characteristics: a) two-step laser polishing of H13 tool steel, b) continuous wave laser polishing of 3D tooling, c) LPed additively manufactured rectangular tube, and d) optical light guide tooling insert

Implementations and Applications of Elliptical Vibration Micro/Nano-Cutting

Nikolai Farrus, O. Remus Tutunea-Fatan
Mechanical and Materials Engineering
Western University
London, Canada
nfarrus@uwo.ca, rtutunea@eng.uwo.ca

Nicolas Milliken, Evgueni V. Bordatchev
Automotive and Surface Transportation Research Center
National Research Council of Canada
London, Canada
{nicolas.milliken, evgueni.bordatchev}@nrc-cnrc.gc.ca

ABSTRACT

Various surface functionality such as friction control, wettability, aerodynamics, aesthetics and well as many others are novel attributes of high-quality/precision parts including tooling components demanded by automotive, biomedical, defense, aerospace, photonics and security applications. Therefore, cost-effective ultraprecision microfabrication technologies for micro/nano-texturing and structuring of functional surfaces are typically regarded as both a “key enabler” and a “game changer” in today’s knowledge-driven advanced manufacturing. Along these lines, elliptical vibration micro/nano-cutting (EVMNC) is one of the emerging technologies that has experienced a significant growth over the past decade. During EVMNC, a single point non-rotational cutting tool (typically with a diamond insert/tip) moves along a 3D trajectory. In this case, the required *slow* motions are provided by the stage of a conventional five-axis machine tool. In addition to them, *fast* externally controlled elliptical vibration motions with micron/under micron amplitudes and high frequencies are applied to the tool, their main role being to ensure a discontinuous contact between tool and workpiece material. This intermittent contact between tool and workpiece - that is in fact one of the EVMNC characteristics – facilitates generation of high quality surfaces on ferrous and/or difficult to cut materials and accounts for reduced cutting forces and tool wear. The EVMNC process can be characterized by several parameters associated with time-space description of the elliptical trajectory, such as: feed rate, frequency, amplitude, and locus spatial orientation. Building on this, the main goal of the present study is to determine options for EVMNC process implementation and suitability for specific applications. The options to be identified will include various possibilities of time-space combinations, synchronization, optimization and control of fast and slow kinematics.

The first implementation to be analyzed is conceived as a combination of linear and/or curved path motions covering a preset 2D area while maintaining a constant linear speed/feed rate and a constant or variable overlap between adjacent moves. The goal of this implementation is to produce a surface with constant and uniform micro/nano-topography. The structure to be generated in this case will be periodical, as defined by the selected combination between feed rate and EV frequency. Typical applications for this implementation are constituted by diffraction gratings for non-paint-based metal color changes, hydrophobic or hydrophilic surfaces for wettability and adhesion control. All other EVMNC process parameters will remain constant in this implementation.

The second implementation is conceived as a modified version of the former one. In this case, feed rate is varied along linear and/or curved path moves over the preset surface area. This approach enables the introduction of an areally-distributed frequency modulation of the periodical texture. As such, this implementation becomes suitable for creating multi-colored surfaces with artistic appearance over a preset area. Moreover, this implementation enables an areal control of surface wettability, in a sense that it facilitates the fabrication of planar microfluidic systems characterized by alternations between hydrophilic and hydrophobic areas.

The former two implementations can be combined into a more complex third EVMNC implementation to additionally control the spatial amplitudes of the tool tip elliptical vibrations along a tool path trajectory. In case of constant feed rate, amplitude modulation of elliptical vibrations changes the cutting depth and therefore a 2½D surface topography is created. By changing the amplitude both along vibration direction (longitudinal) as well as feed direction (bending), fabricated structures can be used for either improved structural functionality or surface quality. The third implementation is primarily suitable for micro/nano-machining of tooling inserts, diffractive and holographic optical structures, picture images, and other functional textured/structured surfaces.

The fourth implementation is a complex and advanced technique involving tuning and control of the spatial orientation of the elliptical vibratory motion of the tool tip. This is achieved by controlling the phase shift between bending direction and longitudinal components of the elliptical vibrations. By adjusting the phase shift value, locus ellipse orientation can be changed from forward-tilted, to vertical and further to back-tilted orientation. This facilitates the control of periodical surface structures shape to represent the basis of directionally-controlled slanted gratings or sawtooth waveforms. The fourth implementation opens a wide range of applications in holographic sensing/security, friction, aero, hydro drag control, and perhaps even yet unknown functionalities.

It is anticipated that a better understanding of the aforementioned EVMNC implementations will create the premise for further development of this advanced micro/nano-fabrication technology to find applications in numerous industries.

Analysis of Remelted Line Formation During Laser Polishing of H13 steel

Srdjan Cvijanovic, O. Remus Tutunea-Fatan

Mechanical and Materials Engineering
Western University
London, Canada

scvijano@uwo.ca, rtutunea@eng.uwo.ca

Evgueni V. Bordatchev

Automotive and Surface Transportation Research Center
National Research Council of Canada
London, Canada

evgueni.bordatchev@nrc-cnrc.gc.ca

ABSTRACT

Laser polishing (LP) is a novel finishing process used to improve surface quality of manufactured parts and functional geometries. In this process, the improvement of surface quality is obtained by remelting initial surface topography and redistributing molten material and its solidification within chosen polished area. Overall, LP has significant advantages over conventional polishing processes, especially since no material is being removed. Additional critical advantages of the LP are: high process speed (several seconds/cm²), accuracy and repeatability (ensured through the use of CNC technology), and versatility (can be applied effectively on both large/small areas and structures as well as on intricate free-form surfaces and deep cavities).

During LP, continuous or pulsed laser beam is focused on the surface to create a pool of molten material. Surface tension and capillary forces are leveling initial profile as surface peaks fill adjacent valleys to form a smoother surface topography. Laser beam travels in accordance with CAM-programmed and CNC-controlled path in order to fully cover the polished area. During these motions, the pool of molten material continuously forms and regenerates with new material drawn through the path-facing melted-solid material interface. Simultaneously, there is a rear solid-melted material interface which is predominantly responsible for the formation of surface quality behind the core laser-material interaction zone. Since little is known about this phenomenon, this study is focused on the analysis of the melted line geometry formation. The approach to be employed will consider an instant cross-sectional profile and volume of the molten material with respect to two main process parameters: travel speed and laser power.

LP experiments were performed on two H13 tool steel samples initially delimited by ground and flat surfaces with an areal average surface roughness of $S_a = 0.75 \dots 0.98 \mu\text{m}$ and $S_a = 0.03 \dots 0.04 \mu\text{m}$, respectively. Sixteen LPed lines were generated on each sample by incrementally varying travel speed (250 mm/s to 1000 mm/s) and continuous wave laser power (25 W to 100 W). 3D topographies of the LPed lines were measured by means of an optical profilometer and digitized for further statistical analysis of the surface parameters. The following characteristics and parameters were calculated for each line: middle longitudinal profile, averaged transversal profile, average surface roughness, effective line width, and averaged bulge height. Moreover, metallographic cross-sections of LPed lines were prepared and their geometric characteristics (instant width, depth and bulge height) were measured by means of a high-resolution digital microscope (250x to 2500 magnification) to determine the instant area of molten material. The novelty of this approach consists in the comparative analysis performed to assess the effect of LP on initially flat and ground surfaces in order to identify the specific contribution of two overlapped mechanisms that are present during LP: solid bulk material melting and surface micro-topography melting.

The initial part of the analysis was focused on determining how LP process parameters (travel speed and laser power) are involved in formation of cross-sectional geometry. While each of the main process parameters contributes to geometry alterations, only their superseded effects are able to determine the final shape of the LPed surface. For example, laser power increases translate into larger heat affected zones, bigger polished line widths, and expanded cross-sectional areas. Conversely, increased travel speeds result in reduced depths, widths and cross-sectional areas. It was also found that the presence of the ground topography does not significantly affect the cross-sectional geometric parameters. This means that the analysis of cross-sections is capable to reveal the effect of just one mechanism associated with LP, namely melting of the solid bulk material. A better understanding of this mechanism is desirable for controlling formation of actual surface quality. Furthermore, to achieve the best and/or intended surface quality, an optimal combination of melting bulk and topography material should be attained. As such, an additional emphasis was placed on the analysis of surface quality formation that was performed by linearly separating contributions from melting bulk and topography material. For instance, it was shown that melting bulk material predominantly contributes to the final quality with high applied laser powers and low travel speeds. More specifically, when a ground sample with $R_a = 0.510 \mu\text{m}$ was polished at 500 mm/s and with 100 W laser power, the contribution of melting bulk material was assessed at $R_a = 0.391 \mu\text{m}$.

Present study was focused on the development of understanding regarding remelting mechanisms to occur during line polishing performed with different process parameters. It is expected that further advancements of this knowledge will be useful in various industrial applications including reconfigurable texturing and structuring for surface functionalization.

INNOVATIVE STRATEGIES FOR ASSESSING THE METROLOGICAL PERFORMANCE OF OPTICAL 3D SCANNERS

Farbod Khameneifar
Department of Mechanical Engineering
Polytechnique Montréal
Montreal, Canada
farbod.khameneifar@polymtl.ca

ABSTRACT

The latest industrial optical 3D scanners are able to sample millions of coordinate data points from an object's surface in a very short amount of time, as low as just a few seconds. That is the reason why 3D scanners are gaining increasing popularity in different industries, including manufacturing, for part inspection and reverse engineering. In spite of that, currently, there exists no international standard that specifies how to evaluate the metrological performance of these 3D scanners. This issue restricts the application of the 3D scanners in highly regulated industries such as aerospace manufacturing, where precision is of paramount importance. The users of the scanners, especially for precision inspection of manufactured parts, need instructions for assessing the accuracy and precision of different systems in measuring specific parameters. They need to perform acceptance tests to verify whether a particular scanner is qualified to be used to inspect the part in question. In this work, innovative strategies are proposed for qualification of 3D scanners based on novel reference artifacts calibrated by a contact probe on a coordinate measuring machine (CMM). The CMM measurement is known to have high accuracy and precision. It is demonstrated that the reference data from CMM-calibrated artifacts can be utilized for benchmarking the 3D scanners. Using the reference data, the proposed procedures realize acceptance tests and yield comprehensive insights into the volumetric errors of the 3D scanners, in particular, structured-light scanners and laser scanners, in measuring surfaces at each point in their working volume. The effects of ambient light, surface color and texture, as well as the influence of scan data processing for stitching multiple scans, outlier removal and noise reduction on the metrological quality of data are discussed. In addition, an uncertainty propagation scheme is presented to estimate the combined measurement uncertainty of the scan data originated from different sources.

Axial Strategy for Ultraprecise Single Point Cutting of V-grooves with Constant Cross-Sectional Cutting Area

Delfim Joao, O. Remus Tutunea-Fatan
 Mechanical and Materials Engineering
 Western University
 London, Canada
 djoao@uwo.ca, rtutunea@eng.uwo.ca

Nicolas Milliken, Evgueni V. Bordatchev
 Automotive and Surface Transportation Research Center
 National Research Council of Canada
 London, Canada
 {nicolas.milliken, evgueni.bordatchev}@nrc-cnrc.gc.ca

ABSTRACT

It is well accepted today that the miniaturization of numerous components has triggered a rapid development of the associated microfabrication technologies. In addition to the inherent focus on obtaining a precise geometry and/or shape of the small-sized products, one of the latest trends in the area is represented by surface functionalization, a procedure through which regular shaped-microstructures are being added on outer/inner surfaces in order to enhance some of their innate physical properties such as optical and/or thermal. Among the broad palette of microstructures used for surface functionalization purposes, V-grooves have found a relatively wide use in mechanical, electronic, photonic, biomechanical and optical applications. However, the manufacturing strategies associated with their fabrication were little investigated so far. To address this, the main goal of the current study was to develop a new cutting strategy capable to generate V-grooves characterized by a constant cross-sectional area to presumably translate into a constant cutting force, a desirable trait of virtually any micromachining process.

As the first step of the strategy development process, sample V-grooves were manufactured in an aluminum block by means of a Kugler Microgantry Nano 5x micromachining center. The preselected geometry of the V-groove has imposed the use of a V-shaped monocrystalline diamond tool characterized by a 0° rake angle, an included angle of 90° (to match the included angle of the V-groove structure), and a clearance angle of 10° on both cutting edges involved in the generation of the symmetric V structure. The motions of the tool were constrained to YZ plane, Y representing the principal feed/cutting direction whereas Z was used to gradually increase the depth of the microstructure until its final value (Fig. 1a). The incremental increases in cutting depth were calculated in advance according to the main objective of the current strategy, namely constant cross-sectional cutting area. It is perhaps important to note here that the underlying geometry of the V-groove implies that the constant areal constraint is correlated with a gradually decreasing depth of cut/chip thickness. According to the implemented strategy, the kinematics of the tool has followed a consistent pattern involving a succession of predefined motions to essentially follow a rectangular trajectory characterized by a constant Y length and an incrementally increasing Z depth. The only out-of-plane motion was the X step over movement required to ensure the positioning of the tool at the beginning of a new V-groove structure. The entire succession of motions was programmed in advance in a custom-built Matlab postprocessor and subsequently validated by means of Vericut software prior to its actual run on the micromachining center. In order to verify the accuracy of the programmed cutting strategy and its physical implementation on the machine tool, the Y component of the cutting force was monitored by means of a Kistler dynamometer mounted underneath the workpiece.

The results presented in Fig. 1b suggest that a correlation exists between the amount of material removed in each pass and the magnitude of F_Y . However, while from a theoretical standpoint it is reasonable to predict that F_Y will remain constant since it is directly proportional with the amount of material removed (of a preset constant cross sectional area), the approximately 20% decay of F_Y that is visible in Fig. 1b seems to suggest that the gradually decreasing chip thickness also plays a less anticipated but also important role on cutting force magnitude.

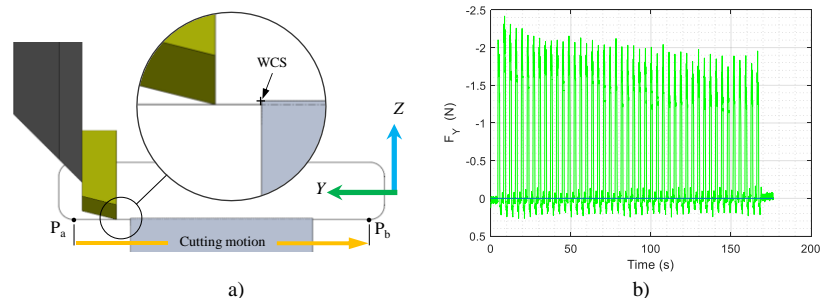


Figure 1. V-groove fabrication: a) cutting kinematics, b) Y component of the cutting force

While the primary objective of this study was attained if not even surpassed (since the cutting force decreases in this cutting strategy), future extensions of this work will aim to include an investigation of the tool wear as well as cutting chips by means of scanning electron microscope (SEM). Beyond that, additional V-groove cutting strategies will be being developed for further comparison and optimization purposes.

Additive manufacturing of electroactive polymer actuated patient-specific arterial phantoms simulating stenosis and dilation

F. Benjamin Holness, Aaron D. Price

Department of Mechanical and Materials Engineering
Western University
London, Canada
fholness@uwo.ca

Tamie L. Poepping

Department of Physics and Astronomy
Western University
London, Canada

ABSTRACT

Fabrication of arterial phantoms is enabled through specially developed additive manufacturing techniques in the Organic Mechatronics and Smart Materials Laboratory to produce high resolution 3D conjugated polymer structures. These techniques have been modified to enable fabrication of soft, flexible structures through direct ink writing of polydimethylsiloxane (PDMS) into a microgel support bath. Employment of this support bath enables the ability to produce true 3D structures with high resolution, and overhanging features. The support bath is composed of a (poly(acrylic acid)) powder hydrated in water to form a viscous dispersion. This bath behaves as a Bingham plastic, fluidizing under an applied shear stress by the print nozzle moving through it, but returning to its solid state after the nozzle passes. This allows the nozzle to move freely through the bath during extrusion but also support the newly extruded ink as it cures and maintaining dimensional accuracy.

Following curing and removal of the PDMS phantom from the support bath, PEDOT:PSS thin films are deposited on the phantom surface. These films have demonstrated significant hygroscopic actuation under an applied electric field. Through controlled joule heating and the volume change associated with the desorption and adsorption of water actuation is observed. Particle image velocimetry (PIV) has been employed to characterize the effect of actively changing vessel geometry. PIV can provide the instantaneous full-field velocity profile and is a well-established technique to characterize flow through phantoms fabricated by conventional casting techniques to provide a standard of comparison. To effectively image the device via PIV, the optical properties of the components must be considered. To this end, PDMS and PEDOT:PSS have been employed due to their favourable transmission properties in the visible spectrum. Additionally, PDMS provides a compliant passive structure to be deformed with relatively low force, reducing the performance requirements of the actuators. While this device focuses on the actuation of phantom vessel geometry, this technique may be extended to other applications in microfluidics to create onboard peristaltic pumping action and vascular networks.

EFFECT OF TOOL PATHS ON FINE & ULTRAFINE PARTICLES EMISSION & DISTRIBUTION DURING POLISHING OF MEDIUM SILICA CONTENT GRANITE

F.B. KOUAMOU, V. SONGMENE, J. KOUAM

Department of Mechanical Engineering
 École de technologie supérieure, ÉTS
 Montreal, Canada,
Victor.songmene@etsmtl.ca

ABSTRACT

Granite processing has intensified in recent decades thanks to the development of mining techniques and the high demand of the market. When polishing granite, particle emissions expose workers to many diseases. In fact, granite contains silica, which can be a chronic contaminant for the body. Studies aimed at reducing particulate emissions during granite polishing have mostly focused on cutting parameters, on lubrication or on ventilation. This study focuses on the influence of the abrasive tool path on particle emissions. Three different polishing strategies (linear, arc and spiral) are used to polish medium silica content granite and their effects on fine and ultrafine particles emission studied. It is found that the tool path has an influence on part quality but also on the emitted dust particle (concentration, size, composition and distribution). The spiral trajectory generates a better surface finish but produces more ultrafine particles, more dangerous, than other tested tool paths.

Keywords-Granite; polishing; dust; silica; tool path,ultrafine particle

Highlights

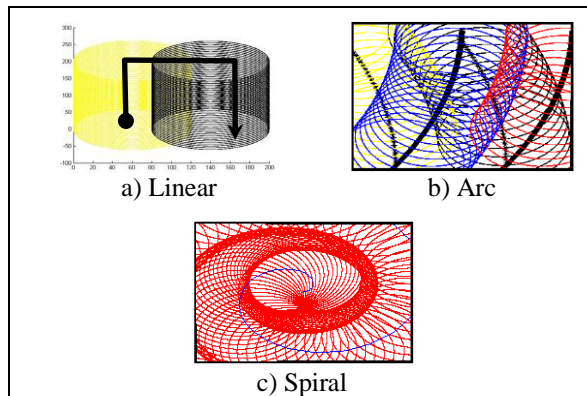


Figure 1. Tool paths tested.

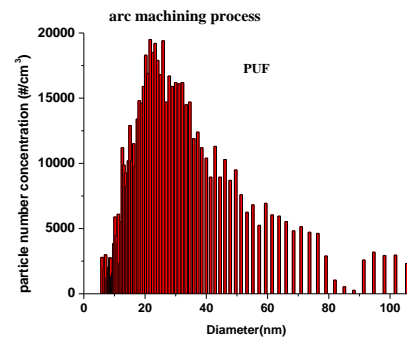


Figure 2: Ultrafine particle distribution obtained while using arc tool path

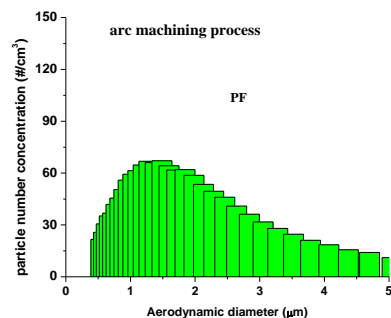


Figure 3: Fine particle distribution obtained while using arc tool path

Machining studies on High Silicon content Al-Si alloys for the aerospace industry

Oluwole A. Olufayo, Oussama Chaieb, Victor Songmene and Mohammad Jahazi

Department of Mechanical Engineering,
École de Technologie Supérieure
Montréal, Canada.

oluwole-ayodeji.olufayo.1@ens.etsmtl.ca

Victor.songmene@etsmtl.ca

Mohammad.Jahazi@etsmtl.ca

A NEW RAPIDLY SOLIDIFIED ALUMINIUM CASE STUDY

To meet the needs for performance of products, the aerospace industry increasingly needs to assess the materials behaviours during manufacturing processes, ensure they possess adequate machinability as well as high performance and an extensive lifecycle. Over the years, industrial research works have focused on developing new alloys with increased conductivity as well as increased strength. High silicon content Aluminum (Al-Si) alloys, due their increased thermal conductivity, low coefficient of thermal expansion and low density, have been identified as suitable materials in space applications. However, there are some problems encountered while machining these alloys. These problems are related to chip formation, machined part's microstructural integrity and the generation of desired surface finish. These challenges are often tied to the machining conditions and strategies but also on workpiece materials.

The advantages the rapid solidification (RS) process offers vis-à-vis material properties enhancements have been seen in numerous research studies. This process which involves the cooling of melted metal at cooling rates in the order of 10^4 – 10^7 K/s and this leads to an alloy with a fine grain microstructure and increased strength. The use of rapidly solidified Al-Si alloys in industrial application under adequate machining conditions can provide industries with a suitable light-weight alloy with improved mechanical properties.

In this study, experimental milling tests were performed on a rapidly solidified Al-Si alloy, with a prominent silicon content (over 50%) to identify inherent machinability challenges faced during industrial production of these alloys. The tests were performed using a polycrystalline cubic boron nitride (PCBN) tool coated with diamond to reduce the effect of wear, material adhesion, surface oxidation and particle diffusion. The effects of cutting parameters such as cutting speed, feed rate and depth of cut on surface roughness, cutting forces, and vibration of the machine were also analyzed. Cutting forces which were measured using a Kistler dynamometer, indicated a marginally stability below 120 m/min. However, these force values increased at higher in cutting speeds due to the material hardness. A marginal effect of feed was observed on measured forces. A laser-digital microscope was used to measure the surface roughness and morphology. A comparative analysis of the surface with conventional Al-T6061 alloy showed an improvement in surface roughness measurements using the RS Al-Si alloy. Lower cutting speed and feed rate on both conventional and rapidly solidified aluminium alloys produced a corresponding reduction in roughness. However, the feed rate possessed a lower influence on the surface condition when compared to the cutting speed. Reduced vibrations were also identified in RS Al-Si alloy, which only displayed notable vibrations levels at cutting speeds above 130m/min.

An overview of the microstructure of the alloy after machining revealed no cracks between the silicon grains and the aluminum matrix, as well as no irregularities or the tearing off of the silicon grains from the alloy. Machinability challenges were mostly associated to poor tool life. This is as a result of the hardness of the silicon present in the matrix promoting a rapid tool wear. In contrast, prominence of silicon in the alloy gives it an added fluidity and low shrinkage, consequently producing good a material with good castability and weldability.

Finite element modeling of fracture in ceramic micro-particles during composite coldspray process

Rohan Chakrabarty

*Department of Mining and Materials Engineering
McGill University
Montréal, Québec, Canada
rohan.chakrabarty@mail.mcgill.ca*

Jun Song

*Department of Mining and Materials Engineering
McGill University
Montréal, Québec, Canada*

ABSTRACT

For developing metal matrix composite coatings by cold spray processing, a mixture of metal and ceramic particles is accelerated at high velocities towards the substrate material. Therefore, understanding the dynamic impact behavior of micron-sized brittle particles is a topic of interest for cold spraying of composite materials. In this work, we examine the fracture behavior of a ceramic particle upon impact on a substrate, using cohesive zone modeling. Simulations have been carried out for different values of ceramic particle grain size and particle impact velocity. The results show that the material failure is found to occur by the linking of wing cracks. The necessary conditions for fragmentation, and retention probabilities of ceramic particles have been studied considering the effects of particle grain size and grain boundary properties. This study contributes important mechanistic knowledge towards understanding and predicting the ceramic retention behavior and composite coating characteristics during metal-ceramic composite cold spraying.

Application of EWMA Control Chart on Volumetric Errors Change Recognition

Kanglin Xing, J.R.R. Mayer, Sofiane Achiche

Department of Mechanical Engineering
École Polytechnique (Montréal)
Montréal, Québec, Canada
kanglin.xing@polymtl.ca

Abstract—The exponentially weighted moving average (EWMA) control chart is a popular tool in quality control and effective in detecting small shifts in the monitored signals. Herein, EWMA has been used for machine tool volumetric errors (VEs) change recognition. To improve its recognition capability, the influence of setup parameters including smoothing coefficient, width coefficient and control limits (time varying control limits and asymptotic control limits) of EWMA are studied using real machine tool faults, pseudo-faults and simulated faults. The results reveal that time varying control limits are better than asymptotic control limits for VEs change recognition when the amount of acquired VEs data is limited. The combination of relatively small smoothing coefficient and width coefficient are recommended for a better VEs change recognition. Finally, the general EWMA input calculation method I is recommended while the proposed EWMA input calculation method II could be used as the second tool for the stability checkup of faulty state data.

Keywords- Machine tool; Volumetric error; Change recognition; EWMA

I. INTRODUCTION

The sudden failure or the degeneration of machine tool could significantly affect their productivity and capability. Condition monitoring is preferred in industry because of its efficient role in improving plant production availability and reducing downtime cost of machine tools. Currently, the monitoring objectives of machine tools are focused on machine tool main components such as feeding systems, spindle system, main mechanical structure, coolant system, etc. as well as the machining process [1]. A non-contact structure monitoring system for machine tools based on the vibration signal results in a reliable monitoring without altering the structure of the machine tool [2]. Spindle bearings condition monitoring via the use of acoustic emission signals and Hilbert–Huang transform analysis reveals good correlation between the AE data and the increase in the preload, the change in the dimensions and geometry of the spindle bearings [3]. In the machining process, tool wear, tool collision and tool life prediction have been monitored by analyzing signals such as the spindle current, feed, forces, acoustic emission, vibrations, etc. [4].

Although machine tool condition monitoring system has been extensively researched, there are still major limitations that

need to be overcome. It is hypothesized that the holistic state of the machine tool condition cannot be reflected just by partially monitoring some of its key components. In addition, it is difficult to link the degradation of the monitored mechanical parts such as the feed drive systems with the quality of machined parts [5]. Even though, the real time monitoring techniques without interference on the normal machining process still makes the current monitoring strategy widely used in industry [4].

Machine tool geometric information is often periodically measured during the maintenance of machine tools. However, this geometric information has rarely been used for accuracy monitoring. Take the machine tool volumetric errors (VEs) as an example, VEs are defined as the relative Euclidian vector between the tool frame and workpiece frame [6]. They are the compact reflection of machine tool geometric condition [7]. Therefore, they are greatly valuable for machine tool condition monitoring. The research on VEs is focused on the VE modeling, prediction and compensation. Recently, it has been used for machine tool condition monitoring. Techniques such as vector similarity measure (VSMs) [8], principal component analysis and K-means have been used for VEs feature extraction and classification [7, 9]. Unlike the physical signals such as power, vibration, force, etc., VEs cannot be acquired by real time measurement but under the normal maintenance period. This drawback makes it impossible to build a large VEs data set in a short time. Therefore, how to precisely use the limited VEs data for machine tool condition monitoring is a crucial question before implementing it as a valuable condition monitoring solution.

The EWMA control chart based on current and past historical data is widely used for industrial quality control, since it is especially suited for detecting small and moderate shifts in a process, and it shows good robustness for the non-normal distributed data for certain values of setup values [10]. In the structure monitoring area, it cannot identify the presence of damage at early stages but also the severity of the damage [11]. Similarly, it reveals good efficiency in detecting and identifying faults such as short-circuit faults, open-circuit faults and partial shading faults in a photovoltaic system [12].

Control limits (time varying control limits and asymptotic control limits) of EWMA control chart and its setup parameters-smoothing coefficient and width parameter can affect its capability in fault change recognition. For example, time varying control limits are useful when the smoothing coefficient is small

(for example, less than 0.3) [13]. In addition, the EWMA control chart is more robust for smaller values of the smoothing coefficient [14]. The objective of the present study is to study the application of EWMA control chart on the VEs change recognition. The effects of control limits (time varying control limits and asymptotic control limits), the setup parameters and the inputs of the EWMA on VEs change recognition is also discussed, and this provides deep understanding of EWMA control chart in VEs change recognition. In this paper, VEs monitoring system is introduced in Section 2. Section 3 presents the EWMA control chart and its application process in VE change recognition. VEs data used for this research is shown in Section 4. Section 5 presents the results and discussion, followed by the conclusions in section 6.

II. VEs MONITORING SYTEM

Volumetric errors (VEs) are dependent on a wide range of machine components which make them potentially be able to provide a broad view of the machine condition. Let us take the five-axis machine tool with the error enriched kinematic diagram shown in Figure 1 as an example. It has axis alignment errors such as the axis location errors (EA0Y, EB0Z, EC0Y, EX0C, EA0B, EA0C, EB0C, EC0B), three linear gains (EXX1, EYY1, EZZ1) and two spindle offsets (EX0S, EY0S) potentially causing VEs in 3D space (Figure 1). Other intra-axis errors, within each individual axis, may also occur.

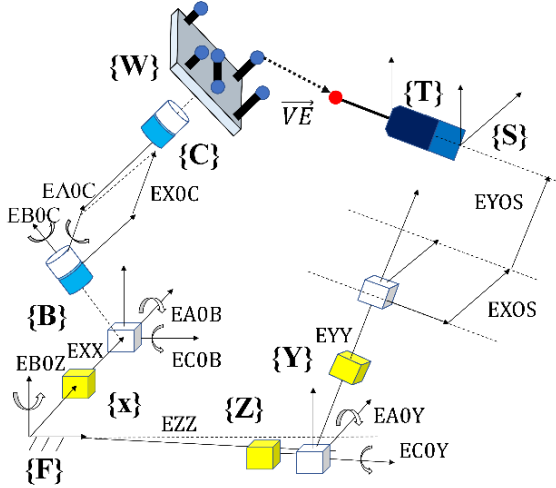
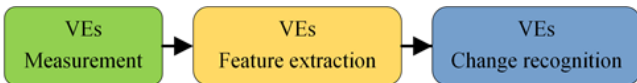


Figure 1. Illustration of the modelled geometric errors of the experimental five-axis machine tool with WCBXFZYST topology, where a tough trigger probe is installed on the spindle and some master ball artefacts are mounted on the machine workpiece table

The flowchart of the VEs data monitoring is shown in Figure 2. During the machine tool maintenance period, machine tool VEs are measured with the scale and master ball artefact (SAMBA) method [15]. Then, the VEs features are extracted with vector similarity measures (VSMs). Finally, these VE features will be processed by the exponentially weighted moving average (EWMA) control chart for VE change recognition. When the fault is detected, possible corrective actions for machine tool will be activated.



Identify applicable sponsor/s here. (sponsors)

Figure 2. Flowchart for VE monitoring system

The general VEs measurement method includes laser trackers, R-test and the Scale and Master ball artefact (SAMBA) method. Owing to its simple setup, the SAMBA method is selected for VEs acquisition. The raw probing data are acquired by on-machine probing a series of master balls and a scale bar with a tough trigger probe. The probing data are then processed using a kinematic model to estimate a set of 13 machine error parameters (EXX1, EB(0X)Z, etc.) as well as the positions of master ball artefacts which are the reference values for VEs calculation. Using our designed SAMBA measurement plan, 27 angular positions pairs of the B- and C-axes and 109 ball center positions are acquired for each test run. For each test run a set of $VE_{(i,j)}$ can be written as:

$$VE_{(i,j)} = (VE_{x(i,j)}, VE_{y(i,j)}, VE_{z(i,j)}) \quad (1)$$

where i stands for the VE measurement positions identifier ($1 \leq i \leq 109$, i is related to the SAMBA measurement plan) and j stands for the measurement repetition thus resulting in a time series of VEs with the “13” machine error model (Figure 3, a).

The concept of similarity refers to how alike two objects are. Usually, each object can be viewed as an N-dimensional vector with its feature components. Their similarity could be compared by vector similarity measures. There are about 60 different similarity measures with the most popular types being: distance-based similarity measure and angle-based similarity measure. The VSMs including the module (Modu) and Cosine (Cos) are used for VE feature extraction (Eq. (1) and (2)). The Modu measure is the Euclidean magnitude of the VE. The Cos measure calculates the angle between the first measured VE and the remaining VEs. They are calculated separately at each VE probing position. The geometric meanings of the two measures are shown in Figure 3. (b). Using these two measures, the VEs features could be extracted considering the magnitude and directional information. These extracted VE features VEF, are the inputs for the EWMA control chart for VEs change recognition.

$$VEF = \text{Modu}_{(i,j)} = \sqrt{VE_{x(i,j)}^2 + VE_{y(i,j)}^2 + VE_{z(i,j)}^2} \quad (2)$$

$$VEF = \text{Cos}_{(i,j)} = \cos(VE_{(i,j)}, VE_{(i,1)}) = \frac{VE_{(i,j)} \cdot VE_{(i,1)}}{\|VE_{(i,j)}\| \cdot \|VE_{(i,1)}\|} \quad (3)$$

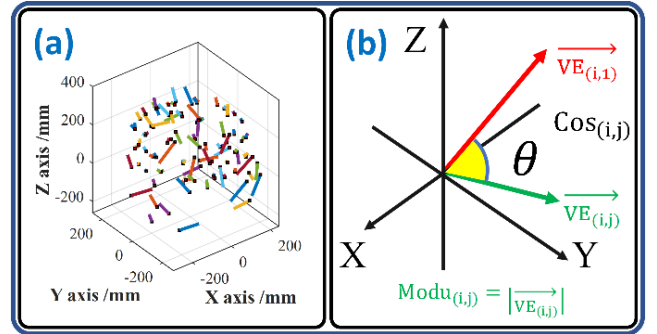


Figure 3. (a) VEs estimated with the SAMBA method in 3D space, amplified 1000x; (b) Geometric meanings of Modu and Cos measures

III. EWMA CONTROL CHART

For VEs monitoring, the objective is to detect the faults with gradual or sudden change. Usually, gradual change faults caused by the degeneration of machine tool account for an important percentage of machine tool faults [16]. The exponentially weighted moving average (EWMA) control chart was first introduced as a good alternative to the Shewhart control chart for detecting smaller shifts in the process parameters [17]. Therefore, an EWMA chart is a promising tool for VEs change recognition. EWMA chart for VEs change recognition is developed using a statistic of the following form:

$$NVEF_i = (1 - \gamma)NVEF_{i-1} + \gamma VEF_i \quad (4)$$

Eq. 4 is presented as a general form for a single VE measurement position, where $i \in [1, n]$, and it stands for the i th VE measurement or simulation and n is the number of VEs features to be monitored. Under the assumption of independent and normally distributed data, the time-varying control limits-upper control limit (UCL) and lower control limit (LCL) are expressed as follows.

$$UCL = NVEF_0 + L\sigma \sqrt{\left(\frac{\gamma}{2-\gamma}\right)(1 - (1-\gamma)^{2i})} \quad (5)$$

$$LCL = NVEF_0 - L\sigma \sqrt{\left(\frac{\gamma}{2-\gamma}\right)(1 - (1-\gamma)^{2i})} \quad (6)$$

where $NVEF_0$ is the expected mean value of the first k VEs features VEF. It is worth noting that the common values for the width parameter L and smoothing coefficient γ are $2.6 \leq L \leq 3$ and $0.05 \leq \gamma \leq 0.25$ [17]. When i increases, this variance can quickly converge to a steady state value. Then the asymptotic control limits can be generated by Eq. (7) and (8).

$$UCL = NVEF_0 + L\sigma\sqrt{\gamma/2-\gamma} \quad (7)$$

$$LCL = NVEF_0 - L\sigma\sqrt{\gamma/2-\gamma} \quad (8)$$

Theoretically, the recognition results of EWMA control chart are related to the selection of the control limits and its setup parameters (the smoothing coefficient and the width parameter). For the selection of control limits, with the increase of the VEs dataset, the asymptotic control limits could be considered by the users to replace the time-varying control limits. For the selection of the setup parameters, in this research, they are selected from the recommended range ($2.6 \leq L \leq 3$ and $0.05 \leq \gamma \leq 0.25$).

Herein, EWMA is used as a supervised learning method including two parts- learning and testing. In the learning process, the EWMA control chart is developed and the control limits-UCL and LCL are determined with labeled VEs features. Owing to the limitation of the number of VEs features, the two kinds of control limits are all applied and discussed. In the testing process, the new acquired VE data $NVEF_i$ will be inputted into the EWMA model for condition monitoring. When $NVEF_i$ is within UCL and LCL, machine tool accuracy condition is considered stable and under control. Otherwise, some necessary maintenance work is needed.

There are several ways to calculate newly acquired VE data $NVEF_i$. The calculation process is shown in Figure 4. Firstly, it can be calculated with the previous VEF_1 to VEF_{i-1} (Figure 4).

This is the usual way for the calculation of $NVEF_i$ (method I). Alternatively, $NVEF_i$ (method II) can also be calculated with the new VEF_i and the k VEF of learning data (Figure 4). For monitoring, the two types of $NVEF_i$ may both functions. The only difference is how many test data $NVEF_i$ uses. $NVEF_i$ (method I) uses $i-1$ VEF while $NVEF_i$ (method II) uses $k+1$ VEF including each newly acquired VE and the learning data ($i > k$). It is worth noting that the application of $NVEF_i$ calculated by method I/II in VEs change recognition are also studied.

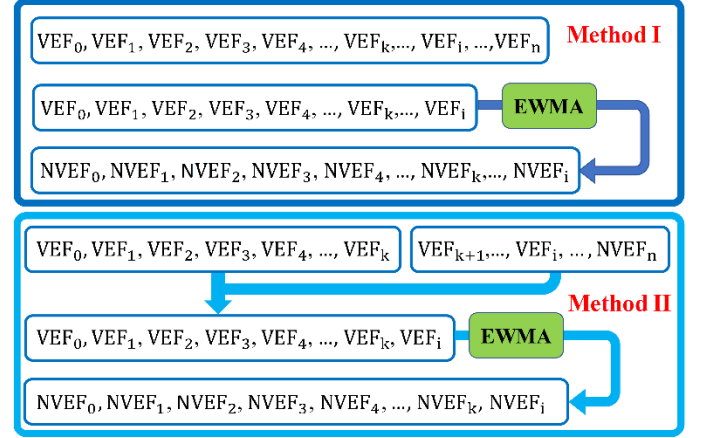


Figure 4. The calculation of $NVEF_i$ using method I and method II

IV. VEs SOURCES

Real and persuade faults are used to discuss the mentioned questions. Periodic experimental tests on the machine tool provide the VEs data with real fault. The raw volumetric information is acquired from the HU40-T machine tool with a MP700 Renishaw touch trigger probe and series of master ball artefacts and scale bar artefact, the coordinates of master ball artefacts are measured under 27 angular positions pairs of the B- and C-axes and 109 ball center positions are recorded (Figure 5). The probing data are then processed with the “13” machine error model for VEs calculation (Procedure a). During the test phase, a fault developed by C-axis encoder causing significant ECC error has been detected. 12 SAMBA testing results before and after this fault are used for C-axis encoder fault representation.

Two pseudofaults are generated based on the same SAMBA model and testing plan of the periodical tests. Machine tool error compensation function is designed to correct the increased geometric errors of one linear or rotary axis. Here, it is used as a non-destructive tool to simulate a fault caused by the change of X axis positioning error. A U-shape error with magnitudes of 35um have been added into the pitch error compensation table to generate the pseudofaults. The SAMBA tests are repeated for six times before and after error injection (Figure 5, procedure b).

Procedure c of Figure 5 reveals the modelling process of the pseudofaults caused by the change of X axis straightness error in Y of the X-axis, EYX. It can make the movement of X axis miss its normal destination position. For the SAMBA measurement, this inaccuracy can be revealed in the Y coordinate of master ball artefacts. The straightness error compensation table is not activated in HU40-T machine tool; therefore, this error is generated by manually modifying the raw probing results. An error (Figure 6) is added to the Y coordinate

of the measured position of the master balls as a function of the X-axis x position. Then, the pseudofaults EYX is generated.

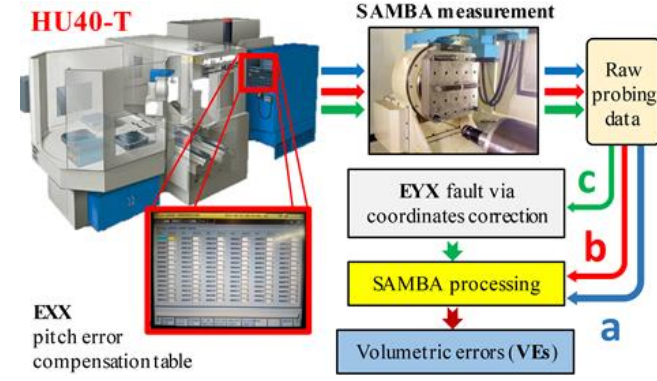


Figure 5. VEs data sources for this research

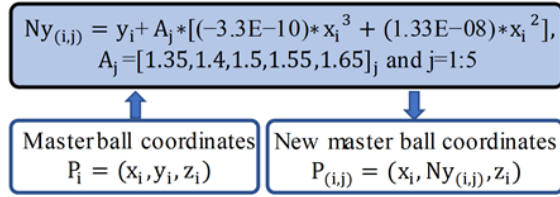


Figure 6. Pseudofaults caused by the change of straightness error by manual correction [6]

To further discuss the effect of the EWMA set up parameters, control limits and the calculation of $NVEF_1$ on VEs change recognition results, simulated faults with gradual changes (type A and type B) are introduced by numerical modification of machine error parameters (ISO230-1). The difference between type A and B is the error factor magnitude and the existing of the change of error factor sign. A SAMBA simulator software including the “13” machine error model is used for this purpose (Figure 7).

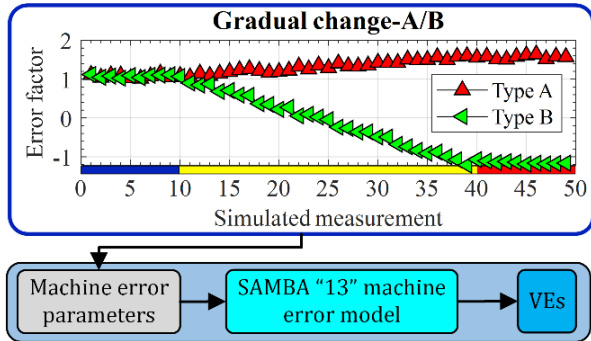


Figure 7. Simulated faults caused by the gradual change (type A and type B) of machine error parameters; The blue, yellow and red bars stand for the normal states, transition states for gradual changes and faulty states, respectively; Some randomness is added for a “more realistic” effect; The error factor is an amplifier of each machine error parameter to control the changing shape of the simulated fault.

V. RESULTS AND DISCUSSION

The acquired VEs are classified as two parts-learning and testing based on the known fault occurrence time. The learning data is used for EWMA modeling, the testing data inputted into the EWMA model for the abnormal change detection. The

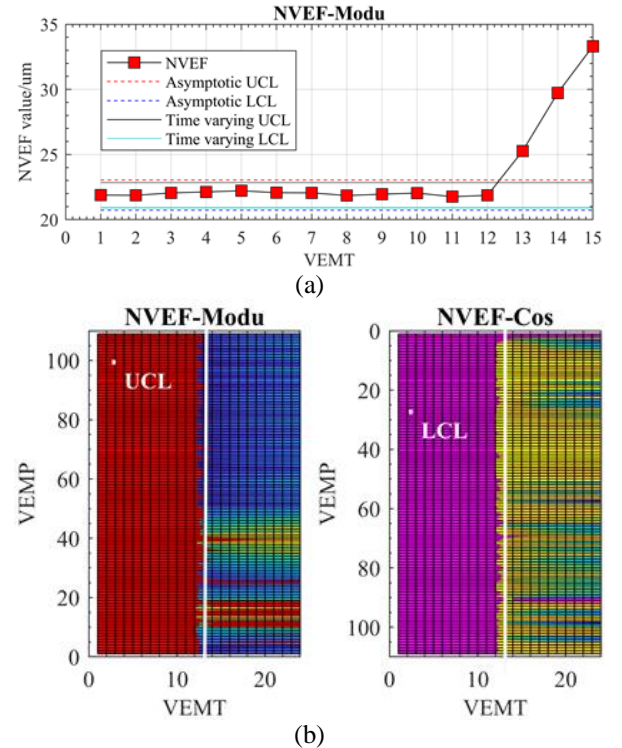
recognition process is repeated in each VE probing position (109 times). The recognition rate (RR) is calculated for the discussion of the effect of the control limits and setup parameters of EWMA on final recognition results. RR is defined as the ratio between the total number of the successful recognition positions of an abnormal behavior (TNSRP), where the detected change occurrence time is equal to the known fault occurrence time, and the total VE measurement positions (TVEMP=109) (Eq. (9)).

$$RR = \frac{TNSRP}{TVEMP} * 100\% \quad (9)$$

A. The effect of control limits on final recognition results

The time varying control limits and asymptotic control limits of the real and pseudo-faults are all calculated for the VEs change recognition. Theoretically, the time varying control limits are suitable for limited amounts of VEs data. With the increase of the amount of VEs data, the time varying control limits will become or even closing to the asymptotic control limits. The VEs features extracted by Modu and Cos are used as the inputs of the EWMA for fault detection.

The first measured 12 VEs are used for EWMA modeling. Figure 8 reveals the control limits of time varying and asymptotic towards the C axis encoder fault where the setup parameter $\gamma=0.05$ and $L=2.6$. Figure 8 (a) shows the recognition process of the C axis encoder fault using two kinds of control limits. When using small size VEs data, the control limits are close but the time varying control limits have narrow monitoring range. This can promote the VEs change recognition. This fault can be mostly detected at the 13th measurement time by visual checking (Figure 8 (b) and (c)), there are no obvious differences in the number of VEs measurement positions where this fault can be successfully detected.



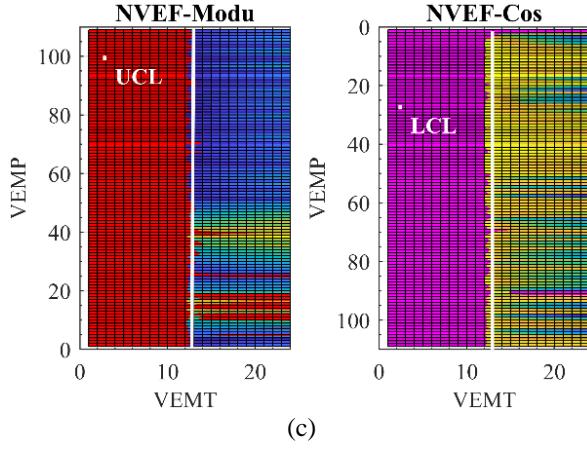


Figure 8. (a) Recognition result of C axis encoder fault in the 1st VE measurement position, in order to show the control limits well, only 15 VEs are shown; VEMT means volumetric errors measurement times and VEMP means volumetric errors measurement position (the same definition for the following figures); (b) Recognition result of this fault using asymptotic control limits; (c) Recognition result of this fault using time varying control limits.

The recognition rates of real and pseudofaults using asymptotic and time varying control limits are calculated and shown in Table 1. Using the same setup parameters and VEs features, the time varying control limits can improve the recognition rate by about 2%~5% when compared with asymptotic control limits. Therefore, the time varying control limits are recommended for VEs change recognition when the learning samples are limited.

Table 1. RR of VEs using different control limits

Faults	Asymptotic control limits		Time varying control limits	
	Modu	Cos	Modu	Cos
C axis encoder fault	70%	93%	74%	95%
Linear positioning error	45%	95%	62%	97%
Straightness error	58%	100%	67%	100%

B. The effect of setup parameters on final recognition results

Smoothing coefficient γ and width coefficient L directly affect the performance of EWMA control chart. Theoretically, the closer γ is to 1, the more the EWMA chart resembles a Shewhart chart (When $\gamma = 1$, the EWMA chart is equal to Shewhart chart). When γ is near 0, a small weight is applied to almost all the past observations, then the performance of the EWMA chart parallels that of a cusum chart. For our dataset containing faults, the time varying control limits are used for the discussion of the setup parameters on final recognition results.

The recognition results of the C axis encoder fault, pseudofaults caused by linear positioning errors and straightness errors are discussed separately using the values of L and γ in the ranges of $2.5 \leq L \leq 3$ and $0.05 \leq \gamma \leq 0.25$ (Figure 9). The step increase of the two parameters- L and γ are 0.05 and 0.01 respectively. For the NVEF-Modu, the change ranges of recognition results are around 4% (68%~72%), 35% (27%~62%) and 30% (39%~69%). For the NVEF(Cos), the change ranges of recognition results are around 4% (95%~99%),

2% (95%~97%) and 0% (100%~100%). Figure 9 illustrates that Modu and Cos measures have different capabilities in VEs change detection. At the same fault, Cos performs better than Modu. In addition, the setup parameters directly affect the final recognition results.

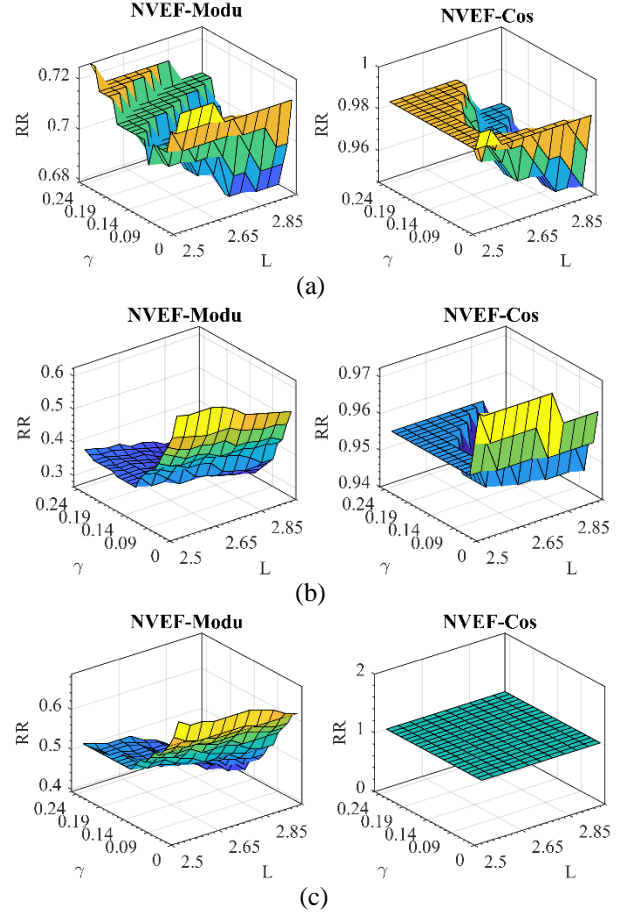


Figure 9. (a) Recognition rate (RR) of C axis encoder fault using VEs feature extracted by Modu and Cos measures; (b) RR of the pseudofault caused by linear positioning error; (c) RR of the pseudofault caused by straightness error.

For the same width coefficient setup, a smaller smoothing coefficient is helpful in improving the recognition rate. Similarly, a smaller width coefficient can also improve the recognition rate. Therefore, in the recommended ranges of the setup parameters $2.6 \leq L \leq 3$ and $0.05 \leq \gamma \leq 0.25$, small smoothing coefficients and width coefficients are recommended for a better VEs change recognition.

C. The effect of NVEF_i calculation on final recognition results

The NVEF_i are calculated according to the above methods I/II (Figure 4). Based on the results of the above analysis, the smoothing coefficient γ and width coefficient L are selected as 0.05 and 2.6. The time varying control limits are used for VEs changes recognition.

Figure 10 reveals the recognition results of the C axis encoder fault using NVEF_i calculated by method I/II. From Figure 10 (a), one can see that there are no differences in the recognition results of VE acquired in the 1st measurement position when using NVEF_i (method I/II). At the 23th VE

measurement position (Figure 10 (b)), the first detected VE change point is the 14th with one-time delay (the exact fault occurrence time is the 13th). After the fault change time, $NVEF_i$ (method I) reveals the VE change clearly. However, for $NVEF_i$ (method II), although the VEs data are measured in the faulty state, they could be wrongly recognized as the normal state (Figure 10 (b)). This could mislead the machine tool user. The recognition results of all VEs measurement positions are revealed in Figure 10 (c and d). At the 13th position, most faults could be detected. However, under some cases (1~10 VEs measurement positions), although the exact fault occurrence time can be detected by $NVEF_i$ (method II) the VEs in the faulty state could be wrongly classified as normal state (Figure 10 (c and d), yellow block).

The recognition rate (RR) of real and pseudo-faults using asymptotic and time varying control limits are the same as the RR value calculated by time varying control limits (Table 1). This means that the two methods of $NVEF_i$ are all effective in the detection of the first change point of the faulty state. However, owing to the presence of the noise in VEs data, the faulty state data could be also be wrongly classified as the normal state when using $NVEF_i$ (method II). For the $NVEF_i$ (method I), the faulty state could be easier detected. In addition, using $NVEF_i$ (method II), the stability of the faulty state data could be revealed. This is its advantage that $NVEF_i$ (method I) does not have.

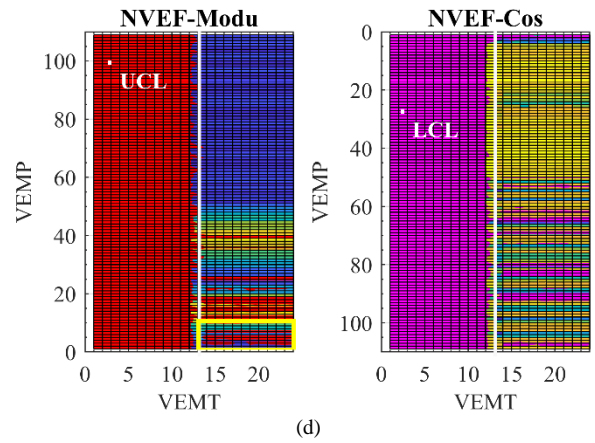
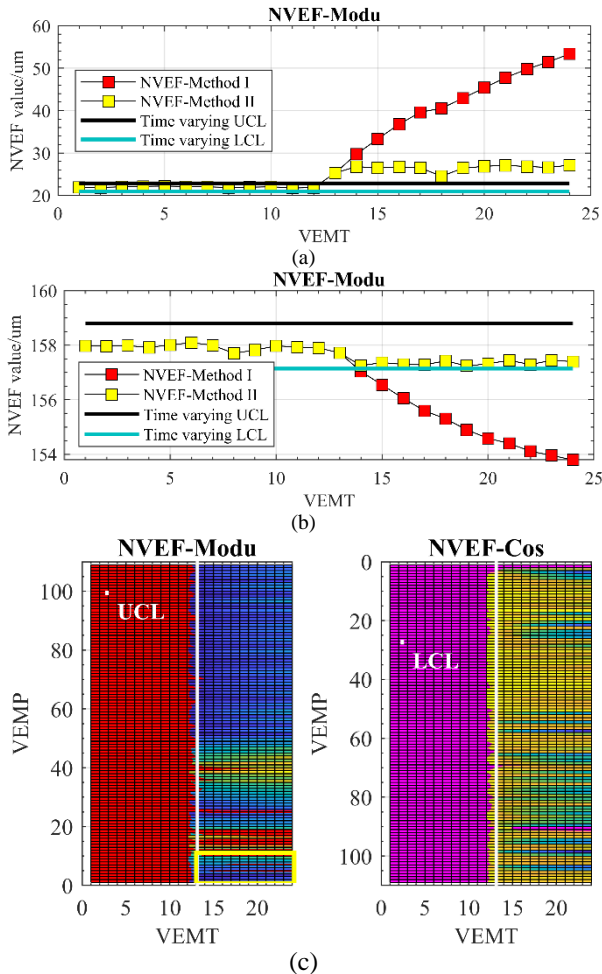


Figure 10. (a) Recognition result of C axis encoder fault in the 1st VE measurement position using $NVEF_i$ calculated by (method I/II); (b) Recognition result of this fault in the 23th VE measurement position; (c) Recognition result of this fault of all VE measurement positions using $NVEF_i$ (method I); (d) Recognition result of this fault of all VE measurement positions using $NVEF_i$ (method II).

D. Verification of the above findings using the simulated faults

The two types of simulated gradual change faults (type A and type B) are separately processed with the time varying control limits and asymptotic control limits, $NVEF_i$ calculated by method I/II and different setups of EWMA setup parameters.

Figure 11 shows the recognition results of the two types of gradual change faults caused by the change of the EBOC and EZZ1 errors using the time varying control limits and asymptotic control limits. The values of the two kinds of control limits are close, and the time varying control limits are narrower than the asymptotic control limits. The EBOC and EZZ1 faults are detected at the 16th and 12th measurement in each probing position. In some VEs measurement positions, the $NVEF_i$ is smaller than the LCL, therefore, the change is not shown in Figure 11 (b) and (d). Figure 11 (e) also reveals that the combination of small smoothing coefficients and width coefficients are helpful for VEs change recognition.

Figure 12 presents the recognition results of the gradual change faults caused by the change of the EZZ1 error using $NVEF_i$ (method I/II). The exact fault occurrence time could be detected by the $NVEF_i$ (method I/II). However, the fault can only be detected at this measurement when using $NVEF_i$ (method I/II). $NVEF_i$ (method I) reveals the fault change at the 13th measurement. Similarly, in some VEs measurement positions, the $NVEF_i$ is smaller than the LCL, therefore, the change cannot be shown in Figure 12 (b) and (c). The change tendency of the $NVEF_i$ (method II) can reveal the transition and the faulty states while it is hard to see this from $NVEF_i$ (method I) (Figure 12 (a)). For the remaining simulated faults, the same data processing has been applied into each fault. The results all prove the above finds.

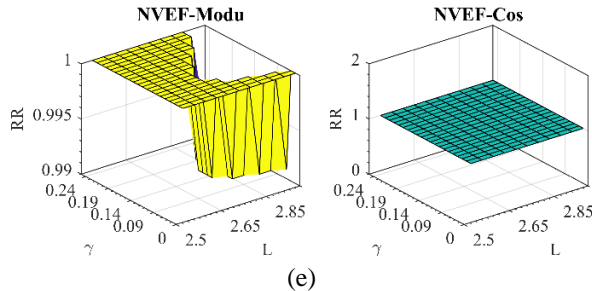
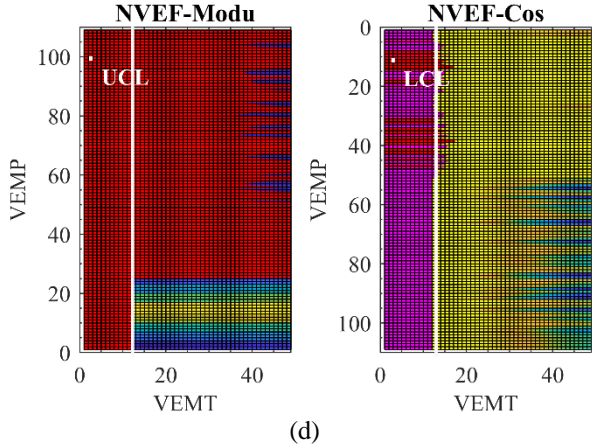
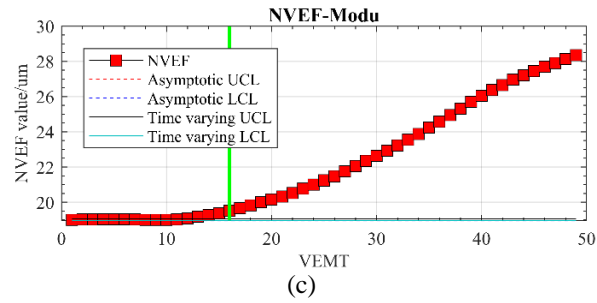
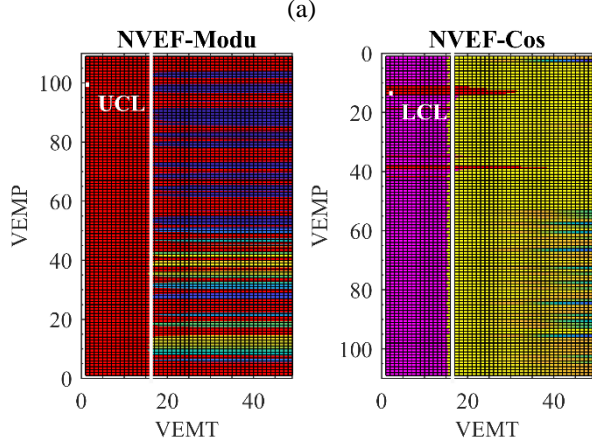
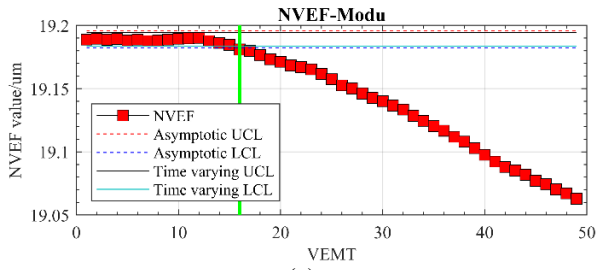


Figure 11. (a) Recognition result of type-A gradual change fault caused by EB0C in the 1st VE measurement position; (b) Recognition result of EB0C fault from all VEs measurement positions with the asymptotic control limits and time varying control limits; (c) Recognition result of type-B gradual change fault caused by EZZ1 in the 1st VE measurement position; (d) Recognition result for EZZ1 fault from all VEs measurement positions with the asymptotic control limits and time varying control limits; (e) Recognition rate (RR) of the simulated EB0C fault with the setup of using the values of L and γ in the ranges of $2.5 \leq L \leq 3$ and $0.05 \leq \gamma \leq 0.25$.

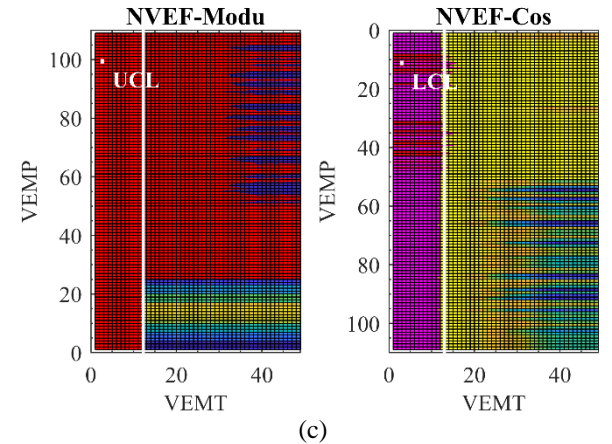
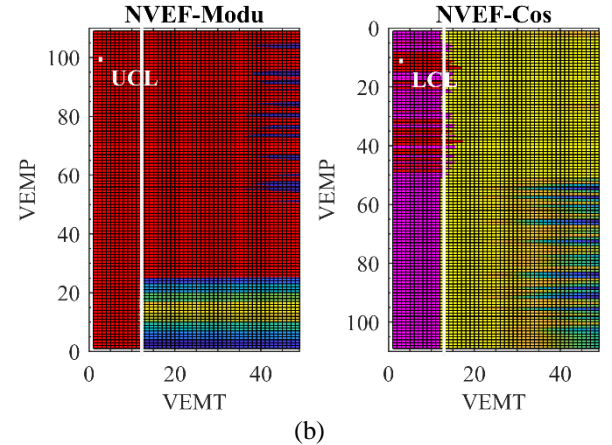
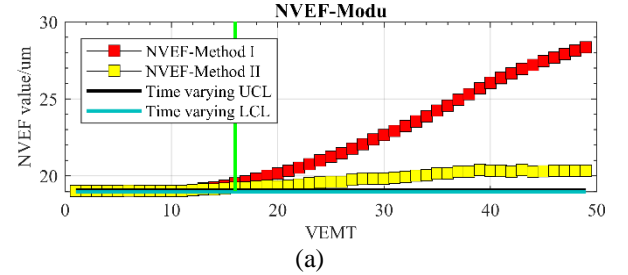


Figure 12. (a) Recognition result of type-B gradual change fault caused by EZZ1 in the 1st VE measurement position using $NVEF_i$ (method I/II); (b) Recognition result of this fault from all VEs measurement positions with $NVEF_i$ (method I); (c) Recognition result of this fault from all VEs measurement positions with $NVEF_i$ (method II).

E. Discussion

The EWMA control chart performed well in VEs changes recognition. It is able to detect the sudden change fault (to some extent, the real and pseudo-faults are with a sudden change tendency) and the simulated gradual change faults (type A and type B). The setup parameters smoothing coefficient and width

parameter can affect the recognition capability of EWMA control chart. Selection of relatively small smoothing coefficient and width parameter can improve the recognition rate. Although the time varying and asymptotic control limits are close, the time varying control limits can promote the recognition rate of VEs faults when using limited VEs data for learning. For the $NVEF_i$ calculated by method I/II, they are all effective in the detection of first fault change position. However, $NVEF_i$ calculated by method II may mistakenly classify the faulty state data as a normal state. This can make maintenance planning more complex for the operators of the machines. So, it is not recommended to be used solely for VEs change recognition. However, under the application of $NVEF_i$ calculated by method I, $NVEF_i$ (method II) can be used as a additional tool to check the stability of faulty states. For the gradual change fault (type B), $NVEF$ -Cos does not perform well because the estimated VEs by the SAMBA simulator have a linear relationship with the modelled machine error parameters. Therefore, the changes are rarely reflected in the Cos measure.

VI. CONCLUSIONS

This study examines the performance of EWMA control chart using asymptotic, time varying limits and different setups of smoothing and width coefficients and two types of $NVEF_i$ calculation for VEs change recognition. We conclude that shift detection ability of EWMA chart can be improved by using time varying limits instead of asymptotic control limits. Smaller values of smoothing parameter and width parameter ($2.5 \leq L \leq 3$ and $0.05 \leq \gamma \leq 0.25$) are recommended for VEs change recognition when the learning data set is small. $NVEF_i$ calculated by method I has good performance in VEs change recognition; $NVEF_i$ calculated by method II could be used as a second tool for the stability checkup of the faulty state data.

ACKNOWLEDGEMENT

The authors would like to thank the technicians Guy Gironne and Vincent Mayer for conducting the experimental part of this work. This research presented in this paper was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) under the CANRIMT Strategic Research Network Grant NETGP 479639-15 and China Scholarship Council (No.201608880003).

REFERENCE

- [1] A. Al-Habaibeh, G. Liu, and N. Gindy, "Sensor Fusion for An Integrated Process And Machine Condition Monitoring System," IFAC Proceedings Volumes, vol. 35, no. 1, pp. 25-30, 2002.
- [2] D. Goyal and B. S. Pabla, "Development of non-contact structural health monitoring system for machine tools," Journal of Applied Research and Technology, vol. 14, no. 4, pp. 245-258, 2016.
- [3] Leh-Sung Law, Jong Hyun Kim, Willey Y.H. Liew, Sun-Kyu Lee, "An approach based on wavelet packet decomposition and Hilbert–Huang transform (WPD–HHT) for spindle bearings condition monitoring," Mechanical Systems and Signal Processing, vol. 33, pp. 197-211, 2012.
- [4] R. Teti, K. Jemielniak, G. O'Donnell, and D. Dornfeld, "Advanced monitoring of machining operations," CIRP Annals, vol. 59, no. 2, pp. 717-739, 2010.
- [5] Gregory W. Vogl, M. Alkan Donmez, Andreas Archenti, "Diagnostics for geometric performance of machine tool linear axes," CIRP Annals, vol. 65, no. 1, pp. 377-380, 2016.
- [6] M. Givi, J. R. R. Mayer, "Volumetric error formulation and mismatch test for five-axis CNC machine compensation using differential kinematics and ephemeral G-code," The International Journal of Advanced Manufacturing Technology, vol. 77, no. 9-12, pp. 1645-1653, 2014.
- [7] Kanglin. Xing, Sofiane. Achiche, J.R.R. Mayer, "Five-axis machine tools accuracy condition monitoring based on volumetric errors and vector similarity measures," International Journal of Machine Tools and Manufacture, vol. 138, pp. 80-93, 2019.
- [8] M. McGill, "An Evaluation of Factors Affecting Document Ranking by Information Retrieval Systems," School of information studies, 1979.
- [9] Kanglin. Xing, J.R.R. Mayer, Sofiane. Achiche, "Machine Tool Volumetric Error Features Extraction and Classification Using Principal Component Analysis and K-Means," Journal of Manufacturing and Materials Processing, vol. 2, no. 3, p. 60, 2018.
- [10] S. W. Roberts, "Control Chart Tests Based on Geometric Moving Averages," Technometrics, vol. 42, no. 1, pp. 97-101, 2000.
- [11] P.Srinivasa Rao, Ch.Ratnama, "Damage identification of welded structures using time series models and exponentially weighted moving average control charts," Jordan J. Mech. Ind. Eng, vol. 4, no. 6, pp. 701–710, 2010.
- [12] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, "A statistical-based approach for fault detection and diagnosis in a photovoltaic system," International Conference on Systems and Control (ICSC), pp. 75-80, 2017.
- [13] Stefan H. Steiner, "EWMA Control Charts with Time-Varying Control Limits and Fast Initial Response AU - Steiner, Stefan H," Journal of Quality Technology, vol. 31, no. 1, pp. 75-86, 1999.
- [14] P. Kritzing, S. Chakraborti, "Robustness of the EWMA control chart for individual observations AU - Human, S. W," Journal of Applied Statistics, vol. 38, no. 10, pp. 2071-2087, 2011.
- [15] J.R.R. Mayer, "Five-axis machine tool calibration by probing a scale enriched reconfigurable uncalibrated master balls artefact," CIRP Annals, vol. 61, no. 1, pp. 515-518, 2012.
- [16] Y. G. Saravanan S, Rao PV, "Machine tool failure data analysis for condition monitoring application," in the 11th National Conference on Machines and Mechanism, New Delhi, 2003, pp. 552–558: Pvt. Limited.
- [17] D. C. Montgomery, Statistical Quality Control: A Modern Introduction, 6th ed. New York: John Wiley & Sons, 2005.

ORTHOGONAL MACHINING OF PURE MAGNESIUM-SILICA SYNTACTIC FOAM

Abdulla Alhourani

Department of Mechanical Engineering
American University of Sharjah P.O. Box 26666 Sharjah,
United Arab Emirates
b00057903@alumni.aus.edu

Sathish Kannan

Department of Mechanical Engineering
American University of Sharjah
P.O. Box 26666 Sharjah, United Arab Emirates
skannan@aus.edu

Salman Pervaiz

Department of Mechanical and Industrial Engineering
Rochester Institute of Technology – Dubai Campus
P.O. Box 341055 Dubai, United Arab Emirates
sxpcad@rit.edu

Abstract — Metal syntactic foams are a new grade of novel composites in which hollow ceramic microspheres are dispersed through the magnesium metal matrix. For a spectrum of applications, in the industries of interest and applicability, the chosen material needs to have good corrosion resistance, mechanical strength and energy absorption capabilities under compression loading. This dire requirement necessitated the need for an investigation on SiO₂ hollow microsphere-reinforced magnesium syntactic foams. The machining of Metal Matrix Syntactic Foams (MMSFs) is very novel and is significant to industrial products. In comparison to polymer matrix syntactic foams and metal matrix composites, MMSFs exhibit lower densities and high impact absorption capabilities. The reinforcement of silica micro balloons in the Magnesium (Mg) matrix significantly improves the material properties such as compressive yield strength and ultimate compressive strength. In order for wide spread application of these novel syntactic foam materials will require significant machining processes to be carried out to produce the final product.

This research presents for the first time, a machining investigation on a pure Magnesium reinforced with SiO₂ hollow bubble metal syntactic foam. This work studies the effect of SiO₂ hollow sphere volume fraction and the cutting parameters such as cutting speed and uncut chip thickness on machining forces generated and chip formation under dry cutting conditions. Orthogonal cutting tests were performed on pure Mg, 5%, 10%, 15% volume fraction SiO₂ hollow bubbles reinforced in pure magnesium matrix. The test results showed an increase in cutting force with increasing undeformed chip thickness and a decreasing trend with increasing cutting speed. An increase in reinforcement volume fraction results in higher

cutting forces with more pronounced discontinuous chip formation at higher volume fractions.

Keywords-Syntactic Foam, Magnesium, Machining

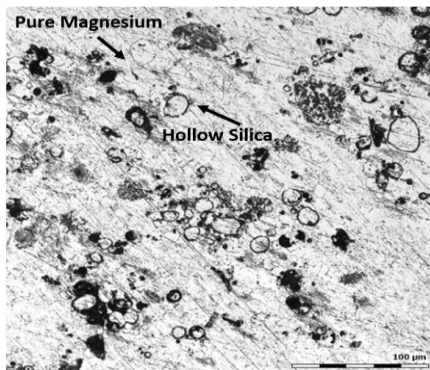
I. INTRODUCTION

MMSFs are closed cell composite foams, which contain microspheres dispersed throughout the metal matrix in appropriate volume fractions Figure 1 [1]. The presence of such reinforcements imparts the required mechanical properties. Several researchers have shown from their work that such material exhibits properties including high specific strength, low density, light in weight and high-energy absorption [1-3]. Primary synthesis and characterization studies on Al, Mg, Ti, Cu, Pb, Zn alloy matrix syntactic foams are available in the published literature [1-5]. Magnesium based syntactic foams find applications in biomedical applications such as bone scaffolds, bone implantations (plates, staples, screws, rods, prosthesis), high temperature energy absorptions, catalysts, crash worthiness, light weight sandwich structures and high temperature heat exchangers [1-5]. Among modern composite materials, magnesium based syntactic foams are of interest to automotive and biomedical sectors for it to be implemented as the hollow spheres increase the material strength [6-9].

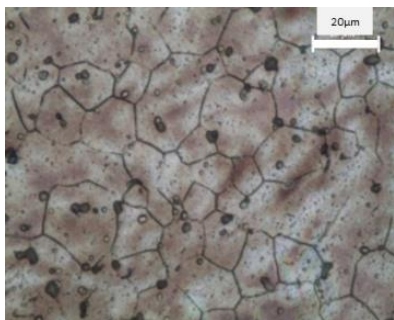
For similar percentage of porosity, the modulus and strength of closed cell syntactic foams are generally higher than open- cell foams [10]. In the recent past, several researchers have studied the synthesis methods, microstructures, mechanical properties and applications of light weight metal foams reinforced with glass hollow

particles, fly-ash cenospheres as filler materials [10-14]. An increase in density of the foam is observed with an increase in modulus or strength for both polymer and metal syntactic foams [10-14]. Carbon, silicon carbide and alumina hollow particles have also been tried to enhance the properties outside of commonly used foams [15-18].

Very few studies have been carried out on the fabrication techniques and characterization of Magnesium- SiO₂ hollow metal matrix syntactic foams. For a wide spread application of these novel material systems and transforming them from their near-net shapes in to a useful final product will require appropriate development of machining methods. The biomedical products manufactured using magnesium syntactic foams will require final conventional machining, such as drilling, tapping, reaming, and milling operations for making features such as bolt holes, threads, screws, grooves, slots etc. The biomedical implants comprise of several threaded holes, slots, pockets, contour milling, profiles, etc. which require highest quality of surface finish and integrity with minimal machining induced defects.



(a) Micro silica hollow bubbles



(b) Nano silica hollow bubbles

Figure 1. SEM Microstructure of a.) Closed cell Magnesium-Silica foam (b) Magnesium- SiO₂ hollow nanosphere syntactic foam manufactured at National University of Singapore (DMD process)

Unconventional processes such as Electrical Discharge Machining (EDM) and Laser Beam Machining (LBM) could cause heat affected zones on the machined magnesium bone implant surfaces. These processes may not be appropriate for finish machining these surface sensitive biomedical products which may affect their functional performance as bone implants inside human body (e.g. accelerated corrosion). This could result in implant loosening or even catastrophic damage

requiring a second operation causing severe discomfort to the patient. Currently, there is no literature available on the machining methods and sustainable lubrication strategies for manufacturing biomedical components from Magnesium-SiO₂ hollow sphere metal matrix syntactic foams. In this paper, machining characteristics of pure Magnesium matrix syntactic foams reinforced with varying volume fractions (5%, 10% and 15 vol%) of hollow silica bubbles is presented. Orthogonal cutting tests were conducted to study the effect of cutting speed, undeformed chip thickness and different volume fractions of silica hollow bubbles on generated cutting forces and chip formation under dry cutting conditions.

II. EXPERIMENTAL SETUP

A. Material

In order to experimentally investigate the machining characteristics of magnesium syntactic foams, squeeze casting method has been developed by SWAMEQUIP, Chennai, India to cast the billets for the machining trials (Figure 2). The Silica hollow bubbles are of average 9 -15 microns in size bought from Sigma-Aldrich.

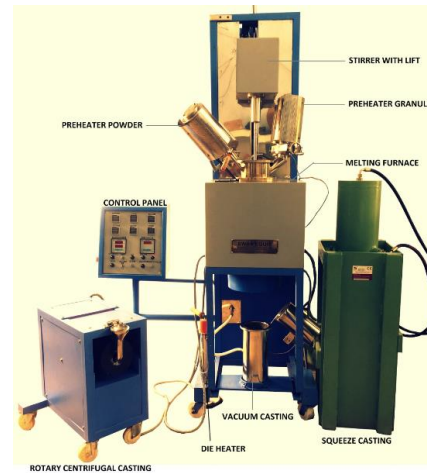


Figure 2. Squeeze casting set up

Material properties of squeeze casted pure magnesium foam billets are presented in Table 1.

TABLE 1. MATERIAL PROPERTIES OF MAGNESIUM FOAM.

Material	Hardness (Brinell, BHN)	Compressive Yield Strength (MPa)	Ultimate Compressive Strength (MPa)	Density (g/cm ³)
Pure Mg- 5% SiO ₂	39.93	165	228	1.812
Pure Mg- 10% SiO ₂	51.4	193	244	1.834
Pure Mg- 15% SiO ₂	60.96	217	261	1.867

B. Orthogonal Cutting

The orthogonal cutting tests were conducted using a three axis DOOSAN DNM 4500 CNC milling machine. Figure 3 shows the machining setup and cutting force measurement. All tests were conducted in a dry environment where the uncut chip thickness, cutting speed and reinforcement volume fraction were varied. KISTLER 9129AA three channel dynamometer was employed alongwith a multichannel charge amplifier type 5080 to measure the machining forces.

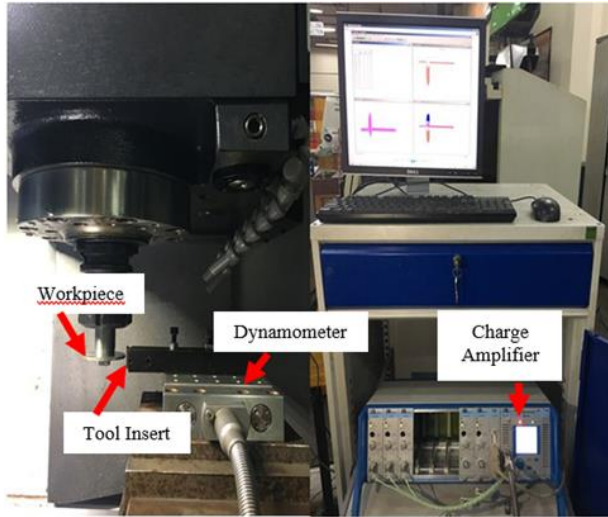


Figure 3. Orthogonal cutting set up and data acquisition

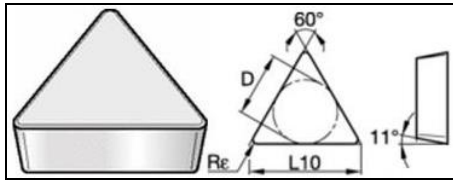


Figure 4. Cutting tool and geometrical specifications of the insert.

Coated carbide cutting insert used for the test was TPGN220404- TPG431- KCU10 manufactured by KENAMETAL is shown in Figure 4. Insert was coated with Aluminum Titanium Nitride (AlTiN) coating by a PVD (Physical Vapor Deposition) process. In Figure 4, the tool insert geometry specifications are presented in which thickness, L10 (edge length), D (diameter), cutting edge radius, Re (nose radius) are 4.76 mm, 22 mm, 12.7 mm, 0.08 mm and 0.4 mm, respectively with a flat top and tool inclination angle of 11°.

Cutting conditions used in this study is shown in the Table 2.

TABLE 2. MACHINING PARAMETERS.

Process Parameters and conditions		
Cutting Speed	m/min	60, 90, 120
Undeformed chip thickness	mm	0.05, 0.1, 0.175, 0.2
Width of cut	mm	3
Coolant	-	Dry
Tool Material	-	Coated Carbide

III. RESULTS AND DISCUSSION

A. Effect of Cutting Speed

Machining force is one of the most important parameters in the orthogonal cutting process which is related to the chip formation process. Magnitude of cutting force helps to determine the energy consumed in the cutting process and understand deformation behaviour of the material system. It also directly affects the friction condition, which is related to tool wear, tool life and has a direct impact on the machined product quality and integrity. Therefore, variation of cutting force will help to analyse the process parameters which is important from the actual part production perspective. Figure 5 shows the effect of cutting speed on the magnitude of thrust force and tangential force generated during cutting pure magnesium reinforced with 15% SiO₂ using a coated carbide insert with a tool cutting edge radius of (~0.08mm). As it can be seen from the results both thrust and tangential forces decrease with increasing cutting speed. This result while cutting metal syntactic foam shows similarity with earlier research conducted on magnesium based alloys [19, 20]. The magnitude of thrust forces were higher than the tangential forces under all cutting speed conditions. The cutting edge radius of the cutting insert used was approximately 30 microns. This could have potentially introduced the ploughing effect along with a considerably smaller undeformed chip thickness values used in the test resulted in higher thrust forces. It can be observed that both forces decrease with increasing cutting speed in the range of 60–120 m/min, and the highest cutting forces are recorded at lower cutting speed. As the cutting speed starts increasing, heat generation in the process due to shearing, friction and ploughing increases resulting in thermal softening of the workpiece with reduction in machining forces.

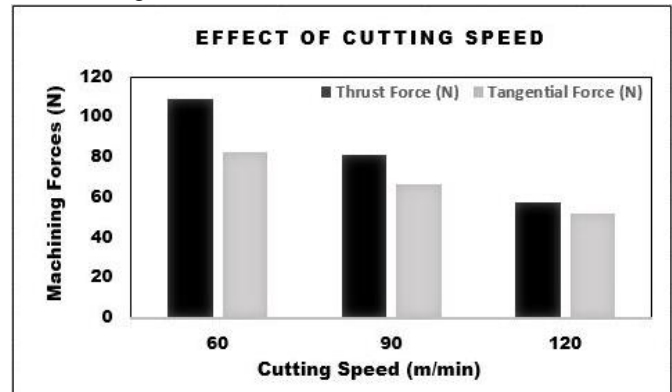


Figure 5. Effect of cutting speed on cutting forces (Undeformed chip thickness: 0.05 mm, Material: Pure Mg-15%SiO₂)

B. Effect of Undeformed Chip Thickness

In order to study the effect of feed rate on the machining forces, 4 different values of undeformed chip thickness were tested under constant cutting speed of 60 m/min and volume fraction of 10% SiO₂.

It is seen from the test results (Figure 6) that thrust and tangential forces increase with increasing feed rate. This confirms with machining similar magnesium based materials [19, 20]. Although this increase is approximately linear in first and second tests, it tends to plateau in third and fourth tests (Fig. 6). In order to map out the trend of increasing forces with increasing undeformed chip thickness, follow on tests will be conducted with higher undeformed chip thickness values from 0.3 to 0.8 mm. The recorded magnitude of thrust forces are higher than the tangential forces which is influenced by the ratio of the cutting edge radius to undeformed chip thickness that causes rubbing between the tool and workpiece and leads to higher thrust force values than the cutting force. The highest value of thrust force was measured as 201 N in the last test while the least value was measured as 100 N in the first test. However, the tangential force values were less than the thrust force values by at least 30N for all tests. The feed motion and the cutting motion occur simultaneously and along with the ploughing effect due to tool cutting edge radius causes the magnitude of thrust force to be more affected by cutting motion during orthogonal metal removal process.

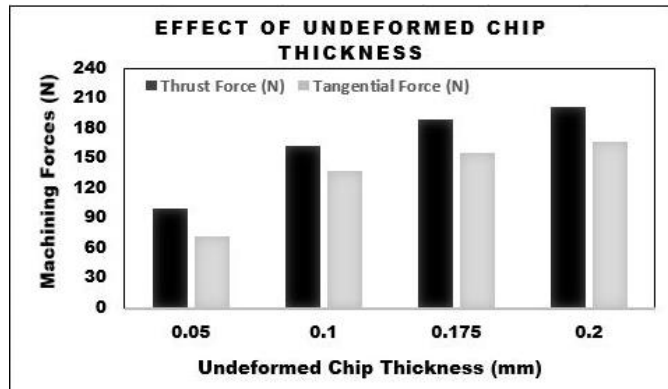


Figure 6. Effect of undeformed chip thickness on cutting forces (Cutting speed: 60 m/min, Material: Pure Mg-10%SiO₂)

C. Effect of SiO₂ Reinforcement Volume Fraction

SiO₂ hollow reinforcements in the magnesium matrix enhances the compressive yield strength, hardness and the compressive strength of the material system (Table 1). As the volume fraction of SiO₂ hollow reinforcements increase, the compressive properties of the material also increase. During cutting metal composites, it has been shown that the majority of the energy is spent to plastically deform the work hardened matrix that is pinned between the hard-ceramic reinforcements in the primary shear zone. A portion of the energy is spent in overcoming the friction on the rake face where the hollow damaged ceramic particles rub against the tool rake face causing secondary deformation. In addition, the presence of tool cutting edge radius and lower undeformed chip thickness

means ploughing effect will be significant in the cutting process while the cutting edge engages with the micron size hollow bubbles (Figure 7). The energy for reinforcement debonding from matrix and fracture energy has been shown to be negligible [21]. As shown in Figure 8, both thrust forces and tangential forces increase with increasing volume fraction of SiO₂ when cutting at constant cutting speed, undeformed chip thickness and width of cut under dry cutting conditions.

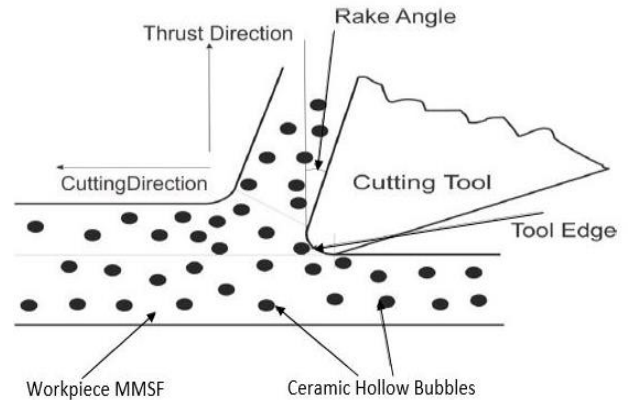


Figure 7. Schematic Diagram for the machining of MMSF

As the volume fraction increases, the number of hollow bubbles increases in the matrix. This enables the workpiece material to resist the compressive forces induced by the tool tip ahead of the shear zone. During this process, the matrix gets pinned between load bearing ceramic hollow bubbles causing an increase in shear strength of the material. In this time period between material being subjected to tool compressive forces and before reaching the critical value of shear stress near the shear zone, a fraction of the bubbles is busted and lose their load bearing capability. In the meantime, as the critical value of shear stress is reached in the shear zone enables the chip to be formed. Higher values of shear stresses are required for chip formation as the number of hollow bubbles in the shear zone increase along with volume fraction leading to higher generated cutting forces (Figure 9). In addition, the tool cutting edge engagement with the hollow bubbles affect the thrust forces.

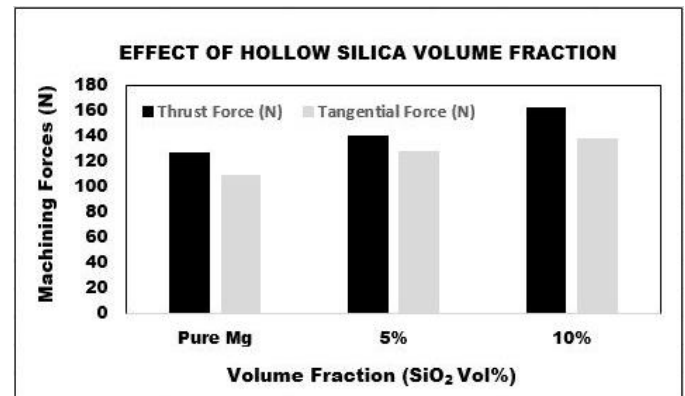


Figure 8. Effect of hollow bubble SiO₂ volume fraction on cutting forces (Cutting speed: 60 m/min, undeformed chip thickness: 0.1 mm)

The radius on the tool cutting edge means the contact area between the cutting edge and hollow bubbles is increased

resulting in higher compressive forces. This effects the hollow bubble to be pushed against hardened matrix without being busted or sheared. This scenario will cause higher normal stresses on the tool. Due to enhanced compressive properties of the hollow bubbles, they resist deformation to great extent causing a higher magnitude of reaction forces on the tool. As a result, thrust forces are more pronounced than the tangential forces.

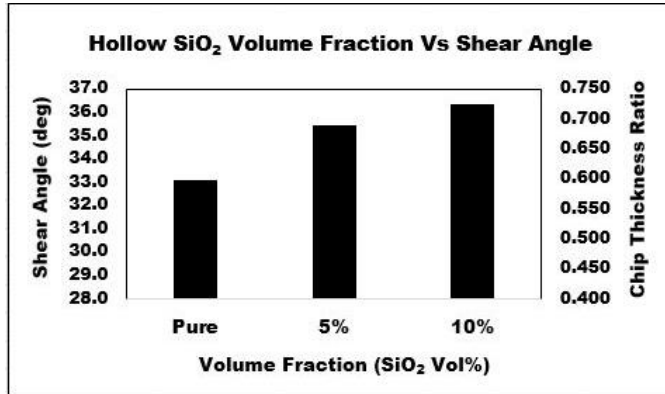


Figure 9. Effect of hollow bubble SiO₂ volume fraction on shear angle (Cutting speed: 60 m/min, Undeformed chip thickness: 0.1 mm)

IV. CONCLUSIONS

This paper presents the dry orthogonal machining characteristics of pure magnesium-SiO₂ hollow bubble syntactic foam. The findings from this experimental investigation are as follows:

1. Cutting speed has been found to have profound effect on the magnitude of thrust force and tangential forces. Machining forces were found to decrease with increasing cutting speed which has been attributed to magnesium matrix thermal softening phenomenon.
2. Higher the undeformed chip thickness, higher the magnitude of thrust and tangential forces. Thrust forces were atleast 30 N higher than the tangential forces for majority of the conditions tested under dry cutting. An increase in undeformed chip thickness results in an increase of number of hollow bubbles ahead of the cutting tool leading to higher cutting forces.
3. An increase in volume fraction of SiO₂ reinforcement leads to an increase in number of hollow particles for same bubble size in the shear zone. Under a constant cutting speed, undeformed chip thickness and width of cut, an increase in volume fraction of hollow bubbles results in increase of both thrust force and tangential force with former being dominant.
4. Further studies will attempt to develop an energy based analytical force model and investigate the chip

formation mechanism for machining metal syntactic foams.

V. ACKNOWLEDGEMENTS

1. Authors sincerely acknowledge the financial support provided by American University of Sharjah-(EFRG18-MSE-CEN-20).
2. Sincere thanks to Mr. D. Venkat Raghaven, (SWAMEQUIP, Chennai, India) for casting the metal foams.
3. Many thanks to Mr. John Memphin for conducting the machining trials at AUS.

REFERENCES

- [1] N. Gupta and P. Rohatgi, "Metal matrix syntactic foams: processing, microstructure, properties and applications," Lancaster, Pennsylvania: DEStech Publications, pp. 1-16, 2014.
- [2] X. Tao, L. Zhang, and Y. Zhao, "Al matrix syntactic foam fabricated with bimodal ceramic microspheres," *Mater. Des.*, vol. 30, no. 7, pp. 2732–2736, 2009.
- [3] D. Lehmhus, et al., "Quasi-static and Dynamic Mechanical Performance of Glass Microsphere- and Cenosphere-based 316L Syntactic Foam," *Procedia Mater. Sci.*, vol. 4, pp. 383-387, 2014.
- [4] D. Luong, V. Shunmugasamy, N. Gupta, D. Lehmhus, J. Weise, J. Baumeister, "Quasi-static and high strain rates compressive response of iron and invar matrix syntactic foams," *Mater. Des.*, vol. 66, pp. 516-531, 2015.
- [5] L. Peroni, et al., "Investigation of the mechanical behaviour of AISI 316L stainless steel syntactic foams at different strain-rates," *Comp. Part B: Eng.*, vol. 66, pp. 430-442, 2014.
- [6] F. Czerwinski, A. Zielinska-Lipiec, P. Pinet, J. Overbeeke, "Correlating the microstructure and tensile properties of a thixomolded AZ91D magnesium alloy," *Acta Mater.*, vol. 49, no. 7, pp. 1225-1235, 2001.
- [7] Z. Zhang, X. Zeng, W. Ding, "The influence of heat treatment on damping response of AZ91D magnesium alloy," *Mater. Sci. Eng. A*, vol. 392, no. 1-2, pp. 150-155, 2005.
- [8] M. Zhao, M. Liu, G. Song, A. Atrens, "Influence of the b-phase morphology on the corrosion of the Mg alloy AZ91," *Corros. Sci.*, vol. 50, no. 7, pp. 1939-1953, 2008.
- [9] X. Xia, W. Zhao, X. Feng, H. Feng, X. Zhang, "Effect of homogenizing heat treatment on the compressive properties of closed-cell Mg alloy foams," *Mater. Des.*, vol. 49, pp. 19-24, 2013.
- [10] P. Rohatgi, N. Gupta, B. Schultz, D. Luong, "The synthesis, compressive properties, and applications of metal matrix syntactic foams," *JOM*, vol. 63, no. 2, pp. 36-42, 2011.
- [11] N. Gupta, D. Luong, K. Cho, "Magnesium matrix composite foams density, mechanical properties, and applications," *Metals*, vol. 2, no. 3, pp. 238-252, 2012.

- [12] A. Daoud, "Synthesis and characterization of novel ZnAl22 syntactic foam composites via casting," *Mater. Sci. Eng. A*, vol. 488, no. 1-2, pp. 281-295, 2008.
- [13] A. Daoud, "Effect of strain rate on compressive properties of novel Zn12Al based composite foams containing hybrid pore," *Mater. Sci. Eng.: A*, vol. 535, no. 1-2, pp. 7-17, 2009.
- [14] P. Rohatgi, J. Kim, N. Gupta, S. Alaraj, A. Daoud, "Compressive characteristics of A356/fly ash cenosphere composites synthesized by pressure infiltration technique," *Compos. Part A: Appl. Sci. Manuf.*, vol. 37, no. 3, pp. 430-437, 2006.
- [15] V. Shabde, K. Hoo, G. Gladysz, "Experimental determination of the thermal conductivity of three-phase syntactic foams," *J. Mater. Sci.*, vol. 41, no. 13, pp. 4061-4073, 2006.
- [16] G. Gladysz, B. Perry, G. Mceachen, J. Lula, "Three-phase syntactic foams: structure-property relationships," *J. Mater. Sci.*, vol. 41, no. 13, pp. 4085-4092, 2006.
- [17] L. Licitra, D. Luong, O. Strbik, N. Gupta, "Dynamic properties of alumina hollow particle filled aluminum alloy A356 matrix syntactic foams," *Mater. Des.*, vol. 66, no. 5, pp. 504-515, 2015.
- [18] M. Labella, V. Shunmugasamy, O. Strbik, N. Gupta, "Compressive and thermal characterization of syntactic foams containing hollow silicon carbide particles with porous shell," *J. Appl. Polym. Sci.*, vol. 131, no. 17, pp. 1-5, 2014.
- [19] N. Tomac, K. Tonnessen and F. Rasch, "Formation of flank build-up in cutting magnesium alloys," *CIRP Ann: Manuf. Techn.*, vol. 40, no. 1, pp. 79-82, 1991.
- [20] D. Carou, E. Rubio, C. Lauro, J. Davim, "Experimental investigation on finish intermittent turning of UNS M11917 magnesium alloy under dry machining," *Int. J. Adv. Manuf. Tech.*, vol. 75, no. 9-12, pp. 1417-1429, 2014.
- [21] H. Kishawy, S. Kannan and M. Balazinski, "An Energy Based Force Model for Orthogonal Cutting of Metal Matrix Composites," *CIRP Ann*, vol. 53, no. 1, pp. 91-94, 2004.

Machinability Studies of Ferritic, Austenitic and Martensitic Stainless Steels

Oluwole A. Olufayo, Osama Elfasi and Victor Songmene

Department of Mechanical Engineering,
École de technologie supérieure
Montreal, Canada.

oluwole-ayodeji.olufayo.1@ens.etsmtl.ca

Victor.songmene@etsmtl.ca

Abstract — Determining the machinability of a material assists industries in assessing possible additional cutting costs incurred in part production. Over the years, several studies have been conducted to assess the machinability of steels due to their wide industrial applications. However, these studies have often been limited to individual material investigations between the most commonly used names. Therefore, a comparative review of machinability across a range of stainless steel grades would guide industries in material selection and in selecting machining parameters. The present work investigates the machinability of three industrial stainless steels grades AISI 409 Ferritic, AISI 304L Austenitic and AISI 410 Martensitic during dry milling. The process parameters considered were cutting speed, feed rate and depth of cut. A statistical analysis of responses, such as surface roughness and cutting forces, was analyzed. Additional machinability performance measures, such as burr formation and chip formation observed during cutting operation, were also investigated. The results indicated improved overall machinability of 409 ferritic stainless steel, as compared to 304L steel and 410 martensitic steel.

Keywords - *Stainless Steels; Milling; Machinability; burr; Surface roughness, Chip formation*

I. INTRODUCTION

Most manufacturing industries, such as aerospace and automotive, are interested in industrial materials providing improved functional performance. However, the performance and service life of a mechanical component are significantly influenced by machining conditions. Cutting factors such as cutting speed, feed rate and depth of cut have a direct impact on machinability performance measures, such as achievable surface finish, cutting forces or power consumption, tool life and tool wear rates, burr and chip formation.

Stainless steel machinability has attracted a lot of attention thanks to its high strength and corrosion resistance. As a material, this steel is a difficult to machine material due to its low thermal conductivity and elevated temperatures in the cutting zones. In addition, machining stainless steels is often associated with elevated tool wear, work hardening, and poor productivity [1]. This hardening and wear phenomenon, which

varies across the different industrial grades, greatly influences the cost incurred in industrial production. Given these challenges, several industrial processes often utilize established steels such as 304 and 316 stainless steels, with known or improved machinability, for part production. The study of the machinability across grades is thus essential to guide industries in selecting materials with improved machining properties.

II. GRADES OF STAINLESS STEELS

Stainless steel alloys have elevated chromium content and are difficult to machine due to their high strength, elevated ductility, average elevated work hardening, low thermal conductivity, and abrasive character. Chromium (Cr) is an essential, and indeed, the main constituent of stainless steel, representing more than 12% of its total makeup. It forms an oxide film on the surface of a material, which resists corrosion, as well as the oxidation of metals, which explains why these alloyed steels are often classified as being “acid-proof”. Molybdenum has the same effect as chromium metal structure, and generally increases the material strength and corrosion resistance. Stainless steels can be classified into four main groups, depending on their structure: ferritic, martensitic, austenitic and duplex steel groups. Their mechanical properties, combined with superior resistance to corrosion, makes stainless steels one of the most versatile material to use [2-4].

A. Ferritic

Standard grades of ferritic stainless steel are mainly comprised of 11.2-19% Cr, with a negligible amount of Ni, No, and max 0.2% C. The most common ferrite steel is the 17% Cr, which has a low carbon content (below 0.10%). In order to improve corrosion properties, molybdenum (0.5 to 2%) is added to ferrite steels. These ferritic grades are referred to as “Cr-steels” and are magnetic due to their ferritic microstructure [4, 5].

B. Austenitic

This is the largest group of stainless steels. The group can be further divided into sub-groups, such as Cr-Mn, Cr-Ni, Cr-Ni-Mo, high performance and high temperature grades. The most common type of austenitic stainless steel is the 18/8 class (e.g., AISI304), corresponding to 18% Cr and 8% Ni. Mo is often added to improve corrosion properties. Most machining

problems seen with austenitic steels are related to built up edge formation, burr, poor surfaces and bad chip formation [4, 5].

C. Martensitic

This group possesses a higher carbon content of about 0.2-1.0% C, which increases the strength and hardenability of the steel. Martensite is often available in an annealed condition consisting of a ferritic matrix with chromium carbides. Sulphur is sometimes added to it as a free-machining additive to improve machinability. Normally, it is machined in this condition and hardening is performed after machining [4, 5].

D. Duplex

Steels falling under this group have a ferritic-austenitic microstructure. The class contain a mixture of ferritic and austenitic elements 20.1-25.4% Cr, 1.4-7% Ni, No, N, plus a little carbon [4].

III. MATERIAL SELECTION

This study investigates the machinability of three major industrial stainless steel grades during a dry milling operation. The grades were selected due to their prominence in common industrial processes. AISI 409 ferritic steel is frequently used in welding processes due to its suitable mechanical properties and high-temperature corrosion resistance [2]. The most extensively used austenitic stainless steel is AISI 304L [6], which is often used in highly corrosive environments both in domestic and industrial settings. It has applications in cooking appliances as well as in the food and dairy industries [7]. Alloy grade 410 is a general-purpose martensitic stainless steel. It is widely used in manufacturing hydraulic machines, pipes, pumps and valves [8].

These three metal alloys under study are low alloy stainless steels characterized by high tensile strength and toughness, and are used in the manufacture of several structural components for the automotive and aerospace industries.

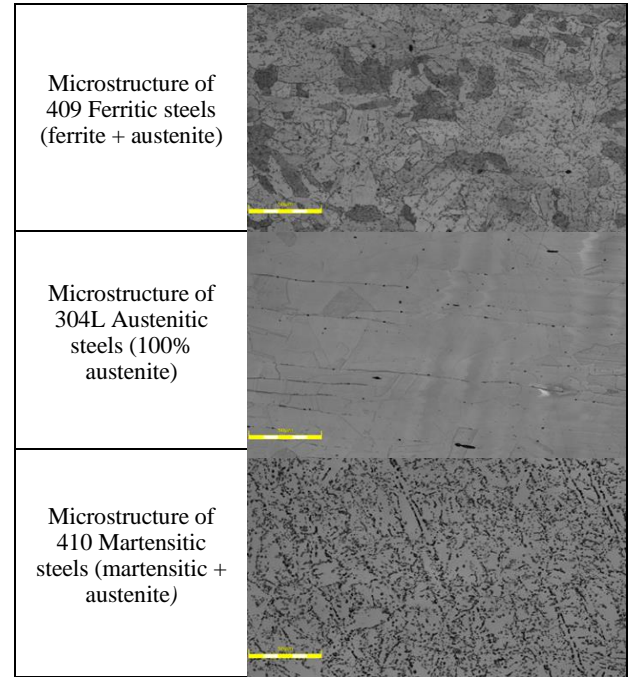
The cutting operation used in the present study was based on its relevance and widespread application. Milling is one of the most commonly used cutting processes in industry. Some inherent properties of AISI 409 Ferritic, AISI 304L Austenitic and AISI 410 Martensitic are known to influence cutting conditions in milling. These properties, which vary with the different stainless steels, could have adverse or positive effects on cutting conditions. The main properties that affect machinability are hardness, strength, shear, ductility, thermal conductivity, and free-machining additive properties of stainless steel, which have been determined to affect machinability. Table I shows the influence these properties have on machinability.

TABLE I. GENERAL IMPACT OF INCREASING VALUES OF DIFFERENT PROPERTIES ON MACHINABILITY

▲ Increasing Condition	Hardness and Strength	negative
	Ductility	negative
	Thermal conductivity	positive
	Work hardening	negative

Table II shows the ferritic microstructure with equiaxed grains and a small inclusion. The annealing twins characteristic can further be seen in the austenite microstructure, while the martensite also shows small dark carbide presence in fine-scale microstructure.

TABLE II. MICROSTRUCTURAL DIFFERENCES BETWEEN STAINLESS STEEL GRADES



IV. MACHINING OF STAINLESS STEELS

The machinability index of different stainless steel alloys differs considerably. In addition to this index, there are other requirements for characterizing stainless steels, such as corrosion resistance and tensile strength, whose values occasionally opposes the machinability [9]. Stainless steels containing 12% Cr are among the hardest-to-cut materials because of their high tendency to harden and their high resistance to heat. Chien and Chou [10] wrote a paper about the development of a predictive model for the machinability of stainless steel. In this model, the ANN theory was used to predict the roughness of the workpiece surface, the cutting force and the tool life. It was found that the maximum metal removal rate (MMRR) increases with the expected surface roughness, it was also noted that the increase of the expected surface roughness is associated with increasing feed rate [10]. In their study on the machinability of PH stainless steel with coated inserts, Palanisamy, Devaraju, Harikrishnan and Manikandan [11] identified an increase in cutting force as the feed rate and nose radius increase. However, although increased feed rates also led to an increase in surface roughness, nose radius and cutting speed increases decreased roughness values. Similar values were found for the surface roughness by Geethanjali, Ramesha and Chandan [12], where a decrease in feed rate and depth of cut improved the achievable surface finish in EN24 steel. However, they identified the

cutting speed as the most influential parameter to machine power requirements.

V. EXPERIMENTAL PLAN

Experimental tests were conducted on a Mazak CNC Vertical Center, NEXUS 410A. The tests assessed the influence of cutting speed and feed rate on measured response parameters (surface integrity and cutting force). A full factorial design for the three-workpiece material (409 Ferritic, 304L Austenitic, and 410 Martensitic) was designed. 27 experimental tests in all were performed on a rectangular workpiece. The experimental tests for each steel material considered only the two cutting factors at three distinct levels ($2^3=9$) (Table III). During machining, cutting forces (F_x , F_y , F_z) were measured using a Kistler 9255B force dynamometer. The workpiece materials were attached with a vice and mounted on the dynamometer. The tool holder selected was the SECO 9930-C3-S12-074, and the cutting tool, the MA FORD TUFFCUT XR 4 FL EM. Fig. 1 shows the tool and its specifications.

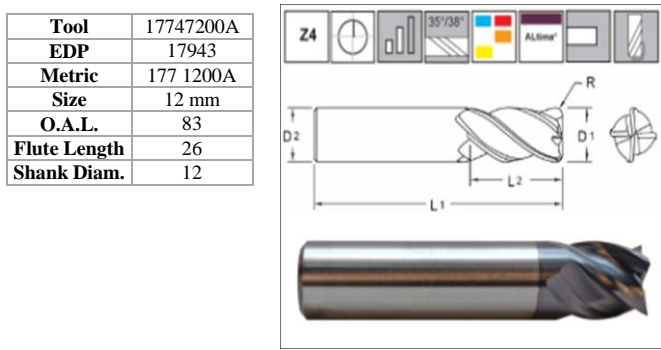


Figure 1. Cutting tool and tool specifications

TABLE III. PARAMETER DATA OF EXPERIMENTS

Factors	Materials	Cutting speed m/min	Feed rate mm/tooth	Depth of cut mm
Low Level	Ferritic 409	50	0.1	1
Mid-Level	Austenitic 304L	100	0.2	
High Level	410 Martensitic	150	0.3	

VI. RESULTS AND DISCUSSION

This section presents an analysis of the cutting force, surface roughness, tool chip and burr observed during the machining tests. The correlations between cutting factors and response parameters was used to identify the material with preferred machinability.

A. Surface roughness

1) Comparative analysis

Surface roughness is frequently used to determine the machinability of a material. Most industrial companies prefer machined products with low surface finish requirements. However, various materials differ in what they require in terms of cutting in order to attain such tolerances. Figures 2, 3 and 4 illustrate the bar plots of the direct relationship of the observed

roughness values (Ra, Rq) with cutting parameters for each material. Both the arithmetic average roughness (Ra) and Root mean square roughness (Rq) responded in the same manner to cutting conditions, with only a slight increase from the Rq parameter.

In these figures, it is clear that increases in roughness parameters (Ra, Rq) corresponds to increases in the feed rate. However, the influence of the cutting speed varies slightly across the samples. For the 409 Ferritic stainless steel in Fig. 2, a minor decrease in roughness was observed with increasing cutting speed. Whereas for the 304L Austenitic stainless and 410 Martensitic steel, a negligible influence of the cutting speed was observed at lower feed rates, but at a feed rate of 0.3 mm/tooth, an increase in speed further amplified the surface roughness.

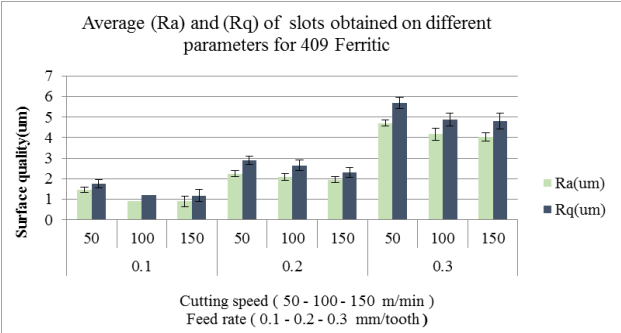


Figure 2. Bar chart of average Ra and Rq for the 409 Ferritic

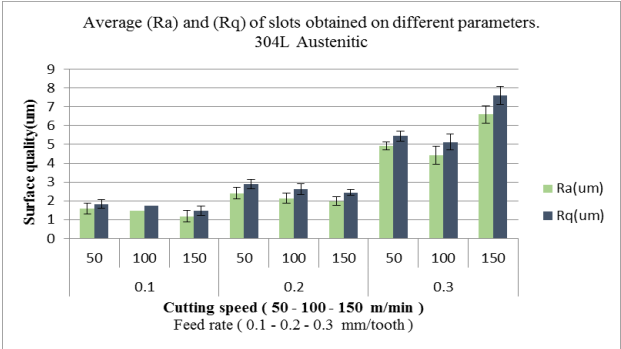


Figure 3. Bar chart of average Ra and Rq for the 304L Austenitic

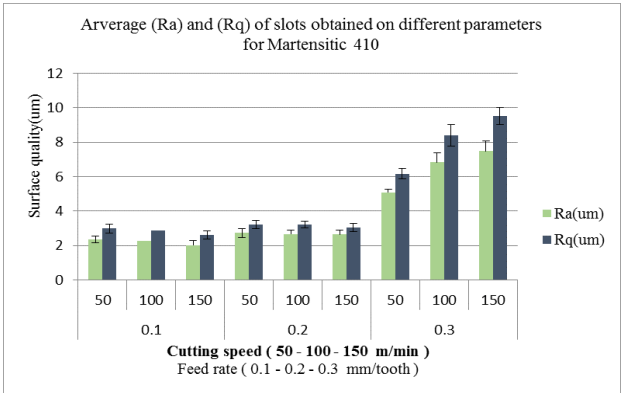


Figure 4. Bar chart of average Ra and Rq for the 410 Martensitic

2) Regression analysis

Figure 5 shows the results of the main effect and the 3D surface plot of the average roughness parameters (Ra) for the three steels. From the figure, we observe that as the feed rate increases, a corresponding increase in roughness value is attained. A negligible effect of cutting speed is seen on the roughness.

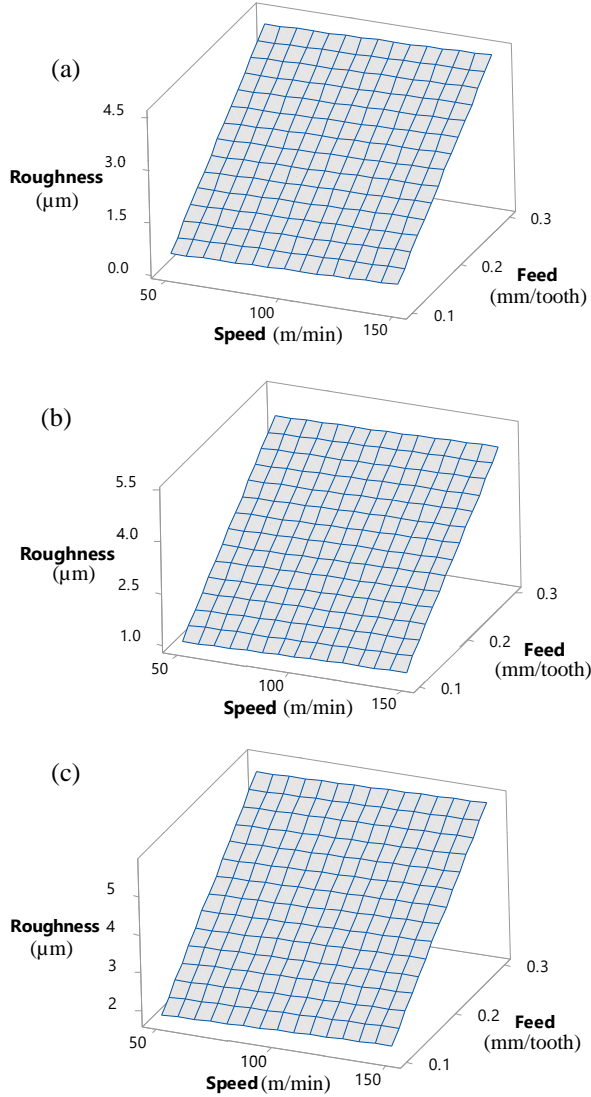


Figure 5. Main effect and 3D surface plot of Surface roughness (Ra) for (a), (b) 304L Austenitic steel, (c) 410 Martensitic steel

Figure 6 illustrates the mean average Ra and Rq during slot milling for the previously described materials machined under the same parameters: feed rate of 0.1 mm/tooth, cutting speed of 150 m/min, and depth of cut of 1 mm. The statistical analysis we performed confirmed that the roughness values of the three tested alloys differed. The best surface roughness Ra were obtained for the 409, which also exhibited the lowest hardness, while the highest values of these parameters were obtained for the 410 Martensitic. The performance of the 410 Martensitic could be related to tool wear (notch wear). This

result is in agreement with those obtained when machining using standard cutting operations, [13]. Excessive notch wear affects the surface texture when finishing, and ultimately weakens the cutting edge [5]. However, the surface roughness values obtained for all the three materials are within acceptable ranges. For a milling operation, a value between 2.018 μm and 0.893 μm is considered acceptable for general applications. There is good agreement between these results and other reported results [14-29]. The most demanding applications require Ra values in a range lower than 0.893 μm .

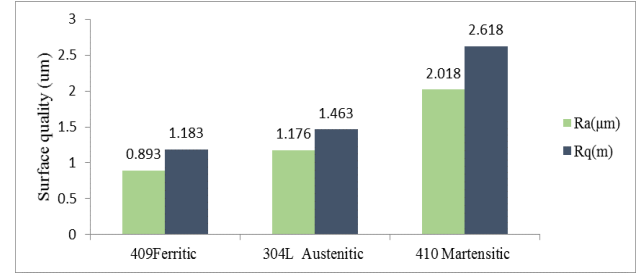


Figure 6. Average Ra and Rq of three different grades of stainless steel at $f = 0.1$ mm/tooth, $v = 150$ m/min, and $ap = 1$ mm

The ANOVA analysis of the regression model provides a significant solution to the problem. A Pearson value correlation with a confidence level of 95% was chosen as the selection criterion. The feed rate was identified as the most significant factor for the surface roughness, with a 73.02% percentage contribution, while the choice of steel material only contributed 8.67% (Table IV). Eqs. 1, 2 and 3 are the regression equations generated for distinct materials. The model displays fit the design data with an r-square value of 81.8%.

TABLE IV. ANOVA RESULTS FOR SURFACE ROUGHNESS RA (MM) AT 95% CONFIDENCE LEVEL

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value	C%
Regression	72.5020	4	18.1255	24.71	0.000	81.79
A:Cutting speed (V)	0.0889	1	0.0889	0.12	0.731	0.10
B:Feed rate (f)	64.7256	1	64.7256	88.24	0.000	73.02
Steel Material	7.6875	2	3.8437	5.24	0.014	8.67
Error	16.1366	22	0.7335			18.20
Total (corr)	88.6386	26				100

Model Summary

R-sq = 81.8%, R-sq (adj.) = 78.49%, and R-sq (pred) = 72.3%

Regression Equation

$$(\text{Austenite}) \text{ Ra} = -0.968 + 0.00141V + 18.96f \quad (1)$$

$$(\text{Ferrite}) \text{ Ra} = -1.440 + 0.00141V + 18.96f \quad (2)$$

$$(\text{Martensite}) \text{ Ra} = -0.149 + 0.00141V + 18.96f \quad (3)$$

B. Cutting force

The comparison of cutting force values for the dry machined workpiece of 409, 304L and 410 stainless steel are depicted in Figures. 7. The resultant cutting force F_R was calculated from the forces (F_x , F_y , F_z). The result reveals that

the cutting force generated in the harder steels reduced by about 2–5% in the less hard material. This reduced cutting force reduced in the lower hardened steel is believed to be due to a reduction in tool-chip contact length and thus chip friction.

The main effect and the 3D surface plot of the average resultant force is illustrated in Fig. 7. From the figure, we also observe that an increase in feed rate also increases the generated cutting forces during operation, however, an increased in cutting speed mitigates these forces. This observation was alike across the various sample materials.

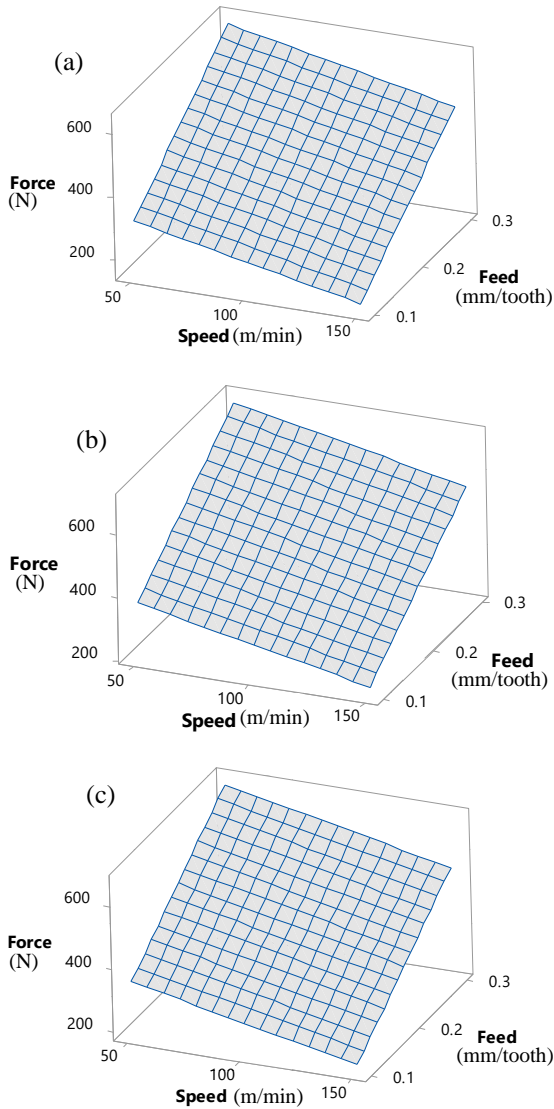


Figure 7. Main effect and 3D surface plot of mean cutting Force (N) for (a) , (b) 304L Austenitic steel, (c) 410 Martensitic steel

The statistical analysis confirmed that the best Force values (N) were attained at a feed rate of 0.1 mm/tooth, and the cutting speed 150 m/min. The 409 Ferritic steel, which has the lowest hardness, exhibited the lowest values of cutting force, but the 410 Martensitic had the highest values of cutting force.

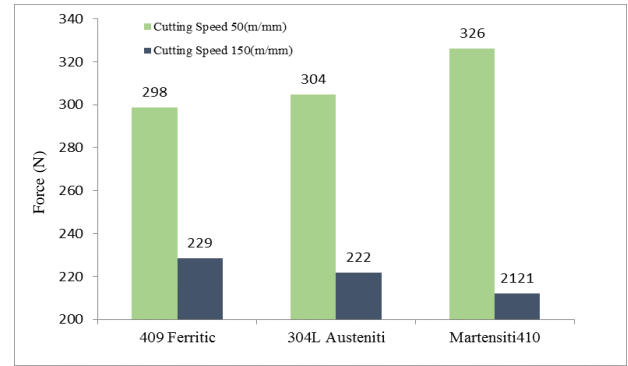


Figure 8. Cutting Forces of three different grades of stainless steel as a function of cutting speeds at $f = 0.1$ mm/tooth, $v = 150$ m/min, and $a_p = 1$ mm

The ANOVA analysis of the regression model presents a significant solution to the problem. A Pearson value correlation with a confidence level of 95% was chosen as the criteria for selection. The feed rate was identified as the most significant factor to the surface roughness with a 74.88% percentage contribution while the choice of steel material only contributed 2.57% (Table V). Eq. 4, 5 and 6 show the regression equations generated for distinct materials. The model displays fits the design data with an improved R-square value of 93.7%.

TABLE V. ANOVA RESULT FOR CUTTING FORCE (N) AT 95% CONFIDENCE LEVEL

Source	Sum of Square	Df	Mean Square	F-Ratio	P-Value	C%
Regression	588772	4	147193	81.76	0.000	93.69
A:Cutting speed (V)	102110	1	102110	56.72	0.000	16.25
B:Feed rate (f)	470528	1	470528	261.36	0.000	74.88
Material	16134	2	8067	4.48	0.023	2.57
Error	39607	22	1800			6.30
Total (corr)	628379	26				100

Model Summary

R-sq = 93.70%, R-sq (adj.) = 92.55% and R-sq (pred) = 90.31%

Regression Equation

$$(\text{Austenite}) \text{ Force} = 286.9 - 1.506V + 1617f \quad (4)$$

$$(\text{Ferrite}) \text{ Force} = 227.5 - 1.506V + 1617f \quad (5)$$

$$(\text{Martensite}) \text{ Force} = 263.4 - 1.506V + 1617f \quad (6)$$

C. Chip analysis

The relationship between the cutting parameters and the chip form is based on many features. In this study, the chip volume, thickness, and length were analyzed. The chip type and size obtained provided some data on the chip development. From the literature, it was established that tool conditions influence the chip morphology in milling operations. [30, 31]. A comprehensive analysis of chips at all cutting conditions was performed. An extract of this analysis covering chips cut at high speeds and under extreme feed rate conditions is shown in Table VI.

TABLE VI. COMPARISON OF THE CHIP FORM GEOMETRY FOR THE STEELS AT 150 M/MIN







Materials	a) 409 Ferritic	b) 304L Austenitic	c) 410 Martensitic
LOW FEED (0.1 mm/tooth)			
HIGH FEED (0.3 mm/tooth)			

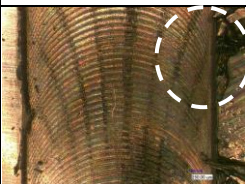

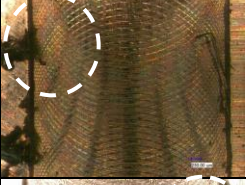

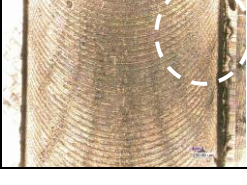

Table VI compares the experimental chip form geometry for the three samples at a 150 m/min cutting speed, a 1 mm depth of cut, as well as 0.1 and 0.3 mm/tooth feed rates. As shown in the figures a slight chip undulation is observed in tests (a) for 409 Ferritic, and (b) for 304L Austenitic, respectively, at low feed rates. Figure (c) with 410 Martensitic stainless steel shows a less prominent curving of the chip form. This is explained by the inherent increased hardness of martensite compared to the other materials. The presence of elevated temperatures at the point of cut further promotes plastic deformation occurring during chip formation. A lamellar surface texture is thus inherently found on the chips. Low speed Ferritic and austenitic chips obtained were mostly arc-shaped, with an amplified spiral twist at a higher speed of 150 m/min. Increasing the feed led to an increase in the chip formation length. With increased hardness of the austenitic steels, serrated edges were produced from the shearing action of tool chips and the workpiece material. An increase in the feed rate further increased the tool chip length and thickness, which in harder materials, leads to a breaking-off of chips when they coil back to machined surface. This is characterized by an increase in surface roughness in these conditions.

A. Burr analysis

Burr formation occurs on the exit edge during the tool cutting passes of experimental tests (Table VII). Research studies have shown that the tool nose geometry holds a significant influence on the formation and characteristics of the burr formed on the exit edges. Material hardness, cutting speed and feed rate also influence its formation in finishing conditions [32, 33].

Table VII shows some sample burr formation observed in the experimental tests of the three steel grades. The burr values for all the three materials were in agreement with those obtained from the literature, and within acceptable industrial ranges for milling operations [33]. The figure for the Ferritic grade produced the longest burr during cutting. A long serrated burr edge exit was obtained with this steel grade. A lower burr formation was seen on the Martensitic stainless grade. A decrease in burr height was observed with increased material hardness [32].

TABLE VII. COMPARISON OF THE BURR FOR THREE STEELS

Materials	Cutting speed: 50 m/min	Cutting speed: 150 m/min
409		
304 L		
410		

VII. CONCLUSION

This study investigates the machinability of three industrial stainless steels, grades AISI 409 Ferritic, AISI 304L Austenitic and AISI 410 Martensitic, during dry slot milling. An experimental plan with varying cutting speeds and feed rates was used to assess the machinability measures, such as the force and surface roughness. The main findings of this research are summarized as follows:

- The surface roughness profile was strongly influenced by the feed rate. This is the most significant parameter to achievable surface finish. A negligible effect was identified from the cutting speed and material selection. A reduction in feed rate produced a corresponding reduction in attainable surface finish.
- The most common chip morphology was a short, arc-shaped chip formed at low feed rates. An increase in the chip length was associated with higher cutting speeds and feeds. Under these parametric conditions, a reduced surface finish was observed due to the recoil and rupture of cutting chips on the machined surface.
- The lowest cutting force values were obtained on Ferritic steels. The feed rate was the most influential parameter to force generation. Optimal parametric combination was at high speed and lower feed rates.
- The improved overall machinability of this study was attained using the 409 Ferritic stainless steel, against the 304L steel grade, and the 410 Martensitic steel. The 409 had improved surface characteristics and generated forces at low feed rates. A challenge faced by this material is associated with the increased burr formation during cutting. However, the burr values experienced are within acceptable tolerances for industrial milling practices.

REFERENCES

- [1]. Wegener, K., et al., *Success Story Cutting*. Procedia CIRP, 2016. **46**: p. 512-524.
- [2]. Ghosh, N., P. Kumar Pal, and G. Nandi, *Parametric Optimization of Gas metal arc welding process by PCA-based Taguchi method on Ferritic stainless steel AISI409*. Materials Today: Proceedings, 2017. **4**(9): p. 9961-9966.
- [3]. Callister Jr, W.D. and D.G. Rethwisch, *Fundamentals of materials science and engineering: an integrated approach*. 2012: John Wiley & Sons.
- [4]. Outokumpu Stainless, A., *Handbook of stainless steel*. Espoo: Outokumpu Oyj, 2013.
- [5]. *Modern metal cutting : A practical handbook*. 1996, Fair Lawn, New Jersey: Sandvik Coromant Publishers. 1 v. (page. multiple).
- [6]. Borek, W., T. Tanski, and Z. Brytan, *Austenitic Stainless Steels-New Aspects*. 2017.
- [7]. Mathiesen, T. and J. Elkjaer Frantsen, *Corrosion aspects for stainless steel surfaces in the brewery, dairy and pharmaceutical sectors*. 2009.
- [8]. Espitia, L., et al., *Cavitation erosion resistance of low temperature plasma nitrided martensitic stainless steel*. Wear, 2013. **301**(1-2): p. 449-456.
- [9]. Altintas, Y., *Manufacturing automation: metal cutting mechanics, machine tool vibrations, and CNC design*. 2012: Cambridge university press.
- [10]. Chien, W.-T. and C.-Y. Chou, *The predictive model for machinability of 304 stainless steel*. Journal of Materials Processing Technology, 2001. **118**(1): p. 442-447.
- [11]. Palanisamy, D., et al., *Machinability Studies on CNC Turning of PH Stainless Steel with Coated Inserts*. Materials Today: Proceedings, 2018. **5**(6, Part 2): p. 14520-14525.
- [12]. Geethanjali, K.S., C.M. Ramesha, and B.R. Chandan, *Comparative Studies on Machinability of MCLA Steels EN19 and EN24 Using Taguchi Optimization Techniques*. Materials Today: Proceedings, 2018. **5**(11, Part 3): p. 25705-25712.
- [13]. Jomaa, W., V. Songmene, and P. Bocher, *Surface finish and residual stresses induced by orthogonal dry machining of AA7075-T651*. Materials, 2014. **7**(3): p. 1603-1624.
- [14]. Philip Selvaraj, D., P. Chandramohan, and M. Mohanraj, *Optimization of surface roughness, cutting force and tool wear of nitrogen alloyed duplex stainless steel in a dry turning process using Taguchi method*. Measurement, 2014. **49**: p. 205-215.
- [15]. Bh, V., S.R. Ch, and P. Vinay, *Effect of machining parameters on tool wear in hard turning of AISI D3 steel*. Procedia Engineering, 2014. **97**: p. 338-345.
- [16]. Uzun, İ., K. Aslantas, and F. Bedir, *The performance Of DLC-coated and uncoated ultra-fine carbide tools in micromilling of Inconel 718*. Precision Engineering, 2015. **41**: p. 135-144.
- [17]. Ezugwu, E.O. and S.H. Tang, *Surface abuse when machining cast iron (G-17) and nickel-base superalloy (Inconel 718) with ceramic tools*. Journal of Materials Processing Technology, 1995. **55**(2): p. 63-69.
- [18]. Aslantas, K., et al., *Cutting performance of nano-crystalline diamond (NCD) coating in micro-milling of Ti6Al4V alloy*. Precision Engineering, 2016. **45**: p. 55-66.
- [19]. Nalbant, M., H. Gökkaya, and G. Sur, *Application of Taguchi method in the optimization of cutting parameters for surface roughness in turning*. Materials & design, 2007. **28**(4): p. 1379-1385.
- [20]. Suresh, R. and S. Basavarajappa, *Effect of process parameters on tool wear and surface roughness during turning of hardened steel with coated ceramic tool*. Procedia Materials Science, 2014. **5**: p. 1450-1459.
- [21]. Elgnemi, T., et al., *Effects of atomization-based cutting fluid sprays in milling of carbon fiber reinforced polymer composite*. Journal of Manufacturing Processes, 2017. **30**: p. 133-140.
- [22]. Xavior, M.A. and M. Adithan, *Determining the influence of cutting fluids on tool wear and surface roughness during turning of AISI 304 austenitic stainless steel*. Journal of Materials Processing Technology, 2009. **209**(2): p. 900-909.
- [23]. Zhang, J.Z., J.C. Chen, and E.D. Kirby, *Surface roughness optimization in an end-milling operation using the Taguchi design method*. Journal of Materials Processing Technology, 2007. **184**(1): p. 233-239.
- [24]. Mantle, A.L. and D.K. Aspinwall, *Surface integrity of a high speed milled gamma titanium aluminide*. Journal of Materials Processing Technology, 2001. **118**(1): p. 143-150.
- [25]. Axinte, D.A. and R.C. Dewes, *Surface integrity of hot work tool steel after high speed milling-experimental data and empirical models*. Journal of Materials Processing Technology, 2002. **127**(3): p. 325-335.
- [26]. Suresh Kumar Reddy, N. and P. Venkateswara Rao, *Experimental investigation to study the effect of solid lubricants on cutting forces and surface quality in end milling*. International Journal of Machine Tools and Manufacture, 2006. **46**(2): p. 189-198.
- [27]. Wang, M.-Y. and H.-Y. Chang, *Experimental study of surface roughness in slot end milling AL2014-T6*. International Journal of Machine Tools and Manufacture, 2004. **44**(1): p. 51-57.
- [28]. Selvaraj, D.P., P. Chandramohan, and M. Mohanraj, *Optimization of surface roughness, cutting force and tool wear of nitrogen alloyed duplex stainless steel in a dry turning process using Taguchi method*. Measurement, 2014. **49**: p. 205-215.
- [29]. Senthil Kumar, A., A. Raja Durai, and T. Sornakumar, *The effect of tool wear on tool life of alumina-based ceramic cutting tools while machining hardened martensitic stainless steel*. Journal of Materials Processing Technology, 2006. **173**(2): p. 151-156.
- [30]. Oxley, P.L.B., *The Mechanics of Machining: An Analytical Approach to Assessing Machinability*. 1989: Ellis Horwood.
- [31]. Wagner, V., et al., *Relationship between cutting conditions and chips morphology during milling of aluminium Al-2050*. The International Journal of Advanced Manufacturing Technology, 2016. **82**(9-12): p. 1881-1897.
- [32]. Chern, G.-L., *Experimental observation and analysis of burr formation mechanisms in face milling of aluminum alloys*. International Journal of Machine Tools and Manufacture, 2006. **46**(12): p. 1517-1525.
- [33]. Lin, T.-R., *Experimental study of burr formation and tool chipping in the face milling of stainless steel*. Journal of Materials Processing Technology, 2000. **108**(1): p. 12-20.

Online Monitoring of built-up edge formation in turning stainless steel using acoustic emission signals

Yassmin Seid Ahmed (*), A.F.M. Arif, S.C. Veldhuis

McMaster Manufacturing Research Institute (MMRI), Department of Mechanical Engineering, McMaster University, 1280 Main Street West, Hamilton, ON L8S4L7, Canada; afmarif@mcmaster.ca; veldh@mcmaster.ca

*Correspondence: seidahmy@mcmaster.ca

Abstract—This paper presents a method of built-up edge (BUE) monitoring by means of the acoustic emission (AE) method. The generation of the BUE during the cutting process is of fundamental importance in machining operations, because it may significantly affect the surface roughness, dimensional accuracy, tool wear, chip formation, and process forces. In this study, the generated AE signals are analyzed during machining AISI 304 stainless steel in the time domain and frequency domain using fast Fourier transformation (FFT) technique. The time domain analysis shows that the root mean square of the AE (AERMS) signals is very sensitive to the formation of the BUE. The increase in BUE height leads to reduce the effective rake angle, which in turn causing an increase in the AE RMS signals. Furthermore, different extracted features from AE signals such as AE mean and FFT can be used to predict the feed and radial forces during the machining process. Predicting the formation of BUE is essential for tool condition monitoring and machining optimisation. Thus, an artificial neural network (ANN) based predictive model of BUE formation has been approached. The results show that the regression coefficients for all ANN architectures are greater than 95% which reflects a good fit of the prediction models and 4-5-1 architecture offers the highest R value (0.9998).

Keywords- *Machining, Built-up edge, Monitoring, Acoustic emission, ANN*

I. INTRODUCTION

Austenitic stainless steels with specific mechanical characteristics, are considered essential when high corrosion resistance is required (Habak and Lebrun, 2011). These alloys are frequently employed in the manufacturing of chemical components, cryogenic vessels, and food equipment (Siddiquee et al., 2014). Austenitic stainless steels are considered difficult-to-machine materials because of their low thermal conductivity and high strain hardening rate during cutting (Habak and Lebrun, 2011). Their low thermal conductivity leads to heat concentration in the cutting zone, and subsequently high localized temperatures. In addition, their high work hardening leads to high adhesion of the workpiece material to the cutting tool, resulting unstable chip and BUE formation (Jomaa et al., 2014).

Monitoring the cutting process enables to detect the process conditions during the cutting process without the need to stop the machine for characterization such as measuring tool wear, BUE, cutting forces, surface roughness, etc. (Li, 2002). Process monitoring is required to gain not only improved product quality and higher productivity, but also, to identify the main problems which may be faced during the cutting process such as chip sticking, especially for sticky material like stainless steels (Bhuiyan et al., 2012).

Considering the gaps in the literature review, this research aims to apply AE and force sensors to detect the behavior of BUE formation when machining AISI 304 stainless steel. In this research, the AE and cutting force signals obtained from actual machining was used to estimate the BUE height. Then, ANN models are presented to predict the BUE height in high accuracy.

II. EXPERIMENTAL TECHNIQUES

The turning operation was performed on a CNC lathe model using a Nakamura-Tome Sc-450. The workpiece used is around bar (diameter 150mm X length 300) of AISI 304 stainless steel. After every 30 m of cutting length in the stable state of wear during machining stainless steel at 60 m/min, the flank wear width was measured using an optical microscope (KEYENCE – VHX 5000) and BUE dimensions (height, area, and volume) were measured using white light interferometry (Alicona infinite focus G5 microscope). The cutting tool was also examined by a scanning electron microscope (SEM Vega 3-TESCAN), coupled to energy dispersive X-ray spectroscopy (EDS) to examine the wear-acting mechanisms on the flank and rake faces of the cutting tools.

Continuous dry cutting was conducted using uncoated cemented carbide inserts with WC 6%Co and the tool holder has ISO specification ISO PCLNL 45165 12HP, supplied by Sandvik. The cutting force components were conducted during the cutting tests using a piezoelectric dynamometer (Kistler-type 9121). Then, the signals were transmitted to a charge amplifier (Kistler-5010 type) and recorded using a LABVIEW software (Fig.2). The acquisition rate was 300 data points per second and sensitivity was 3.85 mV/MU. Moreover, from literature, it was concluded that a piezoelectric sensor with a central frequency in the range 100–300 kHz is found to be the most suitable for the detection of the variation happened in the cutting process like BUE (Prakash et al., 2015). Therefore, a

Klister piezoelectric AE sensor with a response frequency range 40–400 kHz was used in this research.

The cutting conditions were performed with a feed rate of 0.1 mm/rev, depth of cut of 0.5 mm, and cutting speed of 60 m/min. During machining tests, the machine has been stopped every 30 m cutting length, and the flank wear width was measured using an optical microscope (KEYENCE – VHX 5000). According to ISO standard number 3685 (Panda et al., 2012), the tool life is specified to a flank wear of 0.3 mm. The cutting tools were also examined by a scanning electron microscope (SEM) facilities (Vega 3-TESCAN), coupled to energy dispersive X-ray spectroscopy (EDS).

BUE formation on the tool surface was investigated using white light interferometry (Alicona infinite focus G5 microscope) to measure its dimensions (height, area, and volume) during the machining process. Alicona microscope is worked by focus variation to develop 3D images. The lateral resolution is down to 400 nm and the vertical resolution is down to 10 nm. To measure the dimensions of BUE, first a new cutting tool was scanned as a reference measurement. Then, the same cutting tool after machining was compared with the new tool. More details about the technical specification and scanning process using Alicona microscope are available online (“Focus-Variation | Alicona - High-resolution optical 3D measurement,” n.d.).

III. RESULTS AND DISCUSSIONS

A. Tool wear and built-up edge measurements

As shown in Fig.1, with increasing cutting speed, wear is increasing correspondingly on the flank face of the cutting tool. The relation between cutting speed and tool wear is due to the reduction in the tool-chip contact length and the consequent increase in temperature, as well as increase in both normal and shear stresses at the tool tip/cutting region (Da Silva et al., 2013). Furthermore, it is worth to mention that the stable state of wear is only steady with low cutting speed (60 m/min), while the steady stable state of wear cannot be observed with higher speed.

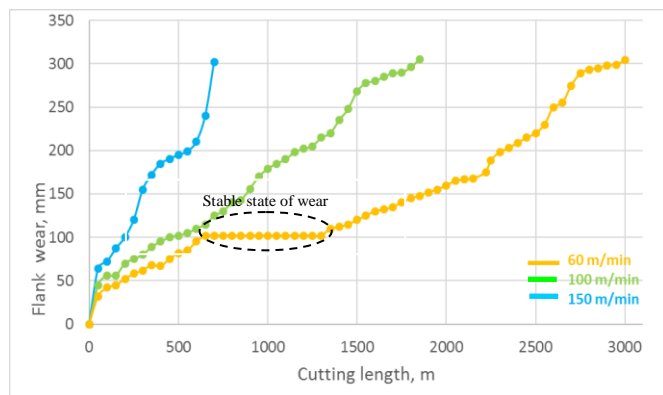


Fig.1 Relation between flank wear and cutting length for different cutting speeds.

Fig.2 presents the results of the determined BUE height, average area, and volume on the cutting tools for different

cutting speeds and cutting lengths. Based on the results a clear tendency of a decreasing height, covered area and volume (BUE dimensions) with increasing cutting speeds. The decreasing of BUE dimensions with increasing cutting speed can be explained as a result of a reduction of the time to dissipate the temperature, leading to higher temperature in the cutting zone. It is also noted that there is no specific relation between the cutting length and BUE dimensions sometimes increases with the increase in cutting length and sometimes decreases: the stability of BUE is very low, it gradually increases, accumulates, and suddenly breaks off and then a new BUE is developed again. These results are consistent with our earlier paper (Ahmed et al., 2017a).

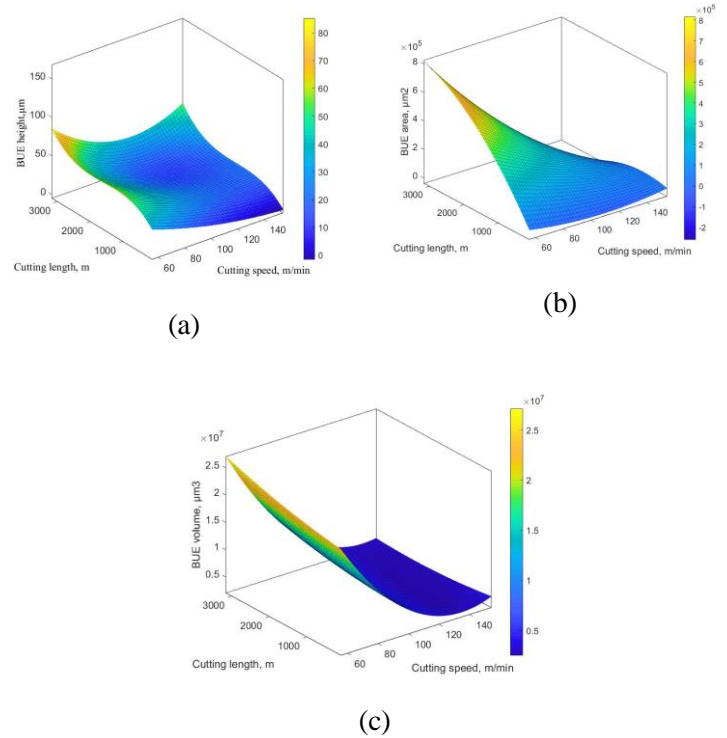


Fig.2. 3D surface plots of BUE dimensions, showing the effect of cutting length and cutting speed on (a) BUE height, (b) BUE area, and (c) BUE volume.

As mentioned previously, the focus of this paper is to monitor the behaviour of BUE during the machining process. For this reason, it is necessary to study its phenomenon in the stable state of wear (steady state zone) to isolate the effect of tool wear and ensure that all conditions are the same. Thus, according to Fig.3, all the analyses are performed with only 60 m/min cutting speed where steady stable state of wear can be observed.

B. Edge radius and cutting force components measurements

In the machining process, the influence of machining conditions on the tool geometry is quite important as stated in (Oliaei and Karpat, 2017, 2016). For that reason, this section evaluates the effect of BUE in the stable state of wear on edge radius. In order to understand the influence of BUE, a 3D

Alicona microscope has been used to analyze the geometry of the cutting tools in the stable state of wear.

Fig.3 shows the profile of the cutting edges at different cutting lengths in the stable state of wear associated with the optical microscope images taken from the cutting tools at the same cutting lengths (L). Based on these images, as BUE increases, the edge radius (r) increases: cutting edge radius is decreased from 37.36 μm at BUE height=110 μm to 34.13 μm at BUE height=15 μm , then edge radius is increased to 35.22 μm at BUE height=80 μm . Increasing edge radius during the machining process may increase the contact between the workpiece and the cutting tools, which in turn, leads to increase the friction conditions at the contact zone, resulting higher process forces.

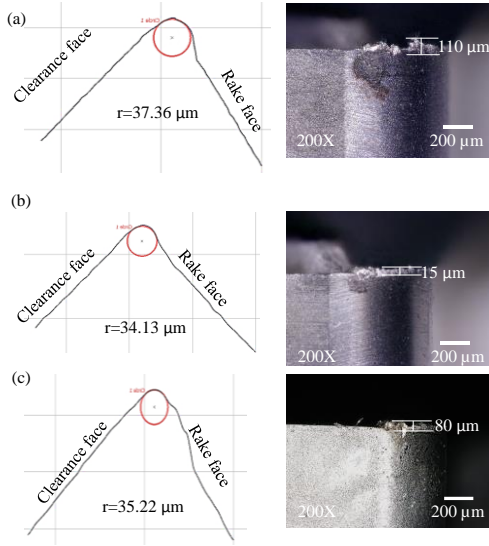


Fig.3. 2D edge profile measurements at different cutting lengths and different BUE heights (a) L=710 m, BUE=110 μm , L=920 m, BUE=15 μm , and (c) L=1010 m, BUE=80 μm .

Thus, measuring forces components during cutting tests enables to understand the influence of BUE on the generated forces. According to the schematic drawing (Fig.4a), F_x , F_y and F_z are cutting force, feed force and radial force, respectively.

Fig.4a demonstrates the changes in the three force components with the cutting length in the stable state of wear (the data are collected every 30 m). As shown, as a result of unstable BUE, the cutting force components varies greatly, sometimes increase and sometime decrease. According to the figure, the feed force in the y direction is the greatest and changes most significantly in the cutting process: fluctuations of feed force and vibration in the y direction are the largest (Fig.4b), while those the z direction (radial force) are the smallest (Fig.4c). The feed force seems to be more affected by the BUE more than other forces. The presence of a BUE is known to be important since a large contact area (larger edge radius) with the workpiece and the cutting tool (in the direction of y axis) as a result of BUE formation. Furthermore, the cutting force in the x direction acts in the same direction as on workpieces resisting the tool and it undergoes smaller

compared to feed force where unstable BUE is normally occurred.

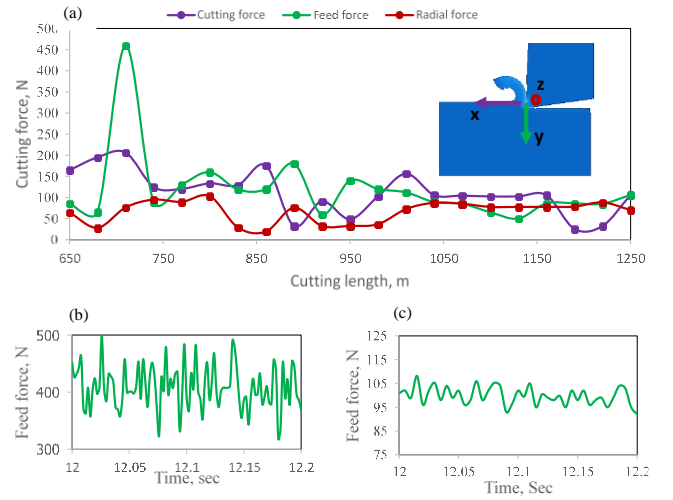


Fig.4 (a) Force components with cutting length in the stable state of wear, (b) feed force wave at 710 m, (c) feed force wave at 920 m.

Furthermore, it is worth stating that BUE formation is consistent with the feed force variation as well. Figs.6b and 6c show feed force waves at 710 and 920 m, respectively. BUE height at 710 m is around 110 μm compared to only 15 μm at 920 m. This explained why there is higher variation in feed force values at 710 m, which indicates that the increase in BUE formation can cause higher friction conditions in the cutting zone during the cutting process, leading to higher feed force variations. This phenomenon is diminished with lower BUE (Fig.4c) where the variations of feed force are smaller.

In order to determine the most sensitive cutting force component for BUE detection, all the dimensions of BUE are measured offline in the stable state of wear and compared to forces components measured by dynamometer, as shown in Fig.5. Here, no clear correlation can be found between the BUE dimensions and cutting forces. To be sure, a correlation analysis is performed using the Pearson correlation coefficients to select the representative force component. The correlation coefficients between the force components and BUE dimensions values are calculated and summarized in Table 1. This table presents the correlation value for each force component for different BUE dimensions. According to Caggiano (2018), he classified the level of correlation according to three classes: weak ($0 < C_p < 0.3$), moderate ($0.3 < C_p < 0.7$), and strong ($C_p > 0.7$). As shown, all the correlation values are lower than 0.7 which means that there is no strong correlation can be identified between the BUE dimensions and cutting forces, however feed force has a moderate correlation value ($C_p = 0.64$) with BUE height. According to the obtained results, it is hard to use forces components as a monitor tool for BUE formation, however they are able to show a general indication if the intensity of BUE is high or low.

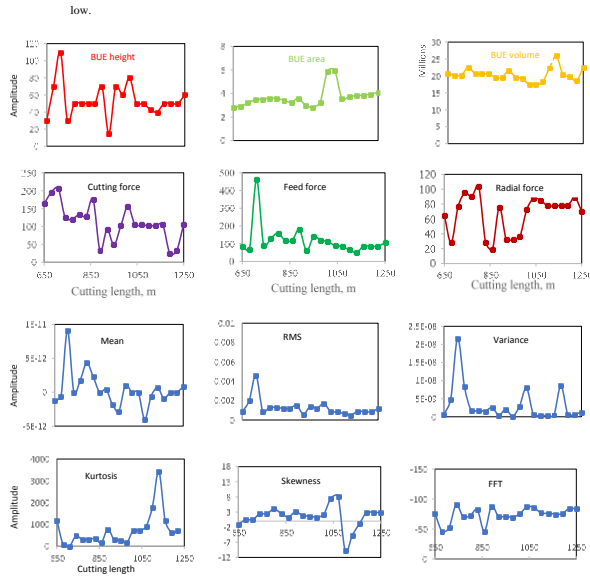


Fig.5. BUE dimensions, force components and extracted features of AE signals.

Table 1: Correlations between force components and actual BUE dimensions.

Component	BUE		
	Height	Area	Volume
Cutting force	0.26	0.214	0.03
Feed force	0.64	0.411	0.15
Radial force	0.02	0.411	0.05

C. AE signal analysis

The acoustic emission system plays an important role in tool condition monitoring during the machining process (Prakash et al., 2015). It can extract more information about the occurrences thus aiding to understand the cutting process accurately (Borghesani et al., 2018). The raw AE signals are the resultant illustration of all occurrences that take place during the turning process (Hase et al., 2014). In turning process, some major occurrences dominate the signal patterns. The tool wear, the plastic deformation and the BUE formation are the key source of the AE signal (Inasaki, 1998). The continuous low amplitude pattern of raw signals represents the tool wear whereas the transient burst pattern of signal shows the BUE formation occurrences (Borghesani et al., 2018; Hase et al., 2014).

Fig. 6 shows the time domain waveforms of the AE signal, where the data is collected on the same cutting lengths as the cutting forces and BUE dimensions (every 30 m in the stable state of wear). It can be seen that there are some differences in the amplitudes of the AE signals at different cutting lengths. Without statistical analysis of the time domain and frequency domain features, the BUE formation cannot be directly identified from the time domain waveform. Therefore, it is necessary to analyze the AE signal in time domain and frequency domain and investigate the provided data in the AE signal. The extracted features from AE signals are calculated

and presented in Fig.5 (mean, RMS, variance, skewness, kurtosis, and FFT).

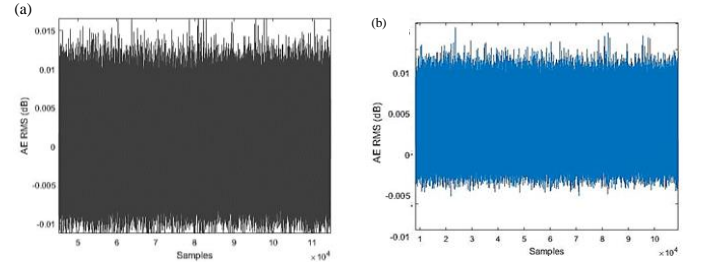


Fig.6 Examples of AE time domain waveforms; (a) at 680 m and (b) at 1010 m.

Starting with BUE height, it seems that AE RMS reveals a good correlation with the BUE height (Fig. 5), which indicate that AE RMS can capture the general behaviour of BUE height very well. To confirm AE RMS is the best feature for BUE height, correlation analysis is performed using the Pearson correlation coefficients to select the representative feature. The correlation coefficients between the extracted features and BUE height values are calculated and summarized in Table 2.

Table 5: Correlations between the extracted features and actual BUE dimensions and forces components.

	BUE			Forces		
	Height	Area	Volume	Cutting	Feed	Radial
Mean	0.613	0.287	0.118	0.43	0.79	0.22
RMS	0.849	0.29	0.198	0.516	0.522	0.051
Variance	0.596	0.39	0.09	0.56	0.557	0.092
Skewness	0.041	0.65	0.260	0.128	0.085	0.12
Kurtosis	0.308	0.14	0.325	0.196	0.37	0.322
FFT	0.38	0.02	0.022	0.615	0.350	0.754

As presented, AE RMS presents the highest correlation coefficient which is 0.849 (>0.7), which confirms that AE RMS can predict BUE height very well. Also, from Table 2, there are a moderate correlation between the skewness and area ($C_p=0.65<0.7$), and a weak correlation between the kurtosis and volume ($C_p=0.32$).

Accordingly, BUE height has the most significant correlation with extracted features. In addition, it was found that BUE height is commonly used in the literature (Childs, 2013; Nomani et al., 2017, 2016; Tomac et al., 2005; Voß et al., 2014). Therefore, the next sections will be only focused on how to utilize AE RMS as method for predicting BUE formation during the machining process.

D. Artificial neural network built-up edge data processing

Artificial neural network (ANN) runs as a human brain where ten layers are connected to each other's and each layer has a number of neurons (Kamruzzaman et al., 2017; Ratava et al., 2017). In the present work, ANN data processing was performed for BUE formation prediction due to its ability to predict tasks through input-to-output vectors mapping (Kamruzzaman et al., 2017).

D.1 Natural network structure

ANN system is capable to propagate unknown data through several factors such as distribution of the database, input and output parameters, number of input values, number of hidden layers, training time, learning function, values of weights and biases, and format of the dataset presentation (Hanief et al., 2017). Consequently, four procedures were performed in the creation of the ANN prediction model which are collection of input and output values, pre-processing of dataset, design and training ANN model, and finally ANN model validation to find the probability of the predicted data will fit with the measured output (Mia and Dhar, 2016). The best ANN architecture was defined after many simulation tests. In this work, a feed forward multi-layer neural network has been employed determine the optimal configuration for the prediction of BUE formation. The outline of the ANN modeling is presented in Fig. 7.

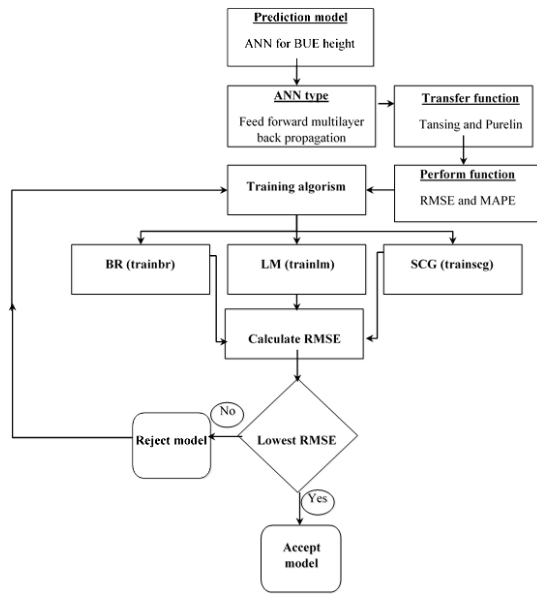


Fig.7. Outline of ANN modeling.

In general, the structure of ANN consists of three layers which are an input layer, hidden layers and output layer where the input and output layers are the first and last layers, respectively (Mia and Ranjan, 2016). The hidden layers can be one or more layers depends on the quality of obtained results (Kamruzzaman et al., 2017). Basically, the hidden layer treats the information received from the input layer. Similarly, the next hidden layer collects the output from the first hidden layer and processes this output to produce the final results (Chen et al., 2017). Increasing number of hidden layers might help to achieve high accuracy, however training time will increase as well as it can cause over-processing problem (Hanief et al., 2017). In this study, the simulation was performed with one hidden layer and a high prediction has been observed. ANN modeling is not only sensitive to the number of hidden layers, but also it is affected by the number of neurons in the hidden layer. Thus, different number of neurons in the hidden layer has been trained. In general, the number of hidden neurons depends on input vector size and the number of classifications of input–output vector space (Neelakanth Gaitonde et al., n.d., 2016).

Too many neurons can lead to over-fitting, whereas few neurons can lead to under fitting (Mia and Ranjan, 2016). For that reason, the number of neurons in the hidden layer was determined by trial and error method and repeated training simulation.

ANN with three different architectures were considered. The first structure has four neurons in the input layer which are cutting length, cutting force, feed force, and radial force. The second structure consists of seven neurons which are cutting length, AE_{mean} , AE_{RMS} , $AE_{variance}$, AE_{FFT} , $AE_{skewness}$, and $AE_{kurtosis}$, while the input layer in the third one has ten neutrons which are cutting length, AE_{RMS} , AE_{mean} , $AE_{variance}$, AE_{FFT} , $AE_{skewness}$, $AE_{kurtosis}$, cutting force, feed force, and radial force. Whereas, the output layer had only one node to predict the BUE height. The three possible structures of ANN model are shown in Fig.8.

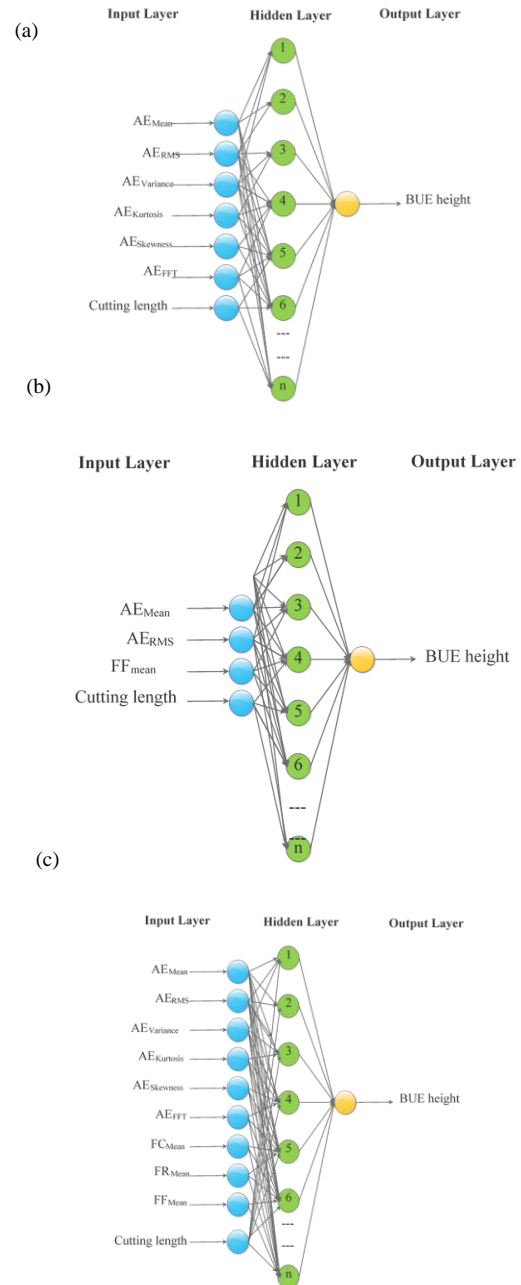


Fig.8. Topologies of the ANN for prediction of BUE height (a) 4-n-1 structure (b)7-n-1 structure, and (c) 10-n-1 structure.

D.2 Neural network training and testing

NNet toolbox of the Matlab software package was utilized to create, train and test the ANN models. The parameters used for ANN are presented in Table 3. To train the network, 30 samples (70%) were used and for validation and testing 9 samples (15%) for each one was used. The 10 X 21 data for inputs (cutting length, AE_{RMS} , AE_{mean} , $AE_{variance}$, $AE_{skewness}$, $AE_{kurtosis}$, FFT, cutting force, feed force, and radial force) and 1X 21 output (BUE height) were utilized to train the model. After building the network, different training algorithms were evaluated by using data (10 X 8).

The three ANN architectures have been trained by using Levenberg–Marquardt (LM), Bayesian Regularization (BR), and Scaled Conjugate Gradient algorithms (SCGA). Each method has its advantages; thus, it is necessary to test which one gives the lowest error and based on the obtained results, the best one is recommended. The performance has been estimated according to root mean squared error (RSME) of the predicted BUE height versus the measured one. LM method, known as “trainlm” is recommended when a moderate size network is used (Addona et al., 2011). It is faster compared with other methods and it can provide the training with validation and test vectors (Kamruzzaman et al., 2017). BR or “trainbr” can deal with the imprecise noisy data and reduces the validation process length. It takes more time to train, however it is more accurate compared to Levenberg–Marquardt algorithm (Mia and Dhar, 2016). SCG is donated by “trainscg” is commonly utilized in ANN which updated weight and bias values based on the scaled conjugate gradient method (Caggiano, 2018).

During modelling, transfer function tangent sigmoid (tansig) and pure linear function (Purelin) were used for hidden layer and output layer, respectively (Hanief et al., 2017). It is worthily to mention that ‘tansig’ was selected due to its symmetric nature. The BUE height prediction performance achieved by the different ANN structures and different training methods was evaluated in terms of (RSME) between the BUE height values predicted (H_{pi}) by the ANN and the BUE height values measured (H_{mi}) during the machining test, which can be calculated using Eq.1. Furthermore, the models have been tested by mean absolute percentage error (MAPE) as shown in Eq.2.

$$RSME = \sqrt{\frac{\sum_{i=1}^n (H_{pi} - H_{mi})^2}{n}} \quad (1)$$

$$MAPE = \frac{1}{n} \sum_i \left(\frac{H_{mi} - H_{pi}}{H_{mi}} \right) \quad (2)$$

Table 3: Training parameters of ANN model.

No. of neurons	Input:1, Hidden:2, and output:3
Initial weight and biases	-1 to 1
No. of iterations	30000
Activation function	Tansig
Learning rate	0.05
Momentum constant	0.95

D3 Modelling of built-up edge trend

In the present paper, many ANN models were generated by employing three different structures and three different training algorithms. These models show accuracy to different degrees. Consequently, to select the best training method, for BUE height prediction in dry turning process, a very accurate analysis of the predicted results and RMSE error analysis have been processed. Before performing the training, number of trials were performed by changing the number of hidden neurons from 5 to 20, to train and test the ANN. Herein, the models were trained with 5, 10, and 20 neurons to achieve a satisfied predicted BUE height with maximum correlation coefficient (R) and minimum root mean squared error (RMSE) value.

Table 4 shows the RMSE in predicting BUE height with different training algorithm and different structures. The models were trained many times and it is found that for each time, different results were generated. Thus, three trials were performed for each structure and the average value of RMSE was calculated (Eq.1). It is previously mentioned that only one hidden layer was used in the predictive model; hence double hidden layer was not tested in the ANN model. From the table, it can be indicated that structure 7-10-1 shows the lowest RMSE values for LM (0.083), whereas structures 4-5-1 (0.058) and 4-10-1 (0.094) reveal the lowest RMSE values for BR and SCG, respectively.

According to that, Fig.9 describes the prediction capability of the ANN model, trained by the three mentioned algorithms, on the basis of correlation coefficient value (R) to select the best algorithm which has the highest R value. It is well known that R is between 0 and 1; if the value is 0, it means that no correlation between the measured values and predicted ones and if it is 1, it indicates that measured and predicted values are perfectly fit (Hanief et al., 2017). Herein, Fig. 12a shows that 7-10-1 architecture trained by LM method has R value equal to 0.98864 compared with 0.9998 and 0.9923 in case of structures 4-5-1 and 4-10-1 trained by BR and SCG, respectively, which are illustrated in Figs 12b and 12c, respectively. These results are inconsistent with Table 4, and consequently structure 4-5-1 trained by BR is the best training algorithm.

ANN Structure		Levenberg-marquardt				Bayesian regularization				Scaled conjugate gradient			
		1 st trail	2 nd trail	3 rd trail	Average	1 st trail	2 nd trail	3 rd trail	Average	1 st trail	2 nd trail	3 rd trail	Average
4-n-1	4-5-1	0.238	0.236	0.256	0.243	0.0562	0.0682	0.0498	0.058	0.168	0.159	0.189	0.172
	4-10-1	0.298	0.298	0.269	0.288	0.2235	0.2350	0.2589	0.239	0.0789	0.0869	0.0855	0.083
	4-20-1	0.359	0.325	0.369	0.351	0.1956	0.1896	0.1789	0.188	0.236	0.2226	0.256	0.239
7-n-1	7-5-1	0.189	0.186	0.166	0.180	0.1789	0.1698	0.1896	0.179	0.255	0.236	0.259	0.250
	7-10-1	0.0895	0.0875	0.0943	0.090	0.0869	0.0866	0.0965	0.090	0.359	0.369	0.389	0.372
	7-20-1	0.1693	0.1589	0.1369	0.155	0.0710	0.0789	0.0799	0.076	0.129	0.139	0.148	0.138
10-n-1	10-5-1	0.1965	0.1896	0.1980	0.194	0.0668	0.0899	0.0956	0.084	0.096	0.099	0.085	0.093
	10-10-1	0.1534	0.1620	0.1689	0.161	0.1436	0.1522	0.189	0.161	0.126	0.125	0.129	0.126
	10-20-1	0.1226	0.1226	0.1223	0.122	0.268	0.298	0.269	0.277	0.189	0.175	0.181	0.181
						0	3	0					

In addition to the regression plots of model, the estimated values of BUE height (H_{pi}) from the ANN model (4-5-1) are compared with the experimental values (H_{mi}). A deviation graph is plotted in Fig.10 and It is very clear that the measured data fit very well the predicted results. Herein the mean absolute percentage error (MAPE) was measured according to Eq.2 and it shows that the error as low as 1.1%, which emphasizes that 4-5-1 ANN model produces good conformity between the measured and predicted BUE height during machining AISI 304 stainless steel.

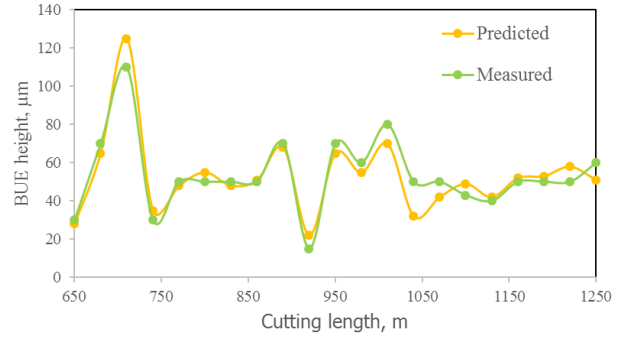


Fig.10. Measured and estimated BUE height for cross-validation data.

4. Conclusions

Acoustic emission generated due to BUE formation during turning of AISI 304 austenitic stainless steel has been studied. The AE signal is analysed in time domain and fast fourier function (FFT). In addition, the statistical parameters, such as Mean, RMS, variance, skewness, and kurtosis have been performed.

AE signals during a turning process were measured using a 200-kHz high-pass filter, and the relationships between the AE signals, BUE formation, and cutting phenomena were examined. The following conclusions were derived from the results and discussion.

1. In order to investigate the effect of AE signals on BUE formation, the analysis is performed at 60 m/min cutting speed where the steady stable state of wear can easily be observed.
2. BUE formation followed a non-uniform trend with respect to the machining length. During the machining process, it gradually increases, accumulates, and suddenly breaks off and then a new BUE is developed again.
3. BUE formation significantly can change the geometry of the cutting tool like edge radius, rake, and clearance angles. Increasing BUE height leads to increase edge radius of the cutting tool, which in turn, leads to increase the friction conditions at contact zone, resulting higher process forces.

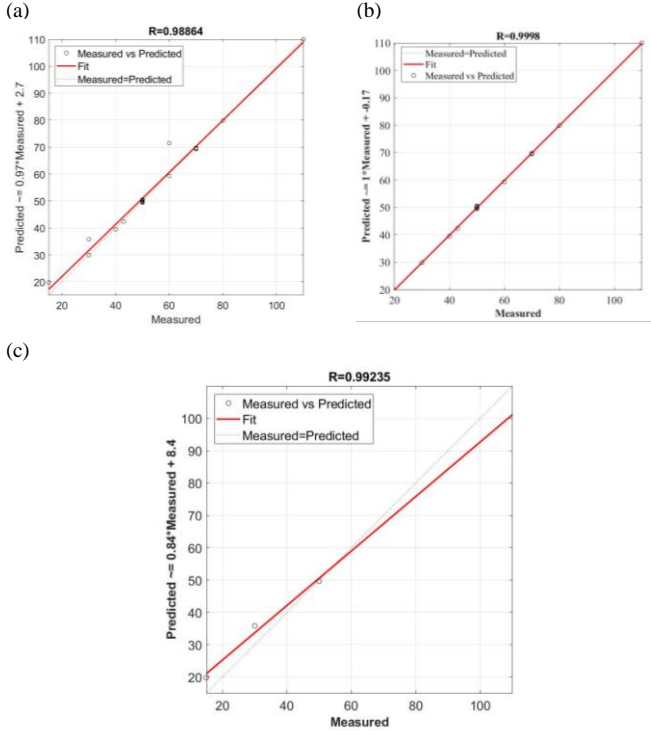


Fig.9. Regression plot for testing data using (a) Levenberg-marquardt, (b) Bayesian regularization, and (c) Scaled conjugate gradient.

4. The correlations between the force components and BUE dimensions are weak. Thus, the force components can be used as indicators of BUE formation, however they can give a general indication about the intensity of BUE.
5. AE RMS, AE mean, and AE FFT are sensitive to the height of the buildup edge, feed force and radial force, respectively.
6. However, skewness and kurtosis have a moderate correlation with BUE area and BUE volume, respectively, they cannot be used as representative features because of the non-uniform trends.
7. In order to monitor the formation of BUE during the machining test in depth, AE RMS is analysed in a very short cutting length (10 m) and it is observed that BUE height are stable at some segments of the stable state of wear. Stability of BUE at these points gives the chance to the BUE to act like a protective layer and prevents the cutting tool from further wear.
8. Similarly, AE mean and FFT are investigated every 10 m cutting length and it is noticed that AE mean has a similar trend as AE RMS signals. In addition,
9. through analyzing FFT signals, it is concluded that the fluctuations of the radial force agree very well with the degree of the sawtooth chips.
10. Thus, Acoustic emission signals are effective in monitoring BUE formation as well as understand the
11. machining performance during machining difficult to cut materials like AISI 304 stainless steels.

Finally, with the aim to monitor the BUE formation during dry turning of AISI 304 stainless steel, the artificial neural network (ANN) has been adopted to develop a predictive model of BUE height. Based on the analysis of ANN models, the performance of the different training methods i.e. Levenberg–Marquardt, Bayesian regularization, scaled conjugate gradient technique was estimated in respect of the lowest root mean square error. The Bayesian regularization with 4-5-1 architecture was found to have the lowest error percentage compared to other methods. In addition, the regression coefficients for all ANN architectures are found to be greater than 98% which reflects a good fit of the prediction models and highest R value (0.998) was obtained with use of 4-5-1.

ACKNOWLEDGE

This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) under the CANRIMT Strategic Research Network Grant NETGP479639-15.

REFERENCES

- Addona, D.D., Segreto, T., Simeone, A., Teti, R., 2011. CIRP Journal of Manufacturing Science and Technology ANN tool wear modelling in the machining of nickel superalloy industrial products. *CIRP J. Manuf. Sci. Technol.* 4, 33–37. <https://doi.org/10.1016/j.cirpj.2011.07.003>
- Ahmed, M., 2017. Optimisation of tool tear and cutting forces on the basis of different cutting parameters. *IJARIII*.3, 613–626.
- Ahmed, Y.S., Fox-Rabinovich, G., Paiva, J.M., Wagg, T., Veldhuis, S.C., 2017a. Effect of built-up edge formation during stable state of wear in AISI 304 stainless steel on machining performance and surface integrity of the machined part. *Materials (Basel)*. 10, 1–15. <https://doi.org/10.3390/ma10111230>
- Antić, A., Kozak, D., Kosec, B., Šimunović, G., Šarić, T., Kovačević, D., Čep, R., 2013. Influence of tool wear on the mechanism of chips segmentation and tool vibration. *Teh. Vjesn.* 20, 279–285.
- Atlati, S., Haddag, B., Nouari, M., Moufki, A., 2015. Effect of the local friction and contact nature on the Built-Up Edge formation process in machining ductile metals. *Tribol. Int.* 90, 217–227. <https://doi.org/10.1016/j.triboint.2015.04.024>
- Bahe, E., Ozel, C., 2013. Experimental Investigation of the Effect of Machining Parameters on the Surface Roughness and the Formation of Built Up Edge (BUE) in the Drilling of Al 5005. *Tribol. Eng.* 15–28. <https://doi.org/10.5772/56027>
- Balaji, M., Murthy, B.S.N., Rao, N.M., 2016. Optimization of Cutting Parameters in Drilling of AISI 304 Stainless Steel Using Taguchi and ANOVA. *Procedia Technol.* 25, 1106–1113. <https://doi.org/10.1016/j.protcy.2016.08.217>
- Bhuiyan, M.S.H., Choudhury, I.A., Nukman, Y.N., 2012. An innovative approach to monitor the chip formation effect on tool state using acoustic emission in turning. *Int. J. Mach. Tools Manuf.* 58, 19–28. <https://doi.org/10.1016/j.ijmachtools.2012.02.001>
- Borghesani, P., Smith, W.A., Zhang, X., Feng, P., Antoni, J., Peng, Z., 2018. A new statistical model for acoustic emission signals generated from sliding contact in machine elements. *Tribol. Int.* 127, 412–419. <https://doi.org/10.1016/j.triboint.2018.06.032>
- Caggiano, A., 2018. Tool wear prediction in Ti-6Al-4V machining through multiple sensor monitoring and PCA features pattern recognition. *Sensors (Switzerland)* 18. <https://doi.org/10.3390/s18030823>
- Chen, Y., Sun, R., Gao, Y., Leopold, J., 2017. A nested-ANN prediction model for surface roughness considering the effects of cutting forces and tool vibrations 98, 25–34.
- Childs, T.H.C., 2013. Ductile shear failure damage modelling and predicting built-up edge in steel machining. *J. Mater. Process. Technol.* 213, 1954–1969. <https://doi.org/10.1016/j.jmatprotec.2013.05.017>
- Da Silva, R.B., MacHado, Á.R., Ezugwu, E.O., Bonney, J., Sales, W.F., 2013. Tool life and wear mechanisms in high speed machining of Ti-6Al-4V alloy with PCD tools under various coolant pressures. *J. Mater. Process. Technol.* 213, 1459–1464. <https://doi.org/10.1016/j.jmatprotec.2013.03.008>

Optimizing Cutting Tool Rake and Clearance Angles based on a Simplified Cantilever Beam Model

C. Hopkins, M. Imad, A. Hosseini*

Department of Automotive, Mechanical, and Manufacturing Engineering
University of Ontario Institute of Technology
Oshawa, Canada

*Corresponding author: sayyedali.hosseini@uoit.ca

Abstract - Optimization of metal removal processes is usually performed by changing the machining parameters that control the process. While it has not been entirely neglected, there have been significantly less research works in the area of optimizing the geometry of cutting tools rather than optimizing the machining parameters. This paper presents a mathematical method to determine the optimum rake and clearance angle for a simple cutting tool in orthogonal cutting. The optimization attempts to minimize the deflection and principal stresses experienced by a cutting tool during orthogonal cutting, with the ultimate goal of maximizing the material removal rate of the tool. Relationships between these angles and the other important cutting parameters are derived and defined such that they can be optimized depending on the material selection of the workpiece and cutting tool. The mathematical model returns the optimal values for the rake angle and clearance angle, the maximum allowable depth of cut, the deflection experienced by the tool tip, and the maximum stresses induced in the tool. The benefits and drawbacks of optimizing the tool geometry and optimizing the cutting parameters are also presented and discussed.

Keywords: rake angle; clearance angle; orthogonal cutting; machining parameters

I. INTRODUCTION

The manufacturing industry can be considered as a major contributor to the global economy. The manufacturing industry is also rapidly growing due to the increased demand of the global economy. While manufacturing has a broad definition and involves numerous processes, this paper is particularly focused on metal removal or machining processes. Metal removal or machining processes are families of operations, where a specific amount of raw material is removed from a workpiece to achieve a desired shape that represents the final product. Numerous cutting tools with specific geometries and made of different materials are available in the market to suit different desires. Although the material from which the cutting tool is made plays a vital role

in the application and performance during machining, geometric features of the tool are also key elements that directly govern the success of a machining operation. This success can be defined in terms of tool life or the quality of the finished product. The quality of the finished product can be measured in terms of surface roughness or dimensional accuracy and is considered to be an essential customer requirement [1].

The optimization of machining parameters and geometrical features of cutting tools have been a trending topic for many years due to the high demand on producing high quality parts while also reducing the manufacturing/production lead time and improving the cutting tool life. According to Reddy [1], geometric parameters of the cutting tools have proven their significance during machining operations. For instance, in a typical turning operation, the radial rake angle has an impact on the machine's power consumption and the chip flow direction; while the nose radius greatly affects the roughness of the newly machined surface. According to recent research by Yussefian [2], the geometry of the cutting edge can impact the tool life, generated forces, surface roughness and power consumption. Its design and optimization has been a topic of interest in the machining community.

Another important factor that determines the success of a typical cutting tool in a very competitive machining industry is the useful tool life. Tool life is highly dependent on machining parameters such as cutting speed, feed rate, and depth of cut. Özler [3] investigated the case of using a sintered carbide tool while machining austenitic manganese steel, and Among the above-mentioned parameters, the cutting speed has the highest impact on tool life followed by the feed rate and the depth of cut.

The aim of this paper is to present a mathematical approach for orthogonal tool design optimization. The objective is to maximize the material removal rate by optimizing the rake and the clearance angles using a genetic algorithm (GA) optimizer.

II. LITERATURE REVIEW

One of the major entities to each machining process that serves in the front line of cutting action is cutting tools. However, as the cutting tools are used, they wear out and ultimately fail, requiring replacement to continue machining. Extending the useful life of cutting tools provides an overall benefit to the manufacturer by reducing manufacturing lead time and tool replacement costs. Although multiple processes such as casting and bulk deformation may be performed on metallic parts to transform them from a block of raw material to the final product, these processes normally lack dimensional accuracy and surface quality. As a result, finishing operations are usually required to adjust to finish an incomplete part and adjust the tolerances. Finishing operations are typically machining operations, the most common of which are milling and turning [4].

Finishing operation optimization have been a trending topic in the literature due to its high impact on the product's quality. The current research in the optimization of metal removal processes can be categorized into two categories: the optimization of cutting parameters and the optimization of the tool's geometrical features. The cutting parameters include cutting speed, feed rate, depth of cut, and width of cut which are controlled by the toolpath of the operation and the capabilities of the machine. Also, the cutting parameters for a specific tool are usually chosen based on the operator's experience or the recommended handbook values [5]. The geometrical features depends on the specific machining operation, and in the case of orthogonal machining, the important geometrical features include the rake angle and the clearance angle.

Researchers have used various optimization approaches to optimize the machining parameters or cutting tools geometrical features, and evolutionary techniques have proven their reliability and robustness. A review of the evolutionary techniques in the optimization of the machining parameters indicates that GA is primarily used as an optimization technique and majority of studies focused on the minimization of surface roughness, machining/production cost, and maximizing material removal [6]. Yang [5] presented the benefits of optimization cutting parameters by implementing the Taguchi method for a turning operation to maximize the tool life and improve the surface roughness, where after testing the optimal cutting parameters a 250% improvement on the tool life and surface roughness were noted. Tandon [7] used particle swarm optimization (PSO) to optimize numerical controlled (NC) end milling cutting parameters and observed a machining time reduction of 35% in a pocket milling case compared to typical cutting parameters.

Reddy [1] investigated the impact of tool geometry and cutting parameters during end milling. A mathematical model was developed and surface roughness was optimized using genetic algorithms (GA). As a result, it was found that the cutting speed and feed rate are the machining parameters that have the highest impact on surface roughness and that the radial rake angle and nose radius are the geometrical parameters that have the largest impact on surface roughness.

Meddour [8] implemented analysis of variance (ANOVA) to explore the impact of the input parameters on the output parameters to minimize the power consumption and minimize surface roughness, while hard turning AISI 52100 steel with a mixed ceramic tool. The input parameters were the cutting speed, depth of cut, feed rate, and tool nose radius. It was concluded that the depth of cut, followed by the feed rate, had the highest impact on the generated forces that affected the power consumption. Also, the nose radius had an impact on the thrust force and on the generated surface, and it was then concluded that the larger nose radius and smaller feed rate resulted in the best surface roughness and vice versa.

Qasim [9] used a statistical approach to optimize the cutting and geometrical tool parameters of a 2D orthogonal tool while machining AISI 1045 steel to minimize the cutting forces and minimize the temperatures. It was indicated that the feed rate and the depth of cut have the highest impact on cutting forces, while the cutting speed has the highest impact on cutting temperatures. Rocha [10] applied response surface methodology (RSM) to turning operations of AISI H13 hardened steel while using a PCBN wiper tool to maximize the tool life, minimize surface roughness and maximize the ratio of material removal rate to the cutting force, by determining the optimum cutting parameters. Hegab [11] utilized NSGA-II multi-objective optimization and response surface methodology (RSM) to minimize power consumption, and surface roughness, while maximizing tool life during cutting of titanium alloys, by selecting the optimum cutting conditions and the optimum percentage of added nano-additives.

Majority of the research in the literature tends to focus on the optimization of the cutting parameters. It is more convenient to optimize the cutting parameters due to the ability to change the parameter values without adding significant cost to the experiment. Researchers can straightforwardly experiment with varying cutting parameters with a fixed geometrical features without the need to change the cutting tool which doesn't add significant cost to the experiment. However in the case of conducting an experiment with varying geometrical features and fixed cutting parameters, multiple tools have to be purchased which obviously increases the cost of the experiment. Also, optimizing the geometrical parameters of a cutting tool means

focusing primarily on the cutting tools manufacturing industry, which is a much smaller industry compared to the metal fabrication industry. The metal fabrication industry, due to economical reasons, tends to be more interested in optimizing the cutting conditions of its already available cutting tools, than investing in the development of new cutting tools.

The aim of this paper is to present a mathematical approach for orthogonal tool design optimization. The objective is to maximize the material removal rate by optimizing the rake and the clearance angles using a genetic algorithm (GA) optimizer.

III. FORCE MODEL

Several inputs are required to model the orthogonal process. These parameters are: Width of cut (w), undeformed chip thickness (t_0), deformed chip thickness (t_c), and the material properties of the workpiece, such as Poisson's ratio (ν_s), ultimate tensile strength (σ_{UTS}), shear strength (τ), strength coefficient (K) and strain-hardening exponent (n). The modulus of elasticity for the cutting tool material (E) is also needed for the deflection analysis. These values can be obtained from a set of material property tables. The first step is to calculate the chip thickness ratio (r). The ratio can be obtained from open literature or through experiments.

$$r = \frac{t_0}{t_c} \quad (1)$$

The next step is to calculate the tool-chip contact length. This is the length over which the chip is in contact with the tool, and therefore is where stress is induced. To begin, the chip compression ratio (ζ) which is the inverse of the chip thickness ratio must be calculated [12].

$$\zeta = \frac{t_c}{t_0} = \frac{1}{r} \quad (2)$$

Next, the exponential constant k must be predicted. The values of this constant have been found experimentally, and are chosen depending on the chip compression ratio [12].

$$k = \begin{cases} 1.5, & \zeta < 4 \\ 1.3, & \zeta \geq 4 \end{cases} \quad (3)$$

Once k and ζ are found, it is possible to calculate the tool-chip contact length [12].

$$l_c = t_0(\zeta^k) \quad (4)$$

The measured force components, called the cutting force (F_c) and the thrust force (F_t), cannot be taken from an existing set of orthogonal cutting data, as their magnitudes will change as the cutting tool geometry changes during the optimization process when the parameters are altered. Astakhov and Xiao

[13] proposed a method for evaluating these forces using an energy based method. The relationship for F_c is given as

$$F_c = \frac{P_{pd} + P_{fR} + P_{fF} + P_{ch} + P_{mn}}{v} \quad (5)$$

Where P_{pd} is power spent on plastic deformation of the surface being cut, P_{fR} is the power spent at tool-chip interface, P_{fF} is the power spent at the tool-workpiece interface, P_{ch} is the power spent forming new surfaces, P_{mn} is the power spent at the minor cutting edge and v is the cutting speed (m/min). For the analysis, P_{ch} and P_{mn} are considered negligible, due the consideration of the orthogonal cutting model and lack of a minor cutting edge respectively. Each of these power components are given as follows:

$$P_{pd} = \frac{K(1.15 \ln \zeta)^{n+1}}{n+1} v(t_0 w) \quad (6)$$

Where K is the strength coefficient and n is the strain hardening exponent of the tool material.

$$P_{fR} = (0.28 \sigma_{UTS}) l_c w \left(\frac{v}{\zeta} \right) \quad (7)$$

$$P_{fF} = (0.625 \tau) \rho_{ce} w \sqrt{\frac{\cos \alpha}{\zeta - \sin \alpha} \left(\frac{1}{\sin \gamma} \right)} \quad (8)$$

This analysis considers the orthogonal cutting model, meaning the tool does not have a radius on its cutting edge (ρ_{ce}). Therefore, the cutting edge radius is assumed to be a small value, like 10 μm [13]. It should also be noted that the cutting speed is not needed for this calculation, as it is cancelled during evaluation. This is one of the simplifying assumptions, as in reality, although cutting speed does not remarkably affect the cutting force, it has a significant effect on tool life due to the temperature it induces.

IV. MODEL OF TOOL AS A CANTILEVER BEAM

Fig. 1 shows a simple model of a cutting tool as a cantilever beam. A side view is also presented in Fig. 2.

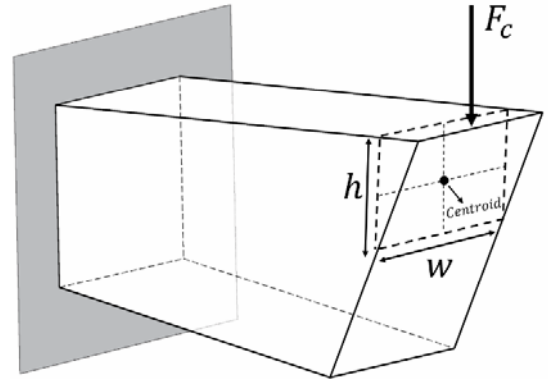


Figure 1. The cutting tool modelled as a cantilever beam and a local cross section with certain width and height

The equation for the deflection of the free end of a cantilever beam is given

$$\delta = \frac{F_c L^3}{3EI} \quad (9)$$

Where L is the distance from the tip of the tool to the fixed end. Instead of treating the tool like a simple cantilever beam, the cutting tool is represented by a more complicated profile. The cross section of the tool at any point along its length can be represented by a rectangle, as seen in Fig 1. The equation for the area moment of inertia of this cross section is given as

$$I = \frac{wh^3}{12} \quad (10)$$

Given the shape of the tool, it is possible to calculate the height of the cross section (h) at any point.

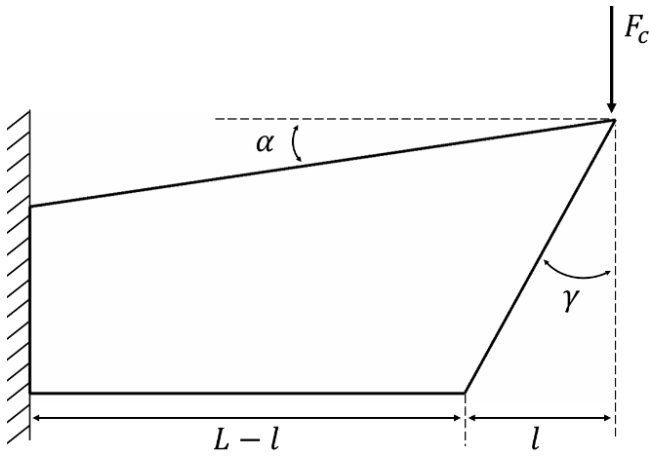


Figure 2. Side view of the cutting tool

The tool shape is divided into two sections as shown in Fig. 3. Each section has its own equation for h . Given the maximum thickness of the tool h_0 , h can be related to the rake angle of the tool in Section 1.

$$h_1 = x \tan \alpha \quad (11)$$

$$h = h_0 - x \tan \alpha \quad (12)$$

For Section 2, h can be related to the clearance angle of the tool. The small height h_1 used in Section 1 still applies in this section as well. It should be noted that this relationship for h is only used if the cross-section is within $0 < x \leq l$, which is also a function of the clearance angle and maximum thickness.

$$l = h_0 \tan \gamma \quad (13)$$

$$h = \frac{x}{\tan \gamma} - h_1 \quad (14)$$

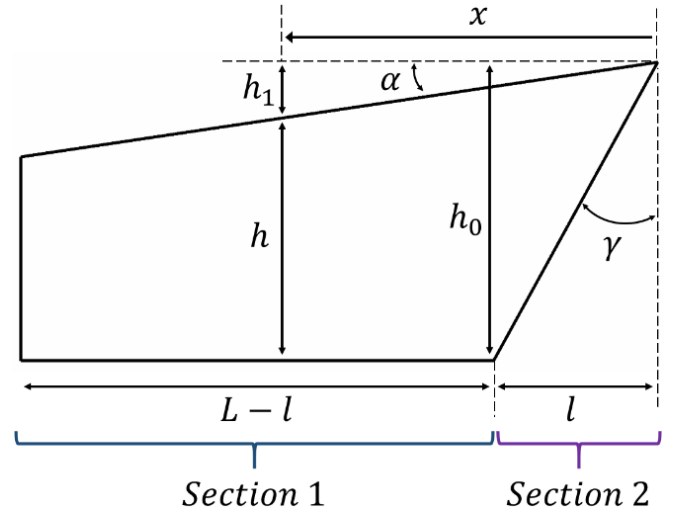


Figure 3. Section 1 and 2 of cutting tool

The complete relationship for h is found as follows:

$$h = \begin{cases} x \left(\frac{1}{\tan \gamma} - \tan \alpha \right) & 0 < x \leq l \\ h_0 - x \tan \alpha & l < x \leq L \end{cases} \quad (15)$$

Note that if the clearance is taken as zero, Section 2 will no longer exist and the formula for Section 1 is used for all points. In addition to the deflection of the tool tip, the stresses experienced by the tool must be considered. Any given solid element of the tool is subject to different stresses. First, the normal stresses induced by bending should be analyzed. The formula for the bending stress induced at a given point is [14]

$$\sigma_{bend} = \frac{Mc}{I} \quad (16)$$

Where M is the moment induced by the applied force measured at a certain distance from the tool tip ($x = 0$), and c is the perpendicular distance from the neutral axis. This means that the bending stress in a certain cross-section will be at a maximum when c is the largest, which occurs at the rake face of the tool, and the stress will be zero at the neutral axis.

In addition to the bending stress, transverse shear stress can also be induced. The formula for shear stress is [14]

$$\tau_{xy} = \frac{VQ}{Iw} \quad (17)$$

Where V is the magnitude of the reaction shear force, Q is the moment of the cross-sectional area either above or below the point in question. While the cross-sectional area of the tool was previously shown to vary along its length, the general shape of the area is always rectangular. For that reason, it is possible to define Q generally for this shape [14].

$$Q = \frac{1}{2} \left(\frac{h^2}{4} - y^2 \right) w \quad (18)$$

Where y is the distance from the neutral axis to the point being analyzed. It can be seen that when the rake face of tool is analyzed ($y = h/2$), the shear stress will be zero. Since no stresses in the y direction or shear stresses are considered, the principal stresses of the element can be simplified.

The principal stresses σ_1 and σ_2 can usually be calculated using the following relationship [14]:

$$\sigma_{1,2} = \frac{\sigma_x + \sigma_y}{2} \pm \sqrt{\left(\frac{\sigma_x - \sigma_y}{2}\right)^2 + (\tau_{xy})^2} \quad (19)$$

But since $\sigma_y = 0$ and $\tau_{xy} = 0$, this can be simplified to

$$\sigma_{1,2} = \sigma_x \quad (20)$$

This stress is then compared to the failure criteria. In the case presented by this paper, no factor of safety is considered, so the tool will be deemed to have failed if the magnitude of the bending stress exceeds the ultimate tensile strength of the tool material.

The last step is to calculate the material removal rate (MRR) of the cutting operation

$$MRR = vwt_0 \quad (21)$$

Where v is the cutting speed used for the cutting operation, w is width of cut, and t_0 represents depth of the cut or feed. This paper treats the cutting speed as constant, so the MRR can be maximized by increasing the width of the cut or the depth of the cut. As these parameters increase, the stresses induced in the tool will also increase. This model provides two constraints for the optimization of the tool geometry: the tool tip cannot be allowed to deflect above a given amount, and the bending stress may exceed the ultimate tensile strength of the tool material. With these constraints in mind, it is possible to find the maximum material removal rate (MRR) that can be achieved before the tool fails.

V. RESULTS

In order to verify the model, simulations needed to be run. The tool material is chosen to be HSS L2 tool steel, and the workpiece material is 1045 hot-rolled steel. The required parameters for the tool are calculated for a set of given input data [14] and typical cutting parameters [15], and checked against the failure criteria. For each run, the depth of cut is increased by 0.01 mm if the tool passes both the deflection and stress checks. The results presented in Tables 2 and 3 show the values taken from the final run where the tool failed one of its checks.

Table 1: Input Parameters for Simulations

Parameter	Value
Width of Cut (m)	0.004
Length of Tool (m)	0.01
Maximum Thickness of Tool (m)	0.008
Cutting Speed (m/min)	30
Chip Thickness Ratio [14]	0.29
Poisson's Ratio of Workpiece	0.32
Modulus of Elasticity of Tool (GPa)	200
Ultimate Tensile Strength of Tool (MPa)	800
Strength Coefficient of Tool (MPa)	530
Hardening Exponent of Tool	0.26
Maximum Allowable Deflection (m)	10^{-5}
Rake Angle Upper Bound (deg)	20
Rake Angle Lower Bound (deg)	-20
Clearance Angle Upper Bound (deg)	45
Clearance Angle Lower Bound (deg)	1

Table 2: Non-Optimized Tool Results

Parameter	Model Value
Rake Angle (deg)	10
Clearance Angle (deg)	10
Depth of Cut (mm)	0.063
Material Removal Rate (m^3/min)	7.44×10^{-6}
Deflection (m)	1.0069×10^{-5}
σ_{\max} (MPa)	164.83

Table 3: Optimized Tool Results

Parameter	Model Value
Rake Angle (deg)	-20
Clearance Angle (deg)	1
Depth of Cut (mm)	0.458
Material Removal Rate (m^3/min)	5.484×10^{-5}
Deflection (m)	1.0013×10^{-5}
σ_{\max} (MPa)	317.02

VI. DISCUSSION

When attempting to optimize the tool geometry for only the stresses, the optimizer showed that the clearance angle (γ) did not affect the results. For each iteration, the optimizer would select a different γ , and the difference in angles between subsequent runs would vary. Thus, the code was changed to optimize the tool geometry to minimize the deflection at the tool tip first, and analyze the induced stresses afterward. By making this change, the results became more consistent. As seen in the previous section, γ would always be selected as 1° (lower bound for the clearance angle), while the rake angle (α) would consistently be chosen very close to -20° (lower bound for the rake angle). In order to minimize the deflection, two factors could be altered: the applied force and the cross-sectional area of the beam. By decreasing α , the force acting on the tool would rise, which seems

counterintuitive to the goal of minimizing deflection. However, the negative rake angle results in a larger h , which in turn results in a larger moment of inertia. The reduction in deflection caused by the larger I outpaces the increase in deflection caused by the larger cutting force, causing a net decrease in the deflection of the tool.

By decreasing γ , the cross-sectional area in Section 2 of the tool would be maximized. Altering the boundaries of these variables would result in different solutions; however, caution must be taken in this regard. From a mathematical standpoint, it is possible to set α to a point where the force on the tool becomes zero or even negative. Solutions in this space are not feasible, since zero or negative cutting force is meaningless.

Another pattern that was noticed was the effect the width of cut (w) had on the deflection and induced stresses of the tool. It was found that increasing w , even to extreme levels ($> 1\text{ m}$) would not cause the tool to fail. Only increasing the depth of cut, also called uncut chip thickness (t_0), would result in an eventual failure. By thoroughly inspecting the equations, it was found that w is cancelled in the calculation of the area moment of inertia (I). For this reason, the decision was made to keep the width of cut constant throughout the analysis. While it is true that increasing the width of cut also increases the MRR, which is the objective function, it is unreasonable to allow such large values that would never be seen in reality.

Several critical assumptions are made for the sake of this purely mathematical model that may cause the results to vary from what is seen in reality. First, this analysis considers the orthogonal cutting model, which is a simple model compared to the reality of most metal cutting operations. While certain operations like parting may closely resemble orthogonal cutting, the oblique cutting model more accurately represents operations like milling or turning.

Next, the tool presented in this paper is treated like a cantilever beam. In reality, a cutting tool is not only supported at the farthest end from the tool tip, but also beneath the tool. This additional support will affect the bending and shear diagrams of the beam, and result in different relationships for the deflection of the tool tip as well as the bending and shear stresses. This is why the deflection of the tool for most cutting processes is not considered to be a significant factor. If a different set of supports were considered, such as supporting the bottom of the tool, the deflection at the tool tip during cutting is expected to decrease.

Another significant assumption this model makes is to neglect the effects of temperature in the tool and workpiece during a cutting operation. The forces involved in cutting metals will release energy in the form of heat, and this heat will affect the performance of the tool. As the temperature increases, the tool material loses its mechanical characteristics due to thermal softening and becomes easier to deform. Consequently, it is expected that the tool will fail under a smaller load compared to a tool in absence of heat. This is one of the many reasons that cutting fluids is used in many metal cutting operations, especially when the workpiece is a material with a low thermal conductivity. Altering this model to

consider the effects of temperature during the cutting process is an avenue for future work.

The future work of this model will be composed largely of including more parameters and considerations to reduce the assumptions made. This will result in a model that more closely resembled the reality of a metal cutting operation. In addition to the previously-mentioned temperature change, other considerations can include the wear that the tool experiences during its useful life. Alternatively, the equations can be changed to consider a different objective than deflection of the tip or the induced stresses.

The optimization of cutting parameters is a more thoroughly investigated field in metal cutting when compared to optimization of tool geometry. It may be more desirable to optimize the cutting parameters due to the methods that can be used. Several techniques, including ANOVA and the Taguchi Method, have been developed to determine the optimum parameters from a given set. These techniques have been researched and verified many times over, so their implementation has become more convenient over time. It is also easier to run experiments to verify the optimization results when the variables are cutting parameters. Cutting speed, feed rate, depth of cut, etc. can be adjusted quickly by an operator or controller. The only limitations of these parameters are those imposed by the limitations of the machine itself. By optimizing the cutting parameters, solutions for a certain objective can be implemented in practice in a timely manner.

The optimization of tool geometry requires the construction of a mathematical model, and in many cases, existing models will not suffice. These models, depending on the assumptions made, can become complex and time-consuming to solve. The benefit of this type of optimization is that certain variables or effects can be investigated that may not be possible by optimizing cutting parameters. Standard tools are used for experiments, but this type of optimization can allow for geometries that may not be available. While it may be more complex, the optimization of cutting tool geometry can lead to solutions that may not have been previously considered, especially when combined with the results gathered from cutting parameter optimization for the same objective.

VII. CONCLUSION

This paper proposes a simplified methodology for finding the optimal rake and clearance angles for a cutting tool, considering the orthogonal cutting model, to maximize the material removal rate of the cutting operation. It was found that negative rake angle and small clearance angles are optimum values for maximizing metal removal rate while stress and deflection are the constraints. The optimized tool was able to withstand larger forces caused by an increased depth of cut when compared to a non-optimized geometry. For every optimization run, it was found that the tool would deflect past the desired level before the stresses would reach the failure criteria. In the future, work will be done to alter the mathematical model to accommodate a more realistic representation of the metal cutting process and the behaviour

of the tool. Needless to mention that the presented model is overtly simplified. The authors plans to further develop the model to include more parameters such as cutting edge micro-geometry, temperature caused by higher cutting speeds etc. The presented methodology paves the way toward that goal.

REFERENCES

- [1] N. S. K. Reddy and P. V. Rao, "Selection of optimum tool geometry and cutting conditions using a surface roughness prediction model for end milling," *The International Journal of Advanced Manufacturing Technology*, vol. 26, no. 11-12, pp. 1202-1210, 2005.
- [2] N. Yussefian, A. Hosseini, K. Hosseinkhani, and H. Kishawy, "Design for Manufacturing of Variable Microgeometry Cutting Tools," *Journal of Manufacturing Science and Engineering*, vol. 140, no. 1, p. 011014, 2018.
- [3] L. Özler, A. Inan, and C. Özel, "Theoretical and experimental determination of tool life in hot machining of austenitic manganese steel," *International Journal of Machine Tools and Manufacture*, vol. 41, no. 2, pp. 163-172, 2001.
- [4] S. Kalpakjian and S. Schmid, "Manufacturing Processes for Engineering Materials–5th Edition," *Agenda*, vol. 12, p. 1, 2014.
- [5] W. p. Yang and Y. Tarng, "Design optimization of cutting parameters for turning operations based on the Taguchi method," *Journal of materials processing technology*, vol. 84, no. 1-3, pp. 122-129, 1998.
- [6] N. Yusup, A. M. Zain, and S. Z. M. Hashim, "Evolutionary techniques in optimizing machining parameters: Review and recent applications (2007–2011)," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9909-9927, 2012.
- [7] V. Tandon, H. El-Mounayri, and H. Kishawy, "NC end milling optimization using evolutionary computation," *International Journal of Machine Tools and Manufacture*, vol. 42, no. 5, pp. 595-605, 2002.
- [8] I. Meddour, M. Yallese, R. Khattabi, M. Elbah, and L. Boulanour, "Investigation and modeling of cutting forces and surface roughness when hard turning of AISI 52100 steel with mixed ceramic tool: cutting conditions optimization," *The International Journal of Advanced Manufacturing Technology*, vol. 77, no. 5-8, pp. 1387-1399, 2015.
- [9] A. Qasim, S. Nisar, A. Shah, M. S. Khalid, and M. A. Sheikh, "Optimization of process parameters for machining of AISI-1045 steel using Taguchi design and ANOVA," *Simulation Modelling Practice and Theory*, vol. 59, pp. 36-51, 2015.
- [10] L. C. S. Rocha, A. P. de Paiva, P. R. Junior, P. P. Balestrassi, and P. H. da Silva Campos, "Robust multiple criteria decision making applied to optimization of AISI H13 hardened steel turning with PCBN wiper tool," *The International Journal of Advanced Manufacturing Technology*, vol. 89, no. 5-8, pp. 2251-2268, 2017.
- [11] H. Hegab, W. Abdelfattah, S. Rahnamayan, A. Mohany, and H. Kishawy, "Multi-objective optimization during machining Ti-6Al-4V using nano-fluids," 2018.
- [12] V. Astakhov and J. Outeiro, "Modeling of the contact stress distribution at the tool-chip interface," *Machine Science and Technology*, vol. 9, no. 1, pp. 85-99, 2005.
- [13] V. P. Astakhov and X. Xiao, "A methodology for practical cutting force evaluation based on the energy spent in the cutting system," *Machining Science and Technology*, vol. 12, no. 3, pp. 325-347, 2008.
- [14] R. C. Hibbeler, *Mechanics of Materials*, Tenth ed. Hoboken, NJ: Pearson, 2015, p. 877.
- [15] G. Ye *et al.*, "Cutting AISI 1045 steel at very high speeds," *International Journal of Machine Tools and Manufacture*, vol. 56, pp. 1-9, 2012.

Multi-objective optimization during sustainable machining of difficult-to-cut materials

A. Salem*, H. Hegab, H. A Kishawy
Department of Automotive, Mechanical and
Manufacturing Engineering
University of Ontario Institute of Technology
Oshawa, ON, Canada
[*Amr.Salem@uoit.ca](mailto:Amr.Salem@uoit.ca)

S. Rahnamayan
Department of Electrical, Computer and Software
Engineering
University of Ontario Institute of Technology
Oshawa, ON, Canada

Abstract— Nickel based alloys have been used in different industrial applications as they offer unique characteristics such as; high thermal fatigue resistance, high erosion resistance, and high melting temperature. However, machining such alloys is still a challenge because of the high heat generated during cutting process. Despite the ability of flood cooling to eliminate the machinability problems, there are governmental restrictions to use it in order to reduce its impacts on the health and environment. Alternatively, the use of minimum quantity lubricant (MQL) reduces the cutting fluid quantity during machining. However, its heat capacity is not same as the flood coolant. Thus, dispersing nano-additives into the base fluid was employed to enrich the heat capacity of the resultant mist through the MQL system. The current study aims to investigate the influence of using minimum quantity lubrication (MQL) technique with vegetable oil based nano-fluid, which includes Aluminum Oxide (Al_2O_3) gamma nano-particles, during machining Inconel 718. The tool wear, surface quality, and energy consumption are studied in the current work. This work offers multi-objective optimization of all studied responses by applying the grey relational analysis.

Keywords- *Inconel 718; minimum quantity lubrication; nano-additives; tool wear; surface roughness; energy consumption; grey relational analysis*

I. INTRODUCTION

Many applications in the field of automotive and aeronautical industries are used the difficult-to-cut materials due to their effective properties which are high corrosion resistance and better strength-to-weight ratio [1]. On the other side, there are characteristics including low thermal conductivity and high hardness, which have bad effect of the machined surface roughness and tool life. An understanding of the influence of different machining design variables (e.g. cutting velocity, tool geometry, coolant strategy, feed rate, depth of cut), has a significant act in improving machinability, productivity, and reducing the total machining cost. Several studies have been conducted with an aim to specify the optimal machining variables and select the better coolant

strategy as machining of hard-to-cut materials is still facing different problems particularly, ceramics, titanium alloys, polymers, nickel-based alloys, composite based materials, and hardened steels [1-4]. However, the unique characteristics of nickel based alloys such as high thermal fatigue resistance, high erosion resistance, and high melting temperature; the disposal of the generated heat during the machining is bad due to the low heat conductivity of these alloys [5, 6]. The increase of the machining generated heat over the critical limit which is $650^{\circ}C$ leads to negative effects on the machined workpiece and the tool [7]. In general, bad tool wear behavior and surface integrity are correlated with the machining of nickel alloys [7, 8]. The usage of cutting fluids during machining is mainly to reduce the between the tool and take out the cutting generated heat. Despite the ability of flood cooling to eliminate machinability problems, there are governmental restrictions to use it to reduce its impacts on the health and environment. Alternatively, the use of minimum quantity lubricant (MQL) reduces the cutting fluid quantity during machining [7, 8]. During machining difficult-to-cut materials, MQL is not a good alternative of the flood coolant. To enhance the wettability and fluid thermal aspects, MQL nano-fluid approach has been used [8-11]. The nano-additives such as aluminum oxide (Al_2O_3) gamma nano-particles have superior cooling and lubrication properties. The effect of the nano-fluids technology on the performance of machining processes has been studied in several articles. For example, the nano-diamond particles were dispersed in paraffin and vegetable oils under MQL during micro drilling of aluminum 6061 [12]; aluminum oxide-based nano-fluid has been used in micro-grinding of tool steel [13]; MQL-nano-fluid with chilly CO_2 gas was implemented during micro-end-milling of Ti-6Al-4V [14]. It should be stated that MQL-nano-fluid showed promising results in enhancing the machining performance in all above-mentioned studies. Furthermore, multi-walled carbon nanotubes based nano-fluids with MQL showed better results in surface roughness [15] and flank wear [16] compared to classical MQL and Al_2O_3 based nano-fluids when machining Ti-6Al-4V. The current study aims to investigate the influence of using minimum quantity lubrication (MQL) technique with vegetable oil based nano-fluid, which includes

Aluminum Oxide (Al_2O_3) gamma nanoparticle, during machining Inconel 718. The tool wear, surface quality, and energy consumption are studied in the current work. This work offers multi-objective optimization of all studied responses by implementing the grey relational analysis. The importance of the proposed hybrid strategy is the replacement the traditional flood coolant technique using synthetic oils with an environment friendly alternative using MQL-nanofluid approach.

II. EXPERIMENTATION

Different cutting tests were carried out to understand and investigate the effects of the dispersed Al_2O_3 nanoparticles on the MQL performance. These cutting tests were performed with different levels of cutting speed, weight percentages of additives nano-additives, and feed rate. The utilized material of the workpiece which was used in these tests is Inconel 718 (ASTM SB 637). Thermal and mechanical properties of Inconel 718 are provided in Table I. The depth of cut and machining length of these tests were 0.2 mm 50 mm respectively. The specifications of the cutting insert is presented in Table II. It should be stated that machining length value was selected carefully to prevent any strike between the nozzle and the chuck. Sensitivity analysis was applied to study the variability effects of the cutting settings on the process (i.e. cutting speed, and feed rate). MQL approach used stand-alone booster system (Eco-Lubric) during these tests to supply the air-oil mixture with a nominal fluid flow rate of 40% and compressed air pressure of 0.5 MPa. ECOLUBRIC E200 (vegetable oil) has been selected a base oil.

TABLE I: THE THERMAL AND MECHANICAL PROPERTIES OF INCONEL 718

Inconel 718 (ASTM SB 637)	
Density	8193.25 Kg/m ³
Modulus of elasticity (at room temperature)	198.57 GPa
Poisson's ratio (at room temperature)	0.3
Tensile strength	930 MPa
Yield strength	482 MPa
Elongation percentage	45%
Hardness	43 Rc
Specific Heat (at room temperature)	435 J/kg °C
Thermal Conductivity (at room temperature)	11.2 W/m °C
Melting Point	1350 °C

TABLE II: THE CUTTING TOOL INFORMATION DURING MACHINING OF INCONEL 718

Inconel 718 cutting processes	
Cutting insert	CNMG 432MMH13A (ANSI)
Tool holder	KENAMETAL MCLNL-164DNC5, clearance angle 0° and rake angle -5°, nose radius of 0.793 mm (Honed edge)

Regarding the aluminum oxide nano-powder, the average diameter was 22 nm size, and the purity was about

91% purity. In order to disperse the nano-additives in the base oil, an ultrasonic device (AQUASONIC-50HT) has been employed for 3 hours at 60°C. Afterwards, the stirring step has been performed for 30 minutes using a magnetic stirrer (Hot Plate Stirrer-3073-21). In terms of the nano-fluid suspension stability, Zeta potential for stability analysis was conducted using a “zeta sizer-nano” device. A toolmaker's microscope has been employed to measure the flank wear after each cutting pass, and the “surface roughness tester” (Mitutoyo SJ.201) has been used to measure the average surface roughness (Ra). In order to measure the energy consumption after each cutting test, the power-sight manager device was used. In the current work, three design parameters were used with 3-levels each. Table III obtains the design parameters and the associated levels. L9 orthogonal array is implemented for the cutting tests design. Tests were repeated 3 times and average values of measured outputs were considered. In this study, the full-factorial array is L27OA (3^3); however, fractional factorial array L9OA depends on Taguchi approach was used. The current cutting tests include nine experiments where the first column was given to the cutting speed, while the second column represents the feed velocity levels, and the nano-additives weight percentage was given at the third column as shown in Table IV.

TABLE III: THE LEVELS ASSIGNMENT TO INDEPENDENT PROCESS PARAMETERS FOR INCONEL 718 CUTTING PROCESSES

Design Parameters	Symbol	Level1	Level2	Level3
Cutting speed (m/min)	A	30	40	50
Feed rate (mm/rev)	B	0.2	0.3	0.4
Nano-additives (wt. %)	C	0%	2%	4%

TABLE IV: THE DESIGN OF EXPERIMENTS OF THE STUDIED CUTTING TESTS

Test #	A: Cutting speed (m/min)	B: Feed rate (mm/rev)	C: Nano-additives concentration (wt. %)
1	30	0.2	0
2	30	0.3	2
3	30	0.4	4
4	40	0.2	2
5	40	0.3	4
6	40	0.4	0
7	50	0.2	4
8	50	0.3	0
9	50	0.4	2

III. RESULTS AND DISCUSSION

Table V shows the results of measuring the tool wear, averages surface roughness and the energy consumption. Regarding the tool life criteria, a maximum tool wear of 0.4 mm is used, otherwise, in case of not reaching the end of tool life criteria, the cutting stopped after nine cutting passes. The test performed with MQL-nano-fluids showed better results compared to the tests done with the classical MQL (i.e. tests 1, 6 and 8). The optimal tool wear was demonstrated at 4wt. %. Additionally, test 8 showed the highest value of flank wear, this test has been done using pure MQL (without nano-additives) at cutting speed of 30 m/min and feed rate of 0.4 mm/rev. In general, the tool wear results revealed that the

Al₂O₃ usage leads to a noticeable decrease in the measured flank wear compared to the classical MQL. In terms of the surface quality, cutting test 7 offered the lowest surface roughness. It can be noticed that better surface roughness results were obtained compared to the cutting tests done with the classical MQL, and these findings confirm the nano-fluids effectiveness in improving the heat dissipation and frictional behavior as discussed in the open literature. Regarding the optimal energy consumption, the lowest value was obtained at test 3 with 4 wt. % of Al₂O₃ nano-particles. The energy consumption results are mainly attributed to the improvements in the progression of tool wear as well as the cooling capabilities of the used nano-fluids which decrease the material thermal softening.

TABLE V: THE MEASUREMENTS OF THE MACHINING OUTPUTS

Test #	Flank tool Wear (mm)	Average surface roughness (μm)	Energy consumption (watt.hr)
1	0.295	3.21	1736
2	0.193	2.86	1650
3	0.176	3.87	1621
4	0.213	1.763	1766
5	0.207	2.45	1745
6	0.32	6.32	1807
7	0.245	1.513	1867
8	0.387	3.85	1963
9	0.235	4.2	1810

IV. MULTI-OBJECTIVE OPTIMIZATION OF MACHINING QUALITY CHARACTERISTICS

In this section, multi-objective optimization technique is employed to select the optimized levels of the design variables. The Grey Relational Analysis (GRA) provides situations (solutions) between black with no information and white with all information. The Grey Relational Analysis technique was previously used in different studies [17-20] to obtain the optimal operating levels associated with multiple performance characteristics in machining processes. The summarized steps of GRA techniques are provided as following:

- Data Normalization;
- Deviation sequence;
- Determination of the grey relational coefficients;
- Determination of the grey relational grades;
- Determination of the optimal process levels (ranking);
- The main effects on mean grey relational grade.

At the first step, the data are normalized by the linear normalization. As the all outputs (tool wear, surface roughness and energy consumption) are to be minimized, the data normalization are carried out by using equation (1).

$$x_{ij}^n = \frac{\max x_{ij} - x_{ij}}{\max x_{ij} - \min x_{ij}} \quad (1)$$

Where $i = 1, \dots, m$; $j = 1, \dots, k$. m is the number of the cutting tests and k is the number of the cutting test outputs. x_{ij} is the original value and x_{ij}^n is the value after normalization. After that, the sequence deviation is calculated by using the equations (2-4).

$$\Delta_{oi} = \|x_i^o - x_{ij}^n\| \quad (2)$$

$$\Delta_{\min} = \min_i \min_j \|x_i^o - x_{ij}^n\| \quad (3)$$

$$\Delta_{\max} = \max_i \max_j \|x_i^o - x_{ij}^n\| \quad (4)$$

Where x_i^o the ideal normalization value. At the third step the grey relational coefficient is determined by using the equation (5).

$$\xi_i(j) = \frac{\Delta_{\min} + \zeta \Delta_{\max}}{\Delta_{oi}(k) + \xi \Delta_{\max}} \quad (5)$$

Where the ζ is the distinguishing coefficient and its value is between 0 and 1. Equation (6) is applied to determine the grey relational grades.

$$r_i = \frac{1}{k} \sum_{j=1}^k \xi_i(j) \quad (6)$$

Where r_i is the grey relational grade. Then, the optimal test levels is selected due to the grey relational grade. Table VI shows the grey relational grades and rankings for all studied responses of using Al₂O₃ nano-fluid. It can be seen from Table VI, the optimal process levels are noticed at run #3 which have been performed at cutting speed of 30 m/min, feed rate of 0.4 mm/rev, and nano-additives weight percentage of 4 wt.%. The result showed that the increasing the nano-additive concentration has a significant effect on the cutting process characteristics. To have more accurate results of the optimal process levels, the main effects on mean grey relational grade are obtained. The mean grey relational grade is obtained through the average of the grades for each design variable at the same level. Figure 1 shows the values of the mean grade relational for each level of the design variables. It can be seen from Figure 1 that the level 1 of the cutting speed (A1) and the feed rate (B1) have the highest mean grade values. For the percentage of the nano-additives, the third level (C3) has the highest value. To obtain the influences of the design variables, the difference between the maximum and the minimum mean gray relational grades are calculated as shown in the table VII. According to these differences, the significance of the design variable changing can be obtained. It was found that the nano-additives percentage was the most significant variable affecting all studied responses.

In terms of the future work, the nanoparticle concentration effects need to be further emphasized in order to

understand the physical effect on the machining quality performance as the nano-particle affects the tribological and heat transfer characteristics of the resultant mist.

TABLE VI: GREY RELATIONAL RESULTS FOR USING Al_2O_3 NANO-FLUID

Run #	Grey relational grade	Grey relational rank
1	0.551	6
2	0.786	2
3*	0.835	1
4	0.729	3
5	0.691	4
6	0.412	8
7	0.672	5
8	0.391	9
9	0.529	7

*The optimal process levels by ranking

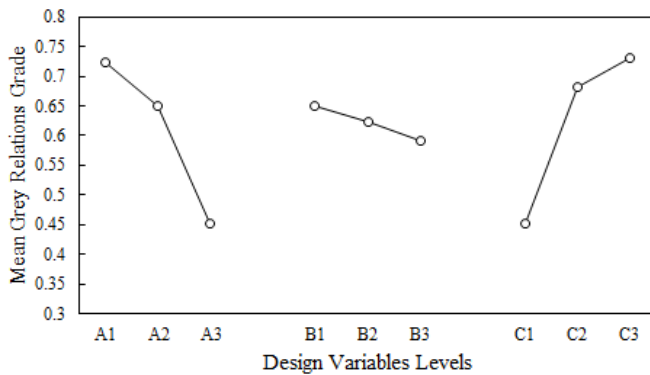


Figure 1. Mean grey relations grade of the design variables levels

TABLE VII: THE MAIN EFFECTS RESULTS FOR USING Al_2O_3 NANO-FLUID

Design variables	Max. - Min.	rank
A	0.195	2
B	0.134	3
C	0.285	1

V. CONCLUSIONS

The current study aims to investigate the influence of using minimum quantity lubrication (MQL) technique with vegetable oil based nano-fluid, which includes Aluminum Oxide (Al_2O_3) gamma nano-particles, during machining Inconel 718. The tool wear, surface quality, and energy consumption are studied in the current work. Different cutting tests were carried out to understand and investigate the effects of the dispersed Al_2O_3 nano-particles on the MQL performance. These cutting tests were performed with different levels of cutting speed, weight percentages of additives nano-additives, and feed rate. The optimal tool wear was demonstrated at 4wt. %. Additionally, test 8 showed the highest value of flank wear, this test has been done using pure MQL (without nan-additives) at cutting speed of 30 m/min and feed rate of 0.4 mm/rev. In terms of the surface quality, cutting test 7 offered the lowest surface roughness. It can be noticed that better surface roughness results were obtained

compared to the cutting tests done with the classical MQL. Regarding the optimal energy consumption, the lowest value was obtained at test 3 with 4 wt. % of Al_2O_3 nano-particles. Regarding the multi-objective optimization using Grey Relational Analysis (GRA), the optimal process levels are noticed at run #3 which have been performed at cutting speed of 30 m/min, feed rate of 0.4 mm/rev, and nano-additives weight percentage of 4 wt.%. In addition, the nano-additives percentage was the most significant variable affecting all studied responses. In terms of the future work, the nanoparticle concentration effects need to be further emphasized in order to understand the physical effect on the machining quality performance as the nanoparticle affects the tribological and heat transfer characteristics of the resultant mist.

ACKNOWLEDGMENT

The authors acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

- Shokrani, A., Dhokia, V., & Newman, S. T. (2012). Environmentally conscious machining of difficult-to-machine materials with regard to cutting fluids. *International Journal of machine Tools and manufacture*, 57, 83-101.
- Abukhshim, N. A., Mativenga, P. T., & Sheikh, M. A. (2006). Heat generation and temperature prediction in metal cutting: A review and implications for high speed machining. *International Journal of Machine Tools and Manufacture*, 46(7-8), 782-800.
- Aggarwal, A., Singh, H., Kumar, P., & Singh, M. (2008). Optimization of multiple quality characteristics for CNC turning under cryogenic cutting environment using desirability function. *Journal of materials processing technology*, 205(1-3), 42-50.
- Sun, S., M. Brandt, and M. Dargusch, Machining Ti-6Al-4V alloy with cryogenic compressed air cooling. *International Journal of Machine Tools and Manufacture*, 2010. 50(11): p. 933-942.
- Jafarian, F., et al., Experimental investigation to optimize tool life and surface roughness in inconel 718 machining. *Materials and Manufacturing Processes*, 2016. 31(13): p. 1683-1691.
- Akhtar, W., J. Sun, and W. Chen, Effect of machining parameters on surface integrity in high speed milling of super alloy GH4169/Inconel 718. *Materials and Manufacturing Processes*, 2016. 31(5): p. 620-627.
- Davim, J.P., *Machining of hard materials*. 2011: Springer Science & Business Media.
- Davim, J.P., P. Sreejith, and J. Silva, Turning of brasses using minimum quantity of lubricant (MQL) and flooded lubricant conditions. *Materials and Manufacturing Processes*, 2007. 22(1): p. 45-50.
- Emami, M., M. Sadeghi, and A.A. Sarhan, Investigating the effects of liquid atomization and delivery parameters of minimum quantity lubrication on the grinding process of Al_2O_3 engineering ceramics. *Journal of Manufacturing Processes*, 2013. 15(3): p. 374-388.
- Amrita, M., R. Srikant, and A. Sitaramaraju, Performance evaluation of nanographite-based cutting fluid in machining process. *Materials and Manufacturing Processes*, 2014. 29(5): p. 600-605.
- Khandekar, S., et al., Nano-cutting fluid for enhancement of metal cutting performance. *Materials and Manufacturing Processes*, 2012. 27(9): p. 963-967.
- Nam, J.S., P.-H. Lee, and S.W. Lee, Experimental characterization of micro-drilling process using nanofluid minimum quantity lubrication. *International Journal of Machine Tools and Manufacture*, 2011. 51(7-8): p. 649-652.
- Lee, P.-H., et al., An experimental study on micro-grinding process with nanofluid minimum quantity lubrication (MQL). *International Journal of Precision Engineering and Manufacturing*, 2012. 13(3): p. 331-338.

14. Kim, J.S., J.W. Kim, and S.W. Lee, Experimental characterization on micro-end milling of titanium alloy using nanofluid minimum quantity lubrication with chilly gas. *The International Journal of Advanced Manufacturing Technology*, 2017. 91(5-8): p. 2741-2749.
15. Hegab, H., et al., On machining of Ti-6Al-4V using multi-walled carbon nanotubes-based nano-fluid under minimum quantity lubrication. *The International Journal of Advanced Manufacturing Technology*, 2018: p. 1-11.
16. Hegab, H., et al., Performance evaluation of Ti-6Al-4V machining using nano-cutting fluids under minimum quantity lubrication. *The International Journal of Advanced Manufacturing Technology*, 2018. 95(9-12): p. 4229-4241.
17. Tzeng, C.-J., et al., Optimization of turning operations with multiple performance characteristics using the Taguchi method and Grey relational analysis. *Journal of materials processing technology*, 2009. 209(6): p. 2753-2759.
18. Kuram, E. and B. Ozcelik, Multi-objective optimization using Taguchi based grey relational analysis for micro-milling of Al 7075 material with ball nose end mill. *Measurement*, 2013. 46(6): p. 1849-1864.
19. Prasanna, J., et al., Optimization of process parameters of small hole dry drilling in Ti-6Al-4V using Taguchi and grey relational analysis. *Measurement*, 2014. 48: p. 346-354.
20. Krishnamoorthy, A., et al., Application of grey fuzzy logic for the optimization of drilling parameters for CFRP composites with multiple performance characteristics. *Measurement*, 2012. 45(5): p. 1286-1296.

Investigation of Hybrid MQL-Nano Fluid Performance When Machining Inconel 718 Super Alloy

S. Chauhan*, A. Eltaggaz
School of Engineering
University of Guelph,
Canada
*shakti@uoguelph.ca
aeltagga@uoguelph.ca

I.Deiab
School of Engineering
University of Guelph,
Canada
ideiab@uoguelph.ca

K. Khanafer
Mechanical Engineering
Department
Australian College of
Kuwait
k.khanafer@ack.edu.kw

Akrum Abdul-latif
Mechanical Engineering
Department, University of
Paris 8, Paris, France
abdul-latif@iu2t.univ-paris8.fr

Abstract— Machining Inconel 718 is a challenging task as it is a difficult to cut material. During machining these types of materials, a high amount of heat is generated at the tool-workpiece interface which affects the quality of the machined surface and the tool life. Therefore, the use of cutting fluid is required to reduce the high generated heat. However, due to detrimental effects of flood coolant on the operator health and environment, there is a constant effort by researchers to eliminate or reduce the use of conventional cutting fluid. An alternative sustainable coolant strategy that can be successfully used is minimum quantity lubrication (MQL). MQL is a more economical and environmentally friendly approach compared to conventional cutting fluid. However, MQL has insufficient cooling capacity. To overcome this, in the current paper, aluminum oxide (Al₂O₃) nanoparticles are added to the MQL base oil in order to enhance its thermal conductivity, heat capacity, and viscosity. The goal of this study is to investigate the effects of MQL-Nanofluid on the surface roughness, flank wear, and cutting forces during machining Inconel 718.

Keywords—minimum quantity lubrication; Inconel 718; nano-fluid; surface roughness; flank wear; cutting forces

I. INTRODUCTION

In metal cutting operations, the use of cutting fluids plays a crucial role in reducing the generated temperature at the tool-workpiece interface and, therefore, increases tool life and improves surface finish [1]. Although the conventional flood cooling method has a considerable ability to reduce the heat in the cutting zone, its negative environmental impact and the cost of the amount of cutting fluid used make it an unsustainable cooling strategy. Therefore, minimum quantity lubrication (MQL) is considered as a sustainable potential alternative cooling strategy to conventional flood cooling[2]. The concept of MQL in machining is using the minimum optimal amount of base cutting fluid. The flow rate for the MQL system generally varies from 6 ml/hr to 100 ml/hr which is very low as compared to conventional flood cooling, which uses up to 12000 l/hr. Therefore, MQL reduces the need to treat or dispose of the used coolant [3, 4].

However, even though the MQL approach satisfies environmental concerns, it may not match the high productivity of flood cooling when machining difficult-to-cut materials due to its insufficient cooling ability. This results in high cutting temperature in the cutting zone and increased the likelihood of tool wear.

To enhance the thermal conductivity and viscosity of the MQL approach, different types of nanoparticles have been added to the MQL base oil to create nanofluids [5]. Many researchers have concluded that nanofluid helps to reduce the friction between the two surfaces because of the ball bearing effect, i.e. such nanoparticles act as a spacer between the cutting tool and workpiece surface, thereby reducing the coefficient friction and rubbing action [5, 6, 7] as shown in Figure 1.

Zhang et al.[8] machined Inconel 718 using minimum quantity cooling lubrication (MQCL) in which they combined cryogenic air with the MQL system to enhance cooling performance during the end milling operation. They found that the tool life increased nearly 1.5 times when MQCL was used compared to dry cutting. Lower cutting forces were also observed under MQCL conditions. From environmental point of view, the use of biodegradable vegetable oil as a base fluid was more sustainable for achieving green manufacturing.

Hegab et al.[9] turned Inconel 718 using two types of nanoparticles (aluminum oxide and carbon nanotubes) to produce nanofluids. They found that using nanofluid cooling showed better results in terms of machined surface quality and energy consumption compared to machining without nanofluids. According to their analysis, they noticed that the percentage of nanoparticles has a major influence on the surface quality and on the energy consumption. The heat dissipation rate was increased using nanofluids leading to improved interface bonding between the workpiece and the cutting tool. They also found that multi wall carbon nanotubes showed comparatively better performance than aluminum oxide in terms of surface quality and energy consumption.

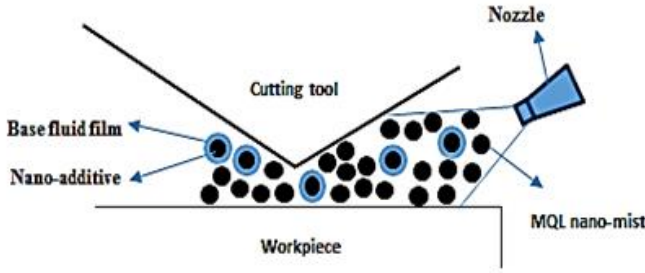


Figure 1: MQL-nanofluid mechanism[5]

Hasan et al.[10] conducted an end milling operation on Inconel 718 using super hard cobalt as a cutting tool. They used MQL vegetable oil at three flow rates (37.5, 25 and 12.5 ml/hr). The MQL cooling at flow rate of 37.5 ml / hr showed lower surface roughness than that at 12.5 and 25 ml/hr. The MQL approach also increased the cutting tool life compared to dry milling. Flank wear was increased in machining at a high cutting speed compared to a low speed, when the feed rate and depth of cut were the same.

Vincenzo et al.[11] machined Inconel 718 superalloy to investigate the best possible cooling approach in terms of cost, environmental impact, surface roughness, and tool life. The machining operation was conducted keeping the same depth of cut while varying the cutting speed. They used cemented carbide as a cutting tool as it has high compressive strength and hardness, low thermal expansion, and low conductivity. The experiments were conducted under five cooling strategies namely flood cooling, dry, minimum quantity lubrication, minimum quantity cooling lubrication, and the use of gaseous refrigerant at cryogenic temperatures. Overall, they found that MQL provided the better results in terms of cost and environmental impact.

Kamata et al.[12] performed machining experiments on Inconel 718 using three different coating carbide tools (PVD, PVD superlattice, and CVD). Two cutting speeds of 60 m/min and 90 m/min were chosen with a feed rate of 0.1 mm/rev and 0.1 mm depth of cut. The CVD coating was found to be the best for machining with MQL in terms of tool life and surface finish. The PVD superlattice showed the second best performance. This paper also discussed the importance of the air pressure, amount of lubricant used and carrier gas used for the MQL oil mist. Better tool life was obtained when the MQL air pressure was 0.40 MPa. The use of argon gas instead of air led to increased tool wear because lubrication characteristics and thermal conductivity of argon gas is poor. Moreover, at higher cutting speeds, the quantity of MQL base fluid played a crucial role in terms of increasing tool life.

In this paper, aluminum oxide (Al_2O_3) nanoparticles were dispersed at a concentration of 4% into the base fluid of MQL. Application of MQL-nanofluid for machining Inconel 718 super alloy was investigated for its effects on tool life, surface roughness, and cutting forces.

II. EXPERIMENTAL METHODOLOGY

Machining experiments were done using the CNC lathe machine (Hass ST-10 CNC). A cylindrical Inconel 718 (ASTM SB 637) with 70 mm diameter and 85 mm length was used as a workpiece. The cutting length was 72 mm, depth of cut was 0.5 mm, and PVD coated carbide was used as the cutting insert. The MQL system, which was supplied a vegetable oil base fluid, was mounted on the top of the CNC machine. The base fluid was supplied at a flow rate of 60 ml/h, while the delivered compressed air had a pressure of 0.5 MPa. High performance, yellow colored, vegetable-ester based micro-lubrication fluid was used as a base cutting fluid in the MQL system. The density and viscosity of this oil are 939 kg/m³ and 347 kg/m·s respectively. The orientation of the MQL nozzle was adjusted by experimental observation in order to avoid being blocked by chips. The flank wear was measured by using a Mitutoyo Toolmaker's microscope after each cutting test to ensure that the tool wear was still acceptable. In this study, the highest allowable flank wear was set to 0.3 mm. The flank wear was utilized to evaluate the cutting tool performance because the flank wear is one the most significant wear mode during cutting nickel-based alloys [13].

Furthermore, to ensure the reliable results, the machining trails were repeated two times for each run. During the experimentation phase, aluminum oxide (Al_2O_3), having 18 nm diameter, 93% purity, and 138 m²/g specific surface area, have been employed for nano-cutting fluid preparations (4% concentration). Furthermore, nanoparticles dispersion into vegetable oil is a significant issue, which affects the thermal conductivity and viscosity of the proposed nanofluid. An ultrasonic device (AQUASONIC-50HT) was employed to disperse the nanoparticles into the base cutting fluid over a period of 3 hr at 60 °C. To confirm fully dispersion of used nanoparticles into the base cutting fluid, a magnetic stirrer (Hot Plate Stirrer-3073-21) was used for the stirring step for 30 min. Sodium dodecyl sulfate (surfactant) of 0.2 gm was used in preparing the nanofluid. The main purpose of adding surfactant is to enhance the stability of the resultant nanofluid to be more hydrophilic, and to increase the nanoparticles surface charges.

After each cutting test, surface roughness (R_a), and cutting forces were collected. Surface roughness tester (Mitutoyo SJ.201) was used to measure the average surface roughness (R_a) after each cutting pass. After each cutting test, the surface roughness tester was used in five random regions along the machined surface and the average value was recorded. Additionally, the Kistler dynamometer 9257B was used to measure the cutting forces after each cutting pass. The experimental setup is shown in Figure 2.

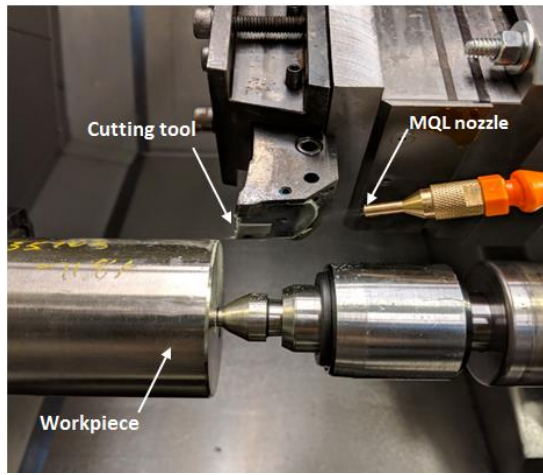


Figure 2: Experimental setup

III. RESULTS AND DISCUSSIONS

Three machining outputs, namely surface roughness, flank wear, and cutting forces, were determined and compared for all the coolant strategies.

A. Surface Roughness

In terms of surface roughness (R_a), the flood cooling showed the lowest value of surface roughness at all the cutting speeds as can be seen in Figure 3. The low value can be attributed to the better cooling efficiency of the flood cooling and the more effective washing away of chips from the cutting zone compared to MQL and MQL-nanofluid. However, the MQL-nanofluid provided better results than the MQL approach. This is because of the ball bearing mechanism of nanoparticles in MQL-nanofluid, which reduces the rubbing action between the tool and machined surface, thereby improving the surface quality.

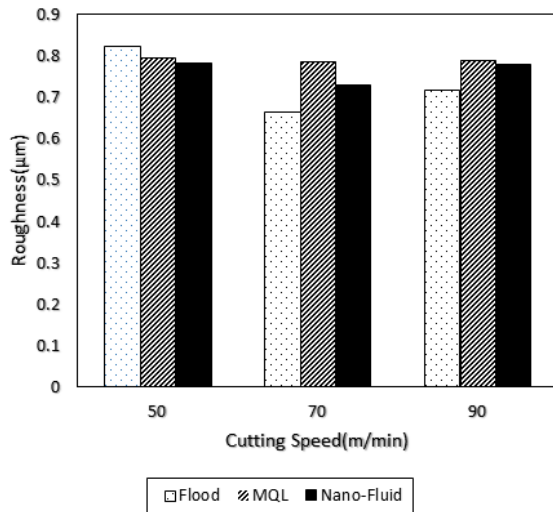


Figure 3: Surface roughness vs cutting speeds under flood, MQL, and MQL-nanofluid

B. Flank Wear

Measurements by Mititoyo Toolmaker's microscope after each pass showed that the use of MQL-nanofluid provided the lowest tool wear rate among the three cooling techniques at all the cutting speeds as shown in Figures 4, 5, and 6. Nanoparticles appear to improve the lubricating and cooling in the cutting zone and the MQL-nanofluid is able to absorb much of the heat during evaporation of that fluid.

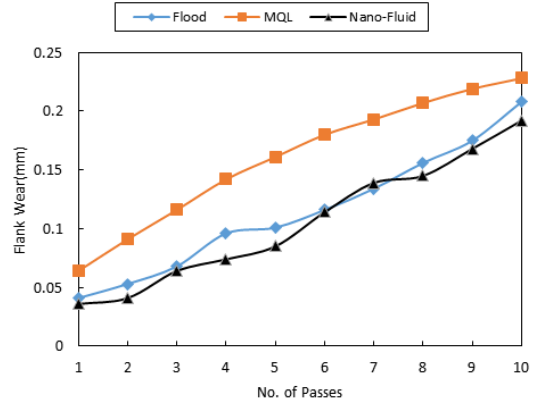


Figure 4: Flank wears vs no. of passes at a cutting speed of 50m/min

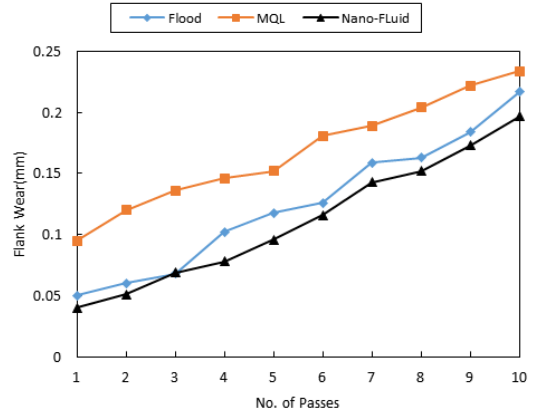


Figure 5: Flank wears vs no. of passes at a cutting speed of 70m/min

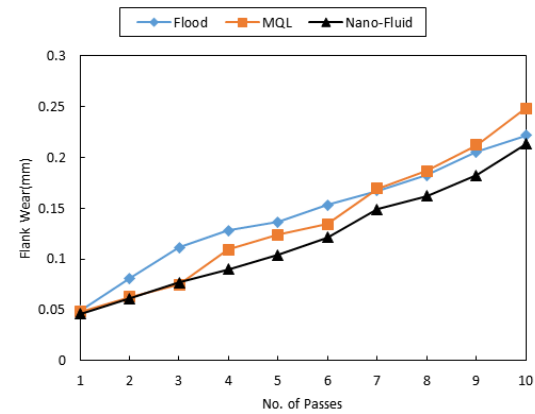


Figure 6: Flank wears vs no. of passes at a cutting speed of 90m/min

C. Cutting Forces

The cutting forces were lowest in machining with MQL-nanofluid due to two aspects of nanoparticles. The first is the thermal aspect, which enhances the thermal conductivity of MQL base fluid, thereby reducing the generated heat. The second is the tribological aspect: nanoparticles act as spacers between the workpiece and the cutting tool, which reduce the coefficient of friction and the rubbing action, thereby reducing the cutting forces as shown in Figures 7, 8, and 9.

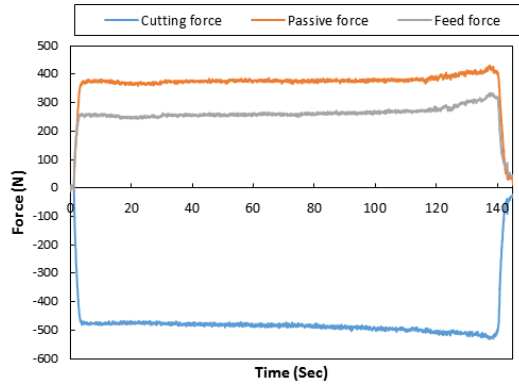


Figure 7: Cutting forces when using flood cooling

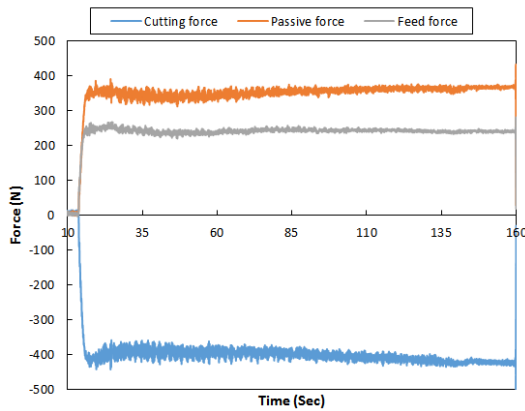


Figure 8: Cutting forces when using MQL cooling

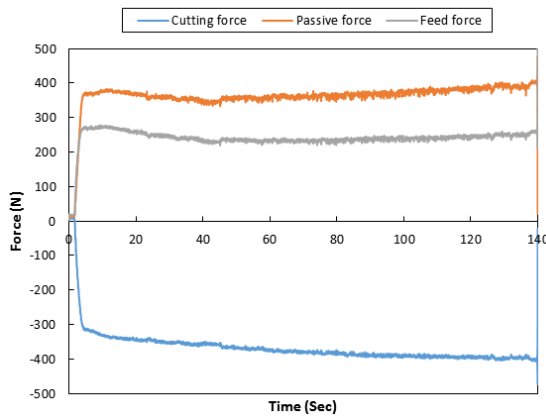


Figure 9: Cutting forces when using MQL-nanofluid cooling

IV. CONCLUSIONS

In this study, the effectiveness of adding nanoparticles to MQL vegetable oil was demonstrated. The experiments evaluated the machining performance of flood cooling, MQL, and MQL-nanofluid. Although the surface roughness was lowest with flood cooling, it was decreased to an acceptable level when applying MQL-nanofluid compared to applying MQL because nanoparticles enhanced the viscosity and thermal properties of the MQL base oil. The MQL-nanofluid is preferred over flood cooling because it is considered a more sustainable cooling strategy. The use of MQL-nanofluid considerably extended the tool life compared to MQL and flood due to better lubrication with MQL-nanofluid cooling. In addition, the use of MQL-nanofluid reduced the cutting forces as a result of the tribological aspect of nanoparticles. This study provides evidence that the MQL-nanofluid strategy can be an effective alternative sustainable cooling strategy to replace conventional cutting fluid strategies.

V. ACKNOWLEDGEMENT

This paper was funded fully by Kuwait Foundation for the Advancement of Sciences under project code: PR18-14SC-01.

VI. REFERENCES

- [1] R. S. G. S. J. Bonvoisin, *Sustainable Manufacturing*. 2017.
- [2] G. Jin and W. Li, "Environmentally Friendly Machining," *Handb. Manuf. Eng. Technol.*, pp. 3405–3435, 2015.
- [3] W. Liew and Y. Hsien, *SPRINGER BRIEFS IN MOLECULAR SCIENCE Towards Green Lubrication in Machining*. 2015.
- [4] D. Hörner, "Minimum Quantity Lubrication Fluids," *Encycl. Lubr. Lubr.*, pp. 1207–1211, 2014.
- [5] A. Eltaggag, H. Hegab, I. Deiab, and H. A. Kishawy, "Hybrid nano-fluid-minimum quantity lubrication strategy for machining austempered ductile iron (ADI)," *Int. J. Interact. Des. Manuf.*, vol. 12, no. 4, pp. 1273–1281, 2018.
- [6] A. K. Sharma, A. K. Tiwari, and A. R. Dixit, "Effects of Minimum Quantity Lubrication (MQL) in machining processes using conventional and nanofluid based cutting fluids: A comprehensive review," *J. Clean. Prod.*, vol. 127, pp. 1–18, 2016.
- [7] M. Garcia Tierno, "Towards sustainability using Minimum Quantity Lubrication technique and Nano-cutting fluids in metal-machining processes," 2017.
- [8] S. Zhang, J. F. Li, and Y. W. Wang, "Tool life and cutting forces in end milling Inconel 718 under dry and minimum quantity cooling lubrication cutting conditions," *J. Clean. Prod.*, vol. 32, pp. 81–87, 2012.
- [9] H. Hegab and H. Kishawy, "Towards Sustainable Machining of Inconel 718 Using Nano-Fluid Minimum Quantity Lubrication," *J. Manuf. Mater. Process.*, vol. 2, no. 3, p. 50, 2018.
- [10] S. Thamizhmanii and R. S. Hasan, "A study of minimum quantity lubrication on Inconel 718 steel," *Arch. Mater. Sci. Eng.*, vol. 39, no. 1, pp. 38–44, 2009.
- [11] V. Tebaldo, G. G. di Configno, and M. G. Faga, "Sustainability in machining: 'Eco-friendly' turning of Inconel 718. Surface characterisation and economic analysis," *J. Clean. Prod.*, vol. 140, pp. 1567–1577, 2017.
- [12] Y. Kamata and T. Obikawa, "High speed MQL finish-turning of Inconel 718 with different coated tools," *J. Mater. Process. Technol.*, vol. 192–193, pp. 281–286, 2007.
- [13] El-Bestawi, M.A.; El-Wardany, T.I.; Yan, D.; Tan, M. Performance of whisker-reinforced ceramic tools in milling nickel-based superalloy CIRP Ann. Manuf. Technol. 1993, 42, 99–102.

Global Statistics Based Method for Denoising of Planar Point Clouds

Cody Berry, Ahmad Barari

Faculty of Engineering and Applied Science

UOIT

Oshawa, Ontario

cody.berry@uoit.ca; ahmad.barari@uoit.ca

Abstract—Reflective surfaces present a difficult challenge in laser-based coordinate metrology. This paper discuss how the point clouds scanned from these surfaces are affected by noise and how the corresponding inspection results are influenced. The paper also presents a novel statistics-based methodology to effectively remove this noise from scanned surfaces. By repeatedly examining the distribution of Euclidian distances of the scanned point to a fit plane and constructing a diagnostic pattern of the distributions, a customized filtration criteria is obtained for the scanned data points. The developed method provides an efficient and effective solution for noise in laser-scanned point clouds from planar surfaces. The developed methodology have been examined on hundreds of data sets scanned from various manufacturing surfaces and highly satisfactory results were achieved.

Keywords—coordinate metrology; noise reduction; planar surfaces; reflective surfaces

I. INTRODUCTION

With the advent of Industry 4.0, there has been a drive in the manufacturing industry towards full automation of facilities. Because of this, careful consideration must be taken to ensure parts are inspected for defects properly, and unaided [1]. Despite best efforts, laser scanning manufactured parts, especially those with highly reflective surfaces, comes with challenges [2]. With these surfaces, as the light from the laser can be reflected in many directions, or light from the surrounding area reflects off the surface back into the receiver, noisy points, or “ghost” points can be generated [3]. These points can greatly impact the accuracy of a scan and create doubt in an otherwise excellent specimen. To keep the spirit of automation, it was important to develop a method that could identify and remove most of this noise without human intervention.

Many noise reduction algorithms have been produced in the past [4]–[8]. These algorithms are differentiated by how they examine the point cloud for noise. Weyrich et al. [4], looked at small neighbourhoods of points and examined the probabilities a point may be noise by examining its relationship to the points in its immediate vicinity. This method was effective at removing noise but can take a long time to finish on data sets

with large amounts of data. Schall et al. [6] also examined the neighbourhood around points, instead generating likelihood functions over the entire surface. Ning et al. [9] deals with local information as well, combining multiple algorithms to output a final denoised point cloud. Our goal was to develop a method that examined the point cloud, with a global method.

Coordinate metrology incorporates the use of three different processes, namely, Point Measurement Planning (PMP), Substitute Geometry Estimation (SGE), and Deviation Zone Evaluation (DZE). In such systems, the inspection accuracy is subject to the issue of uncertainties due to estimating some aspects of a probability distribution on the basis of sampling which adversely affects the accuracy and reliability of the inspection. Selection of location, distribution, and number of sample points using the available information is the objective of PMP process. Finding the best ideal substitute geometry that represent the selected sample points using a desired fitting criteria is performed by SGE process. Developing a comprehensive representation of geometric deviation for any point of the measured surface including the points that are not physically measured, is the goal of DZE process. These three computational tasks were traditionally conducted in a sequential manner. All three tasks have access to three sets of external data including: the desired geometric features which are typically represented by Computer Aided Design (CAD) models, data and information from the corresponding manufacturing operations used to fabricate the work piece, and

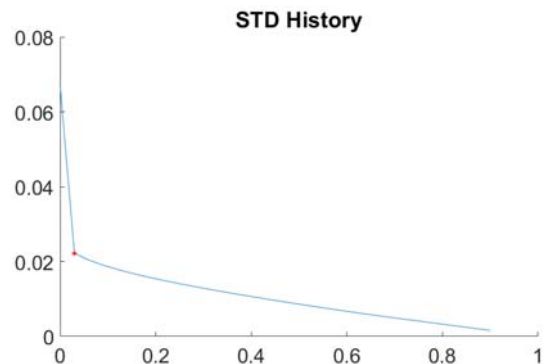


Figure 1. Ideal STD History Graph



Figure 2. Test Samples a) Brushed Aluminum b) Selective Laser Sintered c) Gauge Block

the objective of the inspection process corresponding to the downstream processes that are planned for the measured work piece [10-12]. It is shown in several recent research [13-17] that integration of PMP, SGE, and DZE tasks will result in reducing the computational uncertainties in coordinate metrology.

II. METHODS

Initially, the point cloud is loaded into memory as x, y, and z coordinates representing each point in space. A weight matrix is generated containing all ones, the same length as the point cloud. Then, a weighted total least squares (WTLS) algorithm is used to fit a plane to the entire data set. As the weight matrix is all ones, all points will be used in this calculation. This weight matrix will be utilized as a “decision” matrix, determining whether a point is being used in calculations.

With this initial fit plane determined, the orthogonal distance of each point to the fit plane is calculated using the normal vector $n=[n_1, n_2, n_3]$, point matrix containing all the points $P=[p_1, p_2, \dots, p_n]$, and an offset d , calculated using the following equation

$$d = -w \cdot n \quad (1)$$

where w is the centroid of the point cloud. Using these values, the distance vector, D , is calculated using the following equation.

$$D = (n \cdot P + d) / n \quad (2)$$

With this distance vector calculated, the standard deviation of the distance of all points to the fit plane is calculated. This standard deviation is an important value for this method, as it will form the basis of deciding what percentage of points needs to be filtered. With this standard deviation value, the first

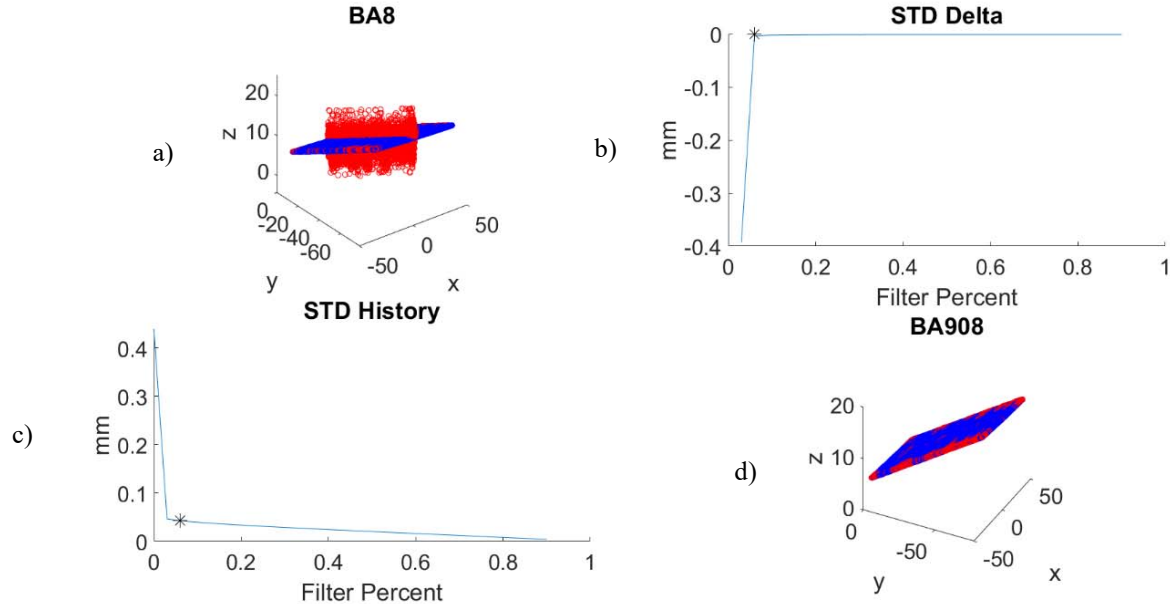


Figure 3. Analysis of Brushed Aluminum Part a) Full Data Set and Filtered Points b) STD Delta graph showing change in STD for each iteration c) STD History graph showing STD for each iteration d) Data set for specimen rotated 90°

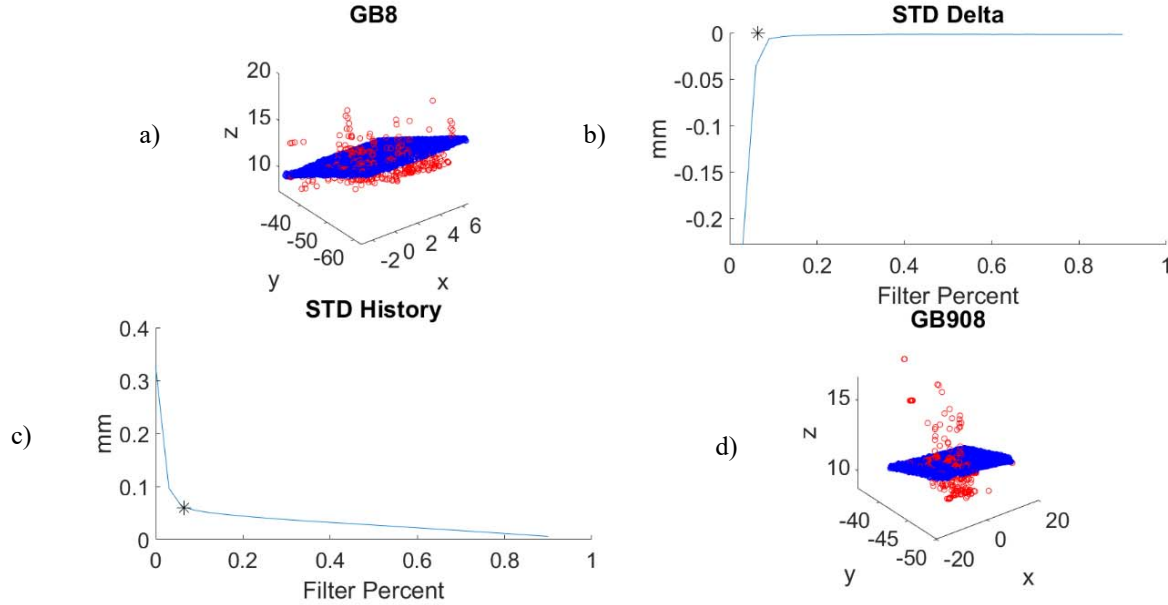


Figure 4. Analysis of Gauge Block a) Full Data Set and Filtered Points b) STD Delta graph showing change in STD for each iteration c) STD History graph showing STD for each iteration d) Data set for specimen rotated 90°

filtering of points is performed. A standard deviation factor (SDF) is chosen, and all points with a distance from the plane greater than the product of the SDF and the standard deviation of the point cloud is given a weight of 0. The percentage of points removed is calculated, and one tenth of this value becomes the reduction step. If this value is less than 3%, it is raised to 3% in order to minimize the run time of the code.

With the reduction step calculated, the filtering begins. Much like the start of the process, a plane is fit to the full data set and the point-plane distance is calculated for each point. Then, an amount of points equal to the reduction percent is removed, starting with the point furthest away. Another plane is fit to this new data set, and the standard deviation of point-plane distances is once again calculated and stored. This process continues until 90% of the points have been filtered out or the change in the change of the standard deviation becomes sufficiently small.

The matrix containing the history of the change of the standard deviations of the point-plane distances is utilized for the final step in the process. By taking the two end points of the matrix and connecting them, a line was formed. The distance of each other value in the matrix to the developed was calculated, and the filtered percentage associated with the value that had the largest orthogonal distance from the line was chosen as the optimal filtered percentage. The full data set was once again filtered using this percentage, and the final data set with decision matrix was returned. The goal of this method is to find the location where the STD of the point set stops changing at a rapid rate, where we assume most of the noise has been removed. The ideal case would have two linear sections, the first section with a steep slope, and the second with a much shallower slope, as shown in Fig. 2. Then, the filtered percentage chosen would be at the vertex where the two lines meet.

III. RESULTS

In order to test the efficacy of the filtering algorithm, three different samples were utilized; a gauge block, a selective laser sintered (SLS) block, and a brushed aluminum plate. These samples were chosen due to their interesting material characteristics. The gauge block is relatively reflective, and so exhibits some of the characteristics the algorithm is meant to handle. The SLS block, due to its powder-based composition, has light scattering qualities that are unpredictable, and form a good test of the algorithm's ability to handle non-systemic noise. The brushed aluminum is a minimally reflective part overall but does have highly reflectively areas located on its surface due to the processing techniques used to create it. Each of the samples used is shown in the Fig 2.

With each of these samples, multiple scans with different parameters were taken. The first parameter was the angle between the scanning head and the surface. This was changed from -25° to +25° in increments of 5°. This allowed for a wide range of scanning angles and possibility of reflection. Another parameter was the orientation of the specimen with regards to the scanning head. After all scans were concluded, the specimen was rotated 90°. This was used to determine whether errors were being induced by the sensor or by the specimen itself. The final parameter changed was the exposure settings of the laser. All parts were also scanned at an appropriate level determined by the camera software, then more scans were taken at a higher and lower exposure. This allowed for a wider range of noise to be captured.

In each of the results figures, four images are presented. The top left image is a representation of the point cloud after it has been filtered. The blue dots are those that were not filtered by the algorithm, and the red dots are those that were. The top right image is of the STD delta as the filtered percentage

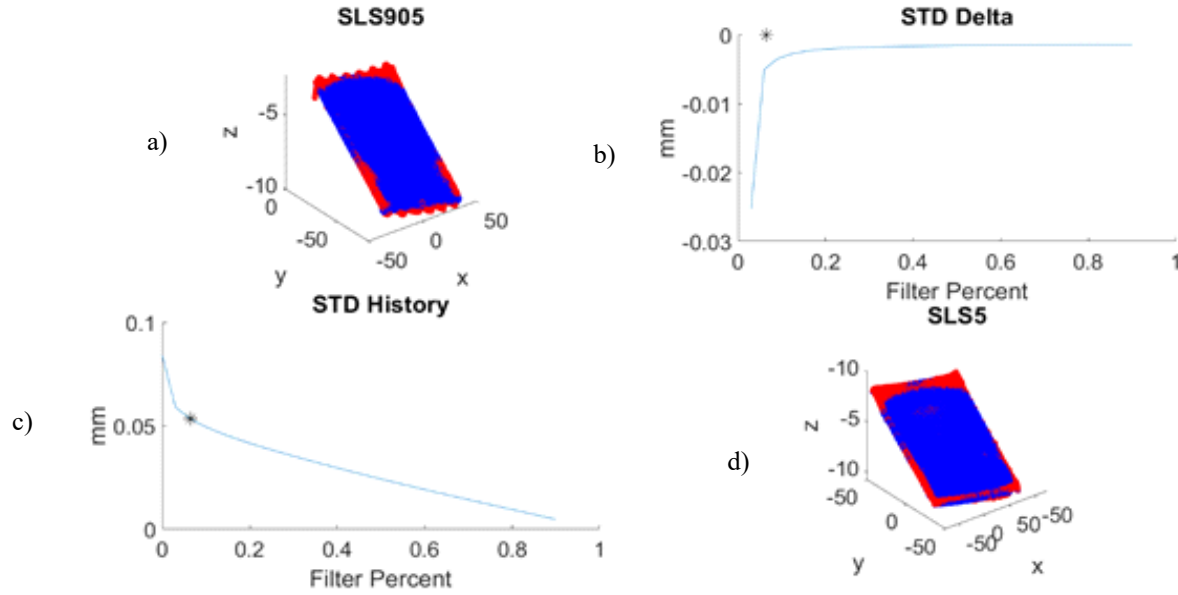


Figure 5. Analysis of Selective Laser Sintered Part a) Full Data Set and Filtered Points b) STD Delta graph showing change in STD for each iteration c) STD History graph showing STD for each iteration d) Data set for specimen rotated 90°

increased, and there is a black star representing the chosen filtered percent. The bottom left image is the history of the STD throughout the filtering process. The final image is a scan of the same specimen with all the same parameters, except the specimen has been rotated 90°.

In Fig. 3, the brushed aluminum sample was scanned. The scanning angle was +5°, the exposure was automatically set by the software, and the specimen had not yet been rotated. With this scan, there was a band of noise along the y-axis, spanning both above and below the real specimen data. The STD Delta and STD History graphs both show the ideal result for this algorithm. This is a two-part linear piecewise function, with a distinct vertex. For this dataset, there was an initial size of 249183 points, and the algorithm removed 6.013% of these values. The initial STD was 0.43903 mm, and the final was 0.04263 mm, a reduction of 90%.

In Fig. 4, the gauge block specimen was used. Like the last example, a scanning angle of +5° was used, and the exposure was set by the software. Again, the specimen had not been rotated. Unlike the previous example, the filtered points in this scan are randomly distributed above and below the surface. There is a greater density of noise below the surface, with sparser patterns of noise above. The curve seen in both the STD Delta and STD History graphs does not exhibit the ideal 2-part piecewise linear curve. Despite this, with a 6.5067% reduction

in points, there was an associated 82% reduction in the STD of the points.

Another example, as seen in Fig 5. uses the SLS part. The scan is at -5°, with an automatically defined exposure, after the piece has been rotated 90°. There are noise patterns on either end of the sampled area with a repeating pattern. The two STD graphs do not exhibit the linear piecewise behaviour that is considered ideal. The curve of the STD History graph is much more gradual overall. For this workpiece, there was a 37% reduction in the STD of the point-plane distances. Table 1 shows the numerical results for these tests.

IV. DISCUSSION

Looking at the graphs and table in the previous section shows some of the patterns experienced by the algorithm. The brushed aluminum and gauge block examples had data sets where the noise was above and below the actual surface, and there was noise relatively far from the surface itself. In both instances, the algorithm was able to effectively remove the noise and provide a more realistic flatness results for these workpieces. However, in the case of the SLS workpiece, the noise wasn't the result of light reflecting off the surface, so the noise was close to the real surface. For this example, the STD history graph did not follow the linear piecewise function that is considered ideal. Instead, this graph shows a result closer to

TABLE I. NOISE REDUCTION RESULTS

Sample	Number	Number of Points	Original STD	Processed STD	Point Reduction Percentage	STD Reduction
BA	8	249183	0.43903	0.042625	6.0131	-90%
GB	8	7271	0.32538	0.060051	6.5067	-82%
SLS90	5	162012	0.084121	0.053362	6.3882	-37%

a linear graph.

This is one of the issues with the algorithm. As it looks only at the number of points a certain distance away, there is no check as to whether those points might be part of the real data set. In scans where there was little to no noise, the STD History graph was nearly linear. Because of this, there would be large reductions in the number of points for a small reduction in the STD of the point-plane distances.

V. SUMMARY AND CONCLUSIONS

In this paper we sought to develop a global, statistics-based algorithm to denoise point clouds of planar surfaces. In the first section, the problem was defined and previous solutions to the problem were examined. The next section developed the methodology for the algorithm. Then, the three samples were scanned repeatedly and processed via the algorithm. Interpretations of the results were then provided.

With the developed algorithm, a reduction the standard deviation of the point-plane distance was achieved in every case. Visual inspection of the data sets showed a reduction in the amount of noise in the data sets. In cases where the data set initially had a very low amount of noise, the algorithm overcompensated and removed real data from the point cloud. This work will be extended to include processing of any surface.

REFERENCES

- [1] C. Berry and A. Barari, "Closed-Loop Coordinate Metrology for Hybrid Manufacturing System," *IFAC-PapersOnLine*, vol. 51, no. 11, 2018.
- [2] Y. Wang and H. Y. Feng, "Modeling outlier formation in scanning reflective surfaces using a laser stripe scanner," *Meas. J. Int. Meas. Confed.*, vol. 57, pp. 108–121, 2014.
- [3] Y. Wang and H. Y. Feng, "Effects of scanning orientation on outlier formation in 3D laser scanning of reflective surfaces," *Opt. Lasers Eng.*, vol. 81, pp. 35–45, 2016.
- [4] T. Weyrich, M. Pauly, S. Heinzle, R. Keiser, S. Scandella, and M. Gross, "Post-processing of Scanned 3D Surface Data," *Symp. Point-Based Graph.*, pp. 85–94, 2004.
- [5] L. Kobbelt and M. Botsch, "A survey of point-based techniques in computer graphics," *Comput. Graph.*, vol. 28, no. 6, pp. 801–814, 2004.
- [6] O. Schall, A. Belyaev, and H.-P. Seidel, "Robust filtering of noisy scattered point data," in *Proceedings Eurographics/IEEE VGTC Symposium Point-Based Graphics, 2005.*, 2005, vol. 75, no. 4, pp. 71–144.
- [7] G. Rosman, A. Dubrovina, and R. Kimmel, "Patch-collaborative spectral point-cloud denoising," *Comput. Graph. Forum*, vol. 32, no. 8, pp. 1–12, 2013.
- [8] X. F. Han, J. S. Jin, M. J. Wang, W. Jiang, L. Gao, and L. Xiao, "A review of algorithms for filtering the 3D point cloud," *Signal Process. Image Commun.*, vol. 57, no. February, pp. 103–112, 2017.
- [9] X. Ning, F. Li, G. Tian, and Y. Wang, "An efficient outlier removal method for scattered point cloud data," *PLoS One*, vol. 13, no. 8, pp. 1–22, 2018.
- [10] A. Lalehpour and A. Barari, "Developing skin model in coordinate metrology using a finite element method," *Measurement*, vol. 109, pp. 149–159, 2017.
- [11] Lalehpour, A., Berry, C., Barari, A., "Adaptive data reduction with neighbourhood search approach in coordinate measurement of planar surfaces", *Journal of Manufacturing Systems*, Volume 45, Pages 28-47, 2017
- [12] Lalehpour, A., Barari, A., "Developing Skin Model in Coordinate Metrology Using a Finite Element Method", *Measurements*, Volume 109, Pages 149-159, 2017
- [13] A. Barari, S. Jamiolahmadi, "Convergence of a finite difference approach for detailed deviation zone estimation in coordinate metrology", *Acta IMEKO*, Vol. 4, No 4, pp. 20 - 25, 2015.
- [14] A. Barari, F. Kaji, "The uncertainty of coordinate metrology as a function of sampling strategy", *Journal of CMSC*, Vol. 9, No 1, pp. 22-27, 2014.
- [15] Jamiolahmadi, S., Barari, A., "Study of detailed deviation zone considering coordinate metrology uncertainty", *Measurements*, In-press 2016
- [16] H. Gohari, A. Barari, "A quick deviation zone fitting in coordinate metrology of NURBS surfaces using principle component analysis", *Measurements*, 2016
- [17] A. Barari, "Inspection of the machined surfaces using manufacturing data", *Journal of Manufacturing Systems*, Volume 32, Issue 1, January 2013, Pages 107-113

Biomechanics

Artificial Intelligent Examination of DXA Image Quality for Automatic Finite Element Assessment of Hip Fracture Risk

Wanshi Yu* and Yunhua Luo

Department of Mechanical Engineering
Faculty of Engineering, University of Manitoba
Winnipeg, Canada

Corresponding e-mail: Yunhua.Luo@umanitoba.ca

INTRODUCTION

Hip fracture has high incidence among elderly people worldwide due to prevalence of osteoporosis in the population. Hip fracture has severe health and economic consequences, but it is preventable via drug treatment or protective measurement. Therefore, assessment of fracture risk for elderly people is an important examination in clinic. Clinicians mainly rely on information extracted from hip DXA (dual energy X-ray absorptiometry) image to evaluate an individual's hip fracture risk. Currently available clinical tools are based on statistical models developed for different race groups. These tools have been reported inaccurate to predict fracture risk in both cross-sectional and prospective population studies. Our group previously developed and automated a finite element tool from an individual's hip DXA image to predict fracture risk [1]. Cross-sectional studies show that the finite element tool has promising performance in discrimination of fracture cases from non-fracture controls [2, 3]. However, the finite element tool does not have a mechanism to examine DXA quality, poor DXA quality leads to inaccurate prediction of fracture risk for the tool.

METHODS

Visual examination method was used in our group to check DXA quality before applying the finite element tool. The main drawbacks of visual examination method are subjective and inconsistency, it is not feasible for large cohort population study. Therefore, we decided to develop an automatic algorithm to exclude unqualified DXA images and incorporate it into the finite element tool. The main challenge in developing the algorithm was to identify and quantify features of common defects in clinical DXA images. Based on our experience of visual DXA examination and requirements on finite element meshes, we identified four defects that commonly occur in clinical DXA images. We quantified features of the four defects. Then we used the quantified features to train a k-nearest neighbors (k-NN) algorithm and an artificial neural network (ANN). The trained k-NN and ANN were then applied to classify new clinical DXA images.

RESULTS

We acquired 11707 clinical DXA images from St. Boniface General Hospital (Winnipeg, Canada) under a research ethical approval issued by the University of Manitoba. We equally split the images into training and validation sets. We used defect features extracted from the training set to train the k-NN and ANN. Convergences of training processes were monitored. The trained k-NN and ANN were applied to identify defective DXA images in the validation set. The accuracy of k-NN and ANN in identifying DXA images of the four defects were 99.0% and 99.1%, respectively.

CONCLUSIONS

The trained k-NN and ANN were able to accurately identify unqualified DXA images with the four major defects. However, there are still other minor defects that affect the accuracy of predicted fracture risk, which will be addressed in future studies.

ACKNOWLEDGEMENT

The reported research has been supported by grants from Research Manitoba and Natural Sciences and Engineering Research Council, which are gratefully acknowledged.

REFERENCE

- [1] Y. Luo, S. Ahmed, and W.D. Leslie. *Computer Methods and Programs in Biomedicine*, 155:75 – 83, 2018.
- [2] S. Yang, Y. Luo, L. Yang, et al. *Bone*, 110:386 – 391, 2018.
- [3] S. Yang, W.D. Leslie, Y. Luo, et al. *Osteoporosis International*, 29:191 – 200, 2017.

Computational Investigation into Materials with Lower Stiffness for Hemiarthroplasty

Carolyn Berkmortel, G. Daniel Langohr, James A. Johnson
Mechanical and Materials Engineering
University of Western Ontario
London, Canada

Graham J.W. King
Saint Joseph's Healthcare
London, Canada

ABSTRACT

INTRODUCTION

Hemiarthroplasty is advantageous over replacement of both sides of the joint as it allows for the preservation of tissue and is less invasive. This results in faster healing and lower costs to the healthcare system. The main concern with hemiarthroplasties is that they cause accelerated wear of the opposing cartilage. It is generally believed that this is due to the stiffness mismatch between the cartilage and the implant material. The purpose of this study was to investigate a range of currently available biomaterials for hemiarthroplasty applications. We employed a finite-element (FE) model of a radial head implant against the native capitellum as a joint model.

METHODS

The FE model was developed in ABAQUS v6.14 (Dassault Systèmes Simulia, USA). A solid axisymmetric concave implant with seven different materials and the native radial head were evaluated; six modelled as elastic materials having varying Young's moduli (E) and Poisson's Ratios (ν), and one modelled as a Mooney-Rivlin hyperelastic material. The materials investigated (listed from highest to lowest stiffness) were CoCr, PEEK, HDPE, UHMWPE, Bionate 75D, Bionate 55D, and Bionate 80A. A load of 100N was applied to the radius through the center of rotation representing a typical load through the radius. The variable of interest was articular contact stress on the opposing native capitellum.

RESULTS

The CoCr implant had a maximum contact stress over 114% higher than the native radial head. By changing the material to lower the stiffness of the implant, the maximum contact stress decreased for each subsequently less stiff material. The maximum contact stress was 24%, 70%, 105%, 111%, 113%, and 113% higher than the native radial head for Bionate 80A, Bionate 55D, Bionate 75D, UHMWPE, HDPE, and PEEK respectively.

DISCUSSION AND CONCLUSIONS

These data show that lowering implant stiffness can reduce cartilage contact stress following hemiarthroplasty, specifically by using a material having a Young's modulus of ~ 100 MPa or less. This has the potential to reduce or prevent degenerative changes of the native articulating cartilage. Low stiffness implant materials are not a novel concept, but to date there have been few that investigate materials (such as Bionate) as a potential load bearing material for implant applications. Our data shows that less stiff materials such as PEEK, HDPE and UHMWPE which have been thought to reduce cartilage wear do not seem to have a marked effect on contact mechanics. Further work is required to assess the efficacy of these materials for articular bearing applications.

Design of a Biaxial Mechanobioreactor for Engineering Pediatric Aortic Valves

Edwin Wong and Craig A. Simmons
Mechanical and Industrial Engineering
University of Toronto
Translational Biology & Engineering Program
Ted Rogers Centre for Heart Research
Toronto, ON
edwinky.wong@mail.utoronto.ca

I. INTRODUCTION

The ability of tissue engineered heart valves (TEHVs) to grow, repair and remodel *in vivo* makes them a potential gold standard for pediatric valve replacement. In theory, designing a TEHV that biomimics the anisotropic mechanical properties and extracellular matrix (ECM) architecture of healthy native valves should yield a construct with the desired functions. Currently, there is a limited understanding of *in vitro* mechanical stimulation protocols used to stimulate cells to produce and remodel the ECM of a construct. For heart valve applications, pressure-based bioreactors have been designed to apply cyclic biaxial tensile strain to cell-seeded substrates. Unfortunately, these devices only have a single input to control bidirectional strain application. This prevents independent modulation of directional strain for tuning of the construct's mechanical anisotropy *in vitro*. An alternate approach is to stretch sheets of tissue biaxially with each stretch axis uncoupled. To date, such existing devices are only designed to apply uniform biaxial strain to small squared specimens (< 25 mm x 25 mm). These designs are not optimized to uniformly strain TEHV constructs (~40 mm x 60 mm) and may lead to poorly controlled tissue remodelling across the tissue sheet. As such, we aim to design and optimize a novel bioreactor to stretch TEHV constructs biaxially in a planar configuration.

II. METHODS

Finite element models (FEMs) with isotropic ($E = 2$ MPa) and anisotropic ($E_x = 16$ MPa, $E_y = 1.2$ MPa) material properties were created in ANSYS 14.0 to: i) design a novel configuration that maximizes biaxial strain uniformity on a tissue sheet; and ii) investigate strain patterns generated due to material and geometric anisotropy. A design of experiments approach was used to optimize geometry and dimensions of each configuration to maximize the uniformity of the applied strain to the model (JMP 14). To evaluate strain uniformity, a region was classified as uniform if the X and Y direction strain magnitudes were within 10% of the strain magnitudes at the center (midpoint) of the specimen. Configurations were compared to the trampoline setup, the literature gold standard for maximum strain uniformity. To evaluate strain patterns, midpoint strains were measured for different side lengths (from 20 mm to 60 mm), side length ratios ($1 \leq L_x/L_y \leq 3$), Young's moduli (1.2 MPa to 16 MPa) and material anisotropy ($1 \leq E_x/E_y \leq 13.3$). Simulations were validated by biaxially testing silicone samples (BioTester, Cellscale, Waterloo, ON) and using digital image correlation to measure strain. Solidworks was used to model prototypes that were then built in-house.

III. RESULTS AND DISCUSSION

A novel "capstan+cruciform" configuration was determined to improve and maximize strain uniformity by 40.9% (isotropic materials) and 39.5% (anisotropic materials) over the standard trampoline setup. The "capstan+cruciform" setup is designed to mount specimens without puncturing them and to maintain tension in corner regions, which the trampoline setup does not do. To our knowledge, these are the first reported values of strain uniformity and optimization of a biaxial strain setup for anisotropic tissues. For strain pattern evaluations, the directional midpoint strain ratios (ϵ_x/ϵ_y) were found to be coupled to the specimen's Young's modulus anisotropy (E_x/E_y) and the side length ratios (L_x/L_y). Increasing the Young's modulus anisotropy led to a decrease in ϵ_x/ϵ_y ($E_x/E_y = 1$, $\epsilon_x/\epsilon_y = 1$; $E_x/E_y = 13.3$, $\epsilon_x/\epsilon_y = 0.67$). Conversely, increasing side length ratios increased ϵ_x/ϵ_y ($L_x/L_y = 1$, $\epsilon_x/\epsilon_y = 1$; $L_x/L_y = 3$, $\epsilon_x/\epsilon_y = 48.9$). Both the individual side lengths and Young's modulus had minimal effect on midpoint strains. Since specimens experience creep from cyclic straining and material properties evolve throughout culture, these results emphasize the need for directional strain modulation to ensure consistency in the strain applied to the cells on the TEHV.

IV. CONCLUSION

We have designed a novel biaxial strain configuration that maximizes the applied strain uniformity for anisotropic tissue engineered constructs inside a bioreactor. We demonstrated that strain magnitudes are coupled with the specimen's material and geometric anisotropy, parameters are expected to evolve throughout culture. For future work, we aim to implement a feedback system that measures the *in vitro* mechanical properties of the specimen and modulates the applied strain as the TEHV develops.

Energy Loss of Aortic Wall in Aneurysms Predicts Tissue Delamination

Edwin Wong, Mingyi Tang, and Craig A. Simmons

Mechanical and Industrial Engineering

University of Toronto

Translational Biology & Engineering Program

Ted Rogers Centre for Heart Research

Toronto, ON

Email: edwinky.wong@mail.utoronto.ca

Jennifer Chung¹, Thomas Forbes², Jagdish Butany³
and Maral Ouzounian¹

¹Division of Cardiac Surgery

²Division of Vascular Surgery

³Division of Pathology

University of Toronto

Toronto, ON

I. INTRODUCTION

The biomechanics of aortic aneurysms are poorly understood, making it difficult to predict which patients should be operated on because their aneurysms are at risk of dissection. Dissections are life-threatening events that occur when blood enters the medial (middle) layer of the aortic vessel wall causing the outer layer (adventitia) to delaminate from the inner layer (intima). Currently, risk is assessed by measuring the diameter of the aortic aneurysm, typically with images obtained from clinical computed tomography. However, this metric leads to missing many patients who then present with aortic dissection. One proposed paradigm for a clinical biomarker is to measure biomechanical properties of these diseased tissues *in vivo*. One such biomechanical parameter is energy loss (ΔU_L): the percentage of mechanical energy per unit volume dissipated between the loading phase and unloading phase of a material. For aneurysmal aortas, ΔU_L has been found to positively correlate with pathological changes to the microstructure of the aneurysmal tissues [1]. Other parameters frequently explored include the tissue's Young's modulus (E) at various loads and the modulus anisotropy index (AI). There have yet to be any studies evaluating correlations between a failure criterion such as the delamination strength of the tissue (S_d) and a biomechanical parameter. Our aim is to determine if biomechanical parameters such as ΔU_L , E and AI can be used to predict the forces needed to dissect aneurysmal aortas.

II. METHODOLOGY

Aneurysmal ascending aortas were collected from patients during surgery ($n = 43$) while non-diseased aorta controls were collected from transplant donors ($n = 8$). Aortic tissues were cut using custom-made cutters to 14 mm x 14 mm square samples. Samples were subjected to 10 preconditioning cycles followed by 3 analyzed cycles under 25% equibiaxial strain (BioTester, Cellscale) to calculate ΔU_L and E at 10% strain in both directions of loading as well as AI for both parameters. 6 mm x 30 mm strips were delaminated by uniaxially peeling the adventitia from the intima. The average force to peel through the strip was normalized by the width of the specimen to determine S_d . Pearson correlation coefficients were calculated using linear regression between potential clinical markers (ΔU_L , E, AI and aneurysm aortic diameter) and S_d (Prism 5, Graphpad). Finally, diseased samples were divided based on whether the patients' aortic valves had two (bicuspid) or three (tricuspid) cusps. One-way ANOVA was performed between bicuspid, tricuspid and healthy control groups for ΔU_L , E and AI to determine if heart valve morphology lead to significant differences in dissection risk in patients suffering from aortic aneurysms.

III. RESULTS AND DISCUSSION

ΔU_L in each axis of loading significantly and negatively correlated to S_d ($r^2 > 0.42$, $p < 0.0001$). However, E in each loading direction as well as AI for both ΔU_L and E were not significantly correlated to S_d ($r^2 < 0.08$, $p > 0.05$). Aneurysm diameters were significantly but poorly correlated to S_d ($r^2 = 0.18$, $p < 0.05$). Our results suggest that ΔU_L maybe a stronger predictive parameter to assess dissection risk compared to E, AI and aneurysm diameter. When stratified based on valve morphology, it was determined that aneurysms associated with bicuspid valves had higher S_d (35.8 ± 1.97 N/mm vs. 27.0 ± 2.4 N/mm, $p < 0.01$) and lower ΔU_L ($6.6\% \pm 1.5$ vs. $9.7\% \pm 2.8$, $p < 0.0001$) compared to tricuspid valve patients.

IV. CONCLUSION

Our study suggests that ΔU_L is a biomechanical marker that is a stronger predictor of aortic aneurysm dissection risk compared to the clinical standard of using aneurysm diameter. For future investigations, we aim to determine the effects of strain rate and strain magnitude on measured ΔU_L values. We further aim to trial clinical imaging techniques to measure ΔU_L *in vivo*.

- [1] J. Chung, et al. "Loss of mechanical directional dependency of the ascending aorta with severe medial degeneration", *Cardiovascular Pathology*, vol. 26, pp. 45-50, 2017.

Vibration Analysis of a Bone-Conduction Hearing Implant

S. Alireza Rohani and Sumit K. Agrawal

Department of Otolaryngology – Head and Neck Surgery
Western University
London, ON, Canada
srohani4@uwo.ca

Hanif M. Ladak

Department of Medical Biophysics
Western University
London, ON, Canada

ABSTRACT

Background

Sound transmission involves the conduction and amplification of sound energy received at the eardrum to the cochlea, the main organ of hearing. Interruptions to this transmission may result in conductive hearing loss. The most common treatment of conductive hearing loss is the use of hearing aids; however some patients do not attain sufficient improvement from, or cannot wear, hearing aids. For these patients, bone-conduction implants may be used. The BoneBridge® (Med-El, Innsbruck, Austria) is a bone-conduction implant used in the treatment of conductive hearing loss by transmitting sound energy to the cochlea through direct vibrations to the skull rather than through the middle ear. For use of the BoneBridge, a transducer is surgically implanted in the skull using two screws. Given that the transducer has 8.7 mm height, in cases where the skull is not thick enough to house the transducer, “lifts” are used to elevate the transducer to avoid pressurization of the brain. These lifts are available in various thicknesses; 1, 2, 3 or 4 mm. Currently, the effect the thickness of the lift has on the successful transmission of vibration to cochlea is unknown. The objective of the present study was to investigate the effect of lift thickness on transmission of sound to the cochlea using the BoneBridge bone-conduction implant.

Materials and Methods

Three cadaveric temporal bones were used in this study. The samples were first embalmed and dried for use. In each sample, a single location was identified as the candidate location for the implant within 5 cm of the cochlea, and this location was drilled to house the implant transducer. The location of the transducer was kept consistent anatomically among samples. To avoid multiple usage of screw holes, four pairs of holes were pre-drilled in a circular pattern, at least 20 degrees apart, with each pair to be used only once by one of the four lifts (1, 2, 3 and 4mm thick lifts). For each lift, the vibration of transmitted sound to the cochlea was measured using a laser Doppler vibrometry (LDV) technique. The measurements were performed on the cochlear wall. The stimulation of the transducer was performed using company-provided software (CONNEX, Version 6.5, Med-El, Innsbruck, Austria) at multiple discrete frequencies, including 0.5, 0.75, 1, 1.5, 2, 3, 4 and 6 kHz. All measurements were performed at a 45 dB hearing level (HL) and averaged after three repetitions. Each time a stimulation was applied, the vibration of the cochlea over the entire 6 kHz bandwidth was measured to record the harmonics produced by the transducer. Measurements were analyzed using a single-factor ANOVA to investigate the effect of lift thickness.

Results and Discussion

Vibration acceleration of the cochlear wall was measured in three samples when four different lift thicknesses were used (1, 2, 3 and 4mm). The measurements were performed with stimulations at eight different frequencies (0.5, 0.75, 1, 1.5, 2, 3, 4 and 6 kHz). A single-factor ANOVA failed to find a significant difference related to lift thickness. Similarly, no statistically significant difference was found in harmonics developed by the transducer when lift thickness varied.

Conclusion

This is the first known study to evaluate the effect of lift thickness on the sound transmission produced by a bone-conduction implant. The results of this study suggest that lift thickness does not affect the sound transmission to the cochlea using the BoneBridge® implant. This information may inform otologic surgical practice, specifically in the use of lifts in bone-conduction implantation. Future studies may benefit from analysis of other surgical choices, such as placement of the transducer on multiple anatomically-significant locations of skull.

PERFORMANCE OF QCT-DERIVED SCAPULA FINITE ELEMENT MODELS IN PREDICTING LOCAL DISPLACEMENTS USING DIGITAL VOLUME CORRELATION

Jonathan Kusins, Nikolas Knowles, Louis Ferreira
 Department of Mechanical & Materials Engineering
 Western University
 London, Canada

Melissa Ryan, Enrico Dall'Ara
 Insigneo: Institute for In Silico Medicine
 University of Sheffield
 Sheffield, United Kingdom

Introduction

Subject-specific finite element models (FEMs) of the shoulder complex are commonly used to predict differences in internal load distribution due to injury, treatment or disease. However, the accuracy of these models is to some extent unknown. Currently, experimental validation of shoulder QCT-based FEMs is limited to the cortical shell, while the accuracy of internal predictions within trabecular bone has yet to be explored. The goal of the current study was to quantify the accuracy of local displacements predicted by subject-specific QCT-based FEMs of the scapula, compared to experimental measurements obtained by combining digital volume correlation (DVC) and mechanical loading of cadaveric scapulae within a microCT scanner.

Methods

A custom-designed six degree-of-freedom hexapod robot was fabricated with carbon fiber struts for compatibility with a cone beam microCT scanner (Nikon XT H 225ST). MicroCT scans (33.5 μm resolution) were acquired of four cadaveric scapulae in a pre- and post-loaded state (500 N compressive applied load, settling time of 20 minutes). BoneDVC was used to derive full-field experimental displacement measures (sub-volume size of ≈ 1 mm). QCT-FEMs were generated for each specimen. Three separate boundary conditions (BCs) were modelled. Experimentally-idealized BCs were simulated by rigidly fixing the medial border of the scapulae and applying either a force (idealized-force BC) or displacement (idealized-displacement BC) to a virtual loading platen. Also, a DVC-derived BC was modelled by directly imposing the local displacements quantified by the DVC algorithm to the top and bottom face of the QCT-FEMs.

Linear regression analysis between local displacements along the x, y, and z directions predicted by the QCT-FEMs and experimental DVC results was performed.

Results and Discussion

DVC-derived BCs resulted in the closest match to the experimental results (best agreement (range of $m = 0.93$ -1.05) and highest correlation (range of $r^2 = 0.83$ -1.00)) compared to the idealized-displacement and idealized-force BCs (Table 1). In addition, a two order of magnitude decrease was observed in RMSE, with the lowest errors found using DVC-derived BCs (average RMSE = 4.7 ± 1.3 μm , $N=4$) compared to the idealized-displacement (average RMSE = 538 ± 239 μm , $N=4$) and idealized-force (average RMSE = 749 ± 379 μm , $N=4$) BCs.

Conclusions

The results of this study show that errors in local displacements predicted by QCT-based FEMs of the shoulder can be minimized using DVC-derived BCs. This has the potential to significantly improve the accuracy of QCT-based FEMs, which could be used to preoperatively optimize patient outcomes associated with a variety of surgical interventions.

Table 1: Linear regression results between local displacements predicted by QCT-based FEMs and DVC experimental results

		Idealized-Force BC				Idealized-Displacement BC				DVC-Derived BC				
		Specimen #	1	2	3	4	1	2	3	4	1	2	3	4
Slope (m)	x-direction		0.52	-0.02	0.01	1.05	2.84	-0.03	0.90	1.39	1.00	0.99	0.99	0.93
	y-direction		0.21	-0.10	0.31	1.66	1.00	0.17	0.96	1.69	0.99	0.99	1.05	0.93
	z-direction		0.11	-0.13	0.10	0.88	0.54	0.08	1.01	1.65	1.01	1.00	1.02	1.02
Coefficient of Correlation (r^2)	x-direction		0.89	0.01	0.00	0.93	0.95	0.00	0.48	0.85	0.97	0.99	1.00	0.97
	y-direction		0.38	0.59	0.50	0.82	0.37	0.01	0.16	0.62	0.98	1.00	0.95	0.83
	z-direction		0.10	0.47	0.48	0.77	0.06	0.00	0.94	0.82	0.99	1.00	1.00	0.99

Material Mapping of QCT-Derived Scapular Models: A Comparison with Micro-CT Loaded Specimens using Digital Volume Correlation

Nikolas Knowles, Jonathan Kusins, Mohammadreza
Faiegi, Louis Ferreira
School of Biomedical Engineering & Department of
Mechanical and Materials Engineering
Western University
London, Ontario, Canada

Melissa Ryan, Enrico Dall'Ara
Insigneo Institute for In Silico Medicine
University of Sheffield
Sheffield, UK

Introduction: Subject-specific finite element models (FEMs) allow for a variety of biomechanical and clinically relevant conditions to be tested in a highly repeatable manner. The accuracy of these FEMs is improved by mapping density using quantitative computed tomography (QCT) and by choosing a constitutive relationship that relates density to the mechanical properties of the bone tissue. Although QCT-derived FEMs have become common practice in contemporary computational studies of whole bones, many of the density-modulus relationships used at the whole bone level were derived using mechanical loading of small trabecular or cortical bone cores [1,2]. These cores were mechanically loaded to derive an apparent modulus, which is related to each core's mean apparent or ash density. This study used these relationships and either elemental or nodal material mapping strategies to elucidate optimal methods for scapular QCT-FEMs.

Methods: Six cadaveric scapulae (3 male; 3 female; mean age: 68 ± 10 years) were loaded within a micro-CT in a custom CT-compatible hexapod robot. Pre- and post-loaded scans were acquired (spatial resolution = $33.5 \mu\text{m}$) and DVC was used to quantify experimental full-field displacements (BoneDVC, Insigneo) [3]. Experimental reaction forces applied to the scapulae were measured using a 6-DOF load cell. Finite element models (FEMs) were derived from corresponding QCT scans of each cadaver bone. These models were mapped with one of fifteen density-modulus relationships and either elemental or nodal material mapping strategies. DVC-derived BCs were imposed on the QCT-FEMs using local displacement measurements obtained from the DVC algorithm [4]. Comparisons between the empirical and computational models were performed using resultant reaction loads and full-field displacements.

Results and Discussion: Reaction forces predicted by the QCT-FEMs showed large variations in percentage error across all specimens and density-modulus relationships when an elemental material mapping strategy was used. The percentage errors were as large as 899%, but as low as 3-57% for the different specimens. Similarly, when using a nodal material mapping strategy, percentage errors were as large as 965%, but as low as 4-59% for the different specimens.

For all specimens, minimal variation only occurred in the slope between the QCT-FEM and DVC displacements in the x and y directions for either elemental or nodal material mapping strategies. Slopes ranged from 0.86 to 1.06. This held true for 3 specimens in the z direction; however, for the remaining 3 specimens more pronounced variations occurred between the QCT-FEM and DVC displacements, dependent on density-modulus relationship. The r^2 values were consistently between 0.82 and 1.00 for both material mapping strategies and density-modulus relationships for all three Cartesian components of the displacement and all specimens.

Conclusions: The results of this study suggest that QCT-FEMs using DVC derived boundary conditions can replicate experimental loading of cadaveric specimens. It was also shown that only slight variations exist when either elemental or nodal material mapping strategies are adopted. Given the recent advancements provided by DVC-derived BCs, this study provides a basis for a common methodology that can be implemented in future studies comparing similar outcomes in all anatomic locations. Expanding the current sample size has the potential to determine if a single density-modulus relationship can exist or if specimen or anatomic location-specific relationships should be utilized.

References: [1] Knowles et al. 2019. JMBBM. **90**:140-145; [2] Helgason et al. 2008. Clin. Biomech. **23**(2):135-146; [3] Dall'Ara et al. 2014. J. Biomech. **47**(12):2956-2963; [4] Costa et al. 2017. Plos One. **12**(7): e0180151

THE EFFECT OF SHORT STEM HUMERAL IMPLANT VARUS-VALGUS POSITION ON BONE STRESS FOLLOWING TOTAL SHOULDER ARTHROPLASTY

Amir Tavakoli
Department of Mechanical Engineering
Western University
London, Canada
Atavako8@uwo.ca

Kenneth J. Faber¹, G. Daniel G. Langohr²
¹Faculty of Surgery, ²Department of Mechanical Engineering
Western University
London, Canada

Introduction: Total shoulder arthroplasty (TSA) is an effective treatment to alleviate pain and restore shoulder function in the case of glenohumeral arthritis [1]. On the humeral side, the articular surface is resected, and the canal is prepared to accept the humeral stem. Stress shielding, which occurs when humeral bone stress is reduced due to the replacement of bone with a stiffer metallic implant, causes bone resorption of up to 9% of the humeral cortical thickness following TSA [2]. Shorter length stems and smaller overall geometries to reduce stress shielding are introduced [3], however the effect of error in humeral short stem varus-valgus positioning on bone stress is not well defined. Thus, the purpose of this study was to quantify the effect of humeral short stem varus-valgus positioning on bone stresses after TSA.

Methods: A 3D model of one cadaveric humeri (age: 68 yrs) was constructed from computed tomography data using MIMICS (Materialise, Belgium). Separate cortical and trabecular bone sections were created, and the resulting bone models were virtually reconstructed three times by an orthopaedic surgeon using an optimally sized short stem humeral implant (Exactech Preserve) that was placed directly in the center of the humeral canal (STD), as well as rotated varus (VAR) or valgus (VAL) until it was contacting the cortex. Bone was meshed using a custom technique which produced identical bone meshes permitting the direct element-to-element comparison of bone stress. Cortical bone was assigned an elastic modulus of 20 GPa and a Poisson's ratio of 0.3. Trabecular bone was assigned varying stiffness based on CT attenuation [4]. A joint reaction force was then applied to the intact and reconstructed humeri representing 45° and 75° of abduction [5]. Changes in bone stress, as well as the expected bone response based on change in strain energy density [6] was then compared between the intact and reconstructed states for all implant positions.

Results: Both Varus and Valgus positioning of humeral short stem size altered the both cortical and trabecular bone stresses from the intact states. Varus positioning had the greatest negative effect in the lateral quadrant for both cortical and trabecular bone, producing greater stress shielding than both the standard and valgus positioned implant. Overall, the valgus and standard positions produced values that most closely mimicked the intact state.

Discussion and Conclusions: Interestingly, varus positioning of the implant produced large amounts of stress shielding in the lateral cortex at both 45° and 75° of abduction but resulted in a slight decrease in stress shielding in the medial quadrant. This may have been a result of direct contact between the distal end of the implant and the varus cortex under loading which permitted load transfer, and hence unloading of the lateral cortex during abduction. Conversely, when the implant was positioned in the valgus direction, significant departures in stress shielding and changes in bone stress were not observed when compared to the ideal STD position. For the valgus positioned implant, the deflection of the humerus under load resulted in liftoff of the distal aspect of the stem from the valgus cortex, hence preventing distal load transfer thus protecting from load transfer. The results of this study seem to show that varus malposition is potentially worse than valgus malposition in terms of expected stress shielding and changes in bone stress.

References:

- [1] Massmini et al., JBJS. 2010; 92(4):916-926. [2] Nagels et al., JSES. 2003; 12(1):35-39. [3] Bobyn et al., CORR. 1992; 274:79-96. [4] Morgan et al., JBiomech. 2003; 36(7):897-904. [5] Bergmann et al., JBiomech. 2007; 40(10):2139-2149. [6] Neuert et al., JBiomech. 2013; 227(9):994-1001.

Impact location affects brain and neck strain during football sport

Kewei Bian

Department of Mechanical and Materials Engineering
 Western University
 London, Canada
 kbian2@uwo.ca

Haojie Mao

Department of Mechanical and Materials Engineering
 Western University
 London, Canada
 hmao8@uwo.ca

Introduction: Traumatic brain injury (TBI) due to sports-related collisions is a severe health problem. Brain responses that are based on strain are considered as the main culprit to cause TBI. Meanwhile, it's acknowledged that impacts from different football-related locations affect regional brain strain in various manners. Hence, this study is focused on investigating brain strain from different impact locations.

Materials and Methods: Based on Global Human Body Models Consortium (GHBM) model and Riddell helmet model (Panzer et al., 2018), we simulated impacts with impactor hitting the frontal, lateral, vertical and rear sites of helmet. Moreover, frontal and lateral impacts to the facemask were also included in this study. The cumulative strain damage measure (CSDM) with the strain level of 20% which indicates the strain distribution was used to analyze brain response. The peak maximum principal strain (MPS) of the neck upper cervical cord, disc and facet joint were also analyzed.

Results: For frontal impact to helmet (Figure 1a), the CSDM20, indicating the percentage of the brain elements with the peak MPS no less than 20%, was 0.04. The peak MPSs for the upper cervical spinal cord, the disc and the facet joint were 0.27, 0.38 and 0.13 respectively. The CSDM20 produced by lateral impact to helmet (Figure 1b) was 0.10, larger than that produced by frontal impact. Lateral impact to helmet induced 0.50 to the upper cervical spinal cord, 0.80 to the disc, and 0.25 to the facet joint. The vertical impact (Figure 1c) induced the lowest CSDM20, with peak MPS to the upper cervical spinal cord, the disc and the facet joint being 0.42, 0.5 and 0.35, respectively. The largest CSDM20 among these six impacts was 0.12 produced by rear impact (Figure 1d). Rear impact caused peak MPS 0.44 to the upper cervical spinal cord, 0.41 to the disc, but very low 0.01 to the facet joint. For impacts to facemask, frontal impact (Figure 1e) induced CSDM20 of 0.01 to the whole brain. The same frontal impact to facemask produced peak MPS of 0.42 to the upper cervical spinal cord, 0.48 to the disc, and 0.12 to the facet joint, respectively. Lateral impact to facemask (Figure 1f) induced CSDM20 of 0.10. It also induced MPS of 0.42 to the upper cervical spinal cord, 0.68 to the disc and 0.18 to the facet joint.

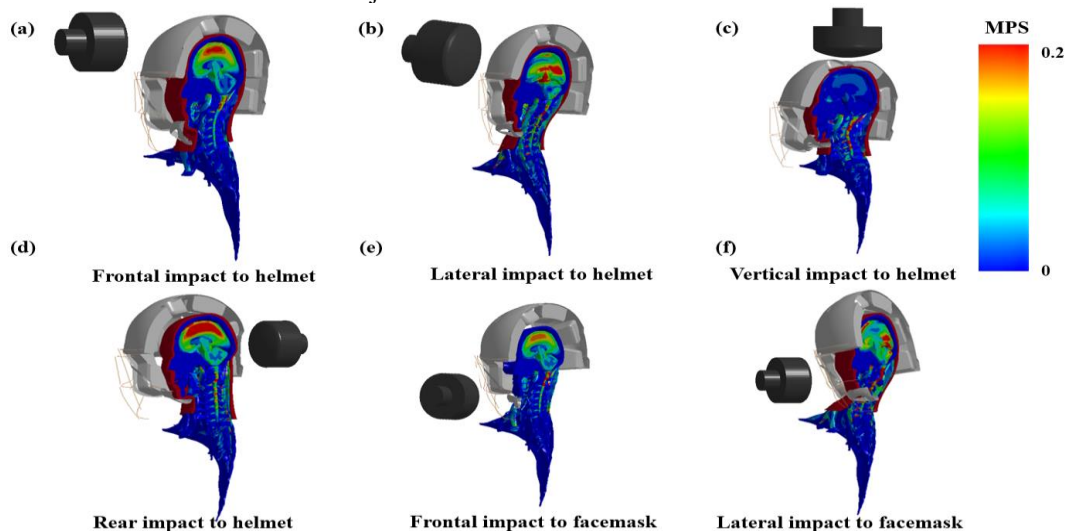


Figure 1 Strain contours in different impact locations. MPS: maximum principal strain

Conclusions and Discussion: For the brain, rear impact induced the largest CSDM20, followed by lateral impact to facemask, lateral impact to helmet, frontal impact to helmet and frontal impact to facemask, while vertical impact induced the lowest CSDM. The upper cervical spinal cord experienced the highest strain among all impacts. Both the disc and facet joint experienced the high strain during lateral and vertical impacts. These results will help improve future helmets to reduce brain strains.

Acknowledgements: We acknowledge NSERC Discovery and Canada Research Chairs program for support.

References: Panzer et al. Journal of neurotrauma 2018, Toronto, Canada.

The Effects Of Demographics on Shoulder Shape and Density for Population-based Design of Orthopedic Implants

Pendar Soltanmohammadi
School of Biomedical Engineering
Western University, London, Canada
psoltanm@uwo.ca

Ryan Willing
Department of Mechanical and Materials Engineering
Western University, London, Canada
rwilling@uwo.ca

Introduction: Statistical shape modeling (SSM) and statistical density modeling (SDM) are tools capable of describing the main modes of variability in the shape and density distributions of bones within a population, expressed as a set of parameters called principal components (PCs). These PCs can be further analyzed to seek correlations with demographic data, improving our understanding of how bone shapes/densities vary to help design more effective population-based orthopedic implants. We hypothesize that the first PC of a developed shoulder SDM, which scales overall bone density, will be inversely correlated with age. We also anticipate that bone density will be greater for males than females. Furthermore, we hypothesize that contralateral shoulders are symmetric in terms of shape and density distributions.

Methods: SSMs and SDMs were developed for scapulae and humeri bones based on 75 human cadaveric shoulders. This set includes 57 male (20 pairs) and 18 female shoulders (1 pair) from 54 donors, with ages ranging between 21 and 94 years (mean 73 ± 13). To develop SSMs and SDMs, surface meshes were previously segmented from computed tomographic (CT) images of cadaveric shoulders using 3D Slicer. Humeri and scapulae bones were co-registered using an iterative closest point algorithm, and a consistent set of vertices was mapped onto the surface of each humerus and scapula and repositioned such that each bone's surface was defined by a deformable but consistent surface mesh. Volumetric tetrahedral meshes with $8.1 \times 10^5 / 4.6 \times 10^5$ nodes and $4.8 \times 10^6 / 2.5 \times 10^6$ elements were defined for one of the humeri/scapulae specimens (respectively) and served as base meshes. These base meshes were morphed to each individual bone's surface in ABAQUS using displacement vectors calculated between corresponding surface vertices on the base mesh and the individual humeri and scapulae surface meshes. This process resulted in a set of 75 corresponding shoulder mesh models. Individual humeri and scapulae models were superimposed upon the corresponding original CT data and CT image intensity (in Hounsfield units – HU) was determined at the spatial location of each node. As a result, a set of 75 shoulder geometry and density distribution models were developed. In Matlab, principal component analyses were performed on the exterior shape and internal density distribution of bones. Pearson correlation coefficients were calculated for age and PC scores. Finally, T-tests were performed to find any differences in PC scores between males and females, and again between contralateral shoulders. In these analyses, p-values of less than 0.05 were considered statistically significant.

Results: For the humerus, the first and sixth PCs of the SDM demonstrated a weak, and moderate (respectively) but significant correlation with age ($\rho = -0.29$, and $\rho = -0.40$, both $p \leq 0.03$). For the scapula, the first and ninth PCs showed such weak but significant correlations ($\rho = -0.31$, and $\rho = -0.32$, both $p \leq 0.02$). Statistically significant differences due to sex were found for PC 2, 3, and 5 of the humerus SDM, with differences in average PC scores of -1.1 ± 0.5 , $+1.2 \pm 0.4$, and -1.2 ± 0.4 standard deviations, respectively, for males relative to the females (all $p \leq 0.04$). For the scapula, PC 2 of the SDM was significantly different between males and females, with average PC scores differing by -1.2 ± 0.5 standard deviations ($p \leq 0.03$). For the contralateral shoulders, statistically significant differences were identified in PC 1 of the humerus SDM, and PC 2, 4, and 5 of the humerus SSM with average PC score differing by -0.3 ± 0.7 , $+0.4 \pm 0.9$, $+0.5 \pm 1.0$, and -0.5 ± 0.9 standard deviations (respectively), for right humeri relative to the left ($p \leq 0.04$). For the scapula, statistically significant differences were observed in PC 2, and 13 of the SDM, and PC 4, and 8 of the SSM with average PC scores differing by -0.4 ± 0.8 , -0.3 ± 0.7 , $+0.4 \pm 0.6$, and $+0.3 \pm 0.5$ standard deviations (respectively), for right scapulae relative to the left (all $p \leq 0.03$).

Discussion: Our results suggest that age has a significant inverse effect on density within the entire bones, especially for females. This implies a natural bone density loss of the shoulder with aging. The second PC of the SDMs for both bones describes a thinning of the cortical shell, which differed significantly between males and females, and indicated that males, on average, have relatively thicker cortical bones than females. Finally, while contralateral bone shapes were symmetric, bone density distributions were not. These findings can help guide the design of population-based prosthesis components.

Significance: This study suggests that age and sex have a significant influence on the density distribution of the shoulder and reveals the main modes by which density distribution can vary among members of a population. The results of this study can pave the way for designing more effective population-based shoulder implants by considering the demographics of their target population.

Comparing Fixation Constructs for Scapula Insufficiency Fractures Following Reverse Shoulder Arthroplasty

Jason S. Lockhart, David Axford, G. Daniel G. Langohr,
 James A. Johnson
 Mechanical & Materials Engineering
 Western University
 London, Canada

George S. Athwal, Francis S. H. Ting
 Hand & Upper Limb Centre
 St Joseph's Health Care Foundation
 London, Canada

INTRODUCTION

Scapular spine insufficiency fractures following reverse shoulder arthroplasty are a debilitating complication, and to date, there is limited literature on their surgical management. The purpose of this study is to compare 4 different fixation constructs to stabilize scapular spine insufficiency fractures. The fixation constructs that were varied are; medial fixation: superior within supraspinatus fossa (Superior) and dorsal subcutaneous border over spine of scapula (Dorsal), lateral fixation: superior plate with and without addition of an inferior supporting hook.

METHODS

Twelve fresh frozen bilateral cadaveric scapulae (average age: 64 ± 6 yrs, $N=24$) were randomly sorted into the four fixation groups for reconstruction. The fixation techniques included dorsal and superior plates, both with and without an inferior supporting hook. A type II scapula insufficiency fracture was simulated for all specimens. Each specimen was optically tracked and loaded cyclically (100 cycles, 1 Hz) on the lateral acromion from 50 N to 700 N in 50 N steps or until failure. A repeated measures ANOVA was used to determine relationships within maximum survival forces and optical tracking data. A binary regression was used to determine if either fixation factor was a significant predictor of failure location (medial or lateral of the fracture site).

RESULTS SECTION

19 of 24 specimen failed before the maximum load of 700 N. The average maximum survival force for medial plate fixation dorsally on the subcutaneous border over the scapular spine (Dorsal) was 550 ± 40 N, significantly greater, compared to 394 ± 40 N for medial fixation superiorly within the supraspinatus fossa (Superior, $p=0.022$). The presence of an inferior supporting hook was found to be a significant predictor of failure location ($p=0.016$). In the presence of the inferior hook failure occurred 8 times medially and 1 time laterally, compared to 6 medially and 4 laterally without the hook.

DISCUSSION

These results suggest medial fixation dorsally to the subcutaneous border of the scapular spine (Dorsal) is ultimately stronger than medial fixation superiorly to the supraspinatus fossa (Superior), and the presence of an inferior supporting hook on the lateral acromion may prevent lateral failure of the repair. However, the results suggest a laterally placed hook does not have a significant effect on the maximum survival force of the entire construct, probably a result of medial fixation being unaffected by the laterally placed support hook. This suggests further measures may be beneficial to increase the strength of the construct's medial fixation.

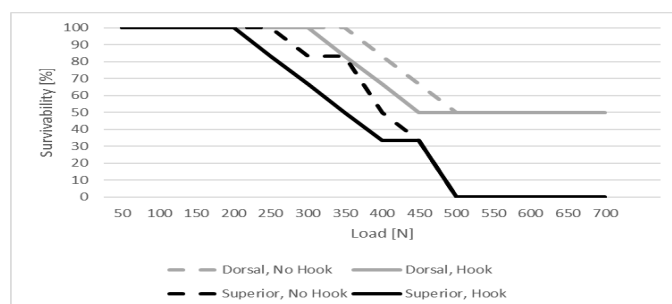


Figure 1 - Percent survivability of the 4 fixation groups as a function of applied load

Using Kinematic Signals Features of Body-Worn Inertial Sensor For Temporal Events Detection of Locomotion

Niloufar Ahmadian¹, Milad Nazarahari¹, Jackie L. Whittaker², and Hossein Rouhani¹

Department of Mechanical Engineering¹, Department of Physical Therapy²
University of Alberta
Edmonton, Canada

na2@ualberta.ca, nazaraha@ualberta.ca, jwhittak@ualberta.ca, hrouhani@ualberta.ca

ABSTRACT

Background: The stance phase duration of gait, defined as the time interval from initial foot contact to terminal foot contact with the ground, can be informative for clinical decision-making and diagnosis. For instance, prolonged and abnormal stance phase durations can be considered as symptoms of frailty or Parkinson's disease in the elderly. Accurate detection of temporal events is also of great importance when analyzing various biomechanical elements of gait (e.g., joint angles and spatial parameters) and similar cyclic activities, since the onsets of these events are essential for segmentation of activity cycles. The most straightforward method of obtaining temporal events is the use of pressure platforms, footswitches, and motion capture cameras. However, the inaccessibility, low durability, and confined measuring volume of these devices make them arguable choices when analyzing natural walking. On the other hand, wearable inertial sensors can be an ideal solution for quantifying temporal parameters in clinics and real-world environments, and do not hinder gait and similar activities.

Objectives: The objectives of this study are: (1) To identify robust features of foot and shank kinematic signals which are relevant to the initial and terminal foot contact instances of gait and are minimally influenced by sensor orientation errors, caused by inaccurate attachment of sensors to body segments (robustness in clinical environments). (2) To assess the accuracy and precision of these kinematic features by validating this temporal events detection algorithm against a reference motion capture system.

Method: Two inertial measurement units (3D accelerometer and gyroscope) were attached to upper-medial shank and forefoot of three able-bodied male participants. Participants were also provided with two reflective markers on their first metatarsal and calcaneal tuberosity bony landmarks. A reference system of 8 motion capture cameras tracked the markers positions while participants performed 40-seconds U-turn walking trials at a self-selected speed. Peak detection and thresholding algorithms were applied to Euclidean norms, sagittal components and time derivatives of foot and shank acceleration and angular velocity signals to obtain 10 kinematic features, relevant to the initial and terminal contact instances of each gait cycle. These features were compared to the camera-based reference temporal events, and the medians and interquartile ranges of their errors were calculated to evaluate the accuracy and precision of initial and terminal contact detection.

Results: All 10 proposed kinematic features successfully identified the initial and terminal foot contact instants of gait with median (interquartile range) errors of less than -50 (30) milliseconds and 50 (70) milliseconds, respectively. Among the 10 investigated features, those pertaining to foot kinematic signals demonstrated the highest accuracies while the shank-related kinematic features demonstrated the highest precisions in temporal events detection of gait. Considering both accuracy and precision, the shank-related features were preferable for the trials with greater dynamic movements.

Conclusion: The 10 proposed features, obtained based on peaks and thresholding of foot and shank kinematic signals, were able to detect the initial and terminal contact onsets of each gait cycle with good accuracy and precision for clinical applications. As most of these features are independent to the orientation of sensors attachment on body segments, they are appropriate to be used in clinical scenarios, without concerns regarding the accuracy of attachment and relatively cumbersome sensor to body calibration procedures. Therefore, quantifying temporal events of gait and more complex dynamic activities with inertial sensors can become a standard part of clinical examinations.

Validity of the Clinician Rated Drop Vertical Jump Scale for Patients following Anterior Cruciate Ligament (ACL) Reconstruction

Morgan Jennings, Dianne Bryant, Trevor Birmingham, Alan Getgood

Health and Rehabilitation Sciences, Faculty of Health Sciences
Western University
London, Canada

Introduction

The Anterior Cruciate Ligament (ACL) is one of the most frequently injured structures within the knee, and often requires reconstruction surgery with extensive rehabilitation. Increasing evidence suggests there is predisposed risk of ACL injury due to poor biomechanical control over the lower limb because of neuromuscular deficiencies. The Drop Vertical Jump (DVJ) is a functional task that has been used to quantify specific movements when performed in a biomechanics lab with a three-dimensional motion capture system. These movement patterns are predictive of ACL re-injury risk. The lack of accessibility, feasibility, efficiency and required level of expertise, prevents the use of motion analysis for clinical use. A Clinician Rated Drop Vertical Jump Scale (CR-DVJS) was developed through expert consensus to allow for objective measurement of biomechanics within a clinical setting.

Hypotheses

The purpose of this study was to evaluate the validity of the Clinician Rated Drop Vertical Jump Scale (CR-DVJS) in a population of patients following anterior cruciate ligament reconstruction (ACLR) using the current gold standard of measuring lower body biomechanics – three-dimensional motion analysis. We hypothesized a strong association ($\rho > 0.5$) between observer scores of scale components and 3D measures of performance for both in-person and video reviewed assessments.

Methods

Patients completed DVJ protocol at 6 and/or 12-months post-operative in the Wolf Orthopaedic Biomechanics Laboratory. The task was recorded using standard video as well as the motion capture system. A designated researcher and clinician served as raters during in-person evaluation using the CR-DVJS. The researcher reviewed and re-evaluated standard video footage in both real time and slow-motion. Scale components were analyzed for association with 3D measures of performance.

Results

The CR-DVJS scores of slow-motion videos were associated with motion capture measures. Specifically, scale components including valgus collapse, trunk flexion, and knee flexion, were correlated with motion capture measures of knee abduction moment and angle, trunk flexion angle, and knee flexion angle respectively. The CR-DVJS in-person rater scores of valgus collapse did not correlate strongly with motion capture measures. However, CR-DVJS scores of trunk and knee flexion did demonstrate an association with 3D measures of trunk ($\rho=0.4-0.5$) and knee ($\rho>0.5$) flexion angle.

Discussion

There is a need for greater standardized and objective criteria to evaluate progress through rehabilitation. New evaluative tools will assist in identifying undesirable movements during functional tasks, as well as monitor changes in performance. This study will provide evidence of the validity of the CR-DVJS in conjunction with video recording to assist with decision-making for ACL rehabilitation and readiness for return to sport. This will facilitate rehabilitation targeting mechanics and neuromuscular imbalances, and decrease the need for expensive, inaccessible equipment for DVJ quantification.

Role of Medial Ligamentous Structures on PCL-Deficient Knee Kinematics During Simulated Clinical Stability Tests and Activities of Daily Living

Alireza Moslemian, Ryan Willing
Mechanical and Materials Engineering
Western University, London, Canada

Alan Getgood
Department of Surgery
Western University, London, Canada
rwilling@uwo.ca

Introduction: The posterior cruciate ligament (PCL) is considered the strongest ligamentous structure of the knee; however, between 3% and 20% of knee injuries include PCL rupture [1]. These injuries usually happen due to a large posterior loading on the tibia which occurs while landing on the knee with the joint flexed. PCL injuries typically occur in combination with injuries to other ligamentous structures. Since whether or not these structures need to be repaired during surgery remains controversial, it is necessary to understand their contributions to the stability of a PCL-deficient knee. The posterior oblique ligament (POL) and deep medial collateral ligament (dMCL) are medial ligaments whose contribution to the stability of the knee have not been thoroughly investigated. The objective of this study is to examine possible changes in joint kinematics due to simulated injury to either of these ligamentous structures in the PCL-deficient knee.

Method: Nine fresh-frozen cadaveric knees were used in this study (aged 24-62 years, 2 male, 4 Female, 3 Pairs). Knees were potted, aligned and mounted onto a 6 degrees-of-freedom joint motion simulator (VIVO, Advanced Mechanical Technologies, Inc.) using custom-designed fixtures. Knees were subjected to prescribed flexion from 0° to 90° with a 10 N compressive force applied along the long axis of the tibia, while all remaining degrees of freedom were unconstrained (baseline). This was then repeated with additional superimposed forces which represented clinical tests for assessing joint stability, including: (1) a 67 N posterior-directed load (2) a 2.5 Nm internal or external moment (in both cases applied to the tibia). In addition to simulated stability tests, knees were subjected to cyclical loading scenarios simulating activities of daily living, including gait, stair ascent and descent. All tests were repeated following dissections of the PCL (arthroscopically) and then medial ligaments. Dissection of medial structures was done in randomized order (5 POL first, 4 dMCL first) through a temporary window on the medial side of the knee. The resulting kinematics of the knee at each stage of dissection during clinical test were recorded at 0°, 15°, 30°, 45°, 60°, 75° and 90° of flexion. During the activities of daily living, the average AP kinematics were recorded and compared. The changes in kinematics were analyzed using paired two-tailed t-test, with a p-value < 0.05 considered statistically significant.

Results: With a posterior-directed load applied, PCL dissection resulted in significantly more posterior positioning of the tibia by up to 7.4 ± 4.7 mm at 90° ($p < 0.000$). Subsequent dissection of either of either medial structure did not alter AP kinematics of the knee. During internal/external loading of the tibia, dissecting the PCL did not change the IE kinematics significantly. With the internal moment was applied, POL dissection caused significantly more internal rotation at all flexion angles in comparison with the PCL-deficient knee, with the maximum of $3.6 \pm 5.2^\circ$ at 0° ($p = 0.015$). dMCL dissection only increased the internal rotation at 0°, 60°, 75° and 90° by up to 1.6 ± 5.6 at 60° ($p = 0.040$). With an external moment applied, POL dissection increased the external rotation at 0° by $0.7^\circ \pm 5.8^\circ$ ($p = 0.018$) in comparison with the PCL-deficient knee and dMCL dissection did not result in a significant change. Interestingly, during activities of daily living, dissection of the dMCL caused the average AP position of the tibia to shift anterior by 0.6 ± 1.3 mm during gait ($p = 0.005$), 0.5 ± 2.0 mm during stair ascent ($p = 0.012$) and 0.5 ± 1.0 mm during stair descent ($p = 0.002$).

Conclusion: When a posterior-directed tibial load was applied, the tibia translated further posterior in the PCL-deficient knee as compared with intact. These results agree with previous studies which investigated the AP kinematics of the PCL-deficient knees under similar conditions. Surprisingly, subsequent dissection of either medial ligamentous structure resulted in a slightly more anterior positioning of the tibia (compared to the isolated PCL injury); however, this change was not statistically significant. Our results show that IE kinematics of the knee were not sensitive to PCL dissection. In the PCL-deficient knee, however, IE kinematics were more sensitive to POL dissection than dMCL dissection. Even though the dMCL dissection did not play a significant role in simulated clinical stability tests, it resulted in a more anterior positioning of the tibia during simulated activities of daily living. We suspect that dissection of the dMCL reduced joint compression, which allowed the posterior slope of the tibia in combination with large joint reaction forces to play a greater role in pushing the tibia anteriorly.

Significance: It is important to understand the role of secondary joint stabilizers in the setting of injuries to main joint stabilizers. Moreover, we need to determine if injuries to these structures result in significant change in kinematics to warrant reconstruction. Our results indicate that, in the case of injury to the PCL, both the POL and dMCL contribute to rotational stability of the knee, and thus they may need to be addressed in PCL reconstructions if damage was sustained.

REFERENCES:[1]Ritchie et al. *J. Sports Med.*, vol. 26, no. 3, pp. 389–394, 1998.

MECHANICAL PROPERTIES OF THE BICIPITAL APONEUROSIS

Emmanuel Ocran, Michele Oliver, Karen Gordon

School of Engineering
University of Guelph
Guelph, Ontario
eocran@uoguelph.ca

Anne Agur

Division of Anatomy, Department of Surgery
University of Toronto
Toronto, Ontario

ABSTRACT

As a biarticular muscle, the biceps brachii supinates the forearm and flexes both the elbow and shoulder, thus allowing the upper to limb to perform a variety of activities of daily living (ADL). The biceps brachii originates on the coracoid apex as well as the supraglenoid tubercle and inserts on the radial tuberosity. At the distal end, the bicipital aponeurosis (BA) provides a transition of the biceps tendon into the antebrachial fascia. Previous work has reported the importance of bicipital aponeurosis in stabilizing distal tendons. Other studies have reported the supination effect that specifically the BA has on the forearm at the radioulnar joint, where it protects the branchial arteries and median nerve (neurovascular bundle). However, despite the important functions fulfilled by this structure, BA mechanical properties are yet to be quantified. Further, due to its morphology (a long and flat tendinous expansion) and location (travelling through the cubital fossa and the proximal end of the forearm fascia), it is entirely possible that the mechanical properties of the BA differ across its length.

In this work, mechanical properties for eight fresh frozen BA specimens are being quantified using a Cellscale Biaxial (Waterloo, ON) testing machine. The length of each full BA specimen is being measured using a Vernier caliper. Three samples (approximately 7X7mm each) are being taken at the 25th, 50th and 75th percentile of each full-length specimen to represent the proximal, middle and distal regions. Using a custom-built micrometer, sample thickness measurements are being averaged from measurements at five random locations. Black graphite particles are being manually applied in a random orientation on the surface of each sample to provide contrast for optical strain measurements captured by the Cellscale Software. While maintaining the orientation of the sample such that the longitudinal collagen fiber orientation is aligned with the X-axis of the tester, samples are then mounted onto the biaxial testing machine for mechanical testing. Samples are first preconditioned with 10 sinusoidal cycles at 10% strain, at a strain rate of 1%/s. Following preconditioning, samples are being uniaxially and biaxially tested at 12% strain at 1%/s. Young's modulus is being quantified for each sample and means for each of the three regions will be compared using a one-way ANOVA. Results from this study will provide input values for future models of distal biceps repair, thus aiding surgical planning by providing insight into the potential load sharing contributions of the BA.

DESIGN OF A 3D PRINTED KNEE WEARABLE FOR BIOMECHANICAL MONITORING OF LOCOMOTOR ACTIVITIES OF DAILY LIVING

Calvin Young, Michele Oliver, Karen Gordon

School of Engineering
 University of Guelph
 Guelph, Canada
 cyoung02@uoguelph.ca

INTRODUCTION

A 3D printable wearable device capable of wirelessly monitoring knee biomechanics during locomotor activities of daily living (LADL) has been developed (Figure 1, A) to study biomechanical changes in knee function during recovery following anterior cruciate ligament (ACL) reconstruction. The 3D printed wearable frame can be customized, ensuring consistent placement of kinematic monitoring sensors. Previous work using knee braces as a framework for gait monitoring instrumentation demonstrated promising results, however commercially available knee braces tended to be too bulky, expensive, and difficult to unobtrusively instrument. To address this a 3D printed wearable frame has been designed. The wearable frame can be customized based on the following 3 frontal plane measurements: knee width, and leg width 150mm above and below the knee joint. This frame is used to hold two inertial measurement units, one anteriorly 150mm above the knee joint and one posteriorly 150mm below the knee joint, and two rotary position sensors for direct flexion/extension angle measurement.

Preliminary tests have been performed using the device to record data during LADL, under real-world conditions. The movements of interest in this study are gait, stair ascent/descent, and stand to sit/sit to stand. These movements represent locomotor activities which would be performed during normal day to day activity. 4 healthy individuals (2 Male, 2 female aged 25.75 ± 0.83 years) with no history of lower body injury presented for data collection on two separate occasions where they performed the aforementioned movements of interest over the course of a short walk ($\approx 500\text{m}$) while accompanied by a researcher.

RESULTS & DISCUSSION

Movement exemplars for gait and stair ascent are presented in Figure 1 B and C where all data were filtered using a fourth order 10Hz Butterworth lowpass filter and normalized to zero mean and unit standard deviation (z-scores). These preliminary results demonstrate the capabilities of the device to record data representative of various LADL, and data is consistent with previous study of LADL using optical motion capture. This device will be applied in two upcoming studies which include validation of the wearable against a gold standard optical motion capture system, and longitudinal recovery tracking in individuals who have had an ACL reconstruction.

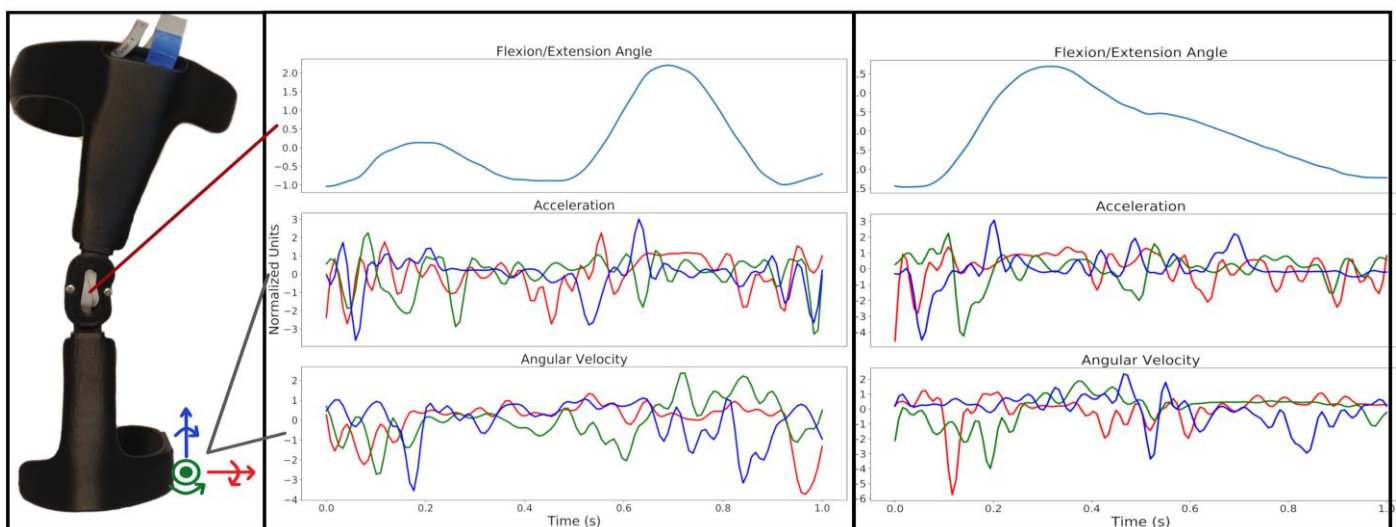


Figure 1: A) The 3D printed wearable device B) 1 gait cycle of kinematic data C) 1 stair climbing cycle of kinematic data (For brevity only the output of one IMU is shown) Positive acceleration and angular velocity directions are indicated in A

Experimental Study of the Effect of Organic and Mineral Content on Bone Mechanical Properties

Oghenerioborue Amromanoh, Yunhua Luo

Department of Mechanical Engineering
University of Manitoba
Winnipeg, MB, Canada

Amromana@myumanitoba.ca,
Yunhua.Luo@umanitoba.ca

Xinyi Wu

Department of Biomedical Engineering
University of Manitoba
Winnipeg, MB, Canada

Wux34518@myumanitoba.ca

Abstract— Mineral and organic protein are the two dominant material phases in bones. The effect of mineral on bone strength has been intensively studied, but much less attention has been paid to the role of organic protein. We investigated the effect of both organic and mineral content on compressive strength of bovine bone specimens. A novel specimen fabrication process was designed to reduce the effect of bone anisotropy on mechanical properties. Mechanical properties of bone specimens, including Young's modulus and peak stress, were characterized by uniaxial compression testing. The experimental results show that both organic and mineral content interaction play important roles in the regulation of compressive strength.

Keywords - bone; compression; Young's modulus; peak stress; anisotropy; mineral; organic protein; ash density

I. INTRODUCTION

Bone fractures occur often in daily life. Many of these incidents can have adverse outcomes depending on the nature of the accident. Hip fracture is one of the most detrimental fractures as mobility is severely diminished. Every year in Canada, approximately 30,000 people experience a hip fracture with the number expected to quadruple by 2030 [1], [2]. The annual cost of care has been estimated from \$1.3 billion in 1993 to \$2.4 billion in 2041 [1]. This represents a snapshot of the burden of the healthcare system in management of skeletal fragility incidents. Research on bone strength is therefore necessary to predict fracture risk and prevent fracture, so that healthcare costs related to bone fractures can be reduced.

Whole bone strength has been found to be affected by microstructural characteristics – with bone tissue strength depending on the level of mineralization of the bone [3]. The mechanical properties of whole bone are widely experimentally determined through destructive tests like compression testing. Various factors however compound the difficulty in the study of bone's mechanical properties – testing setups [4], [5], anatomic location [6], anisotropy [7], specimen geometry [8], strain rate [9], and age [10]. The mechanical properties of both cancellous and cortical bones have been extensively

investigated over the past half a century and summarized by a number of notable reviews [11]–[13]. Helgason et al. [12] reviewed the existing mathematical relationships between bone stiffness and modulus in most of these studies and concluded that methodological differences alone could not explain the discrepancies between these relationships. The mathematical relationships typically utilize mineral density information in the form of ash or apparent density to determine Young's modulus and follows the power law form:

$$E = \alpha \rho_x^\beta; \quad \rho_x \text{ in [g/cm}^3\text{]}, E \text{ in MPa} \quad (1)$$

The utilization of mineral content information however is only able to explain 70% of the variation in Young's modulus [14]. Existing studies – both experimental and theoretical – in order to understand bone properties have therefore, separately focused on mechanical properties with the inclusion of microstructural properties as functions of apparent or ash density, anatomic location, loading direction, strain rate, specimen geometry, age, and diseases [15].

Bone is composed of an inorganic mineral phase (65% vol.), an organic collagen phase (25% vol.), and water (10% vol.) [16]. An examination of the individual mechanical properties of the organic and inorganic component using whole bones suggests that mineral and organic content combine optimally to provide bone's unique qualities. A few studies [15], [17]–[19] have examined the mechanical properties of bone with respect to organic content contribution. Most notably, Novitskaya [19] proved that the weighted sum of bone without mineral content and organic content did not amount to that of similar bone of the same size. This suggests that the contribution of organic content information into the mechanical behavior of bone cannot be understated.

Furthermore, a unique constraint of mechanically testing bone tissues is the anisotropic behavior of bones. Bones behave differently depending on what orientation they are loaded with the highest mechanical properties being found in the longitudinal direction [20]. Bonfield and Grynpas [21] and Ohman et al. [7] proved that a few degrees of misalignment caused reduction in the Young's modulus recorded. To

effectively reduce the effect of bone anisotropy, we designed and implemented a special fabrication process of bovine bone specimens, so that specimens are cut along the longitudinal axis of bovine femora.

The purpose of this work was to therefore explore the effect of organic and mineral content on compressive mechanical properties of bone using a novel approach to control anisotropy. To the best of the authors' knowledge, effect of organic protein on bone Young's modulus and strength has not been considered in the existing bone elasticity-density relationships.

II. METHODS

A. Specimen Preparation

Fresh bovine femur bone samples from 9 heads of cattle were obtained from a local butcher. The age of the cattle ranged from 12 to 18 months. Altogether, 148 cylindrical specimens (32mm \times 8mm) were prepared from the different sites of the femora. The femora were set in a custom-made mold to reduce effect of anisotropy. The anterior side of the bone was placed at the base with the protruding head facing upward. The mold was then confined with clamps in a rectangular contraption (see Fig. 1) to ensure that the longitudinal shaft axis was always parallel to the base – ensuring consistent alignment with the long axis of the bone when the specimens are cut. The mold was then filled with Plaster of Paris, to fix the bone in place relative to the longitudinal axis. The molded femur was cut with a diamond blade into rectangular sections and then cylindrical specimens were cored out of the bone using a trephine. Cylindrical specimens were wrapped in cling film immediately after fabrication and stored in a refrigerator ($T = -20^{\circ}\text{C}$) in an airtight container until mechanical testing.

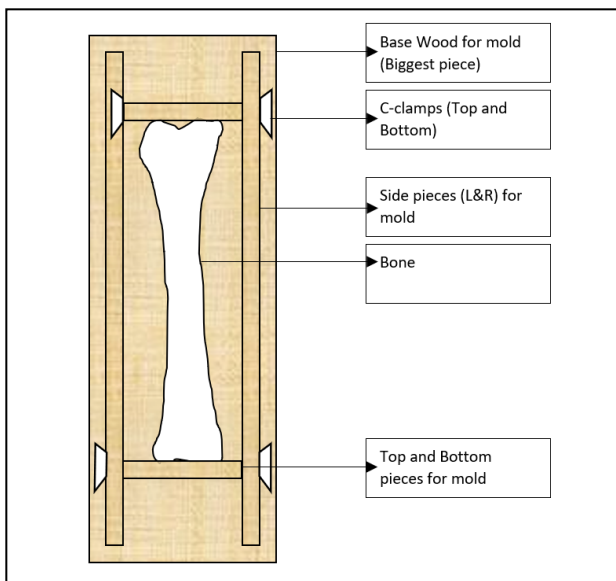


Figure 1. Schematic of the mold setup.

B. Mechanical Testing

The specimens were left to thaw to room temperature before testing. Uniaxial compression testing was performed using an electromechanical material testing equipped with a 30kN load cell (MTS Insight 30, MTS Systems Corporation, Eden Prairie, MN, USA). Specimens were tested at 1.50 mm/min crosshead speed, which translates to a strain rate of 0.0033s^{-1} in unconstrained conditions. The level of displacement was measured by the movement of the crosshead. All samples were loaded till compressive failure. The tests were performed at ambient temperature. The strain was automatically calculated by the accompanying software (TestWorks 4TM, MTS Systems Corporation, MN, USA) from the displacement of the compression platens. The Young's modulus was determined using the average tangential modulus of the initial slope of the stress-strain curve. The yield point was determined by means of 0.2% strain offset.

C. Measurement of Ash and Organic Weight

Immediately following mechanical testing, the ash weight of the specimens was determined by heating the specimens to 700°C for 20 hours. The ash density was obtained by dividing the ash weight with the specimen volume. Organic weight (plus small percentage of water content) was obtained by subtracting the ash weight from the total weight of the specimen. The organic density was then obtained by dividing by dividing the organic weight with the specimen volume.

D. Analysis

All analysis was performed with SPSS (IBM SPSS Statistics, Version 24.0, IBM Corp., Armonk, New York) taking $p < 0.05$ as significant. Relationships between bone mechanical properties (cancellous, cortical and pooled) and mineral density were established following the widely-accepted power law relationships. Relationships between bone mechanical properties and organic/mineral density were also established. The outputs of both relationships were then compared.

III. RESULTS

The average mechanical properties of bone specimens – both cortical and cancellous – in response to compressive loading conditions are summarized in Table 1 below. The mean compressive strength and Young's modulus of cortical bone are higher than those of cancellous bone.

The relationship between ash density and Young's modulus for pooled (cortical and cancellous) bones as well as the correlations ($p < 0.05$) is presented in Fig. 2. The relationships show the most favorable result for the pooled bones ($R^2 = 0.75$) as opposed to that of cortical ($R^2 = 0.33$) and cancellous ($R^2 = 0.28$). The relationship between ash density and compressive strength for pooled bones ($p < 0.05$) as well as the correlation is presented in Fig. 4. The derived relationships show the most favorable result for the pooled bones ($R^2 = 0.80$) as opposed to that of cortical ($R^2 = 0.35$) and cancellous ($R^2 = 0.39$).

TABLE I. AVERAGE PROPERTIES OF TESTED CORTICAL AND CANCELLOUS SPECIMENS.

	n	ρ (g/cm ³)	E (MPa)	σ^{peak} (MPa)	σ^{yield} (MPa)	ρ_{ash} (g/cm ³)	$\rho_{\text{ash}}/\rho_{\text{organic}}$ ratio
Cortical	59	2.073 (0.31) ^a	7191.69 (2412.31)	128.35 (47.34)	122.89 (51.59)	1.234 (0.203)	2.098 (0.297)
Cancellous	89	1.411 (0.24)	1450.01 (643.47)	12.36 (8.63)	12.067 (8.64)	0.490 (0.139)	0.672 (0.219)

a. The numbers in parentheses are standard deviation

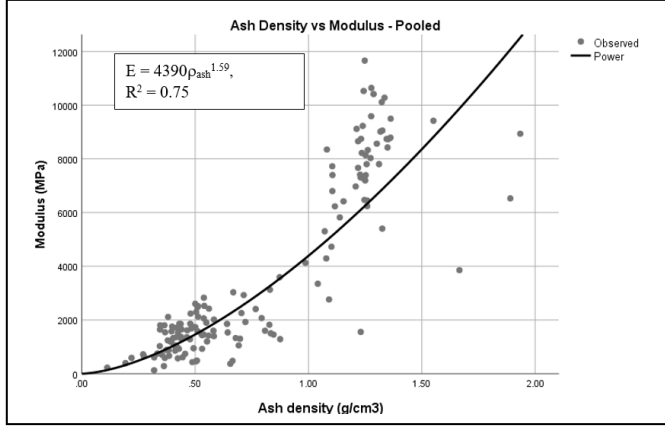


Figure 2. Variation of compressive Young's modulus (MPa) with ash density (g/cm3).

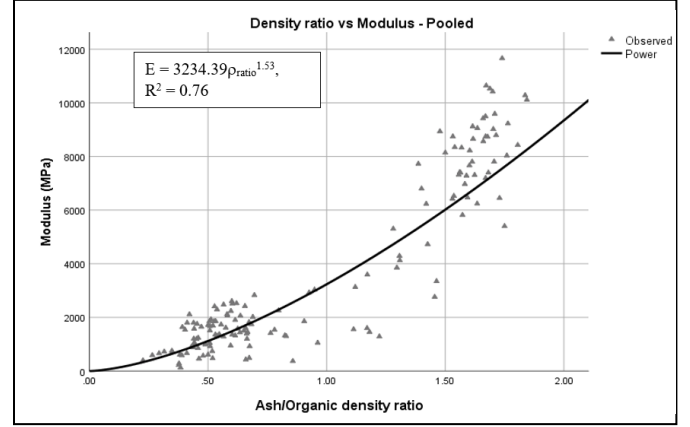


Figure 3. Variation of Young's modulus (MPa) with ash/organic density ratio.

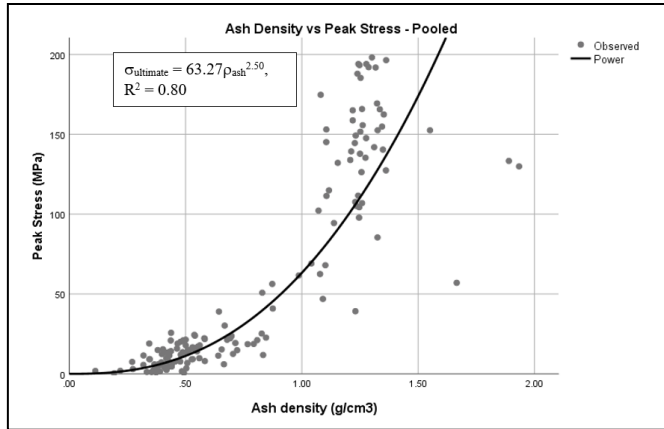


Figure 4. Variation of compressive peak stress (MPa) with ash density (g/cm3).

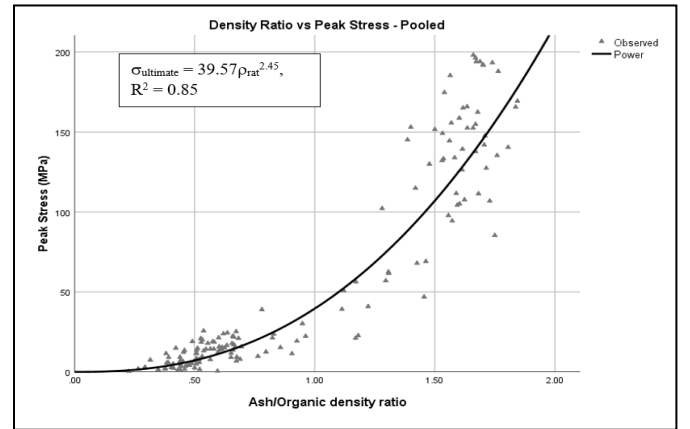


Figure 5. Variation of compressive peak stress (MPa) with ash/organic density ratio.

The contribution of organic content is shown in the relationships generated between density ratio (ash/organic density) and the two properties investigated – modulus and compressive strength. The relationship between density ratio and modulus for pooled bones ($p < 0.05$) as well as the correlation is presented in Fig. 3. The derived relationships show the most favorable result for the pooled bones ($R^2 = 0.76$) while cortical and cancellous bones yielded R^2 of 0.64 and 0.22 respectively. The relationship between density ratio and compressive strength for pooled bones ($p < 0.05$) as well as the correlation is presented in Fig. 5. The derived relationships show the most favorable result for the pooled bones ($R^2 = 0.85$) while cortical and cancellous bones yielded R^2 of 0.64 and 0.46 respectively.

The density ratio relationships yielded higher correlations than relationships utilizing just mineral strength. Notably, the density ratio yielded an 80% increase in accuracy for cortical bone (from $R^2 = 0.35$ to 0.64) and 6 – 10% increase for pooled bones.

IV. DISCUSSION

The goal of this study was to investigate the contribution of organic and mineral integrally on the mechanical behavior of bone. Although the collagen (organic content) behavior have been the subject of a few researches over the past few years, whether its combination with the mineral component yields an increased accuracy has yet to be experimentally resolved. The effect has been studied by isolating organic protein from mineral phase [19], [22]; collagen denaturation through heating

[23]; or simply analyzing the bone matrix behavior [24]. The results of this study corroborate that organic content plays a vital role in bone mechanical properties. While other studies have investigated the contribution of collagen exclusively to arrive at this conclusion, this study investigated the effect of both organic and mineral phase on bone mechanical properties. These findings suggest that existing work may have overlooked this interaction which is crucial to obtaining an increased accuracy in bone mechanical property prediction. Morgan, Bayraktar, and Keaveny [25] proved that the accuracy and precision of modulus-density relationships can only be improved by site-specific relationships since inter-site variations occur across a skeletal architecture. This study provides a more fundamental approach to increase accuracy for site-specific relationships – the consideration of organic content information.

The experimental results for mechanical properties (with respect to the use of mineral content information) obtained from this study are similar to those reported in the literature [20], [25], [26]. They however differ from Chen and Mckittrick [22] who reported modulus values for cancellous bone of about 600MPa in compression. This difference could possibly be ascribed to methodological differences [12] (specimen preparation and geometry specifically). They prepared their sample by water-jetting the bone marrow and air drying after fabrication – a stark contrast to this study’s specimen preparation process which prioritized maintenance of the bone’s integrity as much as possible by leaving the specimen wet and wrapped in cling film unless needed for testing. Other factors could possibly contribute to the variations in mechanical properties such as age and gender; however, quantification of these dissimilarities is beyond the scope of this work at this time.

The results obtained for mechanical properties with respect to the interaction of mineral and organic content suggests that the accuracy of bone mechanical property prediction can be improved by inclusion of organic content information. This is especially useful in subject-specific bone-type prediction scenarios where for example, specialized equations are allocated to cortical versus cancellous bones to increase accuracy. The results of this study do not support the hypothesis that bone type-specific relationships would yield more accurate results. What is more evident is that for cortical bone, the correlation increased more than 80% (from $R^2 = 0.35$ to 0.64) which indicates that organic content information cannot be simply overlooked. The same can be said about the pooled bone relationships as correlation increased about 10%.

The study has some limitations. First, the compression testing was not controlled by the use of end caps as suggested in the literature, which means that vertical misalignment may have not been completely avoided. The cylindrical specimens were placed on the base platen and preconditioned with load using the top platen to minimize this issue. However, errors might have been introduced and caused an underestimation of Young’s modulus.

Secondly, the tissue strength/stiffness equations used in this study relied on sets of constants based on the dataset and were thus empirically derived. Their ability to predict the mechanical

properties of bone is therefore unknown unless they are validated by testing on an independent dataset. The equations used for both mineral and organic-mineral interaction relationships were therefore unique to the dataset under examination.

V. CONCLUSION

The mechanical behaviors of bone in compression – based on the roles that mineral content and organic-mineral content play – were investigated and compared. The results demonstrate that organic content has considerable effect on bone stiffness and strength, which should be considered in bone elasticity-density relationships to improve their accuracy.

ACKNOWLEDGMENT

The reported research has been supported by Research Manitoba, which is gratefully acknowledged. Thanks also direct to Mr. Zeev Kapitanker for his assistance in handling experimental facilities.

REFERENCES

- [1] J.-E. Tarride *et al.*, “The burden of illness of osteoporosis in Canada,” *Osteoporos. Int.*, vol. 23, pp. 2591–2600, 2012.
- [2] S. A. Jackson, A. Tenenhouse, L. Robertson, and C. S. Group, “Vertebral Fracture Definition from Population-Based Data: Preliminary Results from the Canadian Multicenter Osteoporosis Study (CaMos),” *Osteoporos. Int.*, vol. 11, pp. 680–687, 2000.
- [3] E. Seeman and P. D. Delmas, “Mechanisms of Disease Bone Quality-The Material and Structural Basis of Bone Strength and Fragility,” *N. Engl. J. Med.*, vol. 354, pp. 2250–2261, 2006.
- [4] T. M. Keaveny, T. P. Pinilla, R. P. Crawford, D. L. Kopperdahl, and A. Lou, “Systematic and Random Errors in Compression Testing of Trabecular Bone,” *J. Orthop. Res.*, vol. 15, no. 1, pp. 101–110, Jan. 1997.
- [5] X.-X. Wen *et al.*, “Relationship between sample volumes and modulus of human vertebral trabecular bone in micro-finite element analysis,” *J. Mech. Behav. Biomed. Mater.*, vol. 60, pp. 468–475, 2016.
- [6] E. F. Morgan and T. M. Keaveny, “Dependence of yield strain of human trabecular bone on anatomic site,” *J. Biomech.*, vol. 34, no. 5, pp. 569–577, May 2001.
- [7] C. Öhman *et al.*, “Mechanical testing of cancellous bone from the femoral head: Experimental errors due to off-axis measurements,” *J. Biomech.*, vol. 40, no. 11, pp. 2426–2433, Jan. 2007.
- [8] F. Linde, I. Hvid, and F. Madsen, “The Effect of Specimen Geometry on the Mechanical Behaviour of Trabecular Bone Specimen,” *J. Biomech.*, vol. 25, no. 4, pp. 359–368, 1992.
- [9] D. R. Carter and W. C. Hayes, “The Compressive Behavior Porous of Bone Structure as a Two-Phase,” *J. bone Jt. Surg.*, vol. 59, no. 7, pp. 954–962, 1977.
- [10] R. Havaladar, S. C. Pilli, and B. B. Putti, “Insights into the effects of tensile and compressive loadings on human femur bone,” *Adv. Biomed. Res.*, vol. 3, 2014.

- [11] L. J. Gibson, "THE MECHANICAL BEHAVIOUR OF CANCELLOUS BONE," *J. Biomech.*, vol. 18, no. 5, pp. 317–328, 1985.
- [12] B. Helgason, E. Perilli, E. Schileo, F. Taddei, S. Brynjólfsson, and M. Viceconti, "Mathematical relationships between bone density and mechanical properties: A literature review," *Clin. Biomech.*, vol. 23, no. 2, pp. 135–146, 2008.
- [13] J. Y. Rho, R. B. Ashman, and C. H. Turner, "Young's modulus of trabecular and cortical bone material: Ultrasonic and microtensile measurements," *J. Biomech.*, vol. 26, no. 2, pp. 111–119, Feb. 1993.
- [14] S. Nobakhti and S. J. Shefelbine, "On the Relation of Bone Mineral Density and the Elastic Modulus in Healthy and Pathologic Bone," *Curr. Osteoporos. Rep.*, vol. 16, pp. 404–410, 2018.
- [15] P.-Y. Chen and J. McKittrick, "Compressive mechanical properties of demineralized and deproteinized cancellous bone," *J. Mech. Behav. Biomed. Mater.*, vol. 4, no. 7, pp. 961–973, Oct. 2011.
- [16] S. Weiner and H. D. Wagner, "THE MATERIAL BONE: Structure-Mechanical Function Relations," *Annu. Rev. Mater. Sci.*, vol. 28, pp. 271–298, 1998.
- [17] X. Banse, T. J. Sims, and A. J. Bailey, "Mechanical Properties of Adult Vertebral Cancellous Bone: Correlation With Collagen Intermolecular Cross-Links," *J. Bone Miner. Res.*, vol. 17, no. 9, pp. 1621–1628, Sep. 2002.
- [18] M. D. Gryn timer, J. H. Tupy, and J. Sodek, "The distribution of soluble, mineral-bound, and matrix-bound proteins in osteoporotic and normal bones," *Bone*, vol. 15, no. 5, pp. 505–513, Sep. 1994.
- [19] E. Novitskaya *et al.*, "Anisotropy in the compressive mechanical properties of bovine cortical bone and the mineral and protein constituents," *Acta Biomater.*, vol. 7, pp. 3170–3177, 2011.
- [20] D. T. Reilly and A. H. Burstein, "The Elastic and Ultimate Properties of Compact Bone Tissue," *J. Biomech.*, vol. 8, no. 6, pp. 393–405, 1975.
- [21] W. Bonfield and M. D. Gryn timer, "Anisotropy of the Young's modulus of bone," *Nature*, vol. 270, no. December 15, p. 453, Dec. 1977.
- [22] P.-Y. Chen and J. McKittrick, "Compressive mechanical properties of demineralized and deproteinized cancellous bone," *J. Mech. Behav. Biomed. Mater.*, vol. 4, no. 7, pp. 961–973, 2011.
- [23] X. Wang, R. A. Bank, J. M. Tekoppele, and C. M. Agrawal, "The role of collagen in determining bone mechanical properties," *J. Orthop. Res.*, vol. 19, pp. 1021–1026, 2001.
- [24] D. B. Burr, "The contribution of the organic matrix to bone's material properties," *Bone*, vol. 31, no. 1, pp. 8–11, Jul. 2002.
- [25] E. F. Morgan, H. H. Bayraktar, and T. M. Keaveny, "Trabecular bone modulus–density relationships depend on anatomic site," *J. Biomech.*, vol. 36, no. 7, pp. 897–904, 2003.
- [26] R. Hodgkinson and J. D. Currey, "Young's modulus, density and material properties in cancellous bone over a large density range," *J. Mater. Sci. Mater. Med.*, vol. 3, no. 5, pp. 377–381, 1992.

COMPUTATIONAL MODELING OF PASSIVE ANKLE EXOSKELETON BIOMECHANICS

Jin Zhu

Department of Mechanical Engineering
Faculty of Engineering
University of Ottawa
Ottawa, Canada
jzhu068@uottawa.ca

Marc Doumit, Ph.D.

Department of Mechanical Engineering
Faculty of Engineering
University of Ottawa
Ottawa, Canada
marc.doumit@uottawa.ca

Abstract— Passive ankle exoskeleton devices are emerging as a form of wearable robotic systems that offer orthopaedic assist while being light of weight and independent of a power source. With its purpose being to compensate for muscle weaknesses by supplementing a mechanical force, understanding its biomechanical contribution and the behaviour of affected individual muscles during the gait cycle is essential for designing an optimum orthopaedic assist device. While obtaining precise and accurate in-vivo muscle measurements remains the most accurate evaluation for exoskeleton efficacy, this may be challenging and counter productive at the early phase of device development. The use of dynamic simulation software for biomechanical applications can offer a powerful tool that can predict individual muscle behaviour and exoskeleton efficacy subsequent to external mechanical contribution. This study uses OpenSim as a musculoskeletal simulation framework to predict and analyse the effect of wearing a passive ankle exoskeleton (PAE) on the major calf muscles, namely the ankle joint and the knee joint for varying spring stiffnesses. Simulation results showed that not only did the addition of the PAE reduce the soleus and medial gastrocnemius muscle activation force, but also as the spring stiffness was increased from $10 \text{ kN}\cdot\text{m}^{-1}$ to $17.5 \text{ kN}\cdot\text{m}^{-1}$, the muscle activation time decreases as well. The reverse behavior was seen in the knee and ankle vertical reaction forces, where the presence of the biarticular spring force increased the magnitude of vertical joint reaction force in both joints. An increase in joint moment was also observed in both the knee and ankle when the unassisted model was introduced with the $10 \text{ kN}\cdot\text{m}^{-1}$ PAE spring. Thus, the proposed simulation model can be used to validate empirical studies as well as greatly streamline the PAE design process and be customizable to various patient or performance needs.

Keywords—*biomechanics, exoskeleton, dynamic modeling, prosthetics, biomedical engineering, design, computational mechanics.*

I. INTRODUCTION

The musculoskeletal system acts as a complex network of biomechanical elements that interact to execute deliberate motion. As described by Delp et al. [1], there has always been a great interest in studying the biomechanical nature of how orthopaedic assist devices can facilitate various coordinated movements such as in the cases of athletic performance, physical rehabilitation, and gait assist for seniors. As such, a vast collection of studies has been conducted by physicians and scientists to amass a wealth of knowledge regarding the mechanical properties of muscles, bones, and joints [1]. However, the ability to consider all the kinematic properties of an anatomical model and observe the biomechanical contribution of individual elements during any movement remains a significant challenge.

Relying solely on experimental testing to assess biomechanical behavior presents several draw-backs and limitations. Gait studies, electromyography (EMG) measurements and other form of in-vivo experiments are costly, time consuming, and can be very invasive and uncomfortable to the participant [1]. Although they are commonly used for gait analysis, a single EMG sensor can only provide an activation profile for individual muscle elements and the resulting motion produced for one target muscle at a time, which limits the number or muscles that can be analyzed at once. For instance, while several EMG electrodes can register whether a certain group of muscles are excited during a given motion, one is unable to test in a live participant whether the presence of this muscle group is directly responsible for that motion, and what is the extent of contribution, since muscles can also be involved in actuating movement in joints to which they are not directly attached [2]. These challenges in measuring and analyzing the complex biomechanical behavior of the human body creates the need for a computational tool that can model the

musculoskeletal system and allow one to quantify and predict the mechanical roles of different muscles in various motions. Hence, the use of dynamic simulation tools in understanding the cause and effect behavior of individual muscle elements is incredibly valuable in the design process of orthopaedic assist devices.

There are currently several commercially available simulation software packages such as *AnyBody* (AnyBody Technology), *Visual3D* (C-Motion), and *BoB* (Biomechanics of Modeling), in addition to the open source software *OpenSim* (SimTK) [2]. OpenSim is an open source computational software developed by Stanford University [3], which allows users to model, simulate, and analyze dynamic musculoskeletal loading situations. It has been used to model many coordinated movements, such as simulating the biomechanical outcome of surgical procedures [4], jump landings [2], optimizing athletic performance in running [5], and even studying animal gait [6]. Moreover, there is already a growing collection of studies that have used OpenSim to analyze muscle activity under the influence of lower-limb assistive devices for several motions such as running [7], walking [8], and hopping [9]. Uchida et al. [7] modeled running using muscle-driven simulations of 10 participants running at different speeds. They reported a $26 \pm 4\%$ decrease in metabolic power consumption when running at a slower speed of $2 \text{ m}\cdot\text{s}^{-1}$ as a result of virtually adding an ideal ankle plantarflexion/dorsiflexion assistance as an ideal actuator. On the other hand, a higher running speed of $5 \text{ m}\cdot\text{s}^{-1}$ only showed a $19 \pm 4\%$ decrease in metabolic power for the same assistance parameters [7].

Another study by Poppo et al. [8] also investigated the changes in various muscle parameters such as activation, fiber force, and metabolic power of an unassisted OpenSim model compared to a model assisted by a spring component running parallel to the calf. By comparing to a virtual exoskeleton model represented with an added spring stiffness of $180 \text{ Nm}\cdot\text{rad}^{-1}$, they found that for the soleus muscle, the metabolic power was reduced by 6.1%, the fiber force reduced by 10.5%, and the muscle activation was reduced by 7.8% [8]. In addition, the trends bore rough resemblance to a previous study by Collins et al. [10]. However, the models used in Poppo's simulations were simplified down to include only the medial and lateral gastrocnemii, soleus, and tibialis anterior muscles. With all other muscles of the lower limb removed, these simulations do not account for any possible compensation from other muscles that articulate with the ankle or knee joint, such as the quadriceps muscles.

For this study, the OpenSim application will be focused on modeling a passive ankle exoskeleton (PAE). Numerous studies have repeated evidence that ankle exoskeletons with a clutch and spring design is able to partially inject a torque force component to the ankle joint for jumps [11] and walking [10], [12] to reduce the load on the Achilles tendon. The design consists of an elastic component that is anchored at the ankle and knee joint, with a mechanical clutch at the ankle joint that allows the elastic to stretch when engaged. The elastic-spring

component is stretched over the stance phase and recoils during the swing phase of a normal gait cycle, and acts in parallel to the calf muscles. This requires no external power and can be designed to be light of weight and comfortable to wear [10].

To accurately evaluate the effectiveness of these PAE designs, it is necessary to investigate the degree to which individual muscles in the lower limb are off-loaded. Using prototype testing alone is cost and time consuming, particularly the process of building the prototype itself. Even once ready for testing, PAE prototypes can often disrupt the natural kinematics of normal gait in attempt to cooperate with the device and can result in a large variability of results among different test participants [7], [13]. Hence a kinematics-driven simulation offers a valuable predictive tool to compute the biomechanical effects of wearing an exoskeleton, allowing one to determine the mechanical response of each individual muscle during a motion and make comparisons between assisted and unassisted conditions driven by the same kinematics.

Thus, this study aims to propose a computational musculoskeletal model that does not intend to replace experimental study but rather to complement it by providing parametric and predictive models that can accelerate the analysis process. The model will offer the ability to simulate biomechanical effect of adding a passive mechanical device to the leg and provide valuable and immediate insight towards determining the most effective walking-assist design.

II. METHODS

To understand the lower limb muscles and joint response behavior when a PAE is used, a parametric comparison is made between a non-assisted and an assisted musculoskeletal model in OpenSim 3.3. A biarticular path spring is virtually added to the right calf representing the assisted model and imitating the effects of a PAE [11], while an unassisted model represents the natural biomechanics of healthy gait. For both models, solving static optimization and joint reaction loads will be the two main analyses performed through OpenSim to determine the muscle forces and net joint reaction force and torques respectively. The impact observed at the soleus and medial gastrocnemius muscles, the knee joint, and the ankle joint will be compared between the unassisted model and assisted model to determine whether the influence of a PAE reduces the force requirements supplied by the calf or increases the force requirement in attempt to resist the PAE. These analyses will be repeated for different spring stiffnesses to assess its behavior and evaluate optimum results.

A. Static Optimization

The OpenSim static optimization tool is used in this study to determine the muscle activations and forces from a pre-existing motion file made available by SimTK [14], from which the results are further used to compute the joint reaction loads

[15]. The term “static” optimization refers to the process of solving for equations of motion based on the provided input generalized coordinates for position, velocity, and acceleration [16]. This occurs frame by frame, without integrating between each time-step, hence the static method allows for decreased computation time [15]. The algorithms used by OpenSim’s static optimization method assumes that the muscle fiber activation acts along an inextensible tendon fiber, in other words it does not take into consider compliance properties in the tendon [16].

To implement the static optimization function to solve for the muscle forces of the assisted and unassisted model, empirical data was used to define the subject and loading scenario being analyzed. The generic model and kinematic and kinetic data was made readily available by DeMers et al. [14] from a previously conducted gait experiment using motion capture technology, with external loads being obtained from force plate measurements with recordings of ground reaction force. The static optimization force is necessary to estimate muscle activation forces, which in turn are used to compute joint reaction forces [14].

B. Joint Reaction Load Analysis

Subsequent to the Static Optimization, the joint reaction loads are solved through an inverse dynamics process, where unknown reaction forces for each body segment are determined using a force equilibrium equation starting at the foot and repeated for each body segment moving up to the pelvis. A set of 6D Newton-Euler equations of motion is constructed for each body segment, and knowing the mass, acceleration, and velocity components, OpenSim then computes the reaction forces at each joint. To illustrate this, a free body diagram of the tibia has been isolated to show the joint reaction calculations in Fig. 1.

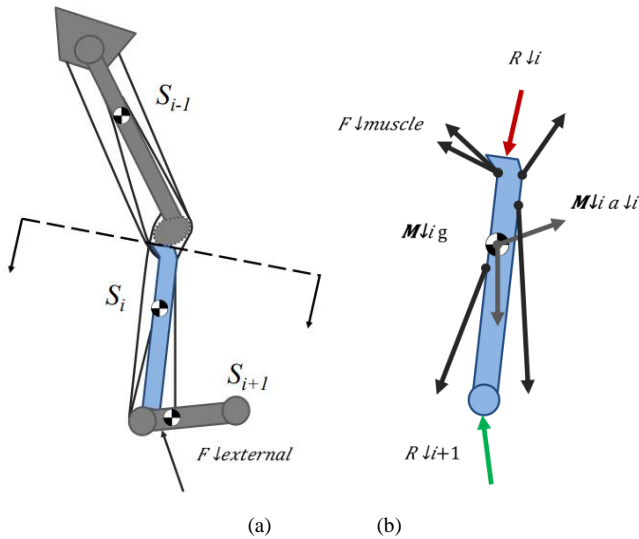


Figure 1. (a) A free body diagram of the lower limb. (b) A free body diagram of the isolated tibia [14]

To achieve an equilibrium system in (1) for the entire body, all the forces exerted on model must sum up to the net inertial forces represented by $M_i a_i$, where M_i is the mass of the body segment, and a_i represents its acceleration. This net inertial force also considers Coriolis and centrifugal forces [14].

$$M_i a_i = \sum F_{external} + \sum F_{muscles} + \sum F_{constraints} + R_{i+1} + R_i \quad (1)$$

$$R_i = M_i a_i - (\sum F_{external} + \sum F_{muscles} + \sum F_{constraints} + R_{i+1}) \quad (2)$$

The $F_{external}$ component represents any external loads applied to the lower limb, in this case would simply be gravity and ground reaction forces measured from the force plate. There are no external forces applied to the tibia in this study, since the exoskeleton will be modeled as an additional muscle as opposed to an actuating force. The $F_{muscles}$ component represents the muscle forces acting on the segment, computed from the static optimization step, and would include the effects generated by the PAE in the assisted model. A loading scenario may also contain constraint forces, which would be accounted for in the $F_{constraints}$ term. Reaction forces from the distal articulating body segment would yield from solving the equilibrium condition for the previous distal body segment, namely the talus bone reaction force acting on the tibia in this case, represented by R_{i+1} . With these known force components, OpenSim will solve for the unknown joint reaction force R_i , in this case at the proximal end of the tibia. These calculations are then repeated up to the pelvis-femur joints.

C. Musculoskeletal model

The default OpenSim 3.3 library contains several different 3-dimensional musculoskeletal models that can be used to determine the muscle and joint forces in the lower extremities. The two generic models are *Gait2392* and *Gait2354*, both representing an unscaled subject of 1.8 m tall with a mass of 75.16 kg, which consists of a head, torso, no arms, a pelvis, and two legs [14]. Both models are compatible with many setup files and motion data available in the OpenSim default download folder [18], however *Gait2354* is a slightly more simplified version of *Gait2392*. The *Gait2392* model includes 23 degrees of freedom and 92 actuator forces to represent 76 muscles in the lower limbs. Hence a modified version of *Gait2392* prepared by DeMers et al. was selected for this study, which can be found in the example files from the Joint Analysis Example Webinar available on the OpenSim website [14].

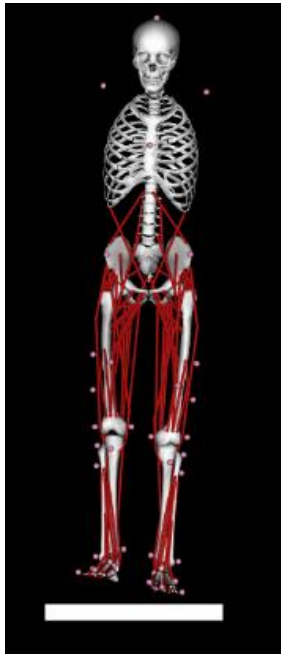


Figure 2. Generic lower limb model Gait2392

D. Modeling the Passive Ankle Exoskeleton

The assisted model includes a biarticular spring model added to the right ankle that acts as the PAE device that works in parallel with the right medial gastrocnemius muscle with the purpose of reducing calf muscle forces during plantar flexion.

To adapt the model, a script written in Python syntax was coded to represent wearing the PAE in the OpenSim's built in scripting shell [19]. In order to accurately elucidate the response behavior of an added spring component, the control model of the generic *Gait2392*, which acts as the base model, was cloned before the spring was added so that spring stiffness is the only virtual anatomical difference between the assisted and non-assisted model. Figure 3 shows a visual representation of the steps taken to edit the original model to include a spring component.

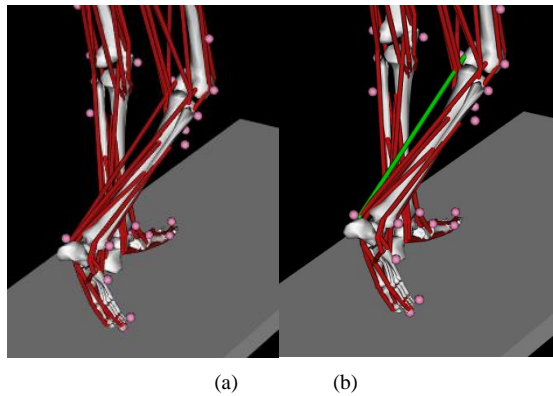


Figure 3. (a) Unassisted model of calf muscles. (b) Added biarticular path spring for assisted model.

Once the base model is loaded, it is cloned and added with a spring component, which consists of assigning the spring properties, namely the rest length, stiffness, and dissipation. A rest length of 0.4 m and a dissipation of $0.01 \text{ N}\cdot\text{s}\cdot\text{m}^{-1}$ will be used throughout all simulation runs, while stiffnesses of $10 \text{ kN}\cdot\text{m}^{-1}$ and $17.5 \text{ kN}\cdot\text{m}^{-1}$ will be used to compare the effects of PAE stiffness. Once the spring properties are defined, the geometry path and insertion points of the spring were based on those of the medial gastrocnemius. In OpenSim, muscles and tendons are modeled as cable-like force components with one insertion point at each extremity and can bend around potential wrap-around points. Adding a biarticular path spring is analogous to adding a supplemental muscle force with an adjustable stiffness and damping component, hence mimicking the purpose of a passive spring-based exoskeleton device. It is important to note that this script does not *replace* the gastrocnemius muscle with a spring, but rather uses the same location and geometry path as the gastrocnemius muscle to define the placement and geometry of the biarticular spring. Because the mass of the spring is very small in proportion to the rest of the body mass, the spring is assumed to be mass-less with negligible effects on the ground reaction forces.

Using the developed assisted and unassisted models, the static optimization function is computed to determine the muscle activation force using the model, kinematics, and external force inputs described previously. The muscle force is used as an input to the Joint Reaction force function to determine joint forces in the right knee and ankle for both conditions. Since this study is interested in determining the biomechanical effects of wearing a PAE device, muscle and joint forces will be computed with a focus on the calf muscles, namely the soleus and the medial gastrocnemius. Although the lateral gastrocnemius is also a major calf muscle as a component of the triceps surae muscle group [20], its behavior is expected to be similar to that of the medial gastrocnemius, hence only one of the gastrocnemius muscles will be studied to minimize redundancy. These results will be compared at a two different spring stiffnesses of $10 \text{ kN}\cdot\text{m}^{-1}$ and $17.5 \text{ kN}\cdot\text{m}^{-1}$ to observe the effects of increasing exoskeleton stiffness on the muscle and joint forces.

III. RESULTS AND DISCUSSION

A. Ground Reaction Forces and Gait Phases

The ground reaction forces are plotted as a function of time to identify the timings of a single bipedal gait cycle. By inspection, the main phases of the gait cycle in Figure 4 are matched to a time frame and tabulated in Table I.

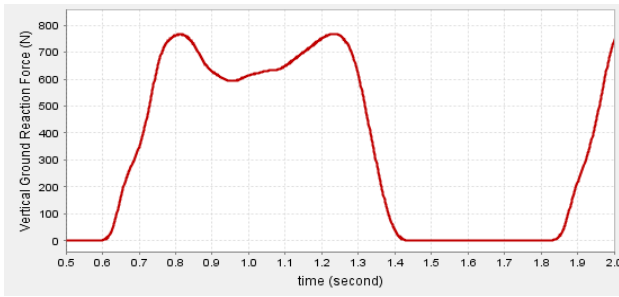


Figure 4. Right foot ground reaction force.

TABLE I. PHASES OF GAIT AND THE CORRESPONDING TIME FRAME FOR THE RIGHT FOOT

Gait phase (right foot)	Time (second)
Initial contact	0.61
Midstance	0.95
Terminal stance	1.24
Toe off	1.42

B. Muscle Activation Force

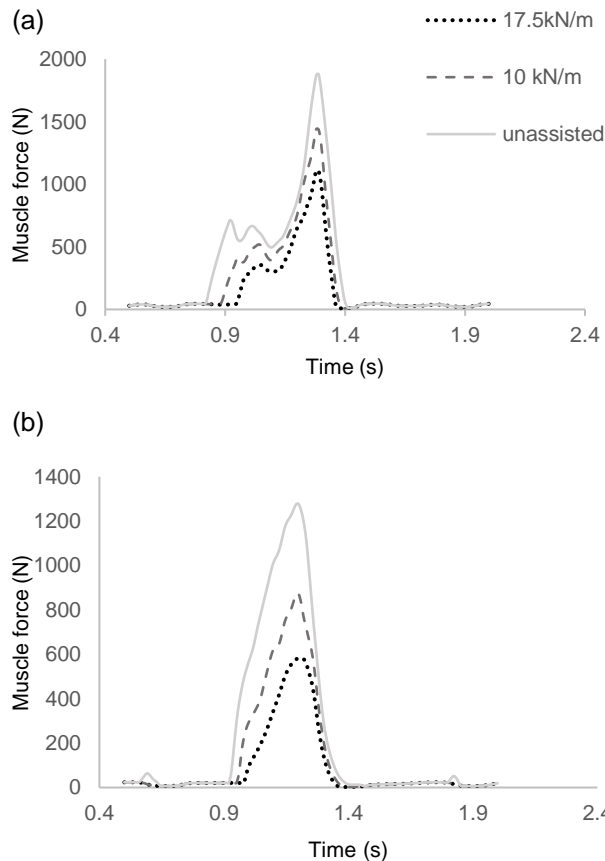


Figure 5. (a) Muscle activation force in the right soleus muscle. (b) The right medial gastrocnemius muscle.

Two different levels of PAE spring stiffness in addition to the unassisted condition are plotted to compare the resultant muscle force in the soleus and medial gastrocnemius. As displayed in Figure 5, both calf muscle forces are highly dependent on the spring stiffness in the PAE worn. As the spring stiffness is increased, there is a notable decrease in muscle activation force. In other words, the unassisted model exerts the highest amount of force, while the model with the 17.5 $\text{kN}\cdot\text{m}^{-1}$ PAE spring experiences the lowest amount of exerted muscle force. All models have a peak muscle force right before toe-off at approximately the 1.3 second frame for the soleus, and at terminal stance (1.2 second frame) for the medial gastrocnemius. Compared to the unassisted model, the 17.5 $\text{kN}\cdot\text{m}^{-1}$ model showed a peak muscle force in the soleus of only 59% of the unassisted model, while the 10 $\text{kN}\cdot\text{m}^{-1}$ model showed a peak muscle force of approximately 75% of the unassisted model. Since the kinetics and kinematics for all three models are fixed, this suggests that a PAE spring can decrease the force requirement from the muscle in order to maintain net joint moments, and that the PAE is able to supplement a parallel actuation force without having the muscles resist against the foreign force.

The stiffness of the spring also appears to affect the timing at which each muscle begins to exert force. In both the medial gastrocnemius and the soleus muscle, the greater the spring stiffness, the greater of delay is caused in muscle force exertion. This is slightly more apparent in the soleus muscle, where the difference in muscle activation start times between the three conditions is approximately 0.1 seconds. This may be attributed to the spring stiffness providing stability, where the soleus and gastrocnemius normally contribute stabilizing the foot between the initial contact and midstance phase of the gait cycle. Conversely, the addition of the PAE shortens the force exertion time between the terminal stance and toe-off phase. Therefore, the results show that the addition of a biarticular spring not only decreases the overall muscle activation force in the calf muscles, but also reduces the duration of force exertion, with the 17 $\text{kN}\cdot\text{m}^{-1}$ model having even lower muscle force exertion and duration than the 10 $\text{kN}\cdot\text{m}^{-1}$ model.

It is also important to note that the muscle force simulation results clearly demonstrate the magnitude of influence of the soleus and medial gastrocnemius at each frame of the gait cycle. Figure 5(b) shows that the gastrocnemius is at its peak activation at terminal stance, where the right heel lifts off the ground, while Fig. 5(a) shows that the soleus peaks right before toe-off, where the entire right foot is lifted off the ground. This suggests that the soleus muscle plays a greater role in lifting the foot off the ground, while the gastrocnemius muscle's greatest contribution is to lift the heel while the toes are still supporting part of the body's weight. The decrease in gastrocnemius muscle activation is also seen to correspond with the rise of the soleus activation force, showing how the two triceps surae muscles work together to execute limb lift-off.

C. Joint Reaction Force and Moment

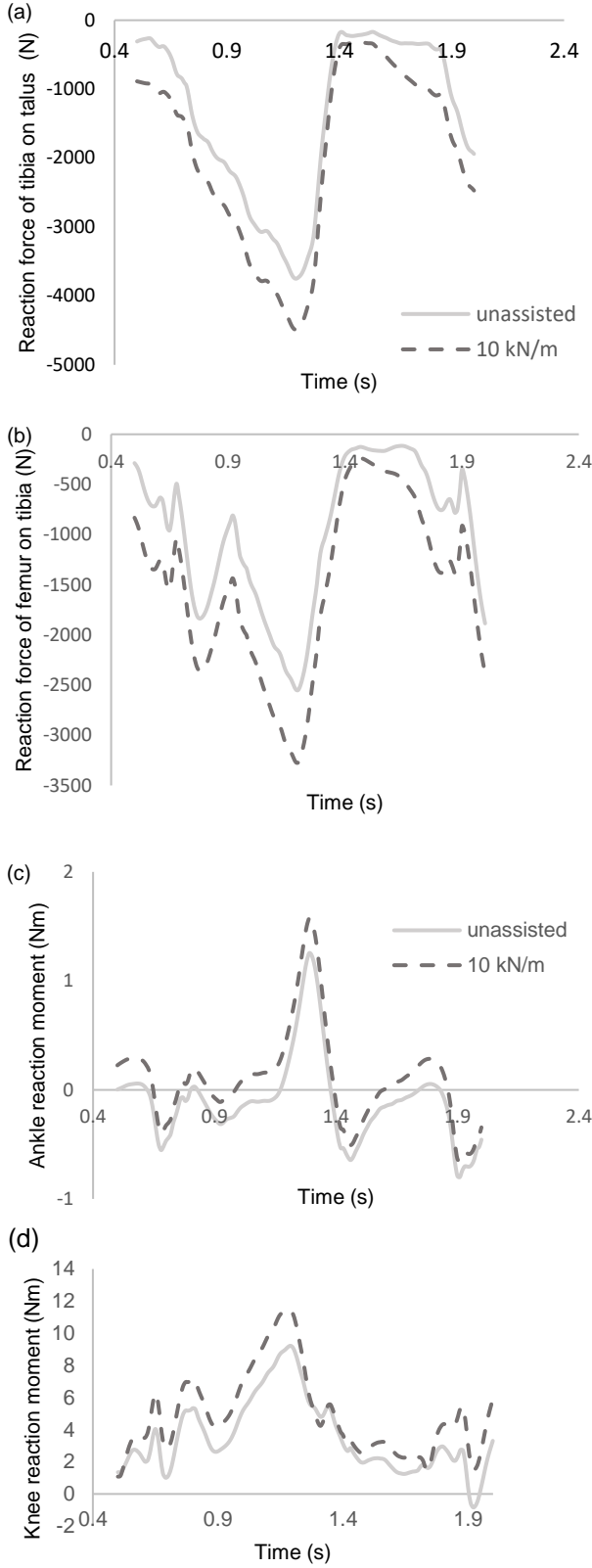


Figure 6. (a) Vertical reaction loads during one gait cycle for the right ankle, (b) and knee. (c) Moment about the Z-axis for the ankle, (d) and the knee.

In OpenSim, a joint is modeled as a collective point between articulating bones [14]. Hence, the vertical joint forces for tibia-on-talus ankle forces and femur-on-tibia knee forces are negative, since ground reaction forces are modelled as positive. From the muscle force comparisons in Figure 5, it is observed that increasing spring stiffness from $10 \text{ kN}\cdot\text{m}^{-1}$ to $17.5 \text{ kN}\cdot\text{m}^{-1}$ results in similar trends of change compared to an increase from the unassisted model to the $10 \text{ kN}\cdot\text{m}^{-1}$, hence why only the $10 \text{ kN}\cdot\text{m}^{-1}$ stiffness is compared to the unassisted model for joint reaction load results. From Figure 6 (a) and (b), the results show that adding a PAE increases the magnitude of vertical reaction force in both the ankle and the knee. At terminal stance, where the vertical reaction force as well as the difference between the two models are at their peak, the PAE increases the joint reaction force by as much as about 10% in the right ankle and 29% in the right knee.

To this point, the presence of a PAE spring has reduced the muscle force and increases vertical joint reaction forces. Hence, it is expected that the moment plots in Figure 6(c) and (d) also show an increase in net joint moment for both the knee and the ankle once the PAE is added. In addition, the moments plot for the knee joint in Figure 6(d) show a slight overlap in the moment values between the unassisted model and the $10 \text{ kN}\cdot\text{m}^{-1}$ model, meaning at certain points of gait the PAE actually creates a lesser joint moment while mostly yielding a higher joint moment for the rest of the gait cycle.

D. Model Validation and Limitations

One of the challenges in using dynamic simulation software is ensuring that the simulation computation results are representative of real-life loading conditions. In this study, the most proximal joint for all models used is the hip joint, which is unconstrained with reference to the ground. This means that the pelvis is free to move in any direction, and that there should not be any reaction loads between the ground and the pelvis as there are no constraints to resist. Therefore, by ensuring the ground-on-pelvis reaction force is zero, this intuition is used to validate whether the model is correct [14].

Nevertheless, it is recognized that there are certain modelling limitations. The simulation model constrains the medial gastrocnemius, which as a bi-articular muscle spans the knee in addition to the ankle joint. This condition may produce passive knee-ankle exoskeleton behavior as the PAE only constrains the ankle joint. It is also noted that the moment arm plays an important role in determining the ankle joint moment stiffness when it comes to PAE design, where small changes to the moment arm can dramatically alter ankle moment stiffness. However, the moment arm for the PAE tested in this study remained constant throughout all trials and can be varied for future studies comparing different PAE designs.

The model is assumed to be mass-less and uses the same ground reaction forces in both the assisted and unassisted condition. It is possible that under real-life conditions, the subject may try to resist the exoskeleton in attempt to walk

normally or change their gait to adapt to the foreign device, which may affect the ground reaction force and kinematic inputs. The model also does not consider muscle fatigue from wearing the device for extended amounts of time. However, some studies have found evidence that suggests this can be trained and users can gradually regain normal gait after experience with wearing the assistive device [21]. Furthermore, the behavior of muscles above the knee and the tibialis anterior are also likely affected by the load and constraints of the assistive device, however the scope of this study is focused on investigating the main posterior calf muscles.

IV. CONCLUSION

The biomechanical effects of wearing a PAE was investigated using computational results from OpenSim. OpenSim's static optimization and joint load reaction analysis capabilities were used to determine the muscle forces and the joint reaction loads respectively. The main actuators in the calf muscle group, the soleus and medial gastrocnemius, were studied along with the ankle and knee joint. Two models were compared for the joint reaction force while three models were compared for the muscle forces, ranging from an unassisted limb, to having a $10 \text{ kN}\cdot\text{m}^{-1}$ spring stiffness, and a $17.5 \text{ kN}\cdot\text{m}^{-1}$ stiffness for the muscle force comparisons only. Joint kinematics, a generic lower limb musculoskeletal model, and ground reaction forces were used as inputs into the simulation and were obtained from the OpenSim library. Joint reaction load computation is based off an equilibrium problem that includes known external forces, constraint forces, muscle forces, and solves for the unknown reaction force of the bone of interest and is iterated from the distal end of the lower limb to the most proximal bone which is the pelvis. Simulations were validated by showing that the reaction force on the pelvis with reference to the ground is zero.

Simulations showed that there was an overall decrease in both muscle forces and an increase in both knee and ankle vertical reaction forces and moments. The highest stiffness model of $17 \text{ kN}\cdot\text{m}^{-1}$ yielded the lowest muscle activation force, followed by the $10 \text{ kN}\cdot\text{m}^{-1}$ spring stiffness model, with the unassisted model having the highest muscle force exertion, suggesting that PAE devices are indeed capable of supplementing muscle force, and offer an effective solution to walking with muscle injuries or weaknesses. These muscle force, joint reaction, and joint moment results demonstrate how OpenSim can serve as a powerful tool to compare the magnitude and timing of each muscle contribution over the span of a gait cycle.

V. REFERENCES

- [1] S. Delp *et al.*, "OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement," *Biomed. Eng. IEEE Trans. On*, vol. 54, pp. 1940–1950, Dec. 2007.
- [2] R. Bakker, "The Effect of Sagittal Plane Mechanics on Anterior Cruciate Ligament Strain During Jump Landing," Thesis, University of Waterloo, Waterloo, 2014.
- [3] "OpenSim - Home." [Online]. Available: <http://opensim.stanford.edu/>. [Accessed: 27-Aug-2018].
- [4] "Tutorial 2 - Simulation and Analysis of a Tendon Transfer Surgery - OpenSim Documentation," *SimTK*. [Online]. Available: <https://simtk-confluence.stanford.edu/display/OpenSim/Tutorial+2+-+Simulation+and+Analysis+of+a+Tendon+Transfer+Surgery>. [Accessed: 12-Sep-2018].
- [5] S. Hamner, A. Seth, and S. L. Delp, "Muscle contributions to propulsion and support during running," *J. Biomech.*, vol. 43, no. 14, pp. 2709–2716, Oct. 2010.
- [6] O. Panagiotopoulou, J. W. Rankin, S. M. Gatesy, and J. R. Hutchinson, "A preliminary case study of the effect of shoe-wearing on the biomechanics of a horse's foot," *PeerJ*, vol. 4, 2016.
- [7] T. K. Uchida, A. Seth, S. Pouya, C. L. Dembia, J. L. Hicks, and S. L. Delp, "Simulating Ideal Assistive Devices to Reduce the Metabolic Cost of Running," *PLOS ONE*, vol. 11, no. 9, p. e0163417, Sep. 2016.
- [8] M. N. Poppo, E. McCain, T. Dick, K. Saul, and G. S. Sawicki, "Dynamic simulation of elastic ankle exoskeleton effects on plantarflexor muscle-tendon neuromechanics during walking," presented at the 41st Annual Meeting of American Society of Biomechanics, Boulder, Colorado, 2017.
- [9] D. J. Farris, J. L. Hicks, S. L. Delp, and G. S. Sawicki, "Musculoskeletal modelling deconstructs the paradoxical effects of elastic ankle exoskeletons on plantar-flexor mechanics and energetics during hopping," *J. Exp. Biol.*, vol. 217, no. 22, p. 4018, Nov. 2014.
- [10] S. H. Collins, M. B. Wiggin, and G. S. Sawicki, "Reducing the energy cost of human walking using an unpowered exoskeleton," *Nature*, vol. 522, no. 7555, pp. 212–215, Jun. 2015.
- [11] S. Kim, Y. Son, S. Choi, S. Ham, and C. Park, "Design of a simple, lightweight, passive-elastic ankle exoskeleton supporting ankle joint stiffness," *Rev. Sci. Instrum.*, vol. 86, no. 9, p. 095107, Sep. 2015.
- [12] J. Leclair, S. Pardoel, A. Helal, and M. Doumit, "Development of an unpowered ankle exoskeleton for walking assist," *Disabil. Rehabil. Assist. Technol.*, vol. 0, no. 0, pp. 1–13, 2018.
- [13] S. Pardoel, "Development and Testing of a Passive Ankle Exoskeleton," University of Ottawa, Ottawa, 2017.
- [14] "Example - Estimating Joint Reaction Loads - OpenSim Documentation," *SimTK*. [Online]. Available: <https://simtk-confluence.stanford.edu/display/OpenSim/Example+-+Estimating+Joint+Reaction+Loads>. [Accessed: 08-Apr-2018].
- [15] "Working with Static Optimization - OpenSim Documentation." [Online]. Available: <https://simtk-confluence.stanford.edu/display/OpenSim/Working+with+Static+Optimization>. [Accessed: 08-Apr-2018].
- [16] "How Static Optimization Works - OpenSim Documentation." [Online]. Available: <https://simtk-confluence.stanford.edu/display/OpenSim/How+Static+Optimization+W+orks>. [Accessed: 22-Feb-2018].
- [17] "Gait 2392 and 2354 Models - OpenSim Documentation." [Online]. Available: <https://simtk-confluence.stanford.edu/display/OpenSim/Gait+2392+and+2354+Models+#+Gait2392and2354Models-ExperimentalDataIncludedwiththeModels>. [Accessed: 27-Mar-2018].
- [18] "SimTK: OpenSim: Downloads." [Online]. Available: https://simtk.org/frs/?group_id=91. [Accessed: 12-Sep-2018].
- [19] "Scripting in the GUI - OpenSim Documentation." [Online]. Available: <https://simtk-confluence.stanford.edu/display/OpenSim/Scripting+in+the+GUI>. [Accessed: 27-Mar-2018].
- [20] C. Hacking, "Triceps surae | Radiology Reference Article | Radiopaedia.org," *Radiopaedia*. [Online]. Available: <https://radiopaedia.org/articles/triceps-surae>. [Accessed: 10-Apr-2018].
- [21] K. E. Gordon and D. P. Ferris, "Learning to walk with a robotic ankle exoskeleton," *J. Biomech.*, vol. 40, no. 12, pp. 2636–2644, Jan. 2007.

Biomedical Engineering

Prosthetic Foot Design – A Simplified Analytical Approach

Stephanie Gibbons
Graduate Student
Faculty of Medicine, Biomedical Science
Memorial University of Newfoundland
St. John's, NL

Sam Nakhla, PhD
Assistant Professor
Mechanical Engineering, Faculty of Engineering
Emergency Medicine, Faculty of Medicine
Memorial University of Newfoundland

Adam Dubrowski, PhD
Professor
Emergency Medicine, Faculty of Medicine
Memorial University of Newfoundland
St. John's, NL

ABSTRACT

A prosthetic foot is without a doubt an essential article that enables amputees reclaim a significant range of their activities. Three major characteristics differentiate various designs of prosthetic feet, namely, biomechanical functions, metabolic demand and shock absorption. The most basic prosthetic foot, the Solid-Ankle Cushion-Heel (SACH) foot is used as baseline design to demonstrate against potential advantages of other designs. In reality, international prosthetic market offers multitudes of designs at various levels of technologies [1]. Multi-terrain or energy storage capabilities are common advertisement claims, although these claims are not empirically confirmed [2]. Also, technical comparisons and studies demonstrating advantages are not standardized nor published. Moreover, gait analysis is the common method in comparing different designs in which the performance of a population of patients is monitored when using alternative foot designs. Meanwhile gait analysis cannot be considered standard testing technique given the fact that individuals are different even in a finite population of patients. In summary, there exists a knowledge gap in methods suited to quantify the mechanical response of a prosthetic foot. Therefore, the aim of this study is to develop simplified solid mechanics analysis procedure of prosthetic foot. Accuracy of developed model is to validate against finite element analysis and test measurements. In particular foot prosthetic response must be based on testing performed by means of standard apparatus/fixtures to eliminate biasing the measurements.

Energy principles are elected to develop simplified analysis hence to remain consistent with major identifying characteristics of prosthetic foot. Initial or base line model is developed for the SACH foot based on idealized geometry, i.e. uniform thickness and straight segments. Consequently, enabling comprehensive study focused on geometry migration accounting for individual effects or advantages of variable thickness distribution and curved segments. In later stage comparisons are held with finite element study conducted at Memorial University. Further comparisons with SACH prosthetic foot testing contribute towards better understanding of prosthetic foot functions.

Using the developed method, the expected outcomes of the study are to qualify and quantify different designs of prosthetic feet.. These processes are necessary during amputee fitting based on targeted terrain. Further extension of developed models enables developing patient-specific prosthetic foot design and selection tools. These tools when coupled with rapid prototyping techniques should enable high quality and affordable healthcare for amputees.

References

- [1] Versluys et al. Prosthetic feet: review. Disability and Rehabilitation: Assistive Technology. 2009.
- [2] Lehmann et al. Comprehensive analysis of energy storing prosthetic feet. Arch of phys med and rehab. 1993.

Aortic Wall Stresses in Correlation to Potential of Rupture

Stephanie Gibbons
Graduate Student
Faculty of Medicine, Biomedical Science
Memorial University of Newfoundland
St. John's, NL

Sam Nakhla, PhD
Assistant Professor
Mechanical Engineering, Faculty of Engineering
Emergency Medicine, Faculty of Medicine
Memorial University of Newfoundland

Adam Dubrowski, PhD
Professor
Emergency Medicine, Faculty of Medicine
Memorial University of Newfoundland
St. John's, NL

ABSTRACT

Abdominal Aortic Aneurysms (AAAs) occur when an area of the aorta dilates causing stress and weakening of the aortic walls. AAAs run the risk of rupture as they are very often asymptomatic. Previous studies utilizing analytical and numerical simulations focused on estimating the forces acting on the stents (device inserted into a vessel to keep the vessel open, or to alleviate stress on the vessel walls in the event of an aneurysm) due to blood flow [1, 2]. In particular, fluid structure interaction simulations enabled evaluating stresses in the aortic walls [3]. However, none of these studies accounted for realistic mechanical properties of the aorta or potential thinning of its walls. Moreover, no specific testing is designed to validate results obtained from simulations. An additional challenge is that the geometry of many arteries can be quite complex making analytical solutions difficult to obtain. For this reason, it is desired to utilize and validate the stresses using simplified geometry.

Current work is planned to investigate the accuracy of simplified analytical and numerical models. For this purpose, testing plan is designed to validate analytical/computational results. The testing plan involves employing physical models resembling artery in question to measure stresses and strains in its walls. Physical models can be obtained from animal tissue (ie. porcine) or artificial vessel developed using rapid prototyping techniques. Essential material characterization phase is established to measure mechanical properties of physical models. In the following phase test setup aims at instrumenting physical models to measure strains in their walls due to fluid flow. These strain measurements establish realistic benchmark for validating analytical and computations models.

Analytical models established based on actual geometry measure of tested physical models are to be used for approximate estimation of wall stresses. Simplified and in-depth calculations of stresses in the walls are to be compared to strains obtained from testing accounting for measured mechanical properties. The accuracy of these analytical models determines the necessity or need to conduct further numerical investigations. Meanwhile numerical investigations in terms of computational fluid mechanics and finite element analysis can establish local zones of extreme stresses within the artery. In general, once strains and stresses in the walls are validated, measures on wall thinning are straightforwardly established.

This comprehensive double-faceted study is necessary to enable understanding the dangers and potential of rupture in AAA. It worth noting, the study involves intricate details imposed by unique characteristics of fluid flow and non-uniform distribution of geometric and mechanical properties of an artery. Successive and careful steps are required to guarantee identifying the effects associated to each of these unique characteristics.

References

- [1] Liffman K, Lawrence-Brown MM, Semmens JB, Bui A, Rudman M, Hartley DE. Analytical modeling and numerical simulation of forces in an endoluminal graft. *Journal of Endovascular Therapy*. 2001 Aug;8(4):358-71.
- [2] Howell BA, Kim T, Cheer A, Dwyer H, Saloner D, Chuter TA. Computational fluid dynamics within bifurcated abdominal aortic stent-grafts. *Journal of Endovascular Therapy*. 2007 Apr;14(2):138-43.
- [3] Molony DS, Callanan A, Kavanagh EG, Walsh MT, McGloughlin TM. Fluid-structure interaction of a patient-specific abdominal aortic aneurysm treated with an endovascular stent-graft. *Biomedical engineering online*. 2009 Oct 6;8(1):24.

Modelling of flow-induced shear stress to predict targeted delivery of cells in the decellularized lung

Jason K. D. Chan

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
jasonk.chan@mail.utoronto.ca

Eric A. Chadwick

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
chadwick@mie.utoronto.ca

Takaya Suzuki

University Health Network
University of Toronto
Toronto, Canada
suzuki.takaya@gmail.com

Golnaz Karoubi

University Health Network
University of Toronto
Toronto, Canada
golnaz.karoubi@uhnresearch.ca

David Romero

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
d.romero@utoronto.ca

Cristina Amon

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
cristina.amon@utoronto.ca

Thomas K. Waddell

University Health Network
University of Toronto
Toronto, Canada
tom.waddell@uhn.ca

Aimy Bazylak

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
abazylak@mie.utoronto.ca

ABSTRACT

Approximately 120,000 patients die from end-stage lung disease annually in the United States. Lung transplantation remains the only definitive treatment for end-stage lung disease. However, many complications arise due to the immune system's response to the foreign organ. Decellularization and recellularization of whole organs is a promising approach for reducing this rejection. In this procedure, detergents are used to remove cells from the donor organ via a process known as decellularization. The organ is then repopulated with recipient-specific cells to reduce the chances of an autoimmune response, thereby reducing the rate of organ rejection. Yet, much of the process still requires optimization to improve the delivery, differentiation and maturation of the seeded cells. One of the many challenges includes uniform distribution of the cells throughout the lung, especially in the distal regions. In this study, computational modelling methods are used to understand the effects of flow-induced shear stresses on cell distribution and viability throughout the lung.

Micro-computed tomography (micro-CT) was used to quantify the internal structure of the decellularized mouse lung vasculature. Using this information, computational fluid dynamic models of representative single blood vessels surrounded by porous media zones representing the extracellular matrix (ECM) were created. Bulk information of fluid velocity and pressure distribution were attained for various representative vessels throughout the lung. This data was then incorporated into an existing vessel network model. Shear-stress was calculated using equations derived by Tarbell-Shi. Finally, contour plots of shear stress were correlated to cell distribution results found in literature.

Contours of fluid velocity and shear-stress within the decellularized lung based on flow parameters and mechanical properties were correlated to cell distribution, thus providing an understanding of the underlying mechanics that affect cell distribution and proliferation. With the knowledge that flow conditions are one of the major factors in cell delivery, differentiation, and maturation, this model presents itself as a powerful tool to be used in lung regeneration. The model developed in this study can be used to inform of flow conditions and set of mechanical properties for optimal regeneration of the lung. For example, the human lung consists of over 40 cell types that control its function. Manipulation of inlet and boundary conditions in the model will help advance methods of targeted delivery of cells to specific vessels and airways. Furthermore, the model can also be adapted to decellularization of the lung. This will help researchers understand how the detergents used in the process affects the integrity of the lung scaffold. Ultimately, this work will lead to effective new therapies which will save the lives of countless patients affected by lung disease.

Low-temperature 3D printing of PLGA: printability study

*Emad Naseri, Haley Butler, Wyatt MacNevin, Marya Ahmed, and Ali Ahmadi**

Faculty of Sustainable Design Engineering,
University of Prince Edward Island,
Charlottetown, Canada

*Email: aahmadi@upei.ca

ABSTRACT

Poly-Lactic-co-Glycolic Acid (PLGA) is a bioabsorbable polymer which has shown promise in bioprinting applications through high temperature extrusion. As the bioabsorbable properties of PLGA can be tuned by altering the lactide and glycolide ratio, there exists potential for PLGA to be 3D printed with the inclusion of pharmaceutical agents, living cells, growth factors, or secondary biomaterials for the development of bioprinted drug delivery systems, smart implants, or living constructs. Current methods of PLGA-based biofabrication depend on the use of increased temperatures or bioincompatible solvents unsuitable for the use with pharmaceutical agents and living cells. Current modalities such as solvent casting requires the use of molds, increased cost and fabrication time, and does not allow for a high degree of spatial control that 3D bioprinting offers. To address these challenges, a novel low temperature bioprinting technique was used to fabricate PLGA constructs using methyl ethyl ketone (MEK). Using a custom-built microsyringe and open-sourced 3D bioprinter, printing conditions and software parameters were studied and optimized to enable low-temperature PLGA extrusion. To assess the quality of the printed constructs, printability, defined as a set of non-dimensional values, was analyzed with respect to varying PLGA co-polymer grade, PLGA molecular weight, and PLGA:MEK concentration. PLGA 85:15 (100-200 K molecular weight) demonstrated an enhanced printability in 80% w/v concentration. Furthermore, PLGA 75:25 and 85:15 grades (molecular weight of 100-200 K) demonstrated improved printability compared to other studied grades of PLGA. A low-temperature solvent removal process to remove the residual MEK solvent in the fabricated scaffold was implemented and proton Nuclear Magnetic Resonance (^1H NMR) was used to demonstrate the complete removal of residual MEK solvent. Through this proposed low-temperature approach to PLGA-based 3D bioprinting, the deposition of pharmaceuticals, growth factors, and/or living cells into PLGA scaffolds can be further developed.

FINITE ELEMENT ANALYSIS OF POROUS TITANIUM ALLOY CONSTRUCTS MATCHING THE MECHANICAL PROPERTIES OF MANDIBULAR BONE: A PILOT STUDY

Khaled M. Hijazi, Yara K. Hosein, Haojie Mao, David W. Holdsworth, S. Jeffrey Dixon, Jerrold E. Armstrong, Amin S. Rizkalla

Faculty of Engineering, Schulich School of Medicine & Dentistry, Bone and Joint Institute
Western University
London, Canada
khijazi@uwo.ca

ABSTRACT

The mandible (lower jaw bone) anchors the lower teeth and plays a critical function in mastication (chewing of food), swallowing and speech. Following removal of tumors, infection or trauma, reconstruction of the mandible may be required. Titanium implants are used in mandibular reconstruction and include rigid or semi-rigid plates or trays, which fix fractured bones or support bone grafts. Titanium 6-aluminium 4-vanadium (Ti6Al4V) alloy is currently used for mandibular reconstruction implants due to its mechanical strength, biocompatibility, and ability to anchor into the bone post-surgery. However, stress shielding remains to be an issue with these implants. Stress shielding occurs as result of the mismatch between the elastic modulus of bone and the adjacent implant. This mismatch results in inadequate mechanical loading of the bone, causing a reduction in bone density around the implant, leading to its loosening. Studies have shown that Ti6Al4V lattice constructs have lower elastic modulus than solid titanium. The porous structure also allows the formation of bone ingrowths. The objective of this pilot study was to optimize the design of porous Ti6Al4V constructs to match the elastic modulus of native mandibular bone.

This pilot study employed a design-optimization approach to determine the required strut size range for porous titanium constructs. We used Simulia Abaqus 2017 (Dassault Systèmes, France) to perform computer-aided design (CAD), coupled with finite element analysis (FEA). A series of CAD models were prepared as either rectangular prisms or dumbbell-shaped constructs with cubic pores for the entire rectangular prism and between the solid dumbbell grips. Each construct was built with a uniform strut thickness that varied between 150 and 650 μm . Solid construct models were used as controls. All models were meshed to hexahedral elements of 125 μm average size. FEA was then used to perform three-point bending, tensile, and compressive loading simulations. All simulations were done with ramp loading that automatically terminated at 4500 N. Boundary conditions were chosen to mimic the mechanical testing protocols detailed in ASTM standards for each of the three mechanical testing simulation groups (ASTM E290 – 14, E8/E8M – 16a, and E9 – 09). True stress-strain and force-displacement results from the simulations provided estimates of the flexural, tensile, and compressive mechanical properties of the designed constructs, which were compared to the mechanical behavior of mandibular bone.

The elastic modulus values derived from the mechanical simulations showed second order polynomial increase with increasing strut thickness. Simulations also showed second order polynomial increase of strength with increase in strut thickness. It was determined that Ti6Al4V cubic porous lattice structures with strut thicknesses ranging between 275 and 343 μm had elastic modulus values matching those of mandibular cortical bone.

The FEA simulations on porous constructs done at this point are the first step to optimize the design of lattice porous Ti6Al4V constructs for novel intraosseous mandibular implants. The next step is to fabricate constructs from CAD models using selective laser melting technology (3D metal printing). Internal microporosities within the fabricated constructs will be imaged using micro-computed tomography (μCT). Finally, the mechanical simulations will be validated through mechanical tests using a Universal Instron Testing Machine.

The use of patient-specific, functionally-graded, intraosseous implants is expected to prevent stress shielding. Findings from this research will be translated to the development of novel reconstructive bone implants for oral and maxillofacial surgery, as well as implants for other parts of the skeletal system, with the intention of improved mechanical performance and better interaction with bone.

A simulation on multi-pixel carbon nanotube field emitter for medical imaging

Jiayu Alexander Liu, Elahe Cheraghi, Yonghai Sun, John T.W. Yeow
Department of Systems Design Engineering, University of Waterloo
Waterloo, Canada
jyeow@uwaterloo.ca

ABSTRACT

Carbon nanotube (CNT) field emitter arrays (FEAs) has been studied in recent years for various applications including X-ray imaging. This technology has a high potential for medical therapeutic and measuring method in clinical treatment.

X-ray imaging such as Computational tomography (CT) produces a detailed three-dimensional (3D) images for patients during clinical treatment and includes diagnosis of cancer and guidance of treatment. However, during the X-ray radiating, not only the tumors are radiated, but many healthy cells are also radiated which potentially damage healthy cells and increase cancer risk because the traditional CT has a huge and uncontrollable radiation area.

In order to reduce the radiation does but keep it still enough for treatment, Multi-pixel patterned CNT FEA cathodes is suggested. Multi-pixel CNT cathode is a type of well-patterned CNT array which leads to controlling the X-ray beam by controlling the FE performance of the CNT array cathodes.

In this study, a 3×3 multiple cathode is designed for multi-pixel X-ray source with one anode. All the cathodes are patterned on a n-type doped silicon wafer and each one consists of 100×100 CNT arrays. The carbon nanotube (CNT) field emitter arrays (FEAs) are synthesized by plasma enhanced chemical vapor deposition (PECVD) method. The main purpose of this study is considering the electric field in multi-pixel X-ray source and investigating the controllability of the power of each X-ray pixel. In this regard, COMSOL Multiphysics was used to simulate the electric field of 3×3 multiple cathodes.

Multiscale fluid flow analysis of decellularized mouse lungs to improve lung regeneration

Eric A. Chadwick

Mechanical & Industrial
Engineering
University of Toronto
Toronto, Canada
chadwick@mie.utoronto.ca

Mohammadali Ahmadipour

Institute of Biomaterials & Biomedical
Engineering
University of Toronto
Toronto, Canada
mohammadali.ahmadipour@mail.utoronto.ca

Takaya Suzuki

University Health Network
University of Toronto
Toronto, Canada
suzuki.takaya@gmail.com

Golnaz Karoubi

University Health Network
University of Toronto
Toronto, Canada
golnaz.karoubi@uhnresearch.ca

David Romero

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
d.romero@utoronto.ca

Cristina Amon

Mechanical & Industrial
Engineering
University of Toronto
Toronto, Canada
cristina.amon@utoronto.ca

Thomas K. Waddell

University Health Network
University of Toronto
Toronto, Canada
tom.waddell@uhn.ca

Aimy Bazylak

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
abazylak@mie.utoronto.ca

ABSTRACT

De- and recellularization of whole organs is a promising approach to increase the survivability of organ transplants. In late-stage lung disease, transplantation is often the only viable treatment for the patient. However, when foreign DNA is introduced into the recipient's system, the body can reject the organ even if the donor and recipient are matches. 10 years after receiving a donor lung, there is an average survival rate of only 10-20%. With de- and recellularization, the donor lung is first cleared of the donor's cells, thereby removing their DNA and leaving behind a collagen scaffold of the organ. During the recellularization step, recipient cells are seeded back onto the scaffold of airways and vasculature and then perfused with a cell culture medium to facilitate the growth and migration of cells. Since the airways do not have natural pathways for liquid flow, this perfusion takes place in the vasculature during recellularization. Previous studies have successfully seeded cells into larger, proximal airways; however, seeding and growing mature cells in the capillary bed of the vascular network has been challenging. In this work, we perform high resolution X-ray tomography imaging of the mouse lung and perform flow simulations to identify the transport pathways of cell culture media through the vasculature of decellularized mouse lungs.

Numerical models of native and decellularized mouse lung vasculature have been created using micro-computed tomography scans of vasculature injected with a Microfil contrast agent. The scanned images were skeletonized and further analyzed using morphological image processing techniques to extract a digital model of the vasculature. Analysis of these models have shown evidence of the tissue damage that occurs during decellularization. This evidence takes the form of a lack of connectivity from the arterial side to the venous side of the pulmonary circuit as well as a lack of circularity in terminal vessels. The former is likely due to the contrast agent bursting or leaking out of these vessels due to their porosity or reduced mechanical structure. This model will be the basis for fluid flow simulations using current perfusion parameters such as input pressure and fluid properties. The simulations will be based on Poiseuille flow through the network of vessels. The outcome of this will be pressure and flow distribution maps throughout the vasculature, thereby elucidating the dominating pathways for cell culture media. It is hypothesized that low flow rates will be found in the capillaries, due to the high transport resistance exhibited by the small vessels. Evidence of fluid leaking during recellularization will also be detectable in the form of irregular shaped vessels with high fluid flow rates and low pressures.

This study will be one of the first to characterize the fluid flow during recellularization in mouse lung vasculature. This work will be used to inform researchers of optimal flow conditions to deliver cells to desired tissues during recellularization and will help to advance the efficacy and viability of whole lung regeneration.

A Novel Approach of Predicting Deep Brain Response Following Traumatic Brain Injury

Yanir Levy
School of Biomedical Engineering
Western University
London, Canada
ylevy3@uwo.ca

Kewei Bian
Mechanical and Materials
Engineering
Western University
London, Canada

Haojie Mao
Mechanical and Materials
Engineering
Western University
London, Canada

INTRODUCTION

The presence of sports related concussions and minor traumatic brain injuries has increased in frequency over the past decade. While a possible reason for the increase is improved diagnosis methods and general awareness, the overall improvement in helmet safety technology has seemingly not had a significant effect on reducing these concussion symptoms. Short term effects of traumatic brain injuries are known to lead to impairment in cognitive functions. This study looks to provide a quantitative predictor for how the direction of head impact leads to a biomechanical response in the deep brain, more specifically the corpus callosum (CC) and how it affects cognitive aptitude. The CC contains the greatest density of fibril tract connections between the cerebral hemispheres of the brain making it the primary neural connector and an important factor for human cognitive function. The existence of previous literature with defined diffuse axonal injury (DAI) thresholds allows for a comparative study and the ability to use a mathematical approach as a preliminary predictor of injury.

METHODS

Using Diffusion Tensor Imaging (DTI) a computed tractography finite element model of the fibril network (Essen et al 2013) was created and inserted into the Global Human Body Model Consortium (GHBMC) head and brain model. Different impact loading cases ($n=4$) were simulated with similar time-history impact curves ($5k \text{ rad/s}^2$ at 10 ms) to recreate artificial rotational impact scenarios. The four cases presented in this study are: flexion, extension, rotation and lateral bend. The results of these mathematical simulations were then analyzed quantitatively for both magnitude and location of greatest maximum principal strain (MPS) as well as largest Cumulative Strain Damage Measure (CSDM).

RESULTS

The GHBMC and DTI fibril network of the CC finite element model showed that the peak MPS was largest in the lateral bend (0.256) that was followed by extension (0.221) and flexion (0.206) and finally rotation (0.192). The Fibril networks had opposing rankings with flexion and extension having an MPS of 0.475 followed by lateral bend at 0.31 and rotation at 0.225. These results were then further analyzed using CSDM to determine more in depth, the extent of the injury.

DISCUSSION AND CONCLUSION

These results help set up the quantitative assessment of the cognitive effects of brain injury, based on previous literature thresholds. The use of computer simulation to both diagnose and identify areas of the brain injured through reverse processing the qualitative symptoms of cognitive tests is a novel approach with many clinical and industry advantages. The future goal of this study is to provide the link between current cognitive tests and the areas affected by traumatic brain trauma while providing a new tool for safety manufactures to quantify the effectiveness of safety equipment designed to protect the brain.

[1] Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., and Consortium, W. U.-M. H., 2013, "The WU-Minn Human Connectome Project: an overview," *Neuroimage*, 80, pp. 62-79.

ASSESSING FINGER FORCE AND GRIP CONFIGURATION VARIABILITY OF VARIOUS STANDARD AND ARTHRTIC GOLF GRIPS IN INDIVIDUALS WITH AND WITHOUT HAND ARTHRITIS

Sara Holland

line 1: Mechanical and Materials Engineering
line 2: Western University
line 3: London, Ontario, Canada
line 4: sholla6@uwo.ca

Emily Lalone and Louis Ferreira

line 1: Mechanical and Materials Engineering; Roth
McFarlane Hand and Upper Limb Centre
line 2: Western University; St. Josephs Health Care
line 3: London, Ontario, Canada

ABSTRACT

With arthritis affecting 1 in 5 adults in Canada, it makes it the most prevalent health condition with, no cure [1]. Hand arthritis is one of the most common locations for arthritis to form and affects a large proportion of the population over 50 years of age. Unfortunately, hand arthritis has significant consequences, not only causing pain but also deficiencies in the two main functions of the hand being range of motion (ROM) and grip strength. Another limitation to hand arthritis is being able to participate in recreational activities such as golf. Currently, few 'arthritic' grips are commercially available that use joint protection principles to 'reduce hand forces and tight gripping'. However, no comprehensive examinations in their effectiveness have been assessed. Therefore, the purpose of this study is to determine the golf grip (material and geometry) that produces the highest and lowest finger forces in healthy and arthritic golfers by analyzing the hand forces produced at the hand-grip interface using a new wearable sensor technology. Previously, a non-invasive video-based joint angle measurement technology and Grip Configuration system was developed to examine grip ROM during various functional tasks. This assessment was used to examine the ROM of a golfers gripping style to determine whether there are deficiencies in arthritic golfers grips. In having a better understanding of the mechanics of arthritis and its relation to sports, the design of more advanced sporting equipment can be developed.

In order to study the grip strength and ROM of a golfers grip, the use of a new wireless technology designed by Pressure Profile Systems Inc., (FingerTPS) [2] and a video based optical tracking system known as Dartfish Movement Analysis Software, was utilized. A Grip Configuration Method system was used to evaluate the joint angles of each participants thumb MCP and CMC and index MCP and PIP of the participant's bottom gripping hand (the hand closer to the head of the club). Of the same hand, the forces at the fingertips of the thumb, index, middle and ring fingers were evaluated using the capacitive sensing technology of the FingerTPS system. These tests was conducted at our Thames Hall Biomechanics Laboratory at Western University, with each participant performing 2 golf shots with a real ball on artificial turf with each of the 12, mid-iron clubs fitted with various types of standard and 'arthritic' golf grips. A cohort of 27 (17 healthy control group: 10 individuals with hand arthritis) golfers of varying skill levels were evaluated. The Grip Configuration results showed that the arthritis group had a less flexed grip than that of the healthy group. From the FingerTPS data, graphs were created to demonstrate the force (Newtons) variation throughout the golf swing. Results showed that the arthritic individual's exerted larger average hand forces than the average maximum golf grip strength in 11 out of the 12 grips tested. It was also observed that both the healthy control group and individuals with hand arthritis had lower overall hand forces in the larger diameter, soft firmness (CP2 Pro) designed grips.

With an increasingly aging population and golf often being recommended by clinicians as a low-impact exercise, this research is crucial not only for the players but for clinicians to be able to provide suggestions on the types of grips that should be used by individuals with hand arthritis. Future development in this research could lead to the development of a new arthritic golf grip and to establish better guidelines that will minimize the damaging forces and need for a more flexed grip to enable individuals with arthritis to continue to golf.

REFERENCES

- [1] The Arthritis Society (2018). The Trust About Arthritis. Retrieved from <http://www.arthritis.ca/about-arthritis/what-is-arthritis/the-trust-about-arthritis>
- [2] Pressure Profile Systems (2018). Capacitive Tactile Pressure Sensors. Retrieved from <http://pressureprofile.com/capacitive-sensors>

3D bio-printing of starch-chitosan scaffolds for neuron cells

Haley Butler¹, Emad Naseri¹, Debra MacDonald², Dr. Andrew Tasker^{2,3}, and Ali Ahmadi^{1}*

¹ Faculty of Sustainable Design Engineering, University of PEI

² Department of Biomedical Sciences, University of PEI

³ Department of Clinical Medicine, Aarhus University
Charlottetown, Canada

*Email: aahmadi@upei.ca

ABSTRACT

In this study chitosan and potato starch scaffolds were fabricated and investigated to determine their viability for use as tissue engineering substrates. Chitosan (Dalian Xinde Chitin), which can be enzymatically extracted from lobster shells, is a naturally occurring, biodegradable, and bioresorbable polymer used frequently in the biomedical field. Potato starch (amylose and amylopectin), a waste product from processing potatoes, is a natural thermoplastic with applications in the biomedical field. Both chitosan and potato starch have been studied for use as tissue engineered scaffolds with methods of fabrication ranging from solvent evaporation, freeze drying and molding techniques. Both materials have little background being used as a 3D bioprinting material. On the other hand, polymers such as polylactic acid (PLA), poly-L-lactic acid (PLLA), and polyglycolide (PGA) are commonly used in 3D bioprinted scaffold design due to their mechanical strength, high level of printability and suitable degradation rates. Using natural materials, such as chitosan and potato starch for 3D bioprinting applications can offer enhanced structure and property combinations that can constitute as a viable alternative for scaffold materials.

In this work, the printability of a starch-based chitosan composite materials for use in tissue engineering applications is evaluated using Whitaker et al. approach for printability. This is done by measuring the filament formation, shape retention after printing the material and the degradation of the material. The composites range from 100% potato starch to 100% chitosan. The results from this research is used to investigate the suitability of the Whitaker printability equation for determining the printability of a wider range of materials over time. Moreover, the cytotoxicity of the suggested materials and the viability of printed cells within the scaffold material is investigated. The cells used for this study are Neuro2A neuroblastoma cells. Two approaches are taken: i) cells are seeded onto the scaffold material; and ii) cells are incorporated into the material prior to printing. The cytotoxicity is determined using a lactate dehydrogenase (LDH) assay to determine the percentage of viable cells, along with calcien blue fluorescence staining for visualizing the cells. A live/dead ratio of the printed cells incorporated into the scaffolds compared to the control group of cells is determined using various immunohistochemical markers and confocal microscopy. The research results can lead to development of reliable, inexpensive *in-vitro* models for pre-clinical neurological disease drug discovery.

Detailed Finite Element Analysis of Prosthetic Foot

Amanda Ricketts
 Undergraduate Student
 Mechanical Engineering, Faculty of Engineering
 Memorial University of Newfoundland
 St. John's, NL

Sam Nakhla, PhD
 Assistant Professor
 Mechanical Engineering, Faculty of Engineering
 Emergency Medicine, Faculty of Medicine
 Memorial University of Newfoundland

ABSTRACT

This work addressed finite element method as analysis tool to identify prosthetic foot efficiency. The initial objective is to identify the required accuracy of finite element simulation to estimate prosthetic foot stiffness. To this end multiple analyses were conducted using finite element analysis commercial code Abaqus. Analyses were conducted over three main phases of gait, namely, heel-strike, foot-flat and toe-off. These particular phases of gait directly associate to major characteristics of a prosthetic foot such as shock absorption, metabolic demand and biomechanical function. The standard prosthetic SACH foot design was elected for all analysis purposes. One-, two-, and three-dimensional models were created in order to delineate the effect of geometric features approximation on estimated stiffness accuracy. In order to validate analysis models and their corresponding accuracies comparisons were conducted with test data obtained by South et al. [1]. Sample results from comparisons of one-dimensional model with test data obtain from [1] are provided in Table I.

Table I. One-dimensional model comparison with test results [1]

Gait Phase	Measured Stiffness (kN/m)	One-Dimensional Model (error %)
Toe-off	48.0	68.0 (42 %)
Heel-strike	20.9	30.5 (46 %)
Foot-flat	114.0	78.3 (31 %)

One-dimensional model enabled capturing stiffness properties of the prosthetic foot. Meanwhile as can be noticed from Table I calculated stiffness values were over-estimated for converged models. Detailed analysis steps and investigations in two- and three-dimensions assisted better understanding of some sources of errors. On the other hand, major challenges were raised by the uncertainty in identifying the exact geometry of tested foot. Further details on finite element models and aspects affecting models accuracy are documented in the full manuscript.

References

- [1] South BJ, Fey NP, Bosker G, Neptune RR. Manufacture of energy storage and return prosthetic feet using selective laser sintering. Journal of biomechanical engineering, 2010.

Validation of a Temperature Sensor for Implantable Infection Telemetry

M.K. Lavdas, D. W. Holdsworth, M. G. Teeter
Robarts Research Institute
University of Western Ontario
London, Canada

B.A. Lanting
London Health Sciences Center
Department of Surgery, Division of Orthopaedics
London, Canada

I. INTRODUCTION

Infections affect 1-3% of Total Knee Arthroplasty (TKA) patients with severe ramifications to mobility. Unfortunately, reinfection rates are high (~15%) suggesting improved clinical instruments are required. A common strategy to remedy TKA infection is to remove the infected implant and install a temporary (6-12 week) “spacer” knee implant while the infection is cleared before the implantation of a permanent revision replacement. Subdermal temperature of over 2°C is documented in infection. The context of this study relates specifically to the selection, characterization and statistical validation of an implant-suitable sensor which we can integrate into an instrumented tibial spacer.

II. HYPOTHESIS

We hypothesize that suitable sensing performance for infection monitoring regarding precision and relative accuracy can be attained using a low power, compact, analog sensor with $<0.1^{\circ}\text{C}$ resolution. Testing was referenced against three averaged manufacturer characterized sensors acting as our gold standard.

III. MATERIALS & METHODS

The TMP235A2DCKR (Texas Instruments (TI), Dallas, TX, USA) was selected for our implanted application due to its extremely low (9 μA) current draw and compact DBZ chip package. Based upon dynamic range it was determined that the analog/digital converter must be a minimum of 11 bits to deliver suitable ($<0.1^{\circ}\text{C}$) resolution. A 12-bit ADC equipped microcontroller was selected. The MCP9808 (Microchip Technology, Chandler, AZ, USA) delivers manufacturer characterized thermal data in decimal strings through serial communication to the same microcontroller. The rated accuracy of the MCP9808 sensors in the required temperature range is max/typ $\pm 0.5/0.25^{\circ}\text{C}$ with a precision of $\pm 0.05^{\circ}\text{C}$ delivered at a resolution of 0.0625°C . Three of the MCP9808 sensors were used simultaneously and averaged. The data from both modalities can be exported as comma delimited text files in real time. Within a thermally insulated chamber with a resistive heating element, the following experiment was conducted: Using empirical plant modelling tools, simulating and then implementing a PI control scheme, with the MCP9808s as reference, the temperature in the thermal chamber was driven to 20 different temperatures between 35 and 40°C for 10 minutes each and sampled at 5 Hz. This trial was repeated three times over three days. Transient data was discarded so as only to evaluate the steady state characteristics, wavelet denoising was applied, and a regression between the reference temperature vs the experimental TI sensor was tabulated in Matlab.

IV. RESULTS

Compared to reference values, the TMP235A2DCKR temperature sensor displayed relative accuracy of $\pm 0.27^{\circ}\text{C}$ (with 95% confidence) and precision of $\pm 0.137^{\circ}\text{C}$ over a 35- 40°C range as determined over 190,212 relevant samples. Note that in practice, the precision is independent of reference, but the absolute accuracy is relative to gold standard's accuracy.

V. CONCLUSION

Infection frequently results in permanent mobility issues in the context of total knee arthroplasty. This has led to an ongoing call for better treatments. Analysis suggests that the TMP235A2DCKR offers reasonable precision and relative accuracy in temperature sensing, substantially tighter than the expected stimulus from infection, while also offering desirable characteristics for implantation including low power consumption and compact form factor. This sensing platform will be integrated into our instrumented tibial spacer in future work but could be suitable in many other implant applications as well.

Wireless Telemetry Load Sensor for Orthopaedic Applications

W. D. Anderson, S. Wilson
Faculty of Engineering
University of Western Ontario
London, Canada

D. W. Holdsworth
Robarts Research Institute
University of Western Ontario
London, Canada

I. INTRODUCTION

Recording real time in-vivo load data has long been a challenge due to requirements in size, data transmission, and power management of load sensors. Load data can allow scientists and clinicians to quantify the forces acting on structures within the human body. This information can improve orthopaedic implant design and patient rehabilitation practices.

II. HYPOTHESIS

The objective of this study is to convert a commercially available wireless telemetric tire pressure sensor into a compact, ultra-low power load sensor for use in orthopaedic applications.

III. MATERIALS & METHODS

The tire pressure sensor (FXTH87, NXP) deforms approximately 600 microns over the full range of its pressure measurement capability (100-450 kPa). A deformation calibration from 0-600 microns was performed to determine the signal value corresponding to the deformation of the sensor. Two tire pressure sensors were used, and five trials were performed for each sensor. Compressive extension was applied to the sensor enclosure using an Instron 3343 and a 3D printed indenter at a rate of 600 microns/min. Compressive load (internal sensor resistance) was recorded using a 50 N load cell. Compressive load and extension were obtained using Instron Bluehill Software. Sensor signal data was recorded using an FRDM KW019032 RF receiver and plotted in real time in a graphical user interface (GUI) developed by NXP Semiconductors. Signal values and internal sensor resistance were plotted against compressive extension in Prism. Sensor dimensions and weight were obtained using a caliper and analytical balance, respectively. RF transmission range through air was determined by extending the sensor from the RF receiver until no signal values were observed in the sensor GUI. The average power draw and battery lifetime were calculated using the average current draw of the sensor package during continuous transmission, powered by a 3V, 130mAh CR1632 lithium coin battery. The average current draw of a single transmission was determined using an oscilloscope and a multimeter current adapter.

IV. RESULTS

The sensitivity of the signal vs. extension plot is 1.25 microns/signal unit over the linear region of the graph. The average standard deviation of internal sensor resistance and pressure signal over the linear region of the graph was 0.05 N and 5.35 signal units, respectively. This corresponds to an uncertainty in position of ± 8.04 microns. The maximum internal sensor resistance/compressive load applied to the sensor was 6.34 ± 0.20 N (mean \pm SD). The dimensions of the tire pressure sensor were 7 mm x 7 mm x 2 mm and weighed 0.3 g. The dimensions of the commercially available sensor package used for testing had a diameter of 17.4 mm, a height of 7.4 mm, and weight of 3.5 g. The mean RF transmission range was 4.71 ± 0.49 m. The average power draw during continuous transmission (the highest power consumption state) was 10.3 mW. The average current draw during a single transmission was 3.46 mA. The lifetime of the sensor package during continuous transmission was 37.6 hours.

V. CONCLUSIONS

The results of this study show that a commercially available tire pressure sensor can be converted into a functioning load cell, as there is a relationship between reported signal and deformation over a range of 0-600 microns. The load sensor can be configured into an ultra-miniature package with sophisticated power management and RF transmission capabilities. To the best of our knowledge, this is the smallest system in the world with the capability of wireless transmission of applied deformation, making it an ideal candidate for implantation within orthopaedic components. This sensor package can also be configured within custom mechanical enclosures that are designed to transduce compression, tension, flexure, and torque into compressive extension. This package could be embedded within intervertebral disc replacements, spinal fusion cages, fracture fixation plates, high tibial osteotomy implants, and more.

A new way to link animal head kinematics to human head motion

Lihong Lu
 Department of Mechanical and Materials
 Engineering
 Western University
 London, Canada
 llu89@uwo.ca

Kewei Bian, and Haojie Mao
 Department of Mechanical and Materials
 Engineering
 Western University
 London, Canada
 kbian2@uwo.ca and hmao8@uwo.ca

Background: Scaling is an important method to transform the biomechanical and pathophysiological outcome of neurotrauma animal models to be applied to the humans. So far two brain injury scaling laws have been proposed: head angular acceleration based on brain mass [1] and the equal stress/velocity based on kinematic loading condition [2]. However, the applicability of animal (e.g. mouse) brain injury to human remains uncertain due to the lack of a comprehensive understanding of what exactly happens to animal and human brains during impacts.

Objective: We evaluated these two scaling laws using validated finite element (FE) brain models of human and mouse and proposed a new kinematic-based mouse to human scaling relationship to study concussion.

Methods: Sinusoidal impact curve with the peak acceleration of 5 krad/s^2 and duration of 10 ms was applied to the human head model (GHBMC). We then scaled human kinematics to mouse using two existing scaling laws. We calculated the cumulative strain damage measure (CSDM) of both human and animal brains to verify scaling laws. We finally adjusted the peak acceleration and duration of the mouse model to get the same CSDM as the human model predicted.

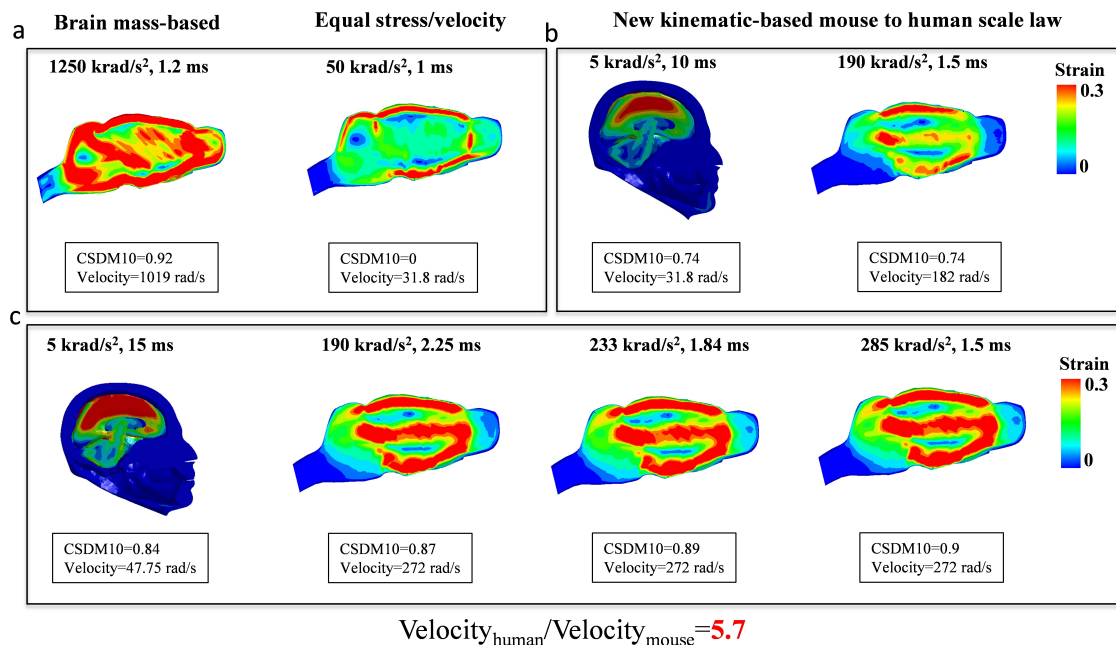


Figure 1. The strain contours based on three scale laws: a. brain mass-based and equal stress/velocity. b and c new kinematic-based mouse to human scale law

Results: Compared with a CSDM10 of 0.74 for the human brain, a CSDM10 of 0.92 and a CSDM10 of zero were found for the mouse brain, using mass-based and the equal stress/velocity-based laws, respectively. Our results demonstrated that existing laws didn't work between human and mouse. We further found that the mouse model with the peak acceleration of 190 krad/s^2 and duration of 1.5 ms produced the same CSDM10 of 0.74. Lastly, we found that the rotational velocity of the mouse head needed to be 5.7 times of that of the human brain to ensure both human and mouse brains were equally stretched.

Conclusions and Discussion: The mass-based and the equal stress/velocity-based scaling laws failed to produce similar brain strains between human and mouse. We proposed a new kinematic-based mouse to human scaling relation, which will be applied on our own animal experiments to help establish effective laboratory animal neurotrauma models.

References:

- 1 Ommaya et al. Stapp Car Crash J 1967 73-80
- 2 Takhounts et al. Stapp Car Crash J 2003 107-134

Carbon-based pH Biosensors for Monitoring Infections

Dogan Sinar, George K. Knopf
 Department of Mechanical and Materials Engineering
 The University of Western Ontario
 London, Ontario Canada
 gkknopf@uwo.ca

ABSTRACT

All biochemical processes in the body including digestion and wound healing are influenced by pH. Saliva, blood and spinal fluid all have a pH of about 7.4, while the stomach is acidic with a pH in the range 1.5 to 3.5. Fluctuations in the corresponding pH levels of these body fluids can indicate the onset of infection or a chronic condition [1]. In this research a mechanically flexible pH sensitive graphene-cellulose interdigitated capacitive biosensor for monitoring the acidity of body fluids [2] is introduced (Fig. 1a). Graphene is a biocompatible carbon allotrope that can be used to create conductive electrodes, passive circuitry and electronic biosensors [3]. The planar interdigitated capacitive (IDC) biosensor is printed on a non-rigid polymer substrate material and coated with a very thin passivation layer to prevent an electrical short with the material under test. The electrically conductive ink is synthesized by using carboxymethyl cellulose (CMC) to suspend hydrophobic graphene (G) sheets in a water-based solvent [3]. Once deposited on the substrate the conductivity of the printed G-CMC IDC electrodes is increased using thermal reduction. To enable the biosensor to respond to changes in fluid acidity, a biocompatible pH-sensitive chitosan hydrogel is immobilized on the IDC electrodes. The simple IDC design (Fig. 1b) takes advantage of the changes in dielectric properties of an electrode surface when exposed to the analyte (i.e. substance undergoing analysis). Both the dielectric constant and surface coverage of an analyte, or dielectric layer, can produce a measurable change in the IDC capacitance. In this regard, a pH sensitive hydrogel is used as the dielectric layer to track changes to the material under test (MUT). Preliminary experiments were performed to test the hydrogel actuated G-CMC biosensor's response to various buffer solutions with different pH values (Fig. 1c). As expected, the volume of the hydrogel would significantly increase (i.e. swell) for low pH and the measured capacitance would be much higher at low frequencies ($< 10\text{kHz}$) than a solution with a high pH. The tests demonstrate that a measurable change in capacitance is observed as the hydrogel swells/de-swells in response to ΔpH . Issues associated with mechanical and chemical stability are also discussed.

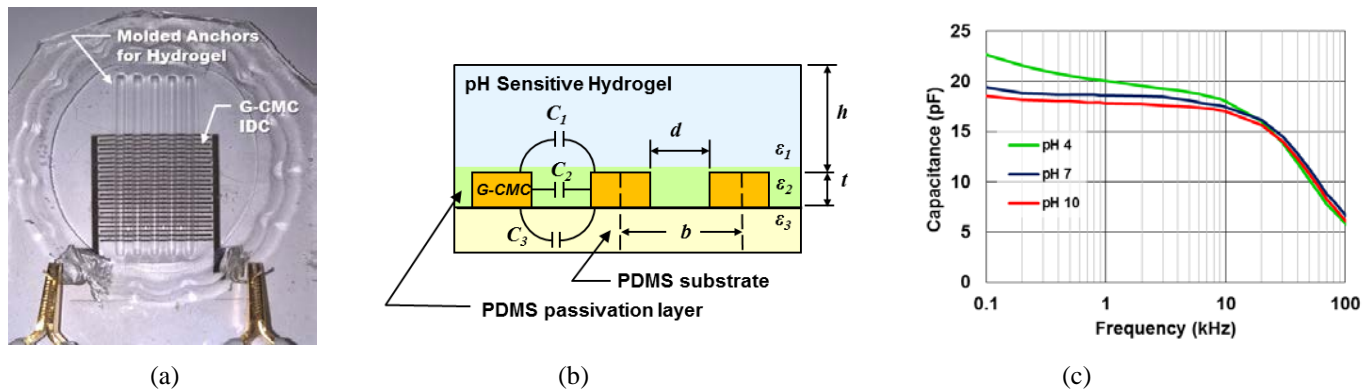


Figure 1. (a) Prototype of a G-CMC printed hydrogel actuated biosensor for monitoring pH in body fluids. (b) Cross-sectional diagram of a planar G-CMC IDC biosensor that responds to dielectric changes in the hydrogel (i.e. MUT). (c) Frequency response of hydrogel actuated G-CMC IDC biosensor to standardized buffer solutions with various pH.

REFERENCES

- [1] Maalouf, N.M. et al. (2007) "Low urine pH: A novel feature of the metabolic syndrome," *Clin. J. Am. Soc. Nephrol.*, 2: 883-888.
- [2] Knopf, G.K., Sinar, D. (2017) "Flexible hydrogel actuated graphene-cellulose biosensor for monitoring pH." *IEEE Symp. Circuits and Systems (ISCAS)*, pp. 1-4.
- [3] Sinar, D., Knopf, G.K. (2018) "Printed graphene derivative circuits as passive electrical filters," *Nanomaterials* 8(2):123-137.

Microfluidic Droplet Transport by Photoelectrowetting

Khaled M. Al-Arife

Department of Mechanical Engineering
 Abu Dhabi University
 Abu Dhabi, UAE
 Khaled.alaribe@adu.ac.ae

George K. Knopf

Department of Mechanical and Materials Engineering
 The University of Western Ontario
 London, ON Canada
 gkknopf@uwo.ca

ABSTRACT

Microfluidic and lab-on-chip devices are complex integrated systems that require very small quantities of fluid to be transported, manipulated, mixed, and separated. Photoelectrowetting on a dielectric film (PEWD) [1] is one approach to transporting single droplets of liquid along microchannels (Fig. 1a) connecting reservoirs and chemical reactors. The basic structure of a PEWD consists of an electrical power source, dielectric film, and photoresponsive layer (Fig. 1b). The operating principle [2] is based on altering the conductivity and/or capacitance using the photosensitive material to induce transitions in the film surface's free energy when it is exposed to a controlled light beam. Changing the localized electrical conductivity influences the electrostatic field that is formed between the activated zones of the dielectric film and the ionic content of the driven droplet. This interaction changes the surface tension locally and, correspondingly, the droplet contact angle. In this research, a novel photoelectric graphene [2,3] doped zinc-oxide (ZnO-G) film deposited on ITO glass is introduced for improved performance. The directed light triggers two forces that enable the droplet to be transported along the substrate. The first arises from the induced hydrophobicity gradient formed across the droplet contact area with the substrate surface. Exposing the ZnO-G film to a broad spectrum white light source alters the surface's electric potential which induces a change in the droplet's contact angle and the associated hydrophobicity. Once the hydrophobicity gradient is generated the droplet will start to move in the direction of the wetting zone. The second force occurs when the absorbed light generates a photoelectric potential that produces a piezo-electrical effect on the ZnO-G film. To exploit the photoelectrowetting phenomenon in microfluidic devices it is often necessary to speed-up the actuation time and reduce the applied voltage difference in an effort to minimize evaporation of the minute quantities of liquid. Preliminary experiments are performed to investigate the photoelectric potential of light activated ZnO-G films. The high resistance piezo-responsive ZnO crystals ($k\Omega\cdot\text{cm}$) doped with highly conductive graphene ($\sim 10^{-6} \Omega\cdot\text{cm}$) show a 2X increase in the photoresponse time (Fig. 1c).

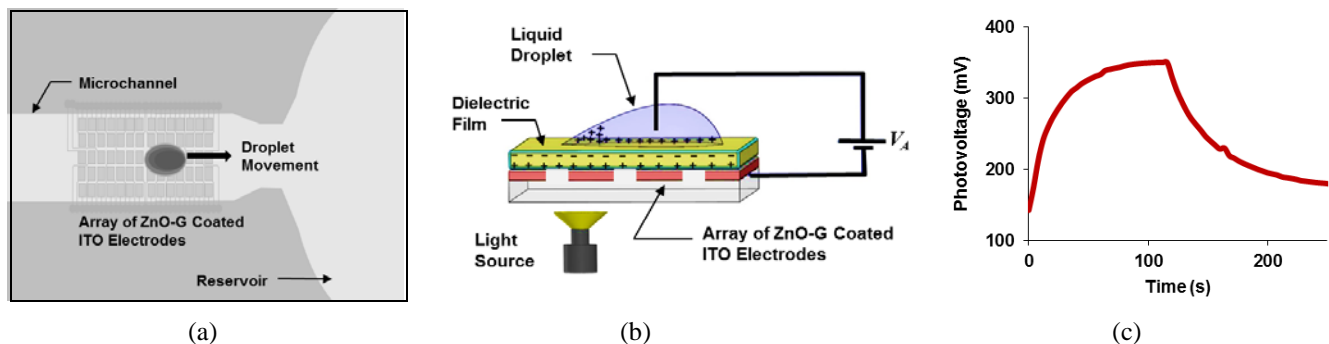


Figure 1. (a) Illustration of light driven mechanism embedded in a microfluidic channel. (b) Operating principle of liquid droplet actuation using photoelectrowetting on a dielectric (PEWD). (c) Measured photovoltage across the ZnO-G coated ITO electrode during light excitation and grounding phase.

REFERENCES

- [1] Ichimura, K., Oh, S.K., and M. Nakagawa, M. (2000) "Light-driven motion of liquids on a photoresponsive surface," *Science* 288(5471): 1624-1626.
- [2] Al-Arife, K., Knopf, G.K. (2017) "Graphene doped ZnO films for photoelectrowetting on microchannels," in *SPIE Microfluidics, BioMEMS, and Medical Microsystems XV*, vol. 10061, p. 1006105.
- [3] Geim, A.K. (2009) "Graphene: status and prospects," *Science* 324(5934), pp. 1530-1534, 2009.

HEMODYNAMICS OF CEREBRAL ANEURYSMS: A COMPUTATIONAL AND EXPERIMENTAL STUDY

Jacob Barrera^{1,2}, Han Hung Yeh^{1,2}, Dana Grecov^{1,2,†}

¹Department of Mechanical Engineering, ²School of Biomedical Engineering
University of British Columbia
Vancouver, BC, Canada
[†]dgrecov@mech.ubc.ca

ABSTRACT

Cerebral aneurysms, which are cerebrovascular abnormalities characterized by the weakening and dilation of a localized area of the wall of a cerebral artery, are a prevalent vascular disorder affecting 2–5% of the worldwide population with the consequence of severe disability or death upon aneurysm rupture [1-3]. Although age and blood pressure are common risk factors for the initiation, growth and rupture of an aneurysm, the precise reasons are not yet well understood. In addition to the hemodynamics in the development of an aneurysm [4], the location and anatomy of cerebral aneurysms are also associated with the risk of rupture [5]. To study the progression of a cerebral aneurysm, computational fluid dynamics (CFD) analysis has been used due to its noninvasiveness and the ability of providing hemodynamics evaluations for blood pressure, shear stress, and velocity distributions, which are all challenging to measure clinically. The benchtop *in vitro* experiments, on the other hand, provide physical and physiological measurements and validation towards a CFD analysis with advance visualization and measurement techniques such as particle image velocimetry (PIV), laser doppler velocimetry, and volumetric three component velocimetry [6-7]. Given that most benchtop systems have a narrow operating peak flow rate under a constrained real-time cardiac output profile, the main objective of the current study is to design a fully automated modular *in vitro* benchtop measuring system in order to mimic the hemodynamics of various sections of the cardiovascular system with adjustable cardiac characteristics.

The system consists of a feedback-controlled servo-pump coupled with flow and pressure sensors through a high frequency data acquisition system and a LabView interface. It allows to be programed with a representative flow rate profile signal to reproduce the specific location of interest. To replicate physiological conditions, a working fluid consisting of a mixture of glycerin (60%) and water (40%) which closely approximates the density and viscosity of human blood is used [8]. The fluid is warm up to maintain a physiological temperature condition of 37°C. The compliance in the system is adjusted to maintain physiologic pressures. The feedback control loop implemented allows the pump to generate different desired flow rate profiles with an error of 2.7% from the target value. The proposed *in vitro* benchtop system reproduces better physiologically realistic flow rate generation accuracy than those in previous studies [6,9]. Moreover, the accuracy is within the range of typical commercial systems ($\pm 2\%$ to $\pm 4\%$). Also, the real-time controlled system was developed to be suitable for integration with other computational programs. In order to measure the flow distributions, silicone elastomer cerebral aneurysm phantoms have been fabricated for PIV measurements. In addition to the development of the benchtop setup, the current study also investigates the influence of the anatomical characteristics of the sidewall cerebral aneurysm to its hemodynamics. For this aspect of the study, the CFD analysis is conducted in parallel with the development of the *in vitro* benchtop system. Multiple idealized and clinically relevant sidewall cerebral aneurysm models are used for the computational study. The CFD hemodynamics analysis consist of the evaluation of blood flow velocity distributions, arterial wall shear stress distributions, as well as blood pressure distributions within the aneurysm sac. Computational model verifications and validations are conducted using PIV measurements with a matching refractive index between the silicone phantoms of idealized cerebral aneurysm and the working fluid.

- [1] Keedy, A. (2006). *Mcgill J Med*, 9(2), 141–146.
- [2] de Rooij, N., et al. (2007). *J. Neurol. Neurosurg. Psychiatry*, 78(12), 1365-1372.
- [3] Shobayashi, Y., et al. (2010). *Med Eng Phys*, 32(9), 1015-1024.
- [4] Sadasivan, C., et al. (2013). *Ann Biomed Eng*, 41(7), 1347–1365.
- [5] Duan, Z., Li, et al. (2018). *Sci Rep*, 8(1), 1–7.
- [6] Tsai, W., & Savaş, Ö. (2010). *Med Biol Eng Comput*, 48(2), 197–201.
- [7] Ford, M. D., et al. (2008). *J Biomech Eng*, 130(2), 21015.
- [8] Bouillot, P., et al. (2018) *Magnetic Resonance in Medicine*, 88(1), 560–572.
- [9] Plewes D., et al. (1995). *Med Phys* 22(1)1111–1115.

A novel microfluidic device for real-time microscopic imaging of endothelial cell responses to laminar and disturbed fluid flow

Daniel Lorusso, Kayla Soon, Hristo N. Nikolov, Jaques S. Milner, John R. de Bruyn,
J. Geoffrey Pickering, S. Jeffrey Dixon, David W. Holdsworth, Tamie L. Poepping.
Schulich School of Medicine & Dentistry, and Faculty of Science
University of Western Ontario

ABSTRACT

I. INTRODUCTION

Mechanotransduction is the process by which cells sense and respond to their local mechanical environment. Fluid shear stress is a form of mechanical stimulation that is critical to the normal functioning of vascular endothelium. Disturbed (non-laminar) fluid flow has been implicated in endothelial dysfunction, which contributes to atherosclerosis. Little is known about how early signals in cellular mechanotransduction contribute to endothelial dysfunction. Our goal was to observe immediate responses of endothelial cells to disturbed flow, but the necessary tools to study cells in these conditions did not exist. In the present study, we developed and validated a novel microfluidic device for applying laminar and disturbed flow to live endothelial cell cultures, while ensuring compatibility with real-time optical microscopy and photometry. Using this device, we performed fluorescence imaging of changes in cytosolic calcium – a mediator involved in transducing mechanical signals in many cell types, including vascular endothelium.

II. METHODS

The microfluidic device was custom designed and fabricated from silicone polymer (PDMS). Medium introduced through the top inlet established laminar flow within the imaging chamber. Medium introduced through the two side inlets generated high-velocity jets that impinged on each other within the chamber to create disturbed flow. Fluid velocities and flow patterns were modeled with computational fluid dynamics (CFD) software and determined experimentally by imaging at high speed with a laser-illuminated micro-particle image velocimetry (micro-PIV) system. To investigate the calcium responses, human aortic endothelial cells were seeded into the device, loaded with the fluorescent calcium probe fura-2 and monitored by fluorescence ratio imaging.

III. RESULTS

Several microfluidic devices were fabricated and tested over a range of flow rates. Using CFD modeling and micro-PIV, velocity maps were generated for our device and used to calculate wall shear stress vectors for both laminar and disturbed flow patterns. A flow rate of 4 mL/min through the side inlets generated oscillatory multi-directional flow at shear stress magnitudes of 3 – 4 Pa. The oscillatory shear index was greater than 0.4 in the region where cell imaging was performed, indicating nearly equal fluctuations in two directions. In preliminary studies, cytosolic calcium transients in primary human aortic endothelial cells were observed in response to changes in laminar and disturbed flow. Cells loaded with fura-2 were imaged within the flow chamber. All cells were first subjected to an initial laminar flow stimulus, which induced a transient elevation of calcium in most cells. Subsequently, changes in laminar flow from 1 and 2 mL/min induced transient elevation of calcium. In contrast, a change from laminar to disturbed flow (2 and 3 mL/min) induced a more sustained elevation of calcium.

IV. CONCLUSIONS

We have developed and validated a microfluidic device capable of – for the first time – delivering clinically relevant laminar and multi-directional disturbed flow to live cells during real-time imaging. This device is compatible with a wide-range of microscopy techniques. Our preliminary studies suggest that endothelial cells may display distinct cytosolic calcium responses to changes in laminar and disturbed flow.

In vitro assesement of aerosol delivery via the Spiriva® Respimat® Soft Mist Inhaler using the ODAPT™ adapter with facemask

Rym Mehri, Abubakar Alatrash, Edgar A. Matida

Department of Mechanical & Aerospace Engineering
Carleton University
Ottawa, ON, Canada

Frank Fiorenza

Product Development
McArthur Medical Sales Inc.
Rockton, ON, Canada

ABSTRACT

Chronic obstructive pulmonary disease (COPD) is a common lung disease that affects the respiratory tract by obstructing the airflow from the lung, hence causing breathing difficulties. For patients suffering from COPD, aerosol drug therapy via inhalation is commonly used to facilitate breathing. Inhaled Pharmaceutical aerosols are delivered by means of different types of inhalers depending on the disease and drug used; metered dose inhalers, dry powder inhalers, nebulizers, and soft mist inhalers. Pressurized metered dose inhalers (pMDIs) have been extensively used in the past and remain one of the most used devices for pharmaceutical aerosols delivery [1]. These pMDIs rely on the use of a propellant to deliver the medication at high speeds (2.0-8.4 m/s, 10 cm away from the nozzle) in a short amount of time (0.15 to 0.36 sec) [2]. As a result, due to particle momentum and turbulent dispersion, a large amount of medication deposits in the mouth-throat region and enters the gastrointestinal tract, hence engendering possible side effects [3] and affecting drug deposition in the lungs. In fact, particle size distribution and aerosol velocity are the most important factors dictating lung depositions. Therefore, a new generation of soft mist inhalers emerged and are currently commercially available. The Spiriva® Respimat® Soft Mist Inhaler (SMI, Boehringer Ingelheim, Ingelheim, Germany) is capable of delivering a metered dose (10-15 µL) of tiotropium bromide monohydrate per puff (2.5 µg of tiotropium per actuation) in the form of aerosols, used in the treatment of COPD. The Respimat® Soft Mist™ Inhaler is a propellant free inhaler that generates and deliver pharmaceutical aerosols suitable for inhalation using mechanical power from a spring, as opposed to the liquid-gas propellant typically used in pMDIs. This inhaler in intended to deliver a slow (initial approximate droplet velocity of 10 m/s) and long lasting (approximately 1.5 s) aerosol mist within the inhalable range, in order to facilitate the coordination of actuation with inhalation for proper medication delivery. The inhaled drug efficiency is strongly influenced by the amount of inhaled drug, the aerosol characteristics (such as particle size distribution), drug deposition, the delivery method, the flow characteristics and the architecture of airways.

Proper inhaler technique is crucial for efficient medication delivery to the lungs. However, misuse of inhalers and poor technique is common among patients such as elderly patients and children. Moreover unconscious patients are not capable of using inhalers to deliver the dose of medication. For these patients, facemasks are required. However, the effect of the addition of a facemask and adapters (add-ons) on drug delivery using the Respimat® SMI is not known. Furthermore, higher average and peak inspiratory flow rates were found with patients misusing the inhaler and were shown to affect lung deposition[4]. Therefore, it is important to determine the effect of the inhalation flow rate on drug deposition using the add-ons. The ODAPT™ soft mist adapter (McArthur Medical Sales Inc., Rockton, ON) was designed to deliver inhaled aerosols via Respimat® SMIs to patients requiring a facemask or tracheostomy application and allows for the use of standard facemasks. In this study, the effect of on medication delivery due to the presence of the ODAPT™ soft mist adapter with a facemask was assessed. For this purpose, the particle size distribution and drug delivery of the Spiriva® Respimat® SMI was investigated *in vitro* under different humidity levels (40-50% and >90%) and different steady inhalation flow rates (28.3 and 60 L/min) with and without the ODAPT™ adapter and facemask, in order to investigate medication losses due to the addition of add-ons to the Respimat® SMI. The particle size distribution and medication delivery was assessed using an Anderson Cascade Impactor and a UV-visible spectrophotometer.

REFERENCES

- [1] W. H. Finlay, "Metered Dose Propellant inhalers," in *The Mechanics of Inhaled Pharmaceutical Aerosols*, Elsevier, 2001, pp. 175–220.
- [2] D. Hochrainer, H. Hölz, C. Kreher, L. Scaffidi, M. Spallek, and H. Wachtel, "Comparison of the Aerosol Velocity and Spray Duration of Respimat® Soft Mist™ Inhaler and Pressurized Metered Dose Inhalers," *J. Aerosol Med.*, vol. 18, no. 3, pp. 273–282, Sep. 2005.
- [3] D. E. Geller, "Clinical Side Effects during Aerosol Therapy: Cutaneous and Ocular Effects," *J. Aerosol Med.*, vol. 20, no. s1, pp. S100–S109, Jun. 2007.
- [4] P. Brand, B. Hederer, G. Austen, H. Dewberry, and T. Meyer, "Higher lung deposition with Respimat Soft Mist inhaler than HFA-MDI in COPD patients with poor technique.," *Int. J. Chron. Obstruct. Pulmon. Dis.*, vol. 3, no. 4, pp. 763–70, 2008.

A Novel Method to Improve Estimation of Tissue Displacement Field in Ultrasound Elastography

Niusha Kheirkhah
 Department of Biomedical
 Engineering,
 Western University
 London, ON, Canada
 nkheirkh@uwo.ca

Ali Sadeghi-Naini
 Department of Electrical
 Engineering and Computer Science,
 York University
 Toronto, ON, Canada
 asn@yorku.ca

Abbas Samani
 Department of Biomedical
 Engineering,
 Western University
 London, ON, Canada
 asamani@uwo.ca

ABSTRACT

Many types of cancer are associated with biological changes in tissue composition and micro-structure that result in tissue stiffening. Elastography, a non-invasive technique to image tissue stiffness, has emerged as a clinically viable technique for cancer diagnosis. In elastography where the stiffness to be mapped is characterized by the Young's modulus (YM) or shear modulus, accurate measurement of the tissue displacement field is required to obtain reliable stiffness images. In quasi-static ultrasound (US) elastography, the displacement field is estimated under a quasi-static mechanical stimulation applied by the ultrasound transducer. The resulting tissue displacement in both axial and lateral directions can be estimated using a pair of the US radiofrequency (RF) frames acquired before and after mechanical stimulation. Finally, the strain images are generated by applying spatial derivatives to the displacement fields. In case a uniform stress distribution can be assumed within the tissue, a region with higher stiffness experiences less strain compared to softer regions. Therefore, strain images can be used to identify a stiff lesion within the tissue. However, having a uniform stress distribution is not an accurate assumption in organs where the tissue is heterogeneous, and the geometry is complex. A more accurate elastography approach accounts for the tissue stress non-uniformity. Elastography image reconstruction algorithms have been developed that account for the stress non-uniformity. These algorithms involve solving an inverse problem, which is formulated based on the theory of elasticity equations, directly or indirectly. The data used in such image reconstruction algorithms are the measured displacement data. The indirect inversion approach can lead to an iterative inversion algorithm which is comprised of two following main steps:

- 1) Tissue mechanical stimulation and acquisition of resulting tissue displacement field
- 2) YM reconstruction using data inversion algorithm

The pre- and post-compression US RF data acquired for elastography is often contaminated with noise. In addition, the lateral resolution in US RF data is often lower than that in the axial direction. These factors frequently affect the accuracy of the displacement field estimation, especially in the lateral direction. In this study, we propose a novel method to reduce the noise effects in estimation of axial and lateral displacement fields by incorporating the physics of tissue deformation. Specifically, the proposed technique incorporates the tissue incompressibility as well as the compatibility equations in elasticity as new constraints into the mathematical formulation of the problem to estimate the displacement fields with higher accuracy. Tissue incompressibility is utilized since soft biological tissue mainly consist of water. In 2D, the incompressibility (1) and compatibility (2) equations are given below:

$$\epsilon_{xx} + \epsilon_{yy} = 0 \quad (1)$$

$$\frac{\partial^2 \epsilon_{xy}}{\partial x \partial y} = \frac{1}{2} \left(\frac{\partial^2 \epsilon_{xx}}{\partial y^2} + \frac{\partial^2 \epsilon_{yy}}{\partial x^2} \right) \quad (2)$$

In the above incompressibility equation, a plain strain situation is assumed where out of plane strain is zero. The characteristics of noise tend to be different in different directions. This technique was implemented numerically while considering the accuracy of data in different directions. The technique was evaluated via phantom studies, including *in silico* and tissue-mimicking phantom studies where in the latter ultrasound elastography was utilized. The highly encouraging results, serving as a strong proof of concept. In the tissue mimicking phantom study, the displacement acquired experimentally were processed using the proposed framework. Results indicated at least 20% improvement in the accuracy of stiffness parameter reconstruction. In conclusion, the results are very encouraging. Further clinical studies are required to demonstrate its effectiveness in the clinic.

Development of a Disposable Single-Nozzle Multi-material 3D Bioprinting System

Tiffany Cameron, Ben MacCallum, Emad Naseri, Ali Ahmadi
Faculty of Sustainable Design
University of Prince Edward Island
Charlottetown, PEI, Canada
aahmadi@upei.ca

ABSTRACT

Generic prescriptions of medicine that do not consider an individual patient's needs can lead to either over- or under-prescription [1, 2], which may ultimately result in a decreased quality of life for the patient and a greater burden on the healthcare system. In recent years, an alternative method to current generic prescription-based treatment to allocate medicine more effectively has emerged. The new method is based on using the genetic profile of individuals [2, 3] to create personalized medicine, such as complex custom-made drug capsule tablets (also known as caplets) [4]. The drug caplets enable drug release profiles that are tailored to the patients' individual physiological needs based on their genomes. Current standard method of creating drug caplets are limited in their ability to create complex geometries and the desired release profile [4]. To address this issue, three-dimensional (3D) printers can be used to create caplets based on a 3D sketch in a computer software. In recent years, 3D printing technologies have been constructed to allow for multi-material printing, using multiple nozzles [5], a custom-designed single nozzle [6], as well as a capillary-bundled single nozzle [7]. However, the current multi-material 3D printers are built for limited materials and do not offer drug incorporation capabilities. Therefore, there is a need for the development of versatile 3D printing systems that are capable of printing multi-material drug releasing caplets.

In this work, a 3D printing system is developed that uses the functionality of a disposable, low-cost (~\$10) multi-fuse IV extension set as the printhead. This printhead is of luer-lock compatibility, and can be connected to multiple material-filled syringe cartridges, where the material can be dispensed using controlled air pressure. For the purpose of this study, chitosan is used based on its biocompatibility, high printability, and biodegradability characteristics [8]. The printing characteristics of the printhead are investigated by altering the applied pressure, the outlet needle size, as well as the material concentration. Particularly the transition length between two materials is characterized using a MATLAB image processing tool.

Having a 3D printing system that is able to operate with multiple materials can lead to the production of drug caplets with a controlled drug release profiles. This would enable healthcare professionals to be able to create drug caplets based on individual patients' physiological needs and genetic information. This will ultimately lead to a more efficient and cost-effective healthcare system.

References

- [1] A. Trasta, “Personalized medicine and proper dosage,” *EMBO reports*, 2018.
- [2] S. Mathur and J. Sutton, “Personalized medicine could transform healthcare,” *Biomedical Reports*, Feb. 2017.
- [3] D. E. Pritchard, F. Moeckel, M. S. Villa, L. T. Housman, C. A. Mccarty, and H. L. Mcleod, “Strategies for integrating personalized medicine into healthcare practice,” *Personalized Medicine*, vol. 14, no. 2, pp. 141–152, 2017.
- [4] W.-K. Hsiao, B. Lorber, H. Reitsamer, and J. Khinast, “3D printing of oral drugs: a new reality or hype?,” *Expert Opinion on Drug Delivery*, vol. 15, no. 1, pp. 1–4, 2017.
- [5] W. Lee, V. Lee, S. Polio, P. Keegan, J.-H. Lee, K. Fischer, J.-K. Park, and S.-S. Yoo, “On-demand three-dimensional freeform fabrication of multi-layered hydrogel scaffold with fluidic channels,” *Biotechnology and Bioengineering*, 2010.
- [6] J. O. Hardin, T. J. Ober, A. D. Valentine, and J. A. Lewis, “3D Printing: Microfluidic Printheads for Multimaterial 3D Printing of Viscoelastic Inks,” *Advanced Materials*, vol. 27, no. 21, pp. 3278–3278, 2015.
- [7] W. Liu, Y. S. Zhang, M. A. Heinrich, F. D. Ferrari, H. L. Jang, S. M. Bakht, M. M. Alvarez, J. Yang, Y.-C. Li, G. T.-D. Santiago, A. K. Miri, K. Zhu, P. Khoshakhlagh, G. Prakash, H. Cheng, X. Guan, Z. Zhong, J. Ju, G. H. Zhu, X. Jin, S. R. Shin, M. R. Dokmeci, and A. Khademhosseini, “Bioprinting: Rapid Continuous Multimaterial Extrusion Bioprinting,” *Advanced Materials*, vol. 29, no. 3, 2017.
- [8] Q. Wu, D. Therriault, and M.-C. Heuzey, “Processing and Properties of Chitosan Inks for 3D Printing of Hydrogel Microstructures,” *ACS Biomaterials Science & Engineering*, 2018.

The Effect of Stem Fit on the Radiocapitellar Contact Mechanics of a Metallic Axisymmetric Radial Head Hemiarthroplasty: Is Loose Fit better than Rigidly Fixed

Szmit, J.

Mechanical and Materials Engineering
Western University
London, Canada
jszmit@uwo.ca

King, G. J. W., Johnson, J.A., Langohr, G.D.G.

Hand and Upper Limb Centre
Western University
London, Canada

Abstract— Radial head hemiarthroplasty is commonly employed to manage comminuted displaced fractures. With regards to implant fixation, current designs vary with some prostheses aiming to achieve a tight 'fixed' fit, and others utilizing a smooth stem with an over reamed 'loose' fit. The purpose of the present study was to evaluate the effect of radial head hemiarthroplasty stem fit on radiocapitellar contact using a finite element model which simulated both fixed (size-for-size) and loose (1, 2 & 3mm over reamed) stem fits. It was hypothesized that a loose stem fit would improve radiocapitellar contact mechanics, with increased contact area and decreased contact stress, by allowing the implant to find its 'optimal' position with respect to the capitellum. Finite element models of the elbow were produced to compare the effects of stem fit on radiocapitellar contact of a metallic axisymmetric radial head implant. Radiocapitellar contact mechanics (contact area and maximum contact stress) were computed for 0°, 45°, 90°, and 135° elbow flexion with the forearm in neutral, pronation and supination. This data suggests that the loose smooth stem radial head implant may be functioning like a bipolar implant in optimizing radiocapitellar contact. Over reaming of 3mm produced a larger amount of stress concentrations on the capitellum suggesting there may be a limit to how loose a smooth stem implant should be implanted. The 'loose' over reamed stem provided optimal contact mechanics of the metallic axisymmetric radial head implant compared to the 'fixed' stem.

Keywords; *Radial Head, Hemiarthroplasty, Elbow, Implant Fixation, Radiocapitellar Joint Replacement*

I. INTRODUCTION

Radial head (RH) hemiarthroplasty is commonly employed to manage comminuted displaced, unreconstructable RH fractures, during which a prosthesis is substituted for the excised RH in an attempt to restore elbow stability and load transfer[1]–[3]. Compared to RH excision, RH replacement reduces the forces on the remaining ulnohumeral joint, potentially lowering the risk of arthritis, which is commonly reported after RH excision. RH replacement also serves to improve elbow and forearm stability in patients with

concomitant ligament injuries, which occur in up to 75% of patients with comminuted RH fractures[2], [4], [5].

In spite of advances in RH prosthesis design, the clinical outcomes are variable and joint contact mechanics remain suboptimal[6]–[8]. The substitution of the cartilage surface of the native RH with a metallic implant with a stiffness multiple orders of magnitude greater than that of the native cartilage may lead to cartilage degeneration and arthritis[1], [9]–[11]. Minimizing peak contact stress and maximizing contact area following RH hemiarthroplasty may well prevent the degradation of the native cartilaginous surface and improve the long term outcome of these devices [12].

Currently, there is no consensus as to the optimal fixation technique of RH prostheses in the proximal radius. Implant designs include textured ingrowth stems with a tight 'fixed' fit, cemented stems, and smooth stems with either a 'size-for-size' or an over reamed 'loose' fit. The effect of stem fit on radiocapitellar contact mechanics has not been previously reported. It has been suggested that a RH prosthesis with a smooth loose stem may function somewhat like a bipolar implant allowing the concave articular surface of the implant to adjust to the relatively spherical capitellum to improve radiocapitellar contact, however this has not been confirmed.

In view of the foregoing, the current study quantified the effect of RH hemiarthroplasty stem fit using a computational model for a series of cadaveric specimens following virtual RH replacement. We hypothesized that a loose stem fit would be advantageous in terms of in radiocapitellar contact mechanics, with a reduction in radiocapitellar joint incongruency, increased contact area and decreased contact stress, by virtue of the ability of the loose implant to find its 'optimal' position against the native capitellum.

II. MATERIALS AND METHODS

A. Specimen Model Generation

A series of three-dimensional elbow models were created with the use of computed tomography (CT) data from 10 fresh-frozen cadaveric elbows (average age: 66.1±18.5yrs, 6M/4F). A GE Discovery CT750 HD scanner (GE Healthcare, Pewaukee, WI, USA) was utilized to obtain CT images at 120kV and 200mA with a slice thickness of 0.625mm (voxel

dimensions 0.624x0.180-0.229x0.190-0.229mm in the axial coronal and sagittal directions respectively) [13]. The native elbows were CT scanned intact, then disarticulated and rescanned in air to establish accurate cartilage geometry for all specimens [13]. Cartilage was modelled as a nonlinear elastic material, with bone material properties assigned based on CT attenuation.

Segmentation from CT data was conducted utilizing Mimics 14.12 (Materialise, Leuven, Belgium) to acquire bone geometry for each specimen using a minimum segmentation threshold for specimen bone and cartilage of +250 and -500 Hounsfield units (HU) respectively. Bone and cartilage geometries were wrapped and remeshed using a radial basis function (FastRBF, FarField Technology, NZ). Subsequently, a quadrilateral mesh was generated on both the bone and cartilage surfaces (NETGEN, RWTH, Germany) applying a maximum element edge length of 0.45mm, previously justified by a mesh convergence study [14]. Using a previously described method, custom code was used to mesh the cartilage using a hexahedral mesh [13]. As reported by Schenck et al., based on the average elbow cartilage aggregate modulus of 0.8 MPa and Poisson's ratio of 0.07 a bulk modulus of 0.31 MPa and shear modulus of 0.37 MPa were assigned [15]. In order to mesh the bone geometry, 1mm linear tetrahedral elements were used with assignment material properties utilizing a Poisson's ratio of 0.3 and a modulus-density relationship ($E = 8345\rho_{app}^{1.5}$) [16]. Cobalt chrome material properties ($E = 230$ GPa, $\nu=0.3$) were assigned to the RH prosthesis following meshing using 0.25mm linear tetrahedral elements.

B. Finite Element Modeling

Finite element analysis (FEA) was performed utilizing ABAQUS v6.14 (Simulia Corp, Providence, RI, USA). A tie constraint was used to bond the backside of the cartilage and subchondral bone of the dissimilar meshes of the native humerus. Using a rigid body tie, the proximal humerus was connected to an individual guiding node to which boundary conditions were applied constraining the humerus in all degrees of freedom. A commonly used axisymmetric implant shape was chosen for this study (Evolve, Wright Medical, Memphis TN), with implant sizing for each specimen being conducted by an upper extremity fellowship trained orthopedic surgeon. The RH implant fixation was varied from fixed with respect to the radius which disallowed motion, or a loose fit with a 1mm, 2mm, or 3mm over reamed (diametral clearance) stem fit. The 1mm, 2mm and 3mm over reamed stem fits corresponded to a maximum allowable angulation of 2.3°, 4.6° and 6.9° based on a 25mm stem length, which is standard on the implant studied (Figure 1) [17]. The distal end of the ulna was then constrained in all degrees of freedom, after which a 100N compressive load was applied to the radius, which articulated it with the humerus and ulna [18], [19]. For the fixed stem fit, the RH implant was not permitted to rotate out of plane, however for each of the loose stem fits, the implant was permitted to rotate out of plane up to the maximum angle specified above for each fit. Surface-to-surface discretization was defined between the RH implant and capitellar surface to establish penalty-based contact. A non-linear modelling approach was employed whereby

assumptions regarding elements, loads, boundary conditions and material stiffness were considered. As reported by Willing et al., this modeling technique was found to yield contact areas within 10% of those acquired experimentally using a cadaveric elbow and a joint casting method [13].

C. Study Parameters & Statistical Analysis

Comparison between fixed, 1mm, 2mm and 3mm RH implant stem fits were made regarding changes in radiocapitellar contact area and maximum contact stress relative to the native intact state (Figure 1). Radiocapitellar contact mechanics (contact stress and contact area) were computed for 0°, 45°, 90°, and 135° elbow flexion with the forearm in neutral rotation. For each flexion/extension angle and implant fixation (fixed, 1mm, 2mm and 3mm) the outcome variables were normalized to the corresponding intact state, with comparison done using a two way (flexion angle, stem fit) repeated measures ANOVA with a level of significance defined as $P<0.05$.

III. RESULTS

As stem fit was loosened from fixed, the RH was observed to angulate in the varus and valgus directions dependent on flexion angle. At 0° of flexion as the stem fit was loosened the RH implant tilted into valgus driven by contact in the trochlear groove, visible as the grey region of high contact stress (Figure 1). At this flexion angle the 3mm loose fit generated a new region of high contact stress at the center of the articular contact, as visible as the grey region near the center of the capitellar surface. A similar result was observed at 45° and 135° of flexion where increased loose stem fit moved the contact from the trochlear groove laterally into the center of the articulation, or on the most lateral edge of contact. Finally, at 90° the RH tilted into varus once loose fit reached 3mm, resulting in regions of elevated contact stress in the trochlear groove (Figure 1).

For all elbows and flexion angles investigated, the fixed stem fit radial head generated a mean contact area of 165.17 ± 28.45 mm². As stem fit was progressively loosened, contact area successively decreased for the 1mm, 2mm and 3mm stem fits producing mean contact areas of 161.81 ± 30.11 mm², 152.83 ± 26.57 mm² and 139.90 ± 24.53 mm² respectively for all angles of flexion investigated. Although there was no significant difference detected between the fixed and 1 mm stem fits ($p\geq0.181$), both the 2mm and 3mm over reamed fits decreased contact area, on average by 12.34 ± 6.17 mm² ($p<0.001$) and 25.27 ± 12.637 mm² ($p<0.001$), respectively, relative to the fixed stem fit (Figure 2). Additionally, increasing the looseness of the stem fit from 1mm to 2mm and 3mm decreased contact area by 8.98 ± 4.49 mm² ($p<0.001$) and 21.91 ± 10.95 mm² ($p<0.001$), respectively.

The fixed stem fit produced higher maximum radiocapitellar articular contact stress compared to the 1mm, 2mm and 3mm over reamed stem fits for all angles of flexion with peak contact stress of 4.10 ± 3.74 MPa for the fixed and 3.02 ± 2.35 MPa, 2.82 ± 2.17 MPa and 3.00 ± 2.15 MPa respectively for the 1mm, 2mm and 3mm fits (Figure 2). On

average, loosening the stem fit from fixed to 1mm, 2mm, and 3mm over reamed decreased maximum contact stress by $1.09 \pm 2.74 \text{ MPa}$ ($p < 0.017$), $1.28 \pm 3.12 \text{ MPa}$ ($p < 0.013$), and $1.10 \pm 3.15 \text{ MPa}$ ($p < 0.033$), respectively. There was no significant difference detected between the 1 mm and 2 mm stem fits ($p \geq 0.41$), or the 1mm and 3mm stem fits ($p \geq 0.94$). However, when the stem fit was increased from 2mm to 3mm, there was a small but significant increase in maximum contact stress of $0.21 \pm 0.015 \text{ MPa}$ ($p < 0.028$, Figure 2). Also, for each of the 'loose' stem fits, for all flexion angles, the out of plane rotation was equal to the maximum permitted value.

IV. DISCUSSION

A better understanding of the effects of RH implant stem fit on radiocapitellar contact mechanics, with specific emphasis on over reamed loose fit fixation, may lead to improved clinical technique, implant designs, and clinical outcomes. Interestingly, as stem fit was loosened, the total articular contact area decreased, which was contrary to our hypothesis. Although the contact area decreased, the corresponding maximum contact stress for the loose stem implants decreased because of a change in RH angulation. In general, a decrease in contact area is associated with increase in contact stress, however in this case the angulation of the implant was thought to be largely driven by impingement of the radial head in the trochlear groove. The resulting RH angulation of the loose stem design alleviated these impingements, but this came at the cost of reduced contact area; particularly with implants with greater stem clearance.

The results of the current study show that for the axisymmetric metallic RH implant investigated, a fixed stem fit produced higher maximum articular contact stresses than all loose stem fits tested (1, 2, 3mm). This higher maximum contact stress produced by the fixed stem fit is thought to be the result of stress concentrations at local incongruities present between the RH implant and the native capitellum during loading due to the implants inability to angulate to alleviate these regions of high contact stress. This was particularly apparent in the trochlear groove where at 0° , 45° and 135° of flexion the loose fit RH implants angulated in the valgus direction in response to impingement. However, at the largest loose fit clearance the implant appeared to have rotated so far in valgus that the contact stress in the center of the capitellum was elevated greater the smaller clearance fits, particularly at 0° flexion. At 90° of flexion, the opposite was observed whereby the implant rotated in the varus direction reducing maximum contact stress for the 1 and 2mm fits, however the 3mm fit permitted too much varus rotation resulting in impingement in the trochlear groove.

Interestingly, for all loose stem fits investigated, under the 100 N applied loading all the RH implants assumed a final angular position that was at, or very close to, the maximum angulation permitted by the respective loose fit (1mm: 2.3 deg ; 2mm: 4.6 deg ; 3mm: 6.9 deg). This is hypothesized to be the reason why the 3mm over reamed stem fit had higher maximum contact stress than both the 1mm and 2mm stem

fits. The extra angulation afforded by over sizing the radial canal permitted too much freedom of motion allowing the implant to rotate into a suboptimal orientation under the applied forearm load. These data seem to indicate that there was an optimal range for loose stem fits; excessive stem over reaming should likely be avoided with this smooth stem implant design.

Radiocapitellar articular contact area was similar between the fixed stem fit and the 1 mm over reamed stem, however decreased with the 2 and 3 mm over reamed fits. The contact area shifted from the lateral portion of the capitellar surface, to a more even distribution with increased contact on the medial portion of the capitellum (Figure 1). Although a decrease in contact area was seen with the 2 and 3mm stem fits, providing a 1 or 2 mm loose stem fit produced a decrease in maximum articular contact stress and reduced local regions of stress concentrations relative to the fixed stem fit. This data suggests that a compromise between contact area and maximum contact stress may be needed when deciding how much over reaming is optimal.

The current study shows that loose stem fits in RH hemiarthroplasty have the potential to reduce maximum articular contact stress for all flexion angles investigated. Comparing all loose stem fits, the 2mm over reamed fit provided the greatest decrease in maximum contact stress, although this was accompanied by a decrease in radiocapitellar contact area. While a reduction in radiocapitellar articular contact area may be a negative consequence, the decrease in maximum contact stress associated with the loose stem fits likely outweighs the reduction in contact area, as preventing local stress concentrations may help preserve native cartilage and improve long term patient outcomes.

Radiographic lucencies are known to commonly occur around smooth stemmed RH arthroplasty implants and may reflect adjustment of the implant to better track with the capitellum [19]. This adjustment of implant position may be directed by the annular ligament or from the interaction of the concave articular surface of the RH with the relatively spherical capitellum. Progression of lucencies around smooth stemmed implants is not typically seen over time and is not correlated with patient symptoms [3] [18]. This data suggests that a loose smooth stem RH implant may in fact be functioning in some respects as a bipolar implant, permitting angulation and rotation of the implant relative to the radius, which may improve radiocapitellar contact.

The strengths of this study include the repeated assessment of each individual specimen and the associated response to each stem fit investigated compared to the native state. Accurate bone and cartilage geometry was also applied using repeated CT data acquisition both intact and in air for the study population of 10 elbows. Region specific bone material properties were also simulated based on CT attenuation.

Limitations of this work include the assumptions associated with finite element analysis including the material models applied to the bone and cartilage, as well as the contact algorithms selected in the modelling process. Although these errors were present in all models, and since all outcome variables are compared to the intact state, we are confident in the comparisons presented in this work. Furthermore, only one smooth stem metallic axisymmetric implant design was studied; further work should extend this to other implant shapes and materials.

V. CONCLUSIONS

The 'loose' over reamed stem provided improved contact mechanics compared to the 'fixed' stem. The 1mm over reamed stem reduced maximum contact stress without significantly affecting contact area. Over reaming of 2mm provided the greatest decrease in maximum contact stress, albeit with a significant reduction in contact area. Over reaming of 3mm produced a larger amount of stress concentrations on the capitellum suggesting there may be a limit to how loose a smooth stem implant should be implanted. A fixed axisymmetric metallic RH implant was shown to be less effective in distributing these regions of high contact stress, and resultantly had the highest maximal cartilage stress. Over reaming a smooth stem axisymmetric RH implant by 1 to 2 mm provided improved elbow contact mechanics and may help to preserve the remaining native capitellar cartilage.

REFERENCES

- [1] J. G. Moon, L. J. Berglund, D. Zachary, K. N. An, and S. W. O'Driscoll, "Radiocapitellar joint stability with bipolar versus monopolar radial head prostheses," *J. Shoulder Elb. Surg.*, vol. 18, no. 5, pp. 779–784, 2009.
- [2] C. VanBeek and W. N. Levine, "Radial Head-Resect, Fix, or Replace," *Oper. Tech. Orthop.*, vol. 20, no. 1, pp. 2–10, 2010.
- [3] I. J. Harrington, a Sekyi-Otu, T. W. Barrington, D. C. Evans, and V. Tuli, "The functional outcome with metallic radial head implants in the treatment of unstable elbow fractures: a long-term review," *J. Trauma*, vol. 50, no. 1, pp. 46–52, 2001.
- [4] A. G. Schneeberger, G. J. W. King, S. W. Song, S. W. O'Driscoll, B. F. Morrey, and K. N. An, "Kinematics and laxity of the Souter-Strathclyde total elbow prosthesis," *J. Shoulder Elb. Surg.*, vol. 9, no. 2, pp. 127–134, 2000.
- [5] R. Grewal, J. C. MacDermid, K. J. Faber, D. S. Drosdowech, and G. J. W. King, "Comminuted radial head fractures treated with a modular metallic radial head arthroplasty. Study of outcomes," *J. Bone Joint Surg. Am.*, vol. 88, no. 10, pp. 2192–2200, 2006.
- [6] C. Chanlalit, D. R. Shukla, J. S. Fitzsimmons, K. N. An, and S. W. O'Driscoll, "Influence of prosthetic design on radiocapitellar concavity-compression stability," *J. Shoulder Elb. Surg.*, vol. 20, no. 6, pp. 885–890, 2011.
- [7] T. Flinkkila, T. Kaisto, K. Sirnio, P. Hyvonen, and J. Leppilahti, "Short- to mid-term results of metallic press-fit radial head arthroplasty in unstable injuries of the elbow," *Bone Joint J.*, vol. 94–B, no. 6, pp. 805–810, 2012.
- [8] G. J. W. King, Z. D. S. Zarzour, S. D. Patterson, and J. A. Johnson, "An anthropometric study of the radial head: Implications in the design of a prosthesis," *J. Arthroplasty*, vol. 16, no. 1, pp. 112–116, 2001.
- [9] V. S. Liew, I. C. Cooper, L. M. Ferreira, J. A. Johnson, and G. J. W. King, "The effect of metallic radial head arthroplasty on radiocapitellar joint contact area," *Clin. Biomech.*, vol. 18, no. 2, pp. 115–118, 2003.
- [10] K. J. Burkhart *et al.*, "Mid- to long-term results after bipolar radial head arthroplasty," *J. Shoulder Elb. Surg.*, vol. 19, no. 7, pp. 965–972, 2010.
- [11] N. Popovic, "Midterm Results with a Bipolar Radial Head Prosthesis: Radiographic Evidence of Loosening at the Bone-Cement Interface," *J. Bone Jt. Surg.*, vol. 89, no. 11, p. 2469, 2007.
- [12] S. E. Irish, G. D. G. Langohr, R. Willing, G. J. King, and J. a. Johnson, "Implications of Radial Head Hemiarthroplasty Dish Depth on Radiocapitellar Contact Mechanics," *J. Hand Surg. Am.*, vol. 40, no. 4, pp. 723–729, 2015.
- [13] R. T. Willing, E. A. Lalone, H. Shannon, J. A. Johnson, and G. J. W. King, "Validation of a finite element model of the human elbow for determining cartilage contact mechanics," *J. Biomech.*, vol. 46, no. 10, pp. 1767–1771, 2013.
- [14] G. D. G. Langohr, R. Willing, J. B. Medley, G. J. W. King, and J. a. Johnson, "Contact analysis of the native radiocapitellar joint compared with axisymmetric and nonaxisymmetric radial head hemiarthroplasty," *J. Shoulder Elb. Surg.*, vol. 24, no. 5, pp. 787–795, 2015.
- [15] G. E. Schenck RC Jr, Athanasiou KA, Constantinides G, "A biomechanical analysis of articular cartilage of the human elbow and a potential relationship to osteochondritis dissecans," *Clin Orthop Relat Res*, vol. 299, pp. 305–312, 1994.
- [16] R. L. Austman, J. S. Milner, D. W. Holdsworth, and C. E. Dunning, "The effect of the density-modulus relationship selected to apply material properties in a finite element model of long bone," *J. Biomech.*, vol. 41, no. 15, pp. 3171–3176, 2008.
- [17] S. Technique, "Explor ® Modular radial hEad Explor ® Modular radial hEad."
- [18] D. Sahu *et al.*, "Influence of radial head prosthetic design on radiocapitellar joint contact mechanics," *J. Shoulder Elb. Surg.*, vol. 23, no. 4, pp. 456–462, 2014.
- [19] H. L. Shannon, S. R. Deluce, E. A. Lalone, R. Willing, G. J. W. King, and J. A. Johnson, "Effect of radial head implant shape on joint contact area and location during static loading," *J. Hand Surg. Am.*, vol. 40, no. 4, pp. 716–722, 2015.
- [20] J. N. Doornberg, "Radial Head Arthroplasty with a

Modular Metal Spacer to Treat Acute Traumatic Elbow Instability,” *J. Bone Jt. Surg.*, vol. 89, no. 5, p. 1075, 2007.

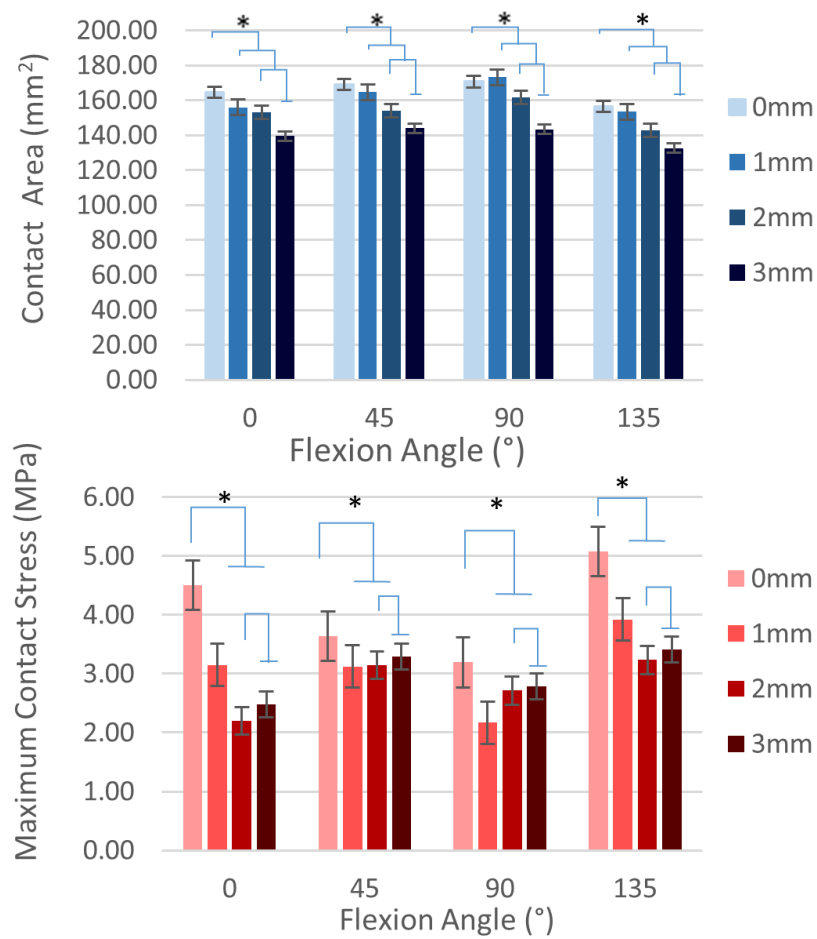


Figure 1. Capiterllar contact stress maps and implant fit for 0mm (fixed), and 1, 2 and 3mm (loose, left to right) over reamed stem fixations

Design of a Custom Bioreactor for the Dynamic Culture of Perfused Cell-Laden Alginate Scaffolds Obtained by 3D Printing

Simon Collin^a, André Bégin-Drolet^a, Julie Fradette^{b,c}, Cindy Jean Hayward^b, Jean Ruel^{a,*}

^a:Département de génie mécanique
Université Laval

Quebec city G1V 0A6, Canada

*Corresponding author. E-mail address:
jean.ruel@gmc.ulaval.ca (J. Ruel)

^b: Centre de recherche en organogénèse
expérimentale/LOEX,
CRCHU de Québec – Université Laval
Quebec city G1J 1Z4, Canada

^c: Département de chirurgie
Université Laval
Quebec city, G1V 0A6, Canada

Abstract—In this paper, a custom bioreactor designed to generate flow through culture chambers containing perfused three dimensional cell-laden alginate scaffolds is presented. This bioreactor was conceived for biological applications, meaning it is compact and robust enough to be placed in an incubator for several weeks and its components can be sterilized through various methods. The culture chambers are designed for a specific geometry of perfused alginate scaffolds, but can be easily customized for various geometries and biomaterials. The scaffold used with the bioreactor are molded using 3D printed rectangular sacrificial constructs containing several layers of alternating transversal and longitudinal self-supporting filaments. This means that the molded scaffolds contain internal channels, which explains its perfused characteristic. The role of the bioreactor is to provide nutrients and oxygen to the cells seeded in the scaffolds and to eliminate cellular waste. Appropriate flow through the perfused scaffolds is imperative to supply these nutrients, thus promoting cell growth and viability leading to functional biological tissues. Live/Dead staining of specimens cultured dynamically in these culture chambers showed excellent viability after six weeks.

Bioreactor; cell culture; design; hydrogel; alginate; 3D printing; vascularization; live/dead staining.

I. INTRODUCTION

Population aging is a growing concern in many countries. For example, in Canada, the median age has reached 40.6 years in 2018, a growth of 10 years since 1984 [1]. This situation is responsible for an escalation of demands in replacement organs, which is a great challenge for the health care system. For this reason, tissue engineering is rapidly positioning itself as one of the most popular research fields. Many approaches are being investigated to create artificial organs, and seeding alginate-based hydrogels with tissue-specific cells is a method that is favored by many researchers for the low cost, biocompatibility and ease of gelation of this hydrogel. With the rapid growth of 3D-printing technologies, innovative techniques are emerging for the fabrication of cell-laden

hydrogel scaffolds. A popular one is bioprinting, which consists of printing alternating layers of hydrogels, cells and extracellular matrix components in order to form a tissue-like construct. However, this technique presents many challenges, such as mechanical shear stress imposed to the cells during the printing process, long-term exposure to crosslinkers during storage and lack of perfusion in the tissues. [2] Indeed, adequate perfusion of thick tissues is primordial to achieve high rates of viability in three-dimensional cell-laden hydrogel scaffolds, because nutrients and oxygen can diffuse through the hydrogel for only a few hundred microns [3, 4]. If the tissues are exposed the media exclusively at their outer surfaces, cells in the core of the tissues have a lesser chance to survive. Another interesting technique, which is exploited in this paper, consists of 3D-printing sacrificial molds subsequently filled with cell-seeded alginate. This approach is interesting because it imposes less mechanical stress to the seeded cells and does not expose them to chemical crosslinkers. Also, molding offers many possibilities for the geometry of the tissues and the design of perfusable channels, especially when the molds are fabricated using self-supporting 3D-printed materials such as carbohydrate sugar. This biocompatible compound has the property to solidify at room temperature, which allows for very complex and self-supported structures to be printed, and can be dissolved using biocompatible solutions such as phosphate-buffered saline (PBS) [5, 6]. For the experiments presented in this paper, rectangular molds filled with internal alternating channels were printed, which allowed for the molded scaffolds to contain internal perfusable channels in which media was circulated, favoring cell viability at the core of the hydrogels. To ensure that media flows constantly and uniformly through the scaffolds, a hydraulic circuit was designed with careful precautions. The goal of this paper is to present a custom bioreactor that contains fully-biocompatible and sterilizable components which can be assembled aseptically and features multiple culture chambers for cell-laden scaffolds.

II. BIOREACTOR DESCRIPTION

The design of the bioreactor was achieved in two steps. First, a culture chamber that could accept a tissue with specific

dimensions, keep it in place and favor media flow through its channel network was designed. Then, the design of a complete bioreactor that supports multiple culture chambers and facilitates flow control was prepared so that parallel experiments could be performed.

A. Culture chambers

The main component of the bioreactor is the culture chamber in which the tissues are grown. The first design consideration for this important component was that it had to be made of an autoclavable and biocompatible material. For this reason, a known biocompatible polymer that can withstand temperature of up to 120°C, was chosen: polycarbonate [7, 8]. The second design consideration was that it had to be able to maintain the tissue in place while forcing media to flow through it. The approach that was employed consisted of including two separate sections in the culture chamber: an entry zone and an exit zone (X and Y sections of figure 1A), separated by a notch in which the cultured tissue is inserted (red section of figure 1A). The entry zone was designed as a pressurized tranquilization chamber producing an even flow of the media through all the micro-channels built in the biological tissue and then to the exit zone.

Additional design considerations were also taken into account. First, the fluid entry port had to be placed in a way that does not favor flow in a direction at the expense of the other. For this reason, it was placed in the bottom left corner of the entry zone (see X section in figure 1A). Other practical aspects were considered, leading to the following characteristics. For manufacturing reasons, the chamber is composed of two pieces: the top component (which includes the entry and exit ports and the walls separating the entry and exit zones) and the bottom component (which includes notches for the walls and for the tissue). The culture chamber is easy to connect to the rest of the bioreactor to facilitate aseptic assembly of the bioreactor. This is achieved with the use of stainless steel barbed-NPT connectors. A groove containing an O-ring is machined in the bottom part to ensure that the chamber is watertight. The components are assembled with stainless screws that press the O-ring to seal the culture chamber. Finally, as shown in figure 1B, positioning pins were added to the design to facilitate the assembly. This particular picture shows one of the culture chambers used in the biological experiment. The tissues cultured in it have dimensions of 10 x 10 x 20 mm (H x W x D) and contain 12 transversal as well as 16 longitudinal micro-channels (figure 1C).

B. Bioreactor

To ensure constant and unidirectional flow in the culture chambers, ergonomic media changes and global robustness, many considerations were taken into account. To improve the clarity of the description, figure 2 shows a schematic representation of the bioreactor with four culture chambers placed in parallel. First, an autoclavable check-valve was connected immediately before the entry port of each culture chamber to ensure unidirectional flow. Two components were set to minimize the time required for media changes and eliminate risks associated with contamination: interchangeable

media bottles connected directly to the bioreactor as well as three-way valves to redirect the flow away from the media bottles to facilitate media changes. Thus, to change media in the bioreactor, 1) a new bottle was filled with media, 2) flow was redirected away from the bottle via a bypass and 3) the old bottle was replaced with the new one. The media bottles caps contain three ports: inlet, outlet and filtered air exchange. The three-way valves were sterilized using peroxide gas. The bypass also contains a filtered air exchange port to facilitate initial filling and emptying of the culture chamber.

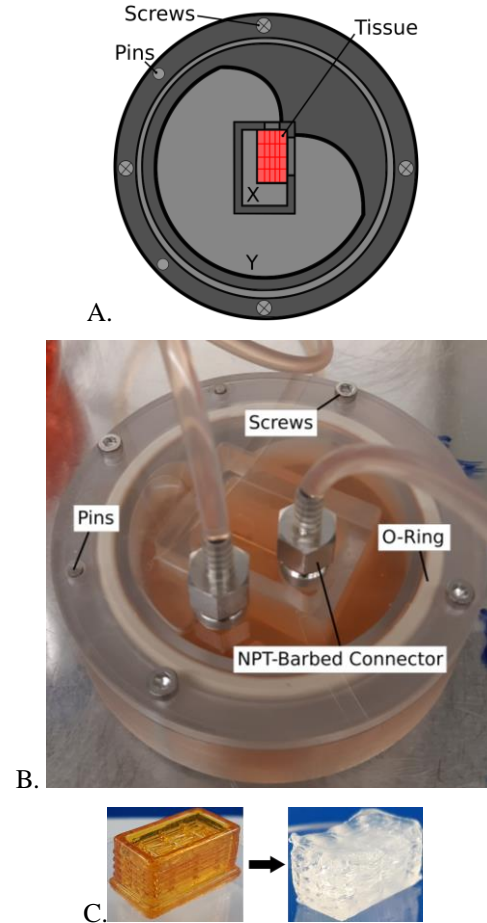


Figure 1: (A) Schematic representation of the culture chamber. Media enters through sub-chamber X and is partially blocked by the cell-seeded scaffold placed in the red section, which builds up pressure and facilitate the passage of media through the micro-channels of the scaffold into sub-chamber Y, where it can exit the culture chamber. The grey circle is where the O-ring was placed, and stainless screws placed outside the seal are used to fasten the top component of the culture chamber to the bottom. (B) Close-up view of a culture chamber. Media enters through the right connector into the entry zone, goes through the tissue, penetrates in the exit zone, and goes out of the culture chamber through the left connector. (C) Left: 3D-printed carbohydrate sugar sacrificial mold. Right: Molded cell-free alginate scaffold with the dimensions and geometry of the cell-seeded scaffolds cultured *in vitro*.

Another consideration was that the setup had to be compact enough to be inserted in a regular incubator, and could withstand a temperature of 37°C, humid conditions and CO₂ proportions of 8%. A peristaltic pump was used to control flow

through the culture chambers (ISMATEC BVK 16 channels). High-purity autoclavable silicone tubing was used to link each component, as well as ISMATEC's Pharmed® BPT tubing for the pump and autoclavable barbed plastic tube fittings. The type of pumps used in our experiments is particularly interesting for this application because it can simultaneously control flow in multiple hydraulic loops. For example, we were able to manage four culture chambers simultaneously in a single bioreactor, and as shown in figure 3, this setup is compact enough to be inserted in a conventional incubator.

III. EXPERIMENTAL VALIDATION

The purpose of this bioreactor was to provide adequate media flow through the perfused alginate scaffolds, in order to maintain cell viability. Thereby, two experiments were conducted to validate the achievement of our goal: a short experiment to observe the flow of a dye solution through a cell-free scaffold tissue placed in the culture chamber and a complete biological experiment during which cell-laden scaffolds were cultured to verify cell viability at different time points.

A. Flow experiment

Adequate flow through the scaffolds placed in the bioreactor's culture chambers was verified with a fairly simple experiment. A buffer solution was dyed green and a cell-free alginate scaffold with the exact dimensions and geometry of cell-seeded scaffolds was placed in the culture chamber. A peristaltic pump pushed the buffer in the entry zone of the culture chamber at the constant rate of 1.5 mL/min. A light source was placed under the culture chamber to enhance the contrast between the scaffold and the solution. Since the scaffolds are relatively transparent, the green-dyed solution was easily observable in the channels. Figure 4 presents the result of this experiment.

B. In vitro experiment

The ability of the bioreactor to maintain cellular viability for many weeks was verified by cultivating cell-laden alginate scaffolds. For our experiments, scaffolds were seeded with 3 million or 10 million human adipose-derived stromal-stem cells (ASCs). Dulbecco Modified Eagle Medium (DMEM) was mixed with fetal bovine serum (10%), penicillin/gentamicin (2%) and ascorbic acid (250 μ M). The culture chambers contained 30 mL of this media, while the rest of a single loop (including the media bottle) contained 80 mL, for a total of 110 mL. Media was changed every two or three days (80 mL out of 110 mL). Two complete bioreactors were used in two separate incubators, meaning that eight tissues were cultured (four with 3 million cells and four with 10 million cells). For each cell concentration, one sample was cultured for three weeks, two for six weeks and one for eight weeks. After completion of the culture period, the samples were sliced perpendicularly to the longitudinal micro-channels, at a thickness of 1-2mm, and stained using the *LIVE/DEAD® Viability/Cytotoxicity Kit*for mammalian cells* (by Molecular Probes, Inc.). This two-color fluorescent cell viability assay uses two reagents, Calcein AM (component A) and Ethidium homodimer-1 (component B), to be diluted in a buffer solution at different concentrations.

Component A is converted from a clear solution to an intensely fluorescent one by an enzymatic reaction caused by intercellular esterase activity in live cells, whereas component B can only enter the damaged cellular membrane of dead cells in which it binds to nucleic acids, enhancing its fluorescence [9]. Figure 5 shows the results of a 10 million cells-seeded scaffold after six weeks of dynamic culture. A slice with a thickness of 2 mm was stained for cell viability, and scanned over a depth of 800 μ m with a confocal microscope (LSM710, Zeiss) in an area near a longitudinal micro-channel (top left of figure 5).

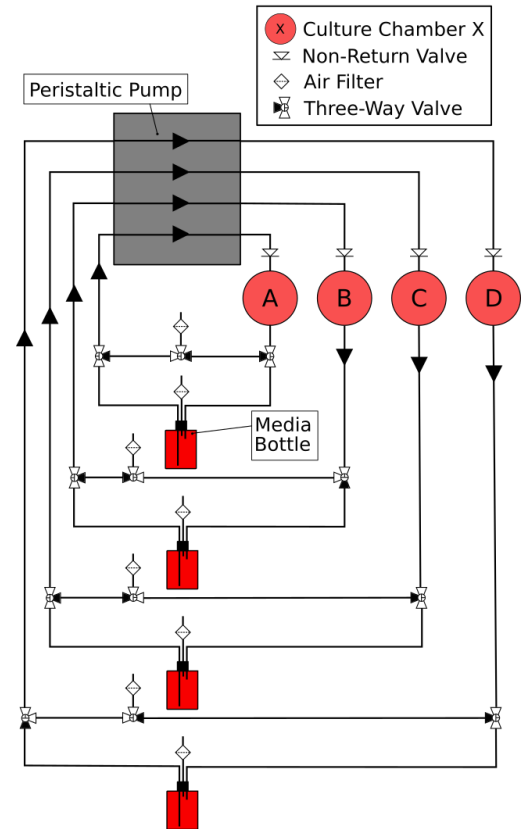


Figure 2: Schematic representation of the bioreactor with four culture chambers. The black part of the three-way valves is blocked. Loop with culture chamber A: Three-way valves are positioned for dynamic culture of the tissue in chamber A. Loop with culture chamber B: Three-way valves are positioned for early filling of culture chamber B. Only air should exit the culture chamber at this stage through the filter in the bypass. As soon as some media starts flowing out of the culture chamber, the three-way valves must be positioned as in A or C. Loop with culture chamber C: This setup is equivalent to Loop A. Loop with culture chamber D: Three-way valves are positioned to empty culture chamber D.

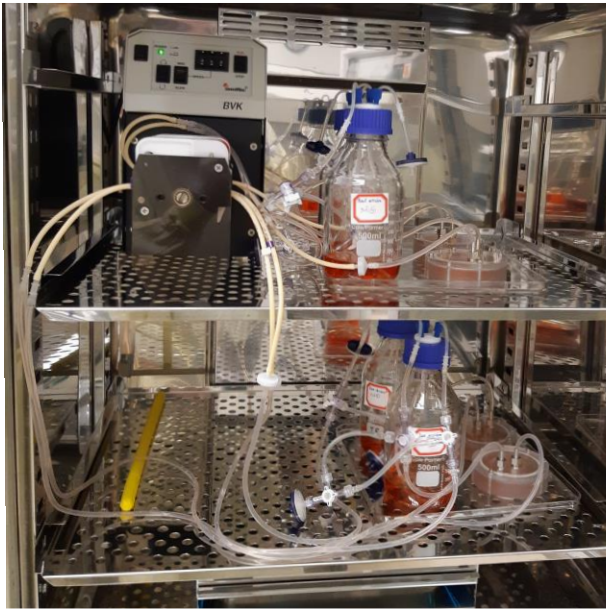


Figure 3: One of the two four-culture chambers bioreactors placed in a conventional incubator. Cell-seeded alginate scaffolds were cultured for three to eight weeks at 37°C, 8% CO₂ in this particular bioreactor. Media was changed every two or three days (80 mL out of 110 mL).

IV. DISCUSSION

The first thing that can be noticed from the flow experiment is the fact that, as seen on figure 4, the transversal micro-channels show darker shades of green than the longitudinal ones, meaning without a doubt that the flows goes more easily in the transversal direction. The first hypothesis comes from the fact that the longitudinal micro-channels are longer than the transversal ones. This means that the pressure level to overcome friction in the micro-channels is higher in the longitudinal direction than the transversal direction, which would clearly favor flow in the later direction. The second hypothesis is that the fluid point and angle of arrival with this particular arrangement may favor the flow in the transversal direction. These two hypotheses, which are not mutually exclusive, provide insight for future improvement of the culture chambers design. Moreover, one of the strength of this particular bioreactor comes from the fact that the culture chamber is an independent component of the system, meaning that it can be replaced or modified at ease without needing to replace the whole system. Thereby, a simple way to force a constant flow in both directions could be to include two entry ports and to separate entry chambers, i.e. one near the transversal channels entrance and one near the longitudinal channels entrance. Another improvement could be to change the micro-channels diameters and the number of micro-channels so that the pressure drop caused by friction is the same in both the longitudinal and transversal directions. The culture chambers used in the *in vitro* biological experiments used a similar design as the one used in the flow experiment, meaning that the flow might have been favored in the transversal direction during the dynamic culture of these tissues.

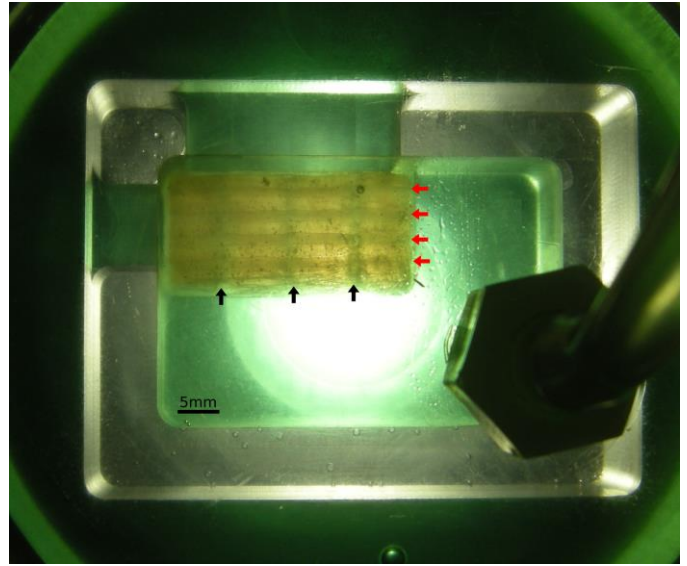


Figure 4: Flow experiment. A cell-free alginate scaffold was placed in a culture chamber that was linked to a peristaltic pump producing a green-dyed PBS flow through the culture chamber at a rate of 1.5mL/min. The green-colored micro-channels in the scaffolds (black arrows) show that the fluid is able to easily flow through the transversal micro-channels. However, it seems harder for the fluid to flow through the longitudinal micro-channels (red arrows), because the green color is less perceptible than through the transversal micro-channels. Scale bar: 5 mm.

The *in vitro* experiments showed consistent results with the flow experiment: cells were still viable after six weeks of dynamic culture in the bioreactor, but in lower proportions than expected. The preliminary results shown in Figure 5 show that a majority of the cells were viable after six weeks. Homogenous proportions of live and dead cells can be seen throughout the sections. However, cell viability nearing 100% would have been expected with an optimal system, which was not the case. This finding suggests that the flow may not be optimal in both direction, and that consequently not all the cells within the core of the tissue were adequately perfused. Furthermore, the homogeneity of the live/dead cells proportion is another unexpected finding. The expected results would have been higher proportions of live cells near the micro-channel (top-left corner of figure 5), and an increasing gradient of dead cells while going away from it, but it is not the case. The explanation for this could either be that the micro-channels are so close from each other that all areas are equally favored, or that the flow in the longitudinal direction is not efficient enough to create a noticeable difference in cell viability. The later hypothesis is more plausible, however, further investigation is required to answer this, i.e. testing different micro-channel arrangements and dimensions and optimizing the flow distribution in the scaffolds. Finally, the fact that no contaminations were observed after six weeks of culture should be mentioned. Such closed-systems cannot be observed by microscopy during culture, which often frightens researchers. The fact that we could culture our tissues for several weeks with no contamination proves the ergonomics of the system, which was one of the goals of our study.

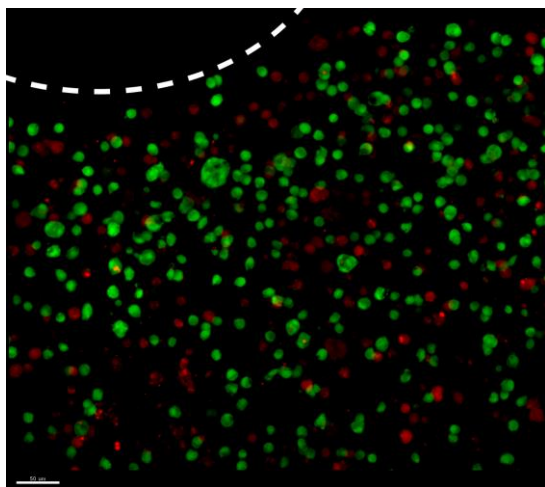


Figure 5: Live/dead staining results of a 10 million cells-seeded alginate scaffolds cultured for six weeks in a bioreactor. A sample with a thickness of 2 mm was stained and scanned over a depth of 800 μm . Live cells are green and dead cells are red. The white dashed line at the top left represents the approximate location of a longitudinal micro-channel. Scale bar: 80 μm .

V. CONCLUSION

This paper presents a custom bioreactor for the dynamic culture of cell-seeded alginate scaffolds as well as the results of a flow experiment and an *in vitro* experiment. The flow experiment showed that the fluid could flow through all the channels and irrigate the scaffolds with the current design, but that improvements could be made to reach an optimal balance of the flow distribution in the longitudinal and transversal directions. The *in vitro* experiments showed that cells could survive in scaffolds cultured dynamically for six weeks in the bioreactor with a majority of viable cells. It also showed that with the current design, constant proportions of live and dead cells are observed through the scaffold near a longitudinal micro-channel. Further investigation is required to determine if, indeed, media flow in the micro-channels improves cell viability in the three dimensional tissues. Another additional experiment that should be conducted is a comparative study of static and dynamic culture of cell-seeded alginate scaffolds.

ACKNOWLEDGMENT

The authors would like to acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) for its financial contribution, as well as the *Centre de recherche en organogénèse expérimentale/LOEX* which made the experimental validation process possible. The confocal imaging system was obtained through the *Fonds des leaders* program from Canada Foundation for Innovation (CFI) to J. Fradette. The authors acknowledge the support of the Quebec Cell, Tissue and Gene Therapy Network – ThéCell (a thematic network supported by the *Fonds de recherche du Québec-Santé*).

REFERENCES

- [1] Statistics Canada, Annual Demographic Estimates: Canada, Provinces and Territories, Catalogue no. 91-215-X, ISSN 1911-2408, p.7, January 2019.
- [2] E. Axpe, M. L. Oyen, "Applications of Alginate-Based Bioinks in 3D Bioprinting," in *International Journal of Molecular Sciences*, vol. 17, issue 12, MDPI, 2016.
- [3] L. G. Griffith, G. Naughton, "Tissue Engineering – Current Challenges and Expanding Opportunities," in *Science*, vol. 295, pp. 1009-1014, February 2002.
- [4] I. S. Kinstlinger, J. S. Miller, "3D-printed fluidic networks as vasculature for engineered tissue," in *Lab Chip*, vol. 16, pp.2025-2043, May 2016.
- [5] J. S. Miller, K. R. Stevens, M. T. Yang, B. M. Baker, D.-H. T. Nguyen, D. N. Cohen, E. Toro, A. A. Chen, P. A. Galie, X. Yu, R. Chaturvedi, S. N. Bhatia, C. S. Chen, "Rapid casting of patterned vascular networks for perfusable engineered three-dimensional tissues," in *Nature Materials*, vol. 11, pp.768-774, September 2012.
- [6] A. Bégin-Drolet, M.-A. Dussault, S. A. Fernandez, J. Larose-Dutil, R. L. Leask, C. A. Hoesli, J. Ruel, "Design of a 3D printer head for additive manufacturing of sugar glass for tissue engineering applications," *Additive Manufacturing*, Elsevier, vol. 15, pp.29-39, March 2017.
- [7] B. Altmann, S. Giselbrecht, K.-F. Weibezahn, A. Welle, E. Gottwald, "The three-dimensional cultivation of the carcinoma cell line HepG2 in a perfused chip system leads to a more differentiated phenotype of the cells compared to monolayer culture," in *Biomedical Materials*, IOP, vol. 3, 10pp, 2008.
- [8] P. M. van Midwoud, A. Janse, M. T. Merema, G. M. M. Groothuis, E. Verpoorte, "Comparison of Biocompatibility and Adsorption Properties of Different Plastics for Advanced Microfluidic and Tissue Culture Models," in *Analytical Chemistry*, ACS Publications, vol. 84, pp. 3938-3944, March 2012.
- [9] Molecular Probes™, LIVE/DEAD® Viability/Cytotoxicity Kit *for mammalian cells*, MP03224, Invitrogen Detection Technologies, December 2005.

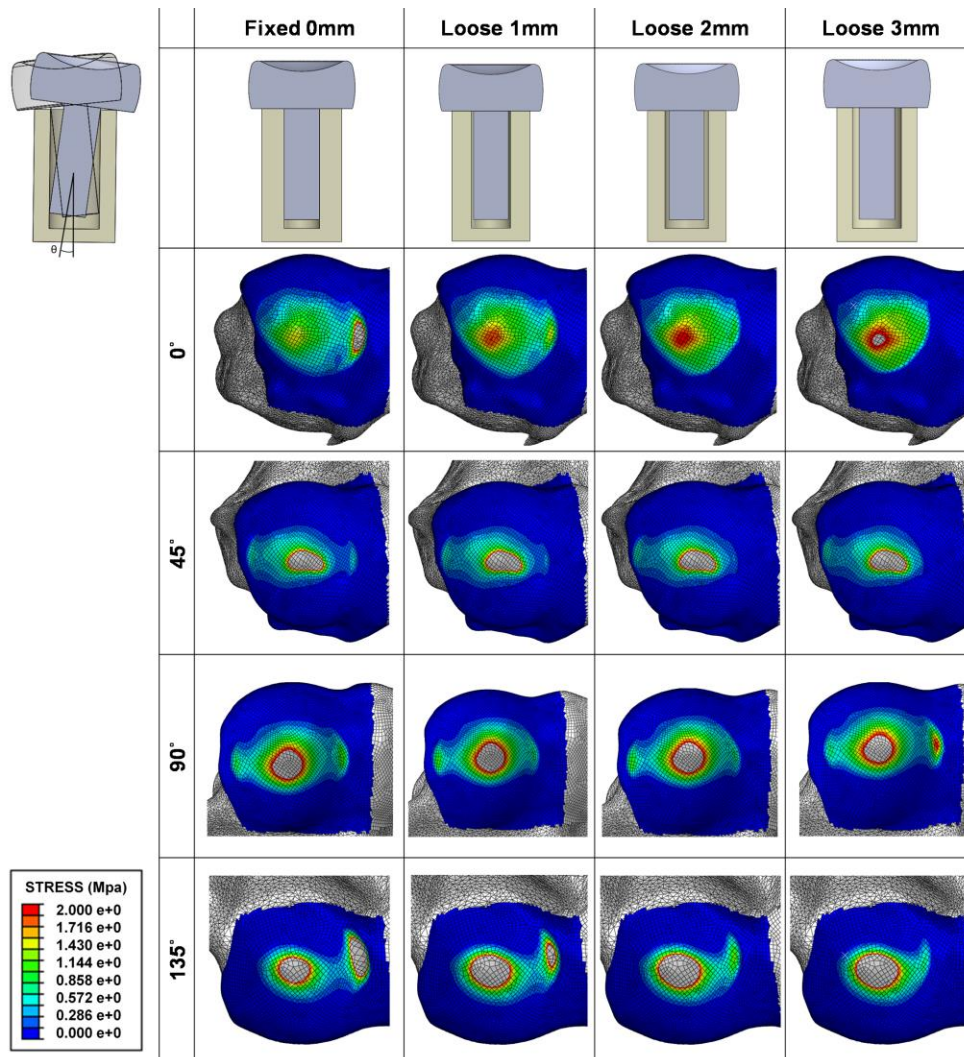


Figure 2. Contact Area (top) and Maximum Contact Stress (bottom) for 0 (fixed) and 1, 2 and 3mm (loose left to right) Stem Fixations (n=10)

Haptic Force Feedback for a Teleoperated Sonography System

Atieh Najafi Semnani, Reza Fotouhi, QianWei Zhang
 Department of Mechanical Engineering
 University of Saskatchewan
 Saskatoon, Canada

Haron Obaid, Scott Adams
 Department of Medical Imaging
 University of Saskatchewan
 Saskatoon, Canada

Abstract— A teleoperated sonography system allows a medical expert to perform an ultrasound examination on a patient remotely from an expert site. With a teleoperated sonography system, the motion of a master manipulator (such as a joystick carrying a virtual probe) is controlled by a medical expert and remotely reproduced at the patient site by a slave-robot carrying an ultrasound probe. Due to the remote application of the system, experts currently have no feedback of the force that the ultrasound probe applies on the patient. As a potential solution, we plan to evaluate forces acting on the patient body; such forces can be simulated in the master/expert site. In this paper, we have presented a virtual reality haptic force feedback simulator compatible with a 5DOF haptic wand to implement haptic force feedback in the expert site. The calculation of haptic force feedback is based on collision detection between two objects using the Gilbert-Johnson-Keerathi (GJK) algorithm. We found that the collision engine can efficiently detect a collision between a wand (representing an ultrasound probe) and a flat surface, providing promising results for future implementation of a virtual reality haptic force feedback simulator into a teleoperated sonography system.

Keywords: *Haptic; GJK algorithm; ultrasound; simulator*

I. INTRODUCTION

Many medical emergencies, shortages of specialist physicians in remote or rural areas, and long distances between health centers could be resolved with the development of a remote diagnostic ultrasound system that can be operated by non-specialists and remotely controlled robots [1]. A remote diagnostic teleoperation system can fulfill the need for full-time medical experts in remote areas.

A remote diagnostic system contains three sections; an expert site, a patient site and a communication network such as servers, fiber optics lines, Wi-Fi, or satellite networks. In the expert site, a master manipulator (a haptic device) with a

certain functionality and mediating communication between the user and the computer is installed and operated by an expert who is responsible for performing an ultrasound examination. The functionality of the device is restricted by the operating robot at the patient's site and is determined by the degrees-of-freedom (DOF) of the system. At the patient's site, a robotic device, combined with an ultrasound probe, is remotely controlled by a medical expert performing the examination [2].

A commercial remote diagnostic tele-sonography system, called Melody, has been developed by AdEchoTech [3]. A qualified sonographer performs an examination remotely on a patient that may be very far away. The sonographer performs the medical examination using a fictive probe at the expert site. The robotic arm holding an ultrasound probe reproduces the exact movement of the fictive probe at the patient site in real time. The robotic arm used in the Melody system is equipped with three rotational degrees of freedom, however, a nurse must be present at the patient's site in order to position the probe holder on the patient's body. The nurse, patient, and sonographer are in contact during the procedure via video conference [3]. However, this system lacks the haptic force feedback simulator in the expert site that would allow medical doctors to have a better understanding of the applied force on the patient during examination. The aim of this research is to develop a virtual reality haptic simulator to incorporate it to a recently developed remote diagnostic musculoskeletal tele-sonography system in the Robotics lab at the University of Saskatchewan. This will make it possible for medical experts to have a more realistic feeling during teleoperation. In this paper, we explain our virtual reality haptic simulator system, which is compatible with a 5DOF haptic wand by Quanser [4].

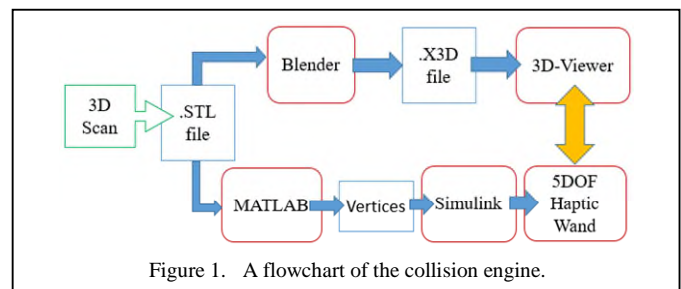


Figure 1. A flowchart of the collision engine.



Figure 2. 5DOF Haptic Wand

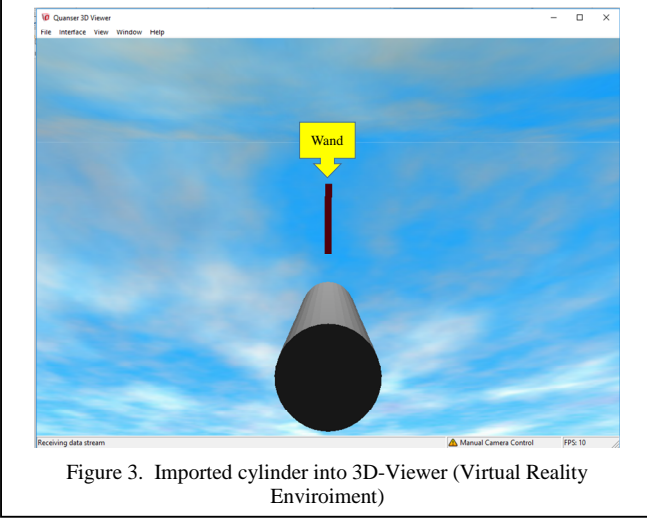


Figure 3. Imported cylinder into 3D-Viewer (Virtual Reality Environment)

II. METHODOLOGY

Our virtual reality haptic simulator system consists of a haptic interface as the hardware and a collision engine as the software.

A. Haptic Interface

The selected haptic device used for this research is a 5DOF haptic wand shown in Fig 2. The 5DOF wand robot is a redundant manipulator that can be used for haptic applications [4]. The interface has five degrees of freedom, which allows for three translations and two rotations (roll, yaw), in this research we only use three translations to define a single point in virtual reality environment.

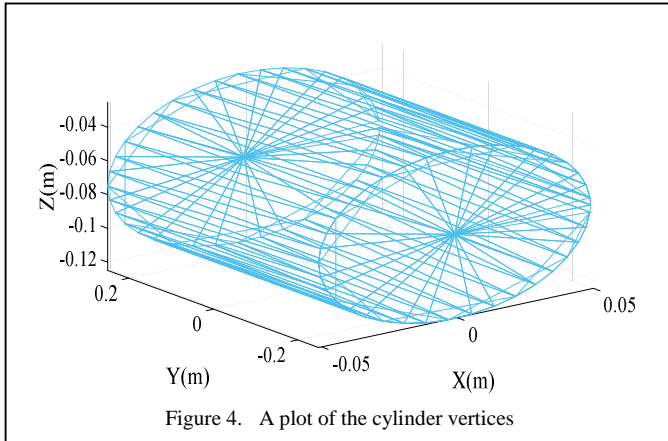


Figure 4. A plot of the cylinder vertices

B. Collision Engine

The purpose of a collision engine is to detect collision between objects, calculate the distance between colliding objects, calculate the collision force, and generate the force in the haptic device [5]. The scope for implementation of haptic force feedback in the expert site via a collision engine is based on Gilbert-Johnson-Keerathi (GJK) algorithm [6]. The GJK algorithm is a method to detect collision between two convex objects, also compute the distance or a pair of closest points between these two objects. Because of fast and simple implementation, the algorithm is widely used in many fields including robotics, real-time haptic rendering, medical surgery and computer graphics. This algorithm is based on the calculation of the Minkowski Difference between two convex objects and determines whether the origin is included in this difference or not. If origin is included, two objects collided to each other and have some points in common [7], [8]. The algorithm is also generalizable to concave objects, using convex decomposition [6].

In this research using the GJK algorithm, we detected collision between the lower tip of the wand and an imported object, also computed the distance or a pair of closest points between these two objects [7]. A flowchart of the proposed collision engine is shown in Fig 1. The presented protocol of the collision engine is as follow:

1. **3D Scan or 3D Drawing:** The object or 3D scan of a patient's body can be created or imported into Solidworks and saved in STL file format.
2. **Convert STL to X3D format:** Using Blender v 2.49, we converted the STL file to X3D format. This conversion is needed to import the object into the 3D-Viewer.
3. **Import in the 3D viewer:** Import object with X3D format into 3D-Viewer. Fig 3 shows the wand and an imported cylinder in 3D-Viewer (virtual reality environment).
4. **Convert STL file to vertices:** In order to perform collision detection analysis, using a variation of code in MATLAB R2018a, we converted the STL file to the object vertices, which are the common endpoint of two or more-line segments. The GJK algorithm basically uses the vertices of the objects. A plot of the vertices of an imported cylinder is shown in Fig 4.
5. **Simulating and Generating Haptic Force Feedback:** We have modified a Simulink model [4] to perform collision detection analysis based on GJK algorithm and calculate the collision force between the wand and the object. The object vertices and the real-time position vector of the wand are the input for the Simulink model. The code calculates the force based on the material properties of the object including damping ratio and object stiffness, as shown on equation (1) [5]. The calculated force is then applied at the haptic wand. The 5DOF haptic wand generates force feedback when the wand in a virtual reality environment collides with the object.

$$F = kx + bv \quad (1)$$

In equation (1), F is the calculated force in N , k is the stiffness of the object in N/m , x is the distance between the external surface of the object and the lower tip of the wand in the direction of a normal vector coming from the center of mass of the object to the contact point between lower tip of the wand and the object in m , b is the damping ratio of the object in $N.s/rad$, and v is the velocity vector of the haptic interface in $N.s/m$. The haptic wand simulator measures the distance x based on the GJK algorithm and calculates the velocity of the wand, v , calculates the contact force using equation (1).

III. RESULT AND DISCUSSION

To examine the developed collision engine, we plan to import a 3D scan of a human forearm, however, due to the complexity of the model and many (17556) vertices, the model was computationally expensive for a PC to process. We simplified the human forearm and replaced it with a cylinder. We imported an STL file of a cylinder with 0.025m radius and 0.1m height and 456 vertices into the Simulink model, as shown in Fig 4. The stiffness of the cylinder for the X, Y and Z direction of the coordinate system manually was set to (160,160,160) $N.s/m$, and damping ratio to (0.8,0.8,0.8) $N.s/rad$. The X3D format of the same cylinder has been imported into the 3D-Viewer, shown in Fig 3. The center of mass of the cylinder was set at (0, 0, -0.075). The initial position of the wand is (0, 0, 0). We simplified the wand with a single point including one vertex and three degrees of freedom. In a virtual experiment, the surface of the cylinder was collided with the wand in a 10s period. Fig 5 shows the distance between the wand and the upper surface of the cylinder in this experiment. When the wand collides with the surface of the cylinder, the calculated distance between the wand and the cylinder must be zero. However, in some proximities, the engine miscalculated the distance, which causes a noisy collision response. The engine has miscalculated distance at 6 instances 3.8, 4, 4.2, 6.5 6.9, 9s, with an average error of $3.44 \times 10^{-2}m$ for distance value. The force during contact between the cylinder and the wand had been measured by the collision engine based on equation (1) in X, Y and Z direction. The simulated forces are shown in Fig 6.

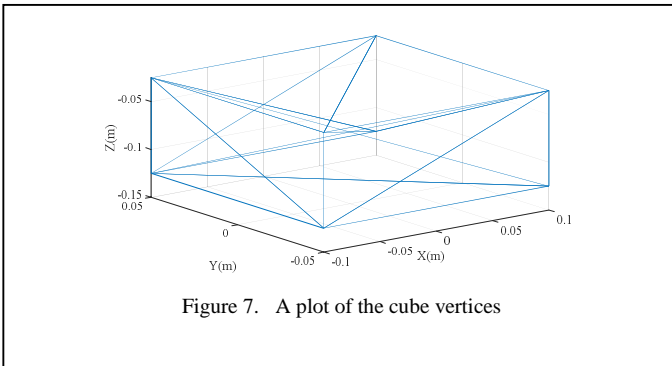


Figure 7. A plot of the cube vertices

The average of force during contact between the cylinder and the wand was 0.3561N, 1.711N and 0.6248N in X, Y and Z

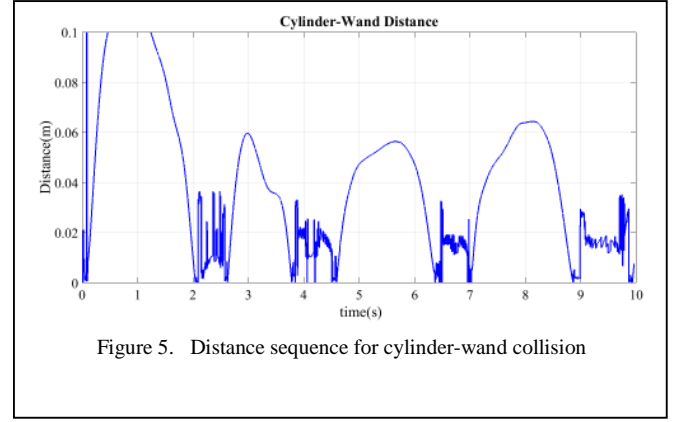


Figure 5. Distance sequence for cylinder-wand collision

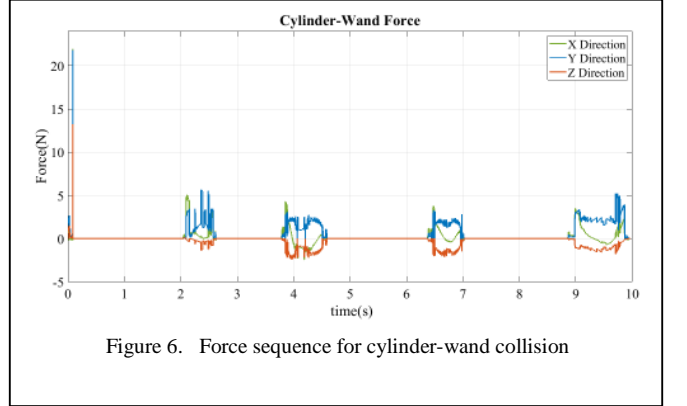


Figure 6. Force sequence for cylinder-wand collision

direction respectively. These average forces were for all 4 contact periods as shown in Fig 6.

In another experiment, we have imported the STL file of a $0.1 \times 0.1 \times 0.2 m^3$ cube with 36 vertices. The stiffness of the cube for the X, Y, and Z direction manually was set to (160,160,160) $N.s/m$, and damping ratio of that had been set to (0.8,0.8,0.8) $N.s/rad$, respectively. The center of mass of the cube has been placed at (0, 0, -0.075), the same place that we placed the center of the cylinder. A plot of the vertices of the cube is shown in Fig 7.

The initial position of the wand is placed at (0, 0, 0). Surface of the cube and the wand were collided for a period of 10s. Fig 8 shows the distance between the wand and upper surface of the cube. In this experiment, the engine has returned a false value for the calculated distance at instance 2.4 seconds, which led to a noisy collision response at this moment, with the average error of $4.176 \times 10^{-2}m$ for distance value. The simulated force during contact between the cube and the wand that was calculated by the collision engine based on equation (1) in X, Y, and Z direction and is shown in Fig 9. The average forces while the wand contacted with the cube in X, Y, and Z direction were about 1.62N, 0.477N, and 0.217N respectively. These averages forces are for all 9 periods of contacts as shown in Fig 9. By comparing the results of two performed experiment and considering the number of errors for the distance points, as shown on Table 1, it can be observed that the developed engine can detect the collision between the wand and a flat surface efficiently in comparison with a curved surface. From this observation, it can be

assumed the presented virtual reality haptic force feedback simulator can be utilized for the flat surface of the patient's body e.g. stomach surface.

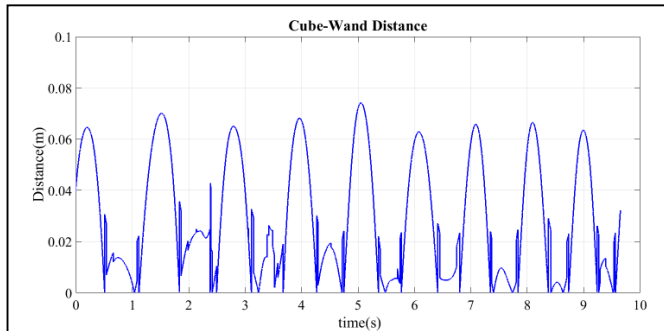


Figure 8. Distance sequence for cube-wand collision

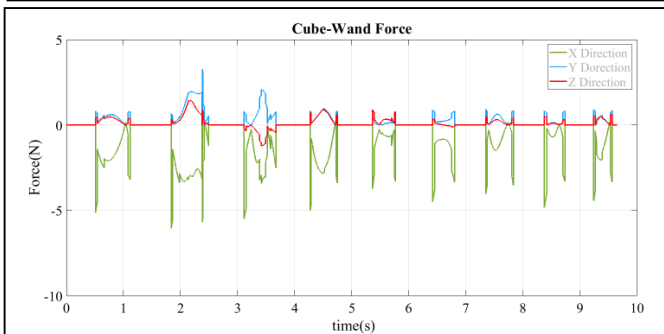


Figure 9. Force sequence for cube-wand collision

TABLE I.
OMPARISON OF THE PERFORMED EXPERIMENT

Object Type	Number of vertices	Average of Distance error (m)	Number of errors
Cube	36	4.176×10^{-2}	1
Cylinder	456	3.44×10^{-2}	6

To verify the generated force based on equation (1) from Simulink, we manually calculated these forces for two random instants of 8.098s and 8.436s in the cube experiment. The distance value between the cube and the wand at 8.098s is 6.645×10^{-2} m, which means that the wand is far away from the cube and it is expected that the force value for this instant is equal to zero. Fig 9 shows that the related force at 8.098s is equal to zero in X, Y, and Z direction. Fig 8 shows that the distance value at instant 8.436 is 9.964×10^{-4} m, which means the wand is very close to cube and is in contact. Table 2 shows the wand velocity, distance vector from the wand toward the center of mass of the cube, and force value for X, Y, and Z direction at instant 8.436s. The calculated force based on equation (1) at this instant is shown in Table 2, 0.3194N, 0.0245N, 0.0148N in X, Y and Z direction which is same as the extracted value of the force from Fig 9.

As Fig 1 shows, the 3D scan of the object needs to be imported in different software including Solidworks,

MATLAB and 3D-Viewer. The relationship between scaling ratios of objects in Solidworks, MATLAB and 3D-Viewer, is shown in Table 3. An object with the dimension of 0.05m in MATLAB has 0.5 mm dimension in Solidworks, and 0.5 in 3D-Viewer. Observed workspace of the 5DOF haptic wand in translation around Y axis is between -240 and +240 [4].

TABLE II.

FORCE CALCULATION AT INSTANT 8.436S

Direction	Wand Velocity (N.s/m)	Distance Vector(m)	Force(N)
X	0.2055	-9.685×10^{-4}	0.3194
Y	0.0087	1.967×10^{-4}	0.0245
Z	0.0069	1.268×10^{-4}	0.0148

TABLE III.

SCALE AND DIMENSION RELATIONSHIP

3D-Viewer	MATLAB	SOLIDWORKS
0.5	0.05m	0.5mm

CONCLUSION

In this paper, a collision engine compatible with a 5DOF haptic wand to simulate force feedback is presented. A limited number of experiments on cube and cylinder is conducted, however, the results show that the employed collision engine simulated the haptic force feedback on the flat surface efficiently, compared with curved surfaces. The future work is to incorporate the haptic force feedback simulator to a recently developed tele-operated sonography system. The virtual reality haptic force feedback simulator investigated in this paper may also have applicability in the development of ultrasound simulators for training and education.

ACKNOWLEDGMENT

This project was funded from the University of Saskatchewan, College of Medicine Research Award.

REFERENCES

- [1] R. Fotouhi, R. Orazi, C. Mondragon, B. Berryman, Development of a Remote Ultrasound Imaging System, 57106-V003T14A008, 2015.
- [2] K. Boman, M. Olofsson, J. Forsberg, S. Boström, Remote-controlled robotic arm for real-time echocardiography: the diagnostic future for patients in, *Telemed J E Health*, 15(2):142-7, 2009.
- [3] www.adechotech.com, visited Jan 1, 2019
- [4] QUANSER, 5DOF Haptic Wand, "Reference Manual.", Markham, ON, Canada.
- [5] S. D. L. A. M. Day, "A Survey of Haptic Rendering Techniques," *Comput. Graph. Forum*, vol. 26, no. 1, 2007.
- [6] H. Noborio, D. K. Sasaki, Y. Kawamoto, T. Tatsumi, T. Sohmura, "Mixed Reality Software for Dental Simulation System," *IEEE Int. Work. Haptic Audio Vis. Environ. Games*, pp. 19-24, 2008.
- [7] C. Ericson, "Chapter 9 - Convexity-based Methods," in *Real-Time Collision Detection*, C. Ericson, Ed. San Francisco: Morgan Kaufmann, 2005, pp. 383-412.
- [8] C. Ericson, Chp 3-A Math and Geometry Primer, in *Real-Time Collision Detection*, San Francisco, Morgan Kaufmann, 23-73, 2005
- [9] B. Unger, N. Sepelri, V. Rampersad, J. Pisa, M. Gousseau, J. Hochmann, Elements of virtual temporal bone surgery Manipulandu, *Laryngoscope Investig. Otolaryngol*, 2(6)358-362, Dec 2017.

Energy Materials

Interfacial Engineering of Solid-State Batteries using Atomic Layer Deposition

Neil P. Dasgupta

Department of Mechanical Engineering
University of Michigan, Ann Arbor
Ann Arbor, MI 48109, USA

ABSTRACT

Solid-State batteries have experienced a recent explosion in research and development, owing to their potential to improve safety and enable higher energy density electrode chemistries. However, it has been widely recognized that all solid-state interfaces present unique challenges compared to traditional liquid electrolytes, including reduction of high interfacial impedances (which can evolve during cycling), accommodation of mechanical stresses due to solid-solid interfacial contact with active materials, and (electro)chemical instabilities that can arise from localized gradients in ionic and electronic concentrations.

To address these challenges, our group focuses on gaining new fundamental insights into the coupled phenomena occurring at interfaces, and applied this knowledge to rationally design interfacial chemistry to address the root cause of performance limitations. The key enabling technology that will be discussed in this talk for surface modification is Atomic Layer Deposition (ALD). This is a gas-phase deposition process capable of conformally coating high aspect-ratio structures with sub-nanometer control in thickness. This atomic-scale modification of surfaces allows for precise control of interactions at heterogeneous interfaces, including (electro)chemical stability, interfacial kinetics, wettability, and mechanical load transfer.

In this talk, examples will be presented in both bulk solid-state battery interfaces, and thin film electrolytes deposited by ALD. By studying interfacial chemistry across length scales ranging from atoms to millimeters, we have been able to systematically identify the mechanisms of interfacial degradation and answer the question of why a particular interface exhibits the observed electrochemical behavior. At bulk length scales, the coupled chemical/electrochemical/morphological evolution of a range of solid electrolytes will be discussed, including metal oxides, sulfides, and polymers [1,2]. To modify these surfaces, ALD of solid lithium electrolytes is demonstrated [3,4], which enables tunability of surface chemistry, defect structure, morphology, and stability. These ALD electrolytes are demonstrated in both high power density thin-film battery architectures, and as interfacial layers at bulk solid electrolyte-electrode interfaces. Through this interdisciplinary approach of fundamental chemistry and applied engineering, strategies to address future interfacial challenges will be addressed.

1. A. Sharafi, E. Kazyak, A. L. Davis, S. Yu, T. Thompson, D. J. Siegel, N. P. Dasgupta, J. Sakamoto, *Chem. Mater.* **29**, 7961 (2017)
2. A. Gupta, E. Kazyak, N. Craig, J. Christensen, N. P. Dasgupta, and J. Sakamoto, *J. Electrochem. Soc.* **165**, A2801 (2018)
2. E. Kazyak, K.-H. Chen, K. N. Wood, A. L. Davis, T. Thompson, A.R. Bielinski, A. J. Sanchez, X. Wang, C. Wang, J. Sakamoto, N. P. Dasgupta, *Chem. Mater.* **29**, 3785 (2017)
4. E. Kazyak, K.-H. Chen, A. L. Davis, S. Yu, A. J. Sanchez, J. Lasso, A. R. Bielinski, T. Thompson, J. Sakamoto, D. J. Siegel, N. P. Dasgupta, *J. Mater. Chem. A* **6**, 19425 (2018)

Computational Design of 2D Nanomaterials for Sustainable Energy

Chandra Veer Singh

Department of Materials Science & Engineering, University of Toronto
184 College St., Suite 140
Toronto, ON M5S 3E4, Canada
Email: chandraveer.singh@utoronto.ca

ABSTRACT

Increasing energy demand along with an increased risk of global climate change has led to an enhanced significance of materials science research. In recent decades, a multitude of novel low dimensional materials, such as graphene, graphene oxide, and their composites, have been synthesized to overcome this challenge. Although still in infancy, these low dimensional materials, have shown great promise for the development of new technologies that will help meet our future sustainable energy and transport. Nevertheless, developing and implementing new materials technologies takes many years. Computational materials science can reduce this materials discovery-deployment cycle by providing fundamental atomistic insights, quick computational screening of promising material designs, and guide experimentation.

In this talk, we would describe the role of computational materials science to design low-dimensional materials for sustainable energy applications. The following representative examples will be discussed: (a) design of champion materials for solar-driven CO₂ reduction, thereby mimicking photosynthesis [1-3]; (b) role of 2D materials in improving capacity and performance of lithium-ion and lithium-sulfur batteries [4-7]; (c) strong size dependent mechanical behavior of graphene oxide monolayer and nanosheets [8-9]; and (d) defect-engineering of graphene and other low-dimensional materials for photocatalytic water splitting and hydrogen storage [10-11]. These works will also serve as an example of how we can combine atomistic modeling with experimental testing to provide a more coherent and in-depth understanding.

References:

- [1] Qian, C. et al. "Catalytic CO₂ reduction by palladium-decorated silicon-hydride nanosheets." *Nature Catalysis* 2.1 (2019): 46.
- [2] Ghuman, K. K. et al. "Photoexcited surface frustrated Lewis pairs for heterogeneous photocatalytic CO₂ reduction." *Journal of the American Chemical Society* 138.4 (2016): 1206-1214.
- [3] Sun, W. et al. "Heterogeneous reduction of carbon dioxide by hydride-terminated silicon nanocrystals." *Nature communications* 7 (2016): 12553.
- [4] Li, Lu, et al. "Phosphorene as a Polysulfide Immobilizer and Catalyst in High - Performance Lithium - Sulfur Batteries." *Advanced materials* 29.2 (2017): 1602734.
- [5] Gao, Jian, et al. "Vertically oriented arrays of ReS₂ nanosheets for electrochemical energy storage and electrocatalysis." *Nano letters* 16.6 (2016): 3780-3787.
- [6] Grixti, S., Mukherjee, S., & Singh, C. V. (2018). Two - dimensional boron as an impressive lithium - sulphur battery cathode material. *Energy Storage Materials*, 13, 80-87.
- [7] Li, Lu, et al. "A foldable lithium-sulfur battery." *ACS nano* 9.11 (2015): 11342-11350.
- [8] Cao, C et al. "Nonlinear fracture toughness measurement and crack propagation resistance of functionalized graphene multilayers." *Science advances* 4.4 (2018): eaao7202.
- [9] Cao, C., Daly, M., Chen, B., Howe, J. Y., Singh, C. V., Filleter, T., & Sun, Y. (2015). Strengthening in graphene oxide nanosheets: bridging the gap between interplanar and intraplanar fracture. *Nano letters*, 15(10), 6528-6534.
- [10] Ghuman, K. K., Yadav, S., & Singh, C. V. (2015). Adsorption and dissociation of H₂O on monolayered MoS₂ edges: Energetics and mechanism from ab initio simulations. *The Journal of Physical Chemistry C*, 119(12), 6518-6529.
- [11] Yadav, S., Zhu, Z., & Singh, C. V. (2014). Defect engineering of graphene for effective hydrogen storage. *International journal of hydrogen energy*, 39(10), 4981-4995.

Highly Efficient Protection by Atomic Layer Deposition/Molecular Layer Deposition for Lithium and Sodium Metal Anode

Yang Zhao, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

Li metal anode is considered as the promising candidate for next generation Li-metal batteries. However, it is still a crucial problem of mossy or dendritic growth of Li occurs in the repetitive Li stripping/plating process with an unstable solid electrolyte interphase (SEI) layer, which may create short circuit risks, resulting in possible burning or explosion. Atomic layer deposition (ALD) can be one of the promising technique with excellent coverage and precise control over coating thickness to stabilize the SEI layer and longer the life time for Li metal anode [1]. Two different groups have demonstrated ultrathin ALD Al_2O_3 coating film as protective layer for metallic lithium with the prevention of corrosion and reduced dendrite growth [2]. As an analogy of ALD, molecular layer deposition (MLD) can be employed to produce polymeric thin films with many advantages such as tunable thermal stability, and improved mechanical properties [3]. Herein, we firstly demonstrate MLD alucone (Al-EG) coating as a protective layer for Li metal anode for improved stability and life time [4]. alucone layer can stabilize the SEI film and further reduce dendrite growth, which leading to the better performances than Al_2O_3 coating. Na-ion batteries and Na-metal batteries are expected to be used as large scale energy storage devices. Na metal anode show a high theoretical specific capacity of 1166 mAh g⁻¹ and lowest electrochemical potential. However, similar problems including mossy and dendritic Na growth is one of the challenge for Na metal anode. Here, we demonstrated the successful application of the ALD Al_2O_3 protective coating on Na metal anode in ether based electrolyte to achieve long lifetime Na metal anode [5]. Furthermore, we firstly demonstrate the inorganic-organic coating MLD Alucone as coating layer for Na anode in carbonate based electrolyte with suppressed dendrite growth and enhanced electrochemical stability [6].

In conclusion, we demonstrate ultrathin ALD Al_2O_3 and MLD alucone protective coating layers for both Li and Na metal anode. The results show that the formation of mossy and dendrite-like Li/Na can be effectively suppressed and life time is significantly improved with protective layer. It is also worth to mention that we have first introduced a strong tool of Rutherford Backscattering Spectrometry (RBS) in the Li/Na protection area. It is believed that our design of ALD/MLD coating on Li/Na anode opens up new opportunities to the realization of next-generation high energy density Li/Na metal batteries.

- [1] X. Meng, X. Sun, et al, *Advanced Materials*, 2012, 24, 3589
- [2] A. Kozen, M. Noked, et al, *ACS Nano*, 2015, 9, 5884; E. Kazyak, N. Dasgupta, et al, *Chemistry of Materials*, 2015, 27, 6457.
- [3] Y. Zhao, X. Sun et al, *ACS Energy Letters*, 2018, 3 (4), pp 899–914
- [4] Y. Zhao, X. Sun et al, *Small Methods*, 2018, 2, 1700417
- [5] Y. Zhao, X. Sun, et al, *Advanced Materials*, 2017, 29, 1606663
- [6] Y. Zhao, X. Sun et al, *Nano Letters*, 2017, 17, 5653

Organic/Inorganic Composite Electrolyte for Dendrite-free Solid-state Lithium Batteries

Hanyu Huo, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

Solid-state batteries (SSBs) show significant advantages over traditional liquid-electrolyte-based lithium batteries. The large electrochemical window of solid-state electrolytes (SSEs) can enable the use of Li metal anodes and high-voltage cathodes, rendering a high energy density for SSBs. In addition, the flammability and leakage concerns related to the conventional liquid electrolytes are eliminated by replacing with SSEs. Despite of these beneficial intrinsic properties, practical SSEs should meet the following requirements to ensure good electrochemical performance for SSBs: 1) Ionic conductivity above 10^{-4} S cm⁻² at room temperature, 2) stable and robust interface between SSEs and electrodes for smooth Li⁺ transport, and 3) mechanically strong and flexible to suppress Li dendrite growth and allow ease of handling. [1]

Polymers of polyethylene oxide (PEO) have been studied as hosts of solid polymer electrolytes (SPE) for several decades since Wright et al. discovered ionic transport of PEO complex with alkali metal salts. However, the low conductivity of approximately 10^{-7} S cm⁻¹ at ambient temperature attributed to the low effective carrier mobility in the crystalline region below melting temperature (~60 °C) limits its technical application. Herein, the composite electrolytes consisting of Li-salt-free PEO and 200 nm-sized Li_{6.4}La₃Zr_{1.4}Ta_{0.6}O₁₂ (LLZTO) particles interfacially wetted by [BMIM]TF₂N of 1.8 μm cm⁻² have been prepared. [2] Such wetted ionic liquid remains the solid state of membrane electrolytes and decreases the interface impedance between the electrodes and the electrolytes. The interfacially wetted membrane electrolytes show the conductivity of 2.2×10^{-4} S cm⁻¹ at 20 °C, which is one order of magnitude greater than that of the membranes without the wetted ionic liquids. In order to suppress the lithium dendrite growth through the SPEs, the interfaces between Si and SPEs based on poly(propylene carbonates) (PPCs) and LLZTO have been fabricated by deposition of the Si layer on the SPEs at 25 °C. [3] Introduction of LLZTO powders into the PPCs promotes complete dissociation of lithium salt as well as enhances the migration of Li⁺, achieving the conductivity of 4.2×10^{-4} S cm⁻¹ at room temperature. Taking advantages of flexible Si/SPE interfaces, the Li/SPE/Si cells exhibit excellent cycle stability and low charge-transfer resistance after 200 cycles. The reason can be attributed to the effective alleviation of internal stress from volume variation during repeatedly lithiation and delithiation progress. In addition, a novel cationic metal-organic framework (CMOF) is proposed in SPEs to immobilize anions and guide Li⁺ uniform distribution for constructing dendrite-free SSBs. [4] The CMOF grafted with -NH₂ group protects the ether oxygen of polymer chains by hydrogen bonds, which extends the electrochemical window to 4.97 V. Such CMOF tethers anions by electrostatic interaction of charge carriers and the specific surface area as high as 1082 m² g⁻¹ further strengthens the effect of anion absorption on the surface of CMOF, leading to a high Li⁺ transference number of 0.72. With the anion-immobilized composite electrolyte, the Li symmetrical cells can continuously operate for 400 h at 0.1 mA cm⁻² and 200 h at 0.5 mA cm⁻² without discernable dendrites, respectively.

In conclusion, we demonstrate the composite electrolytes consisting of PEO/PPC and LLZTO ceramic particles, which improve the ionic conductivity over 10^{-4} S cm⁻¹ at room temperature. The Si anodes in SSBs are expected to be an alternative material to suppress the Li dendrite growth as well as enhance the energy density. Anion immobilization via electrostatic interaction opens a new avenue to understand the growth and suppression of Li dendrites in SSBs. In addition, design of novel cationic fillers for anion immobilization can be easily scalable to other MOF structures, which are expected to achieve the same or even better electrochemical performances.

[1] J. Zhang, X. Guo et al, *Nano Energy*, 28 (2016) 447–454

[2] H. Huo, X. Guo et al, *Journal of Power Sources*, 372 (2017) 1–7

[3] H. Huo, X. Guo et al, *Journal of Power Sources*, 383 (2018) 150–156

[4] H. Huo, X. Sun et al, *Energy Storage Materials* (2019), <https://doi.org/10.1016/j.ensm.2019.01.007>

Stabilization of Solid-state Electrolyte by Atomic Layer Deposition for High Performance All-solid-state Li-S Battery

Jianneng Liang, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

Lithium-sulfur (Li-S) batteries is considered as a promising candidate for the application in electric transportation due to the high energy density, low cost and nature abundant of sulfur. However, several challenges inhibit the wide application of Li-S batteries. Firstly, the charge/discharge intermediate products, polysulfides, lead to the shuttle effect causing the loss of active material and low coulombic efficiency [1], [2]. Secondly, the application of ether-based liquid electrolyte makes Li-S batteries suffer from serious safety issues due to the flammability and the volatile of the organic solvents [3].

All-solid-state Li-S batteries (ASSLSBs) is proposed to be able to address the challenges in Li-S batteries. The application of solid-state electrolyte (SSE) in ASSLSB can eliminate the safety problems because SSE is naturally inflammable and not volatile. The charge/discharge intermediate products, polysulfides, is in soluble in SSE, thus, the shuttle effect can be effectively inhibited. However, one the most commonly used SSEs, $\text{Li}_{1.4}\text{Al}_{0.4}\text{Ti}_{1.6}(\text{PO}_4)_3$ (LATP), is reported to be unstable toward reductants such as lithium metal and polysulfides [4], [5]. Thus, it is very important to protect LATP SSE in ASSLSBs from being reduced by polysulfides for achieving a high performance ASSLSB.

As a novel deposition technique, atomic layer deposition (ALD) has the capability to growth uniform, controllable thickness thin film materials at relative low temperature compared to other deposition techniques such as chemical vapor deposition (CVD), physical vapor deposition (PVD). Thus, using ALD to protect SSE for ASSLSBs is promising because the thin film structured material will not brock the Li^+ transportation, and the moderate deposition temperature will not destroy the SSE subtract. Most importantly, ALD-derived Al_2O_3 is highly stable toward reductant. Thus, ASSLSBs with ALD Al_2O_3 coated LATP SSE show significant improved cycling performance compared to the bare SSE without coating [6].

In conclusion, LATP SSE with ultrathin ALD coating is highly stable in ASSLSBs, which significantly improve the cycling stability of ASSLSBs. The results show that bare LATP SSE suffered from serious reduction of LATP and collapse of the structure after extensive charge/discharge cycles. With ALD protection, LATP SSE illustrated stable structure and the reduction of LATP is much smaller than that of bare LATP after extensive charge/discharge cycles. It is believed that our design of ALD coating SSE opens up new opportunities to the realization of high energy density Li-S batteries.

- [1] X. Yang, X. Sun, et al., *Electrochemical Energy Reviews*, 2018, 1-55.
- [2] X. Li, X. Sun, *Advanced Functional Materials*, 2018, 1801323.
- [3] J. Yue, Y. Guo, et al., *Advanced Functional Materials*, 2018, 1707533.
- [4] Y. Liu, X. Sun, et al., *ACS Appl. Mater. Interfaces*, 2018, 10(17), 14641-14648.
- [5] S. Wang, A. Manthiram, et al., *ACS Energy Letters*, 2016, 1(6), 1080-1085.
- [6] J. Liang, X. Sun, et al., *Journal of Materials Chemistry. A*, 2018, 6, 23712-23719

High-performance All-solid-state Li-Se and Li-SeS_x batteries

Xiaona Li, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

All-solid-state Li batteries (ASSLBs) using inorganic solid electrolytes are considered as one of the most promising energy storage technologies due to high safety and high energy densities [1]. However, unlike electrolytes in conventional Li-ion batteries that can easily wet the electrodes to ensure effective electrochemical reactions, the build-in ionic and electronic conductivities of solid-state electrodes and electrode/electrolyte interfaces in ASSLBs are crucial. Although S cathode has an ultrahigh theoretical capacity of 1675 mAh g⁻¹, the electronic/ionic insulating nature of S (0.5×10^{-27} S m⁻¹) as well as the high interfacial resistance between solid-state sulfide electrolyte and S cathode ($\sim 10^{-11}$ S cm⁻¹) severely hinder the kinetics in ASSLBs.[2,3] Thus, S cathodes have been suffering from low S utilization, low loading of S cathode, poor cycle life, and low rate performance in ASSLBs. Therefore, developing simple and effective approach to mitigate these challenges beyond optimization of the mixing recipes for S cathodes in all-solid-state Li-S battery systems is urgent.

High performance Se and SeS_x cathodes were introduced into sulfide-based ASSLBs, taking advantage of high electronic conductivity of Se (1×10^{-3} S cm⁻¹) and their good compatibility with sulfide solid electrolytes. Moreover, high interfacial ionic conductivities in the order of 10^{-6} - 10^{-5} S cm⁻¹ are achieved between those active materials and commercial Li₃PS₄ electrolytes. [4,5] As a result, the Se and SeS_x cathodes delivered stable highly reversible electrochemical performance in ASSLBs. It's worth mention that the assembled ASSLBs could even obtain ultra-high capacity up to 12.6 mAh cm⁻².

In conclusion, the results demonstrate all-solid-state Li-Se, Li-SeS_x batteries as promising high-energy storage system, deepen the understanding of Se-S solid solution chemistry in ASSLBs systems, and provide new insights of the solid-state electrode/electrolyte interfacial chemistry.

- [1] Y. Kato, R. Kanno, et al. *Nat. Energy*, 2016, 1, 16030.
- [2] X. Ji, L. Nazar et al, *Nat. Mater.*, 2009, 8, 500.
- [3] C. Yu, M. Wagemaker et al, *J. Am. Chem. Soc.*, 2016, 138, 11192-11201.
- [4] X. Li, X. Sun et al, *Energy Environ. Sci.*, 2018, 11, 2828-2832.
- [5] X. Li, X. Sun, et al, *Adv. Mater.*, 2019, in press.

Sulfide solid-state electrolyte protection layer *in-situ* formed on the surface of metallic Li for high-performance Li batteries

Jianwen Liang, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

Li metal anode is considered as the promising candidate for Li-S, Li-air and solid-state Li batteries, due to its high theoretical capacity and low redox potential. However, it is still a crucial problem of unavoidably reacts of Li with electrolyte and the non-uniform distribution of electrochemical active sites for Li plating/stripping during cycling, which would induce dendrite formation and unfavorable reaction between Li and electrolyte. [1]

Formation of a thin and stable protection layer on the surface of Li metal can be one of the promising technique to provide uniform and high Li ion flux. Among the various candidates, sulfide solid electrolyte is regarded as an attractive choice as protection layer, considering the high ionic conductivity even comparable to that of liquid electrolytes. [2] While, how to obtain thin sulfide solid-electrolyte layer with adjustable thickness on the surface of Li metal is still a huge challenge. Based on new phosphorous sulfide of P_4S_{16} synthesized recently, [3] we further developed a solution-based synthesis of a thin ionic conductive Li_3PS_4 solid-state electrolyte protection layer on the surface of Li metal electrode via an *in-situ* and self-limited reaction between P_4S_{16} and Li. [4] The thickness of the Li_3PS_4 protection layer can be adjusted when change the concentration of P_4S_{16} in solution. Due to the high ionic conductance and low electrochemical activity of Li_3PS_4 , the intimate protection layer of Li_3PS_4 between Li metal and liquid organic electrolyte can not only restrain the formation of Li dendrite, but also reduce the SEI formation and prevent further corrosion of Li metal during battery cycling.

Thus, stabilized Li metal anodes with long term cycling and high rate performance have been demonstrated in symmetrical Li- Li_3PS_4 , Li- Li_3PS_4 /S and Li- Li_3PS_4 /LiCoO₂ cells. In conclusion, the capability to *in-situ* formed of thin and adjustable Li_3PS_4 solid-state electrolyte protection layer on the surface of Li electrode may provide new opportunities to address the intrinsic problems of metallic Li batteries.

- [1] D. Lin, Y. Cui, et al, *Nat. Nanotechno.* 2017, 12, 194-206.
- [2] Y. Kato, R. Kanno, et al., *Nat. Energy*, 2016, 1, 16030.
- [3] X. Li, J. Liang, Y. Qian et al, *Angew. Chem. Inter. Ed.*, 2017, 56, 2937-2941.
- [4] J. Liang, X. Sun et al, *Adv. Mater.*, 2018, 30, 1804684.

Organic Polyurea Thin Film for Stable and Long-Life Lithium Metal Anode

Yipeng Sun ^{a, b}, Tsun-Kong Sham ^b, Xueliang Sun ^{a, *}

^a Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

^b Department of Chemistry, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: tsham@uwo.ca

ABSTRACT

It is urgent to develop efficient energy storage system for the increasing energy demands of human society. Batteries with high energy density is one of the most promising solutions. Metallic Li anode is believed to be the most promising electrode candidate for high-energy-density batteries due to its ultrahigh capacity and low electrochemical potential [1]. However, it has still not been successfully commercialized because of several great challenges, including poor safety and dendrite growth in plating/stripping process. Highly reactive nature of Li metal anode towards organic electrolyte leads to severe side reactions and prompt the formation of solid electrolyte interphase (SEI). Although SEI can passivate the Li surface and serve as a protective layer on the electrode, it cannot prevent the dendrite growth due to its heterogeneous and unstable properties [2]. In recent years, there are reported works on building artificial SEI to reduce side reactions and regulate uniform Li ion flux. Various inorganic materials (such as metal oxides and lithium compounds) and hybrid inorganic-organic materials (such as alucone) have shown their advantages to protect Li metal anodes [3-5]. However, the study on mechanism and principle of organic materials as artificial SEI for Li metal anodes have rarely been reported.

Herein, for the first time, we report a novel ultrathin organic film of polyurea via Molecule Layer Deposition (MLD) as protective layer for Li-metal anodes [6]. MLD is applied as a powerful tool to fabricate the coating layer due to its excellent coverage and accurate control over thickness at nanoscale. Compared to bare Li, the polyurea coated Li shows highly improved cycle life and stability. This electrically nonconductive film can confine the Li deposition beneath the film and suppress the dendrite growth as a protective barrier. Meanwhile, the nitrogen-containing polar groups in polyurea can effectively regulate the Li-ion flux and lead to a uniform Li deposition. Owing to these advantages, the Li metal anode coated with polyurea layer enables greatly prolonged lifetime in the symmetric cells at different current densities and capacities. The full cells were tested using LiFePO₄ (LFP) as the cathode and showed improved capacity retention and rate performance. Therefore, this work sheds new light on the design of protective layer for Li-metal anodes for high-energy-density next-generation batteries.

- [1] D. Lin, Y. Cui, et al, *Nature Nanotechnology*, 2017, 12, 194-206
- [2] X. Cheng, Q. Zhang, et al, *Chemical Reviews*, 2017, 117, 10403-10473
- [3] A. Kozen, M. Noked, et al, *ACS Nano*, 2015, 9, 5884
- [4] D. Lin, Y. Cui, et al, *Nano letters*, 2017, 17, 3731-3737
- [5] Y. Zhao, X. Sun et al, *Small Methods*, 2018, 2, 1700417
- [6] Y. Sun, X. Sun, et al, *Advanced Materials*, 2019, 31, 1806541

3D-Printed Sulfur Cathode for Lithium Sulfur Batteries

Xuejie Gao, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

Lithium-sulfur (Li-S) batteries have been considered as one of the most perspective candidates due to their superior theoretical energy density of 2600 Wh kg^{-1} and low cost, naturally abundant, environmental friendly active materials as well.[1] Nevertheless, there are still such series of intricate challenges as low Li^+/e^- conductivity of sulfur and $\text{Li}_2\text{S}/\text{Li}_2\text{S}_2$, remarkable volumetric expansion during lithiation and “shuttle effect” remaining to be solved.[2] Recently, tremendous efforts have been focused on suppressing PSs (PSs) shuttle between anode and cathode via physical confinement and chemical absorption. Based on these attempts, excellent cycling performance of Li-S batteries up to 1000 cycles have been achieved. Despite the significant achievements, it should be noted that almost long-term cycling life is based on the batteries assembled with low sulfur loading cathodes less than 2 mg cm^{-2} . To achieve high practical energy density, sulfur cathodes with a high sulfur loading more than 3 mg cm^{-2} are essential for practical high-energy density Li-S batteries. [3] Given the thickness of cathode related to sulfur loading, the electrochemical performance of Li-S batteries is highly depended on the thickness of cathode due to the worsening Li^+ transport in thick cathodes, resulting in low capacity output, fast capacity decay and large over potential. So in the first work, a self-standing 3D-printed sulfur/carbon cathode with grid structure was proposed, which could facilitate Li^+/e^- transport at the macro, micro and nano scale in Li-S batteries. When the batteries assembled with the 3D-printed/carbon cathode with a sulfur loading of 3 mg cm^{-2} , it delivers a stable capacity of 564 mA h g^{-1} within 200 cycles at 3 C . [4] But one important problems in most of sulfur cathode is that, the capacity would decrease when increase sulfur loading. So in the second work, a thickness-independent electrode structure is firstly proposed via converting thick electrode into thousands of vertically aligned thin electrode. And each thin electrode delivers a constant thickness of around $20 \text{ }\mu\text{m}$, which is not affected by the intrinsic thickness of electrode as well as sulfur loading. The electrochemical reactions occur on the thin electrode with similar Li^+/e^- transport distance for both low sulfur loading and high sulfur loading cathodes. With this in mind, almost same cycling performance and rate performance and rate performance are exhibited for $250 \text{ }\mu\text{m}$ electrode and $750 \text{ }\mu\text{m}$ electrode, corresponding to sulfur loading of 2 mg cm^{-2} and 6 mg cm^{-2} . [5]

In conclusion, 3D-printing is a powerful technique that is capable of fabricating electrodes with high active material loading and improved ion/electron conductivity, and is thereby a promising method to improve the energy and power density of energy storage systems. And this technique offers a fresh viewpoint in designing high loading cathodes and will arise interest in other energy storage devices such as Li-ion batteries, Li-air batteries, etc.

- [1] X. Ji, et al, *Nat. Mater.* 8 (2009) 500-506.
- [2] Y. Wang, et al, *Adv. Funct. Mater.* 27 (2017) 1703140.
- [3] R. Yi, et al, *Adv. Funct. Mater.* 24 (2014) 7433-7439.
- [4] X. Gao, X. Sun et al, *Nano energy*, 2019, 56, 595-603
- [5] X. Gao, X. Sun, et al, *Advanced Funct. Mater.* 2019, 1806724

Novel Cell Configurations for High-Performance Superoxide Na-O₂ Batteries

Xiaoting Lin, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9,
Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

Na-O₂ batteries have recently attracted an extensive amount of attention due to their high round-trip energy efficiency, good reversibility, and clean chemistry. In addition, the low cost and abundance of sodium metal promise for a wider range of application in the future [1]. However, the poor cycling performance of state-of-the-art Na-O₂ batteries is one of the major challenges facing its future development. Firstly, the parasitic reactions caused by O₂ crossover on the Na metal is a serious problem [2]. Another source of detrimental side reactions on Na metal is induced by the O₂^{•-} migration due to the solution-mediated path for NaO₂ formation [3]. The O₂^{•-} radical is an aggressive species which is prone to attack electrophilic sites, their influence of promoting the parasitic reactions at the alkali metal is profound despite of comparably low in quantities. Except for the Na degradation caused by O₂/O₂^{•-} crossover, the implementation of sodium metal anode in Na-O₂ battery is associated with dendrite formation [4].

To address those issues, we successfully developed different Na-O₂ cell configurations. Firstly, a novel Na-O₂ battery using electrically connected carbon paper (CP) with Na metal as a protected anode is presented [5]. The CP demonstrates great effectiveness in addressing the fatal issue of cell short-circuiting by altering the growth and penetration of Na dendrite into dense Na deposition. On the other hand, the electrochemical potential of the CP gains a pseudo-equal potential when in contact with Na metal, and the side reactions induced by O₂/O₂^{•-} crossover preferentially occur on the CP instead of Na surface. The Na corrosion can be alleviated to some extent compared with the cells with bare Na, and a satisfying cycling performance at a low charge overpotential can be achieved. On the other hand, the O₂/O₂^{•-} crossover from the cathode to anode mainly through the dissolved O₂/O₂^{•-} in the electrolyte, and a physical barrier to retain the O₂/O₂^{•-} in the cathode would be effective. Therefore, for the first time, we successfully develop a novel hybrid solid electrolyte Na-O₂ battery composed of solid-state electrolyte (SSE) and coupled with a protected Na anode. The dense structure of SSE renders a great suppression on the O₂/O₂^{•-} crossing over, which simultaneously eliminates the Na degradation and decreases the capacity loss of the Na-O₂ cells. More importantly, the SSE is very stable against the strong oxidizer O₂^{•-}, enabling high discharge capacities with excellent Coulombic efficiency as well as stable cycling performance.

In conclusion, the importance of addressing challenges with regards to the Na dendrite growth and O₂/O₂^{•-} crossover were confirmed. Although more future work is needed to make Na-O₂ battery system commercially viable, the results presented here provide guidance for fabricating advanced Na-O₂ batteries with longer lifespans and better cycling performance.

Reference

- [1] H. Yadegari, et al., *Advanced Materials*, 28 (2016) 7065-7093.
- [2] X. Bi, et al., *Chemical Communications*, 51 (2015) 7665-7668.
- [3] I. Landa-Medrano, et al., *ACS Applied Materials & Interfaces*, 8 (2016) 20120-20127.
- [4] L. Medenbach, et al., *Energy Technology*, 5 (2017) 2265-2274.
- [5] L. Xiaoting, et al., *Advanced Functional Materials*, 0 1801904.

Interface Engineering of Sulfide Electrolyte-based All-Solid-State Lithium Batteries

Changhong Wang, Xueliang Sun

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9,
Canada,
E-mail: xsun9@uwo.ca

All-solid-state lithium batteries (ASSLBs) with intrinsic high safety and high energy density have gained intensive attention worldwide.¹ As a necessary component in ASSLBs, various solid-state electrolytes have been well developed over the past decade. So far, the ionic conductivity of solid-state electrolytes (particularly sulfide electrolytes (SEs), i.e. $\text{Li}_{10}\text{GeP}_2\text{S}_{12}$) is close to that of conventional liquid electrolytes.² However, the electrochemical performance of SE-based ASSLBs is hindered by the large interfacial resistances between electrodes and sulfide electrolytes. The main reasons are the significant interfacial reactions and the large contact resistance between electrodes and electrolytes in ASSLBs.³ In our research, taking the advantages of atomic/molecular layer deposition techniques, an interfacial layer with designed functionalities can be conformably interposed between the electrodes and solid-state sulfide electrolytes, aiming for overcoming the interfacial challenges. At the anode interface, a artificial solid electrolyte interphase (SEI) has been engineered to suppress the interfacial reactions and lithium dendrite growth between Li metal and $\text{Li}_{10}\text{SnP}_2\text{S}_{12}$, successfully enabling the use of Li metal in SE-based ASSLBs.⁴ At the cathode interface, the dual-shell interfacial architecture was built, in which the inner shell LiNbO_3 suppressing the interfacial reactions, while the outer shell $\text{Li}_{10}\text{GeP}_2\text{S}_{12}$ providing an intimate electrode-electrolyte contact. As a result, the dual-shell structured LGPS@LNO@LCO cathode exhibits a high initial specific capacity of 125.8 mAh.g^{-1} (1.35 mAh.cm^{-2}) with an initial Coulombic efficiency of 90.4% at 0.1 C and 87.7 mAh.g^{-1} even at 1C. Furthermore, the kinetics of interfacial ion transfer was also revealed for the first time, which puts the new insight into the interfacial challenges of ASSLBs. These works not only demonstrate new strategies to enable high-energy-density ASSLBs but also provide new insights into the interfacial challenges of ASSLBs.

References:

1. Y. Zhao, K. Zheng and X. Sun, *Joule*, 2018, 1-22.
2. A. Manthiram, X. Yu and S. Wang, *Nat. Rev. Mater.*, 2017, **2**, 16103.
3. C. Wang, Q. Sun, Y. Liu, Y. Zhao, X. Li, X. Lin, M. N. Banis, M. Li, W. Li, K. R. Adair, D. Wang, J. Liang, R. Li, L. Zhang, R. Yang, S. Lu and X. Sun, *Nano Energy*, 2018, **48**, 35-43.
4. C. Wang, Y. Zhao, Q. Sun, X. Li, Y. Liu, J. Liang, X. Li, X. Lin, R. Li, K. R. Adair, L. Zhang, R. Yang, S. Lu and X. Sun, *Nano Energy*, 2018, **53**, 168-174.

Phase Inversion: A universal method to create high-loading porous electrodes for energy storage devices

Xiaofei Yang, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

High areal active materials loading and high active materials utilization are two critical parameters to ensure the energy storage devices with high energy density. Unfortunately, preparing ultra-high active materials loaded electrodes, such as S/C electrodes with high sulfur loading up to 10 mg cm^{-2} are still limited by the surface tension based on traditional blade-casting technique, especially for the nano-materials with high specific surface areas and large pore volumes. [1] For instance, Ketjen black (KB), a kind of commercially available 0-D carbon, is one of promising candidates due to its low cost, high conductivity, nano sized nanoparticles and abundant pores distribution. However, one problem is that it is extremely hard to bundle the original KB-based electrode materials together with achieving high loading of active materials, and the electrochemical performance of the thick electrode is not satisfied either. [2] Hence, how to solve aforementioned issue is of significance for propelling the practical application of high-energy-density energy storage devices. In the first work, we for the first time, proposed a phase inversion method for fabricating tri-continuous structured electrodes via a simple, convenient, low cost and scalable process. During this process, the binder networks, electron paths and ion channels could be separately interconnected, which simultaneously achieve excellent binding strength and ion/electron conductivity. This is verified by constructing electrodes with sulfur/carbon (S/C) and $\text{Li}_3\text{V}_2(\text{PO}_4)_3/\text{C}$ (LVP/C) nanoparticles, separately delivering 869 mA h g^{-1} at 1 C in Li-S batteries and 100 mA h g^{-1} at 30 C in Li-LVP batteries, increasing by 26% and 66% compared with the traditional directly drying ones.[3] In addition, in the second work, taking flexible and shapeable S/C electrodes as a model, sulfur loadings 24 mg cm^{-2} based on 30 nm S/C particles are, for the first time, successfully achieved via phase inversion method. It can deliver an ultra-high areal capacity of $27.1 \text{ mA h cm}^{-2}$ and a capacity retention of 64.1% after 100 cycles with low electrolyte to sulfur ratio (E/S) of $5.3 \mu\text{L mg}^{-1}$. In addition, flexible and shapeable Li-ion battery electrodes based on 17 mg cm^{-2} 0-D $\text{Li}_3\text{V}_2(\text{PO}_4)_3$ (LVP) electrode could also be obtained, achieving excellent C-rate performance and cycling stability, which is 94 mA h g^{-1} at 5 C and nearly 100% capacity retention at 1C during 100 cycles. [4]

In conclusion, phase inversion, as a universal and scalable method to create high-loading porous electrode with excellent electrochemical performance, paving the way for practical application of high-energy density energy storage devices.

- [1] X. Yang, X. Sun et al, *Electrochemical Energy Reviews* 1 (2018) 239-293.
- [2] D. Lv, et al, *Advanced Energy Materials* 5 (2015) 1402290.
- [3] X. Yang et al, *Advanced Functional Materials* 26 (2016) 8427-8434.
- [4] X. Yang et al, *Nano energy*, 56 (2017), 595-603.

Optimizing porous transport layer structures for polymer electrolyte membrane electrolyzers using stochastic and pore network models

Jason K. Lee

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
leekeonh@mie.utoronto.ca

Aimy Bazylak

Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
abazylak@mie.utoronto.ca

ABSTRACT

Optimizing the counter two-phase flow of liquid water and oxygen gas in the anode porous transport layer (PTL) is vital for improving the performance of the polymer electrolyte membrane (PEM) electrolyzer. At the anode catalyst layer, reactant liquid water is consumed while oxygen gas is generated and transported through the PTL towards the anode flow channel. In this work, we studied the effect of PTL structure on the gas saturation and the two-phase permeability of the PTL using stochastic and pore network models. The stochastic model was used to generate PTLs with varied porosity and titanium powder diameter, and the pore network model was used to simulate the drainage of liquid water in the PTL. Next, we determined the gas saturation and permeability of liquid water following the gas invasion process. We observed that an increase in the titanium powder diameter led to an increase in the PTL pore size and thereby an increase in the gas saturation. In addition, the PTLs with higher porosity exhibited greater quantities of gas saturation.

Pt Single Atom Electrocatalysts by Atomic Layer Deposition for Fuel Cell Related Catalytic reactions

Lei Zhang, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

Platinum is a key component of the catalysts in a proton-exchange membrane fuel cell (PEMFC), which exhibited good oxygen reduction reaction (ORR) activity. Notwithstanding the superior properties of Pt catalysts, the low abundance of Pt, and subsequent high cost have created a major barrier for the large-scale commercialization of Pt catalysts in PEMFCs. One effective route for reducing the cost of Pt is to increase the atom utilization efficiency by reducing the particle size to clusters or even single atoms catalysts.[1] Herein, we will demonstrate our recent approaches about the preparation of Pt single atom catalysts by atomic layer deposition method (ALD). Nearly 100% Pt single atom catalysts were successfully deposited on N-doped carbon nanotubes. Pt-Pd single atom alloys were prepared through ALD method and applied in electrochemical oxygen reduction for the first time. We also successfully prepared Pt-Ru dimers on NCNTs through ALD process. All these different types of single atom catalysts exhibited superior activity in ORR and HER reactions.

In conclusion, we demonstrate a series types of Pt single atoms, including high-quality Pt single atom, Pt-Pd single atom alloy and Pt-Ru dimer structures. The results show that the formation of single atom structure can effectively improve the electrochemical performance. It is also worth to mention that we have first introduced ALD technique in the single atoms. It is believed that our study opens up a new avenue of developing new types of Pt-based catalysts for fuel cells and brings new understanding about catalytic performances of single atoms.

[1] L. Zhang, X. Sun et al, *Energy Environ. Sci.* **2019**, <http://dx.doi.org/10.1039/C8EE02939C>

Ultra-stable anode interface performed by incorporating LiF in Li₆PS₅Cl-based sulfide electrolytes

Feipeng Zhao, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

All-solid-state Li-metal batteries (ASSLMBs) have been attracting increasing attentions because of their high specific energy density and advantage of safety compared with conventional liquid-based Li-ion batteries (LIBs).[1] Solid-state electrolytes (SSEs) play a crucial role for the ASSLMBs performance. Sulfide-based SSEs exhibit a very competitive ionic conductivity compared with oxides, polymer-based, and other kinds of SSEs. In addition, sulfide-based SSEs can show a medium mechanical stiffness, intimate contact property with electrode materials, as well as negligible grain boundary resistance. Therefore, sulfide-based SSEs are viewed as one of the most promising SSEs candidates to commercialize high-performance ASSLMBs.[2] However, the poor Li anode/sulfide SSEs interface largely hinders the development.[3] Surface modification on Li can effectively prevent the interface side reactions and the Li dendrite growth to some extent in sulfide-based ASSLMBs.[4] But the improvement is still insufficient to meet the even higher requirement for current density and cut-off capacities.

Recently, LPS-based (Li₃PS₄) SSEs incorporating with LiI additives can show much stronger endurance (1 mA/cm², and 1 mAh/cm² @ 25°C) to suppress Li dendrite formation on the Li/ LPS-LiI interface.[5] Compared with LiI, LiF is a much more popular compounds that is widely employed as an essential component in stable SEI layers. What is more, according to the computational simulation result, LiF possess the largest stable voltage window (0 ~ 6 V vs. Li/Li+) among all common Li-containing compounds.[6] Therefore, LiF is expected to present better performance if can be introduced compatibly in sulfide SSEs. Herein, we take argyrodite Li₆PS₅Cl sulfide SSE (LPSCl) as the host material, and successfully incorporate LiF by replacing LiCl in the precursor to prepare 'LPSCl-xF' sulfide SSEs. The optimized 'LPSCl-70F' SSEs can deliver an ultra-stable Li plating and stripping behavior up to over 200 h under high current density of 12.74 mA/cm² and high specific capacity of 10 mAh/cm² @ 25 °C. This performance can be even comparable to the best reported performance in the liquid-electrolyte counterparts. Furthermore, the excellent LPSCl-70F/Li interface is applicable to realize high performance in Li/LPSCl-70F//LPSCl//LCO@LNO/LPSCl full batteries. Various characterizations illustrate that highly dense morphology and the high F-containing chemical composition of the in-situ formed interface layer are essential to reach the high performance.

[1] Kato, Y.; Hori, S.; Saito, T.; Suzuki, K.; Hirayama, M.; Mitsui, A.; Yonemura, M.; Iba, H.; Kanno, R. *Nature Energy* 2016, 1, (4), 16030.

[2] Manthiram, A.; Yu, X.; Wang, S. *Nature Reviews Materials* 2017, 2, (4).

[3] Zhang, Z.; Shao, Y.; Lotsch, B.; Hu, Y.-S.; Li, H.; Janek, J.; Nazar, L. F.; Nan, C.-W.; Maier, J.; Armand, M.; Chen, L. *Energy & Environmental Science* 2018, 11, (8), 1945-1976.

[4] Cheng, X. B.; Zhang, R.; Zhao, C. Z.; Zhang, Q. *Chemical reviews* 2017, 117, (15), 10403-10473.

[5] Han, F. D.; Yue, J.; Zhu, X. Y.; Wang, C. S. *Advanced Energy Materials* 2018, 8, (18).

[6] Zhu, Y.; He, X.; Mo, Y. *ACS applied materials & interfaces* 2015, 7, (42), 23685-93.

Investigating the impacts of two-dimensional liquid water evolution in the PEM fuel cell cathode gas diffusion layers on mass transport performance via equivalent circuit modeling and synchrotron X-ray radiography

Nan Ge

Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Canada
nan.ge@mail.utoronto.ca

Aimy Bazylak

Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Canada
abazylak@mie.utoronto.ca

ABSTRACT

Improving the mass transport performance of the polymer membrane electrolyte membrane (PEM) fuel cell is an important opportunity for cost reduction, as this strategy enables higher power output with the same amount of expensive catalyst. However, at high current density, water tends to condense and accumulate in the cathode gas diffusion layer (GDL), thereby hindering oxygen transport to the reaction site and leading to low mass transport performance. The trend in state-of-art fuel cell development has been to control the water saturation distribution since the fuel cell performance has been observed to be dependent on the water accumulation location. While previous studies focus on the relationship between bulk GDL structural properties and the liquid water transport, further optimization of the water saturation requires the understanding of liquid water evolution at higher spatial and temporal resolutions.

In this study, we investigated the impact of the in-operando liquid water evolution in a heterogeneous cathode GDL on mass transport performance by using an analytical equivalent circuit model. We first developed an equivalent circuit model, which incorporated the Tafel slope, to analyze our electrochemical impedance spectroscopy measurements. Next, the two-dimensional liquid water evolution pattern in the GDL was determined via in situ synchrotron X-ray radiography, where the GDL was segregated into four regions based on the pore structure (carbon fiber and microporous layer) and location (rib and channel). Specifically, this pattern illustrates the sequence of water accumulation and evaporation in the GDL. Finally, by matching the modeling result and the visualization data, we found that the mass transport performance was least sensitive to water accumulation in the carbon fiber-rib region, but most sensitive to water accumulation in the microporous layer-channel region.

Local and electronic structure study of black phosphorus by XAFS

Weiham Li^{1, 2, #}, Minsi Li^{1, 2#}, Tsun-Kong Sham, ^{2 *} Xueliang Sun^{1*}

*Corresponding Author: Xueliang Sun, Tel: +1-519-661-2111 Ext 87759;
Tsun-Kong Sham, Tel: +1-519-661-2111 Ext 86341;
E-mail address: xsun9@uwo.ca (X. Sun); tsham@uwo.ca (T.K. Sham)

ABSTRACT

Black phosphorus (BP) has attracted intense research attention as one promising anode material for lithium-ion batteries and sodium ion batteries due to its high theoretical capacity of 2595 mA h g⁻¹. However, several challenges impede the practical application of black phosphorus, including volume expansion, degradation in the air and unsatisfied electronic conductivity. In the case of the challenges in the degradation and electronic conductivity, the local and electronic structure play a crucial role, which can be studied by X-ray absorption fine structure (XAFS).

Here, we report X-ray near edge spectra (XANES) and Extended X-ray absorption fine structure (EXAFS) studies of the local and electronic structure of oxidized BP and heteroatom-doped BP with a focus on the better understanding of the degradation and the band structure variation. At first, we for the first time apply synchrotron-based X-ray photoelectron spectroscopy (XPS), X-ray absorption near-edge structure (XANES) and scanning transmission X-ray microscopy (STXM) for the nanoscale chemical imaging of phosphorene degradation. The result shows clear picture of the degradation process of phosphorene under ambient condition and how the morphology affect the degradation process. Additionally, we use EXAFS combined with DFT calculation and R-space fitting to clarify the crystal structure of heteroatom-doped BP. This is the first time to show the clear structure of heteroatom-doped BP and how the dopant affect the electronic structure in detail.

Dual-layer Li_3PO_4 Coating Supported Ni-Rich Layered Cathodes for All-Solid-State Li-ion Batteries

Sixu Deng, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

All-solid-state lithium-ion batteries (SSLIBs) are promising candidates for electric vehicles due to their intrinsic safety characteristics and the potential to achieve high energy density [1]. Among the developed SSEs, sulfide-based inorganic solid-state electrolytes, in particular, have demonstrated high ionic conductivity (10^{-2} - 10^{-4} S cm^{-1}), which is almost comparable with liquid-based Li-ion electrolytes [2]. However, the state-of-the-art SSLIBs with sulfide-based SSEs still suffer from severe issues such as limited cycle life and rapid performance degradation, which are primarily related to the unstable interface between electrode materials and SSEs that are detrimental to Li-ion transport [3]. Some studies have reported the interfacial phenomenon between lithium metal oxide cathode materials and sulfide-based electrolytes, such as the space charge effect, the diffusion of transition metals into SSEs, the decomposition of sulfide SSEs to form the inactive interlayer, self-cracking of cathode materials from volume change during the lithiation-delithiation process, etc [4]. To suppress the side-reactions and stabilize the cathode/SSE interface, ionic conductive coating materials, such as LiNbO_3 , Li_3PO_4 , LiTaO_3 , have been developed for Li-ion cathodes to improve the electrochemical performance of sulfide-based SSLIBs [5-6].

In this study, we develop a novel dual-layer Li_3PO_4 coating supported Ni-rich $\text{LiNi}_{0.8}\text{Mn}_{0.1}\text{Co}_{0.1}\text{O}_2$ cathode (NMC) via atomic layer deposition for sulfide electrolyte based SSLIBs. The dual-layer Li_3PO_4 coating is proposed to (1) prevent the side-reactions between Ni-NMC and sulfide electrolytes and (2) reducing surface cracking of cathodes from volume change. The developed Li_3PO_4 coated NMC demonstrated significantly improved cycling capacity and prolonged battery life in SSLIBs. Meanwhile, the influence of the developed coating layer on interfacial reactions and phase evolution of cathodes has been further revealed by variously advanced characterizations, such as synchrotron-based X-ray characterizations.

- [1] J. C. Bachman, S. Muy, A. Grimaud, H.-H. Chang, N. Pour, S. F. Lux, O. Paschos, F. Maglia, S. Lupart, P. Lamp, L. Giordano, Y. Shao-Horn, *Chemical Reviews* **2016**, *116*, 140-162
- [2] E. Quartarone, P. Mustarelli, *Chemical Society reviews* **2011**, *40*, 2525-2540
- [3] B. Wu, S. Wang, W. J. Evans Iv, D. Z. Deng, J. Yang, J. Xiao, *Journal of Materials Chemistry A* **2016**, *4*, 15266-15280
- [4] Y. Seino, T. Ota, K. Takada, *Journal of Power Sources* **2011**, *196*, 6488-6492
- [5] P. Yan, J. Zheng, J. Liu, B. Wang, X. Cheng, Y. Zhang, X. Sun, C. Wang, J.-G. Zhang, *Nature Energy* **2018**, *3*, 600-605
- [6] K. H. Park, Q. Bai, D. H. Kim, D. Y. Oh, Y. Zhu, Y. Mo, Y. S. Jung, *Advanced Energy Materials* **2018**, *8*, 1800035

Nanomechanical Study of Atomic Layer Deposited Solid-state Electrolytes Using AFM

Maedeh Amirmaleki, Changhong Cao, Teng Cui, Yu Sun, Tobin Filleter
Department of Mechanical and Industrial Engineering
Email: m.amirmaleki@mail.utoronto.ca

Biqiong Wang, Yang Zhao, Xueliang Sun
Department of Mechanical and Materials Engineering
University of Western Ontario
London, Canada

Jaosn Tam
Department of Materials Science and Engineering
University of Toronto
Toronto, Canada

ABSTRACT

By increasing Li-ion batteries (LIBs) applications from consumer electronics to electric vehicles, athermanous attention has been directed toward improving the safety, energy density, and power density of LIBs. Unlike conventional liquid electrolytes in LIBs, solid-state electrolytes (SSE) are safe, non-flammable, and stable [1]. Miniaturized solid-state batteries (SSBs) embedded in the 3D architecture shown improvements in energy and power density by increasing the surface capacity [2]. However, the formation of microcracks and eventually mechanical breakdown due to the internal and extremal forces in the interface of SE/electrodes leads to reduce ionic conductivity and battery capacity drop. Furthermore, microcracks at can provide a pathway for Li dendrite growth and cause shortcut in the cell [3]. Therefore understanding the mechanical properties of SSEs are necessary for improving the battery performance [4]. Unfortunately, the knowledge of the mechanisms behind the mechanical breakdown is very limited and inconclusive. In this research, I studied the mechanical behavior of two promising amorphous SSEs, Lithium Tantalate (LTO) and lithium phosphate (LPO) thin films prepared using atomic layer deposition(ALD) method [5-6]. ALD is a powerful deposition technique to fabricate uniform, crack and pinhole free ultra-thin films for 3D structures [5-7]. Thin film SSE are deposited over supporting single layer graphene suspended over a holey structure. Mechanical behavior of free-standing ALD prepared SSEs studied using atomic force microscope (AFM) indentation technique. Different thicknesses of SSEs from 5nm to 30 nm are studied using this method in order to investigate the influence of the thickness on mechanical properties of SE films. Elastic behavior studies show that LTO is stiffer than LPO films at all thicknesses. Failure behavior of SSE films is also studied by indenting the films until failure. LTO showed brittle failure while plastic deformation was observed for LPO prior to failure. These results suggest that LTO can perform better as a SSE in order to suppress the crack formation at the SSE/electrode interface and improves the overall performance of the battery. The mechanical testing techniques used in this study can also be applied to other 2D and 3D battery materials at the nanoscale to reveal their potential for application in SSBs and ultimately help with materials selection and design in SSBs.

REFERENCES

- [1] N. Kamaya et al., Nat. Mater., 2011 ,10, 682. , [2] M. Létiche et al., Adv. Energy Mater., 2017,7,1601402., [3] Y. Suzuki et al., Solid State Ionics, 2015, 278,172. , [4] A. Sakuda et al., Sci. Rep., 2013, 3, 2261., [5] J. Liu et al., J. Phys. Chem., 2013, 117, 20260., [6] B. Wang et al., Nanotechnology, 2014, 25, 504007., [7] C.Bae et al., MRS Bull., 2011, 36, 887.

High Performance 3D Li Metal Anodes Enabled by Lithiophilic Cu Nanowire Networks

Keegan R. Adair, Xueliang Sun*

Department of Mechanical and Materials Engineering, University of Western Ontario, London, Ontario, N6A 5B9, Canada, E-mail: xsun@eng.uwo.ca (X.L. Sun)

ABSTRACT

The Li metal anode is considered as an ideal candidate for next-generation battery systems due to its ultra-high specific capacity (3860 mAh g^{-1}) and low electrochemical potential (-3.040 V vs. standard hydrogen electrode). However, the commercialization of Li metal batteries (LMBs) has been hindered due to the parasitic side reaction, large volume fluctuations, and dendrite growth associated with Li metal. Among the proposed methods of stabilizing the Li metal anode, 3D host structures have been shown to minimize the problematic volume fluctuations and enable improved rate capabilities at high current densities [1]. Nevertheless, 3D hosts with prestored Li often require a lithiophilic surface coating in order to enable molten Li infusion into the structure, which often adds additional manufacturing complexity [2-4].

Herein, we show the facile development of a lithiophilic 3D Cu nanowire host that can infuse molten Li into the structure [5]. It is discovered that the surface of the 3D host undergoes a structural transformation upon contact with molten Li, leading to the formation of Cu-Li alloy crystallites on the surface. The prestored Li metal host (3D Li@CuLi) shows outstanding electrochemical performance in carbonate-based electrolyte and can display more than 200 cycles in symmetrical cells at a high current density of 10 mA cm^{-2} . The plating/stripping mechanism of the 3D Li@CuLi is explored and characterized at different stages of cycling. Furthermore, the 3D Li@CuLi host was coupled with LiFePO_4 cathodes, showing excellent cycling stability for over 400 cycles at a C-rate of 2C with negligible capacity fade. This work displays a scalable approach towards the fabrication of high performance 3D Li metal hosts for next-generation battery systems and makes significant steps towards LMBs for fast charging electric vehicle applications.

- [1] C.P. Yang, Y.X. Yin, S.F. Zhang, N.W. Li, Y.G. Guo, *Nat. Commun.*, 6 (2015) 8058.
- [2] Y. Zhang, W. Luo, C. Wang, Y. Li, C. Chen, J. Song, J. Dai, E. Hitz, S. Xu, C. Yang, Y. Wang, L. Hu, *Proc. Natl. Acad. Sci. U. S. A.*, 114 (2017) 3584-3589.
- [3] L. Wang, X. Zhu, Y. Guan, J. Zhang, F. Ai, W. Zhang, Y. Xiang, S. Vijayan, G. Li, Y. Huang, G. Cao, Y. Yang, H. Zhang, *Energy Storage Materials*, 11 (2018) 191-196.
- [4] C. Yang, L. Zhang, W. Luo, B. Liu, Y. Li, S. Xu, T. Hamann, D. McOwen, J. Dai, W. Luo, Y. Gong, E. Wachsman, L. Hu, *Proc. Natl. Acad. Sci. U. S. A.*, 115 (2018) 3770-3775.
- [5] K.R. Adair, M. Iqbal, C.H. Wang, Y. Zhao, M.N. Banis, R.Y. Li, L. Zhang, R. Yang, S.G. Lu, X.L. Sun, *Nano Energy*, 54, (2018) 375-382.

Engineering Analysis and Design

The Motion of a Curling Rock: Predictions using Ice Topography

Megan E.R. Balsdon, Jeffrey T. Wood

Mechanical and Materials Engineering
Western University
London, Canada
mbalsdon@uwo.ca

ABSTRACT

The sport of curling is played on an ice surface that is unique to the sport – water droplets are sprinkled onto flat ice to create raised ice pebbles. The curling rock has a mass of 19kg and is made of granite with a circular running surface with a diameter of 12cm and a thickness of 6mm. The rock is released down the ice with an initial translational and angular velocity, where an applied clockwise rotation will result in a lateral deflection or “curl” to the right, and a counter-clockwise rotation will curl to the left.

The mechanism of curl or lateral deflection has been discussed in the literature; however, current published theories do not align with experimental findings. Among the theories and models are the thin water-layer model (Shegelski, 1996), snowplow model (Denny, 2002), evaporation-abrasion model (Maeno, 2010) and the scratch-guiding theory (Nyberg, 2013). The scratch-guiding researchers completed a series of experiments on professionally prepared curling ice which included delivering smooth curling stones on pre-scratch ice using sandpaper. Rocks were released with an applied rotation, entering the pre-scratched ice where the rock’s path showing a distinct deviation along the scratch direction; however, rocks that were polished did not deviate on the pre-scratched ice. This experiment proved strong evidence that the roughness of the stone’s running band is crucial to the amount of curl in a game of curling; however, the roughness was not measured, nor was the depth of the scratches in the ice.

The only way to affect a rock’s trajectory after it has been released is by sweeping the ice surface ahead of its path. The effect of sweeping on a rock has been found to raise the ice temperature 1-3°C, straighten the projected curved path and increase the travelled distance by up to 7 feet. Recent anecdotal evidence has demonstrated that certain brooms can affect the rock in similar ways to the scratch-guiding theory, depending on the direction in which the sweeping occurs. Sweeping from the center of the running surface *with* the direction of rotation and curl is thought to increase the amount of curl whereas sweeping *against* or in the opposite direction of the curl is thought to decrease the amount of curl. The purpose of this study is to prove experimentally that these different broom materials are (in fact) scratching the ice, which will support the overall physics model, aiding in the prediction of a curling rock’s trajectory.

Curling ice preparation included scraping, pebbling and nipping by the ice technicians. Vinyl Polysiloxane (VPS) (Heraeus Flexitime® Medium, Heraeus Kulzer GmbH) was used to create replicas or ice moulds. Nine total conditions were replicated: normally prepared ice with nipped pebble, rock traversing the ice, and seven broom conditions. A competitive curler and athlete (male, 20 years, 250lbs, 6’1”) was asked to sweep as hard as possible as if it were a draw shot (slow translational velocity), in between two pylons. Following this amount of sweeping, VPS was immediately placed and then pressed into the ice. This process was repeated for the number of conditions described above. Samples were then investigated through an optical microscope (Nikon Eclipse L150), and then with an optical profiler (Sensofar).

Preliminary results showed that scratches in the pebbled ice ranged from 1.1 –15.3µm between the various broom materials. The brooms that created the deepest overall scratches measured mean scratch depths of 4.58µm, 7.64µm, and 8.13µm, for the IcePad Pro, hair broom and IcePad TourElite, respectively, where rock scratches showed only a mean depth of 2.7µm. Scratch width and scratch density will be measured to expand the results. The measured depth of the rock’s scratches provide concrete data for the predictive physics model where one of the main factors affecting the magnitude of lateral deflection is the roughness of the rock’s running surface. Further analysis will be performed in the coming months, adding strength to the study, both experimentally and theoretically.

Denny, M. (2002). Curling rock dynamics: Towards a realistic model. *Canadian Journal of Physics*, 80(9), 1005–1014

Maeno, N. (2010). Curl mechanism of a curling stone on ice pebbles. *Bulletin of Glaciological Research*, 28, 1–6

Nyberg, H., Alfredson, S., Hogmark, S., & Jacobson, S. (2013). The asymmetrical friction mechanism that puts the curl in the curling stone. *Wear*, 301(1-2), 583–589.

Shegelski, M. R. A., Niebergall, R., & Walton, M. A. (1996). The motion of a curling rock. *Canadian Journal of Physics*, 74, 663–670.

An experimental cooling tower model for performance evaluation and 3D airflow measurement

Lisa Clare, Marc Secanell, and Morris R. Flynn
of Mechanical Engineering
University of Alberta
Edmonton, Canada
eclare@ualberta.ca

Brad Vickers
Engineering & Estimating
International Cooling Tower
Edmonton, Canada

ABSTRACT

Cooling towers are an effective, energy-efficient, and comparatively compact means for industrial facilities and power plants to reject waste heat generated as a result of a chemical process or thermodynamic cycle. In industry, cooling towers are typically sized and rated using either a 0D or a 1D model, often the semi-empirical Merkel method or one of its variants. Although this simple approach has been used extensively in industry since the 1940s, it cannot account for effects which present themselves in two or three dimensions. These include inlet effects, nozzle arrangements, maldistribution, and the effects of structural features on air and water distribution. CFD-based models may be well-suited to investigating these effects, and applying these models to numerical optimization promises to produce cooling tower designs with lower capital and operating costs while simultaneously achieving improved performance. CFD-based models however require validation to ensure that appropriate closures are applied. *In operando* data from industrial sites or from experimental models, generally yields only overall performance metrics such as cold water temperature or total pressure drop. Therefore, in literature, CFD simulations of cooling towers are usually validated against these overall performance metrics for lack of more detailed experimental data (Klimanek, *Arch Comput Methods Eng*, **20**, 61-109, 2013). These overall metrics however do not provide any details of the flow inside the tower, and therefore the predicted velocity profiles from CFD simulations cannot yet be considered to have been fully validated.

An alternative approach for CFD model validation is to employ a reduced scale model that is small enough to fit within a standard research laboratory but large enough to match the anticipated industrially-relevant flow regimes. This contribution describes the design process that led to the construction of a research experimental cooling tower that is capable of collecting not only cooling performance and pressure drop data, but also air velocity profiles by means of internal probe measurements in 3D space.

Further to an explanation of design and data collection considerations, we will also present an overview of relevant experimental measurements. Efforts to incorporate such experimental data into the development of CFD models and, by extension, a multi-objective optimization scheme that incorporates the inherently 2D/3D flows within and outside of cooling towers, will additionally be discussed.

Funding acknowledgment: International Cooling Tower and NSERC (through the CRD program).

DYNAMIC ANALYSIS AND DESIGN OPTIMIZATION OF AN OUTDOOR FITNESS MACHINE

Victor Hatai and Eric Lanteigne
Department of Mechanical Engineering
University of Ottawa
Ottawa Canada

Abstract— Outdoor exercise equipment are machines designed to provide accessible healthy living to communities, schools, and other public areas. The design of such equipment are currently based on best practices and do not reflect the capabilities of similar indoor machines, which have typically been extensively optimized for performance. This work presents a method for improving the force-displacement output of the planar four-bar mechanism used as the main structural component of a user-weight based outdoor exercise machine. The method uses a function generation approach based on closed-loop equations to generate the desired motion curve by controlling the increment of the angle displacement for each link at given precision points. Candidate solutions are then evaluated against a set of criteria in order to select the geometry whose output most closely matches the capabilities of a professional indoor machine.

Keywords- *fitness equipment; 4-bar mechanism optimization.*

I. INTRODUCTION

The main goal of this work was to revisit the design of an outdoor chest press machine produced by Active Playground Equipment to improve the user experience during the execution of the exercise. Improvements to the machine were achieved by taking into account anthropometric data of average users and market standards of comfort for motion and resistance. The outdoor chest press machine to be optimized is produced by Active Playground Equipment (APE) and is shown in Fig. 1(a). At its core, it consists of a four-bar mechanism as depicted in the kinematic diagram of Fig. 1(b).

There are several methods used for conducting a kinematic analysis of a mechanism. Russell et al. 2016 detailed a practical and straightforward method to obtain the kinematic equations of the mechanisms by summing its link vectors into a closed loop [1]. The equations associated with the closed-loop method define the displacement, velocity, and acceleration of the mechanism. The method adopted in this work did not require the application of Type Synthesis since the mechanism type had already been defined by the manufacturer. Indeed, in order to limit manufacturing costs, the optimized mechanism solution had to keep a design similar to the one already existent on the market.

With the predetermined mechanism type, it was possible to directly apply the dimensional synthesis method for the studied design problem. This method can be divided into three functional tasks: motion, path and function generation. The work of Wandler detailed each one of these functional tasks, while characterizing their output by using invariant descriptors to generate a solutions database [2]. One of the functional tasks mentioned, the function generation, was previously described by Kota and Chiou as problem for which a description of input and output motion and a set of constraints, such as kinematic and dynamic operational constraints, cost constraints, and reliability are given, and one needs to find One or more mechanism configurations that fulfill the prescribed functional requirements within the constraints [3]. In the case of the design problem studied in this work, the design requirements established by the manufacturer could be conveniently translated into kinematic operational constraints and used to define a relationship between the input and output motion of the machine. Thus, the problem definition was in parallel with Kota's, hence the design of the machine was regarded as a function generation functional task. The three most commonly used approaches for solving function generation problems: graphical, analytical and numerical optimization. The graphical approach utilizes two or three precision points to graphically establish the location of the ground joints and determine the link dimensions. This approach is a straightforward method that delivers a quick solution by sacrificing precision, mainly due to drawing error. Hence, it may have to be repeated several times before achieving suitable results.

The numerical optimization approach requires the formulation of an objective function that incorporates the primary design goals. This function is then minimized in order to find the global minimum, which represents the optimal design solution. This approach is flexible and allows the designer to apply any number of constraints, such as link dimensions or mechanism position, during the synthesis of the mechanism. However, Li et al. listed some of the drawbacks of this approach, which included: 1) elevated difficulty in the obtainment of a globally optimal solution, since no method can ensure that the obtained solution truly is a global solution, and 2) high dependency on the initial search point, which may lead to the search loop trapped in an optimal parameter subspace [4]. These made the numerical approach hard to apply it in this work,

because of the need to have the certainty of finding the global solution and the risk posed by high dependency on the initial search point, which, if wrongly defined, could cause a bias and prevent the finding of the globally optimal solution.

The analytical approach is based on mathematical representations such as closed-loop equations, granting it high precision to design mechanisms that do not present any structural errors at prescribed precision points. This approach has been successfully implemented in the synthesis of mechanism for strength training. In [5], the analytical approach using closed loops equations to synthesize a machine design is applied to produce desired resistance force. Although the method was under the framework of the numerical approach, this work demonstrate that it could have been applied under the analytical one. Both authors devised a force-generation synthesis method to design a bicep-curl weightlifting machine that produces a prescribed resistance curve. Both also used penalty functions that integrated relevant design requirements and were solved by the direct-search method. Their optimization routine was chosen due to the nonlinearity trait of the objective function formulated and the method's ability to handle a wide variety of design parameters.

Scardina developed a four-bar mechanism synthesis technique that satisfies force and torque requirements [7]. The method was based on a static force analysis and considered all links dimension and assembly angles as its design variables. This elevated number of design parameters granted the optimization routine more flexibility to devise a mechanism that better matches the desired force curve. However, the application of this method is unpractical for two main reasons: 1) it yields to an odd mechanism topology, and 2) it assumes massless links and negligible dynamic effects, a serious simplification for the study of an exercise machine.

Rundgren also developed a four-bar mechanism synthesis technique that considered all links dimension and assembly angles as its optimization variables [6]. However, his included a dynamic interaction of structural forces, thus reducing the odds of inaccurate results. However, the resulting mechanism presented a rather unpractical shape because of the method's disregard of the mechanism's topology as one of the free-choice design parameters. Despite the usefulness of the force-generation method, it results in non-premeditated machine shapes, which made the method less suitable to the design case studied in this work, given the importance of controlling the resulting mechanism shape. Alternatively, this analysis adopted a function-generation motion-oriented synthesis method, hereinafter also referred to as the optimization method. This method, contrarily to the force-generation one, allows for the input of the manufacturer's existing machine design as a free-choice design parameter, thus resulting in a predictable mechanism shape.

In this work, the outdoor chest press machine of Fig. 1(a) is optimized without considering a full anatomical study by analyzing the motion of the professional indoor chest press machine shown in Fig. 2 and using the resulting motion of Fig. 3(a) as a basis model.

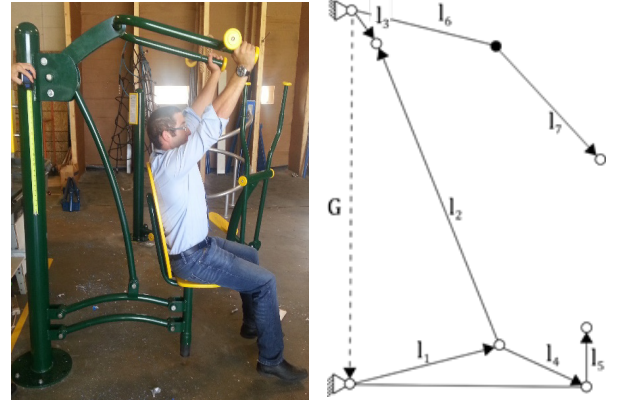


Figure 1. a) outdoor chest press machine to be optimized; and b) the kinematic model



Figure 2. Basis mechanism: professional indoor incline chest press machine produced by Hammer Strength

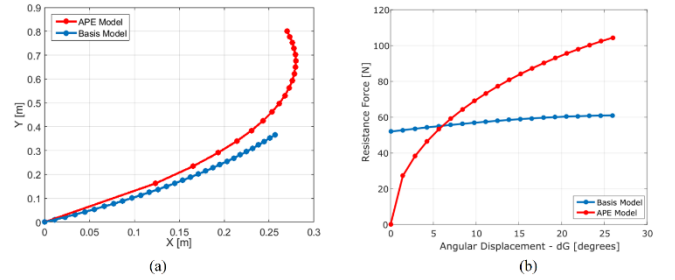


Figure 3. The (a) motion and (b) resistance comparison between the Active Playground Equipment model and Hammer Strength basis model

II. MECHANISM SYNTHESIS

The chest press machine can be reduced to a four-bar mechanism consisting of the lengths l_{1-3} as depicted in the kinematic diagram of Fig. 1(b). The handle link consisting of lengths l_{6-7} and the seat bars l_{4-5} were incorporated to the diagram for they present vital role in the optimization process of the equipment since they must be designed for user comfort.

A. Design constraints

The first design constraint considers the anthropometric data of an average person to dictate the dimensions of the handle and the seat's position. The average chest size of 31.2 cm (difference

between forward reach and chest reach) and the average mid-shoulder height of 53.3 cm were determined the relative x and y-position of the seat with regards to the handle link's extremity in the resting position [8]. In addition, the average popliteal height (from the sole of the foot to the crease under the knee between the upper and the lower leg) of 43.9 cm dictated the absolute y-position of the seat's bottom from the ground [9]. The handle bar length was taken from the basis model. This lead to the pre-optimized model dimensions shown in Fig. 4(a)

The second constraint involved the dimensional modifications of the outdoor chest press machine necessary to align its motion curve to that of the basis model. The motion curves of both the outdoor and indoor basis models were obtained by measuring the link and cable lengths of both devices and computing the handle bar positions at equally spaced intervals of handle bar angular displacements up to their maximum respective positions. Using the motion profile, the resistance curves were then produce using the 50th percentile user weight for the outdoor machine and a medium resistance setting for the indoor machine assuming that the force inputted by the user is perpendicular to the handle bar link. Only the first 26° of handle bar motion of the basis model was used in the analysis as this was determined the useful motion for the average user. The resulting motion and resistance curves are shown in Fig. 3.

Respecting the second constraint led to the upward and forward displacement of the pivoting point. Fig. 4(b) illustrates this displacement by overlaying the original and the modified model that incorporated link dimensions derived from the human factor data and the basis model.

The third constraint was a set of dimensional restrictions which prevented the synthesis of impractical mechanism shapes. For the purpose of this work, an unpractical mechanism was defined as having a shape that greatly differs from the original design since large design variations could lead to increased manufacturing costs. The dimensional restrictions were based on visual interpretation and initial results of the synthesis, and are given by:

1. $l_{1x} \leq 140$ mm - this restriction prevented the synthesis of a lengthy seat link.
2. Transmission Angle $\leq 180^\circ$ - this restriction prevented the synthesis of mechanism that presented a branch-defect, which means that the synthesized mechanism achieved the prescribed precision points without the need for disassembly, or a change of its original assembly configuration [1].
3. $l_3 > 60$ mm - this restriction prevented manufacturing complications.

B. Chest press machine kinematics and statics

The closed loop equations for the four-bar chest press machine in the j^{th} position are given by:

$$l_1 e^{i(\theta_1 + \phi_j)} + l_2 e^{i(\theta_2 + \alpha_j)} = l_3 e^{i(\theta_3 + \psi_j)} + G \quad (1)$$

where G is the ground link, θ_{1-3} represents the initial angular positions of links 1 to 3 and ϕ , α , and ψ are their respective angular displacements.

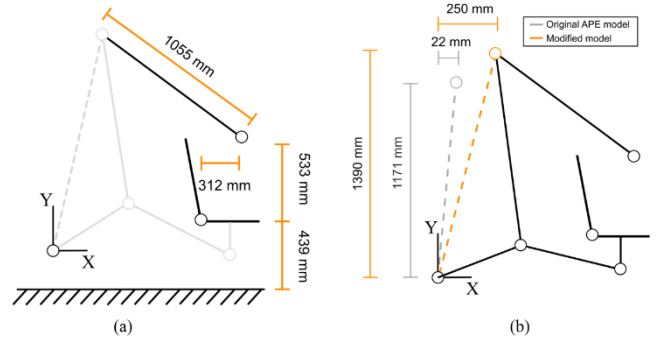


Figure 4. The (a) motion and (b) resistance comparison between the Active Playground Equipment model and Hammer Strength basis model

The function generation was chosen to calculate the optimized link dimensions, for it uses the input and output angle displacements as relevant variables and accommodates any given link dimension. Accordingly, the output angle was based on the angle displacement of the basis model, the input angle was the current angular displacement of seat link, and the stationary link was the accommodated link. Rearranging (1) in its rectangular form by applying the Euler's formula to the exponential term resulted in:

$$\begin{aligned} l_1 \cos \theta_1 \cos \phi_j - l_1 \sin \theta_1 \sin \phi_j + l_2 \cos \theta_2 \cos \alpha_j - l_2 \sin \theta_2 \sin \alpha_j - l_3 \cos \theta_3 \cos \psi_j + l_3 \sin \theta_3 \sin \psi_j &= G_x \\ l_1 \sin \theta_1 \cos \phi_j - l_1 \cos \theta_1 \sin \phi_j + l_2 \sin \theta_2 \cos \alpha_j - l_2 \cos \theta_2 \sin \alpha_j - l_3 \sin \theta_3 \cos \psi_j + l_3 \cos \theta_3 \sin \psi_j &= G_y \end{aligned} \quad (2)$$

Grouping real and imaginary terms, (2) becomes:

$$\begin{aligned} l_{1x} \cos \phi_j - l_{1y} \sin \phi_j + l_{2x} \cos \alpha_j - l_{2y} \sin \alpha_j - l_{3x} \cos \psi_j + l_{3y} \sin \psi_j &= G_x \\ l_{1y} \cos \phi_j - l_{1x} \sin \phi_j + l_{2y} \cos \alpha_j - l_{2x} \sin \alpha_j - l_{3y} \cos \psi_j + l_{3x} \sin \psi_j &= G_y \end{aligned} \quad (3)$$

The static equations for the outdoor chest press machine are given by:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -l_{1y} & l_{1x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & l_{2y} & -l_{2x} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -l_{3y} & l_{3x} & 0 & 0 & -(l_{6y} + l_{7y}) & (l_{6x} + l_{7x}) & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & l_{7x} & l_{7y} & 0 \end{bmatrix} \begin{bmatrix} F_{g1x} \\ F_{g1y} \\ F_{p1x} \\ F_{p1y} \\ F_{p3x} \\ F_{p3y} \\ F_{g3x} \\ F_{g3y} \\ F_{ux} \\ F_{uy} \end{bmatrix} = \begin{bmatrix} 0 \\ -W \\ -(l_{1x} + l_{4x})W \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (4)$$

where F_g represents the joint forces at the ground link connections and F_p represents the joints forces at the connections to link 2, F_u represents the user force, and W is the user weight.

C. Function generation synthesis

The dimensional synthesis (3) for three position points in its combined matrix form is given by:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ \cos\phi_2 & -\sin\phi_2 & \cos\alpha_2 & -\sin\alpha_2 & -\cos\psi_2 & \sin\psi_2 \\ \sin\phi_2 & \cos\phi_2 & \sin\alpha_2 & \cos\alpha_2 & -\sin\psi_2 & -\cos\psi_2 \\ \cos\phi_3 & -\sin\phi_3 & \cos\alpha_3 & -\sin\alpha_3 & -\cos\psi_3 & \sin\psi_3 \\ \sin\phi_3 & \cos\phi_3 & \sin\alpha_3 & \cos\alpha_3 & -\sin\psi_3 & -\cos\psi_3 \end{bmatrix} \begin{bmatrix} l_{1x} \\ l_{1y} \\ l_{2x} \\ l_{2y} \\ l_{3x} \\ l_{3y} \end{bmatrix} = \begin{bmatrix} G_x \\ G_y \\ G_x \\ G_y \\ G_x \\ G_y \end{bmatrix}$$

Cramer's Rule was used to solve for the unknown link lengths dimensions. The function generation synthesis method used to calculate the links dimensions requires three precision points that describe the mechanism motion. These precision points were defined by the angular displacement of the three non-stationary links: l_1 , l_2 and l_3 .

These precision points can be translated into the six following synthesis variables, considering that the angular displacements of the first precision point is zero:

1. The total angular displacement of the handle link: d_G
2. The total angular displacement of the seat link: d_B
3. The total angular displacement of the coupler link: d_A
4. The fractional angular displacement of the handle link: f_G
5. The fractional angular displacement of the seat link: f_B
6. The fractional angular displacement of the coupler link: f_A

The variables d_G and d_B correspond to the input and the output motions, respectively. The variable d_A is measured with respects to the joint that connects link l_1 and l_2 . The input motion (d_G) is the only known variable and is used as a reference to the fractional displacements. The fractional displacement refers to the amount of the total angular displacement that each link has travelled when d_G is at its midpoint ($f_G = d_G/2$). The variables f_A and f_B correspond to a fraction of the total angular displacement of d_A and d_B , and do not have a linear relation with f_G . In other words, at the mid-angular displacement of d_G , f_A and f_B do not necessarily correspond to 50% of d_A and d_B . The hypothetical kinematic diagram of Fig. 5 was used to establish the solution space of consistent system of equations, as it provided visual aid to approximate the maximum and the minimum for each variable, while respecting the predetermined design constraints. The limits of the boundaries were graphically represented by a set of locations of the connection joint of links l_1 and l_2 . The boundaries were limited by the graphical location of l_1 from the minimum x position of 0 to the first design criteria.

TABLE I. VARIABLE BOUNDARIES

Variable	Min	Max
d_A	-6.45°	-4.55°
d_B	1.8°	4.0°
f_A	0.46 d_A	0.48 d_A
f_B	0.26 d_B	0.34 d_B

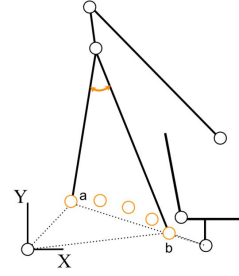


Figure 5. Hypothetical Kinematic Diagram

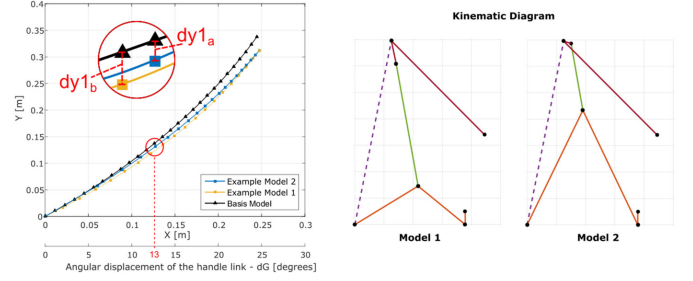


Figure 6. Position deviation and effect of variable f_B on motion curve with resulting models

The kinematics synthesis was programmed into MatLab routines. The first subroutine contains the synthesis method which iterates through the ranges of the variables (d_A , d_B , f_A , f_B) to output different mechanisms geometries using (4). The second subroutine performs a kinematic and a static force analysis to generate the motion and resistance curves for each geometry and to calculate the vertical deviations at the travel midpoint of 13°, denoted by dy_{1p} , and at the maximum travel of 26°, denoted by dy_{2p} . An example of the vertical deviation is shown in Fig. 6, where modified model example 1 and modified model example 2 correspond the lowest and highest values of variable f_B , respectively. It is easy to notice that model example 2 presents an unpractical mechanism shape. The choice of Model 2 for this example was not coincidental. It serves to show the side effects of conducting the optimization synthesis without considering the whole set of constraints.

The third subroutine refines the set of solutions by applying the design constraints. Upon completion of the synthesis routine, 27 candidate solutions that respected the design constraints were pre-selected. These candidates were then ranked according to the following criteria in order of importance:

1. A motion curve that best matches the objective motion curve
2. A resistance curve with the smallest slope
3. Manufacturing considerations – although all the candidate solutions' mechanism shapes were functional, some were unpractical. These mechanisms were filtered out based on the visual assessment by the designer. It is important to note that this criterion was used only to discriminate between two or more mechanisms that equally satisfied the aforementioned criteria.

As previously discussed, the adoption of resistance curve properties, as opposed to motion curve properties, as the main optimization constraint is a common alternative for mechanism synthetization of strength training equipment. This is due to the significant impact that a great resistance force variation during the execution of the motion has on the user's comfort. Force optimization considers force consistency as its main criterion, while maintaining the desired motion curve, whereas motion optimization considers mechanism shape, again while maintaining the desired motion curve. The kinematic diagrams of the optimized solution models are shown in Fig. 6. To achieve the optimized force solution, the geometrical practicality of the mechanism was compromised. The third design constraint (dimensional restrictions) had to be disregarded to generate the optimized resistance force curve shown in while matching the objective motion curve. Although the optimized force solution presented a more constant resistance curve and a better matching motion curve, its mechanism shape increases manufacturing costs since it requires longer link dimensions for the seat bar, which makes this solution impractical. The optimized motion solution presents a slightly superior force variation and maximum force compared to that of the original chest press mechanism. This is mainly due to the resulting dimension and position of the l_3 link, the driver link. In the optimized model, the l_3 is shorter and travels over a more horizontal path.

III. DISCUSSION

One of the limitations of the developed synthesis method is that it did not incorporate design constraints in the first step of the synthesis routine, namely, the function generation synthesis. Variable boundaries had to be defined before initiating the function generation capable of synthesizing functional mechanism solutions. Incorporating the design constraints would not require the variable boundaries to be determined empirically and would prevent human-induced faulty boundary definitions.

This work has also established that the optimized motion solution was the best suited for the design optimization of the chess press machine. Although the optimized force solution provided a constant force curve and a motion that was closer to that of the *Hammer Strength* machine, the generated machine design presented an odd geometry which would have required significant re-design and re-tooling. On the other hand, the optimized motion solution yielded a practical design, and motion curve that more closely resembles professional chest press machines and, while the optimized motion solution presented a higher maximum force, this does not represent a major drawback. The design modification enables the user to correctly activate the targeted pectoralis major muscle group which is considerably larger than the secondary anterior deltoid muscle group activated during the motion. This better replicates the motion of an incline chest press exercise and therefore targets the pectoralis major while limiting excessive involvement of the anterior deltoid [10]. More importantly, the motion-optimized machine targets the general population, and

thus having a lower resistance force at startup is positive since it makes it accessible to all levels of users, including the elderly.

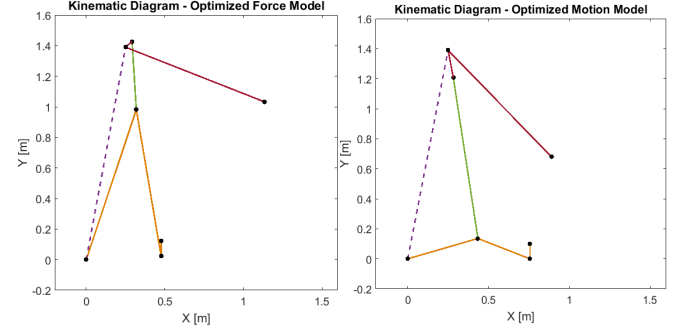


Figure 7. Kinematic Diagrams of the Optimized Force and Optimized Motion models

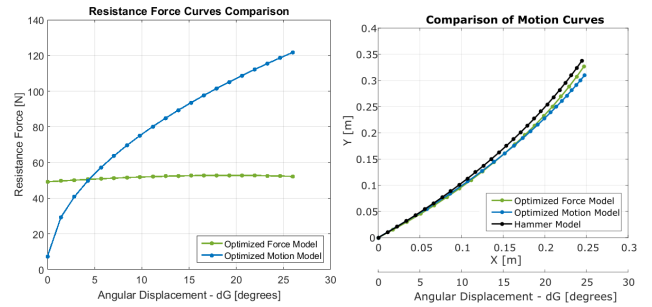


Figure 8. Comparison of the optimized resistance and motion curves for both models

IV. CONCLUSIONS AND FUTURE WORK

This work has presented the full development of the procedure applied to synthesize a four-bar mechanism to be used as the main structural component of an outdoor chest press machine. A key step in the method, the function generation synthesis, was applied to calculate the mechanism's link dimensions necessary for the mechanism to reach three prescribed precision points, which were dictated by the angular displacement of the driver link (handle bar) and follower link (seat bar). These precision points were taken from the motion curve of a professional indoor machine which was chosen as a model for its ergonomic standards.

The method developed in this work differs from previous works in that it uses the function generation synthesis to produce a prescribed mechanism shape by controlling the increment of the angle displacement for each link at each precision point. Usually, the function generation method is applied in order to achieve a certain angular displacement range without concern for a controlled angular increment of the links. Function generation synthesis method can only be used to solve a limited class of problems since it requires the definition of precision points, which often requires substantial intuition on the part of the designer. Also, in precision point syntheses, the number of optimization parameters must be equivalent to the number of points to fully define the problem and, when the number of precision points increases, the governing equations

present high nonlinearity and become more difficult to solve. The use of an optimization algorithm that contains an objective function could allow for the implementation of an increased number of constraints without increasing the complexity of the synthesis problem. An objective function, if well formulated, could be able to effectively conciliate constraints based on motion and force optimization simultaneously. These constraints would then be minimized by an optimization routine, such as the Hooke and Jeeves or Genetic Algorithm [11]. The integration of these two optimization constraints would result in an absolute mechanism synthesis method powerful enough to accommodate the complexity involved in the synthesis of a user-weight based exercise machine.

ACKNOWLEDGMENT

The authors would like to acknowledge the feedback provided by Active Equipment.

REFERENCES

- [1] Russell, K., Shen, Q., Sodhi, R.S., "Kinematics and Dynamics of Mechanical Systems: Implementation in MATLAB® and SimMechanics®", CRC Press, 2016.
- [2] Wandling, G.R., "Synthesis of mechanisms for function, path, and motion generation using invariant characterization, storage and search methods", Iowa State University, 2000.
- [3] Kota, S. & Chiou, S.-J., "Conceptual Design of Mechanisms Based on Computational Synthesis and Simulation of Kinematic Building Blocks", *Research in Engineering Design*, Volume 4, Issue 2, pp 75–87, 1992.
- [4] Li, X., Wei, S., Liao, Q., Zhang, Y. "A novel analytical method for function generation synthesis of planar four-bar linkages", *Mechanism and Machine Theory*, Volume 101, pp. 222–235, 2016.
- [5] Soper, R.R., "Synthesis of Planar Four-Link Mechanisms for Force Generation", Master's thesis, 1995.
- [6] Rundgren, B.T., "Optimized Synthesis Of A Force Generating Planar Four- Bar Mechanism Including Dynamic Effects", Thesis, 2001.
- [7] Scardina, M.T., "Optimal synthesis of force-generating planar four-link mechanisms", Thesis, 1996.
- [8] Woodson, W.E., "Human factors design handbook : information and guidelines for the design of systems, facilities, equipment, and products for human use", *McGraw-Hill*, 1981.
- [9] Helander, M., "Guide To Human Factors And Ergonomics", *London CRC Press*, 2004.
- [10] Howley, E. & Thompson, D., "Fitness Professional's Handbook: Sixth Edition 6th ed.", *Human Kinetics*, 2012.
- [11] Mehdigholi, H. & Akbarnejad, S., "Optimization of Watt's six-bar linkage to generate straight and parallel leg motion", *International Journal of Advanced Robotic Systems*, vol. 9, pp.11–16, 2012.

Experimental measurements of thermal distribution of belt drive system: Application to vehicle front engine accessory drives

Xingchen Liu, PhD Candidate

Department of Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
xc.liu@mail.utoronto.ca

Kamran Behdinan, Professor

Department of Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada
behdinan@mie.utoronto.ca

Abstract— This paper presents the procedure for measuring the thermal distribution of a belt drive system for the temperature prediction of the whole belt drive system. First, this work introduces the theories of heat generation, heat transfer within the system and a thermal prediction algorithm. Then, it expresses the experimental preparation, setup and measurement. Finally, it exhibits temperature results and comparisons between the experimental results and thermal predicted values of the belt drive system.

Keywords—belt drive; thermal measurement; thermal analysis; vehicle engine;

I. INTRODUCTION

The multi-pulley serpentine belt drive system has been commonly used in the engine system as a front-end accessory drive (FEAD, figure 1) for decades to transmit the crankshaft power to accessories like water pumps or alternators. Nowadays, potential materials such as fiber-reinforced plastic (FRP) are being investigated for use in reducing the pulley weight as well as its moment of inertia to improve the efficiency of the whole belt drive system. However, the high ambient temperature in the under-the-hood environment is a barrier for the application of this material because heat accumulates inside the material due to its low thermal conductivity, eventually causing the thermal deterioration of the pulley structure (Figure 2) [1]. The thermal deterioration of a belt drive system is becoming a growing challenge for researchers. It is necessary to study heat behaviors in the FEAD system under various operating conditions and perform experiments to measure the temperature distribution in the system. The work presented in this paper provides a method to measure temperatures at several locations under simulated operating conditions, with the purpose of validating a belt drive system thermal prediction model used to select appropriate thermal resistant FRP materials to avoid the overheating issue based on the engine operating conditions.



Figure 1. Car engine and its front end accessory drive belt transmission [2].



Figure 2. FRP structure failure due to overheating (Courtesy of Litens Automotive Group).

The available literature concerning thermal analysis of belt drive system is relatively poor. However, relevant research such as power loss and belt thermal analysis is abundant. There is some work related to the power loss of the belt drive system because it is highly related to the heat generation within the system. Gerbert [3] analyzed the belt movement and classified five types of power loss generation mechanisms. Manin *et al.* [2] then applied this method to a multi-pulley serpentine belt transmission system. Silva *et al.* investigated the belt hysteresis power losses generated by bending, stretching, shear, flank and radial compression of belt rubber [4]. Regarding the belt thermal analysis, Abe *et al.* [5] developed a thermal analysis model for a timing belt, taking belt deformation as the primary source of heat generation in the system. Merghache [6] provided an experimental and numerical study of heat transfer on the AT10 timing belt. Qin *et al.* have developed an analytical model for the friction damping of round clamp band joints, which can predict the energy dissipation caused by the friction contact between the joint components [7]. Moreover, Gerbert [8] developed an alternative static thermal model for a simple V-belt drive system in which heat is generated from both belt deformations and friction between the belt and pulleys.


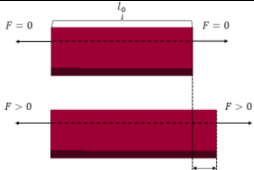

II. FUNDAMENTAL THEORY

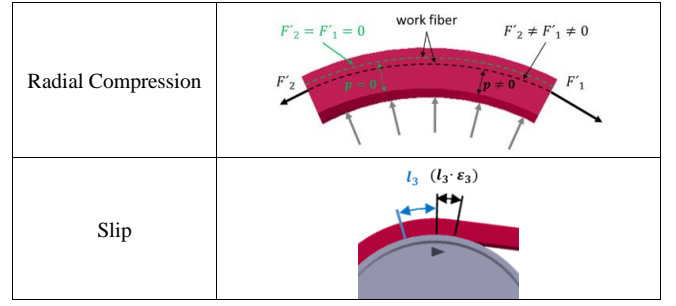
Thermal behaviors within the belt transmission system change dramatically according to operating conditions. However, like other general thermal analysis, it incorporates two sections: the thermal generation as well as the heat transfer and exchange within the system. It is necessary to investigate thermal behaviors within the system before measuring the temperature.

A. Thermal Generation within the System

The heat generation within the belt drive system is equal to the power losses because there is no external heat source in the system. The most prevalent power loss theory has been developed by Gerbert [3]. In general, the power loss is the difference between the input and output power within the system. However, there are five detailed forms of energy losses based on belt drive movement mechanisms as shown in Table 1 [9]. First, the belt bending power loss is from the energy dissipated when the belt is bent as a beam. It depends on the bending stiffness of the belt and radius of the pulley. Second, the belt stretching happens because of the different belt tension at the tight and slack side of each pulley, causing a periodical deformation at each belt section. It is proportional to the belt speed and preset tension. Third, the belt shear loss is due to the shearing deformation of the belt due to a torque transmission on the pulley. Dissipated energy is depending on the shear stiffness. Fourth, the belt radial compression is that the belt is compressed at the belt-pulley contact and energy is dissipated. It is related to the belt tension, radial stiffness and belt speed. Fifth, the belt slip loss is the frictional loss caused by the belt stretching on the pulley surface. This energy loss is dependent on the pulley torsional load, pulley radius, belt tension and pulley-belt coefficient of friction (COF). There is no dominant form of power loss, and generated heat flux of each power loss varies dramatically with different operating conditions and belt drive layouts.

TABLE I. POWER LOSSES WITH BELT CONCERNS

Power Loss Type	Schematic
Bending	
Stretching	
Shear	



a. Courtesy of CONTECS engineering services GmbH

With the purpose of simplifying the thermal calculation, they can be classified as two kinds of heat sources according to their heat generation locations during the belt high-speed movement and pulley rotations, namely belt internal power loss P_h and each belt-pulley contact surface power loss P_f . The belt internal power loss includes the first four types of power loss, while the belt-pulley contact surface power loss is the belt slip loss against the pulley. Therefore, in order to calculate these two heat sources P_h and P_f , two sets of parameters are required. The operation dependent parameters are belt speed and pretension, pulley-belt COF, speed and torque load of each pulley. And the operation independent parameters such as dimensions and thermal material properties of the belt and each pulley, belt longitudinal and bending stiffness. The equation set (1) developed by Manin shows the algorithm of power loss [2].

$$\begin{cases}
 P_{h,b} = \sum_j \sum_i \int_0^{H_j} \pi \left(a_i \frac{x}{R_j} + b_i \right) \left(\frac{x}{R_j} \right)^2 dx \cdot B \cdot R_j \cdot w_j \\
 P_{h,rc} = \sum_j \sum_i \left[\pi \left(a_i \frac{p_v}{E_z} + b_i \right) \left(\frac{p_v}{E_z} \right)^2 \right] \cdot H_i \cdot B \cdot R_j \cdot w_j \\
 P_h = P_{h,b} + P_{h,rc} \\
 P_f = \sum_j \frac{V \left(1 + \frac{c}{k_b R_j^2} \right)}{2c} (F_T - F_s)^2
 \end{cases} \quad (1)$$

B. Thermal Transfer and Exchange within the System

The heat conducts from the belt and pulley-belt engaged surfaces through interior structures of the belt system and finally dissipates to the environment through outer exposed surfaces to the environment. As for each pulley, it dissipates the heat to the environment, the pulley surface temperature is gradually reduced from outer to inner diameter. As for the belt, it maintains a uniform temperature because of its thin layer. Besides, there is also some heat exchange between the belt and each pulley. This is because the heat dissipation capacities of the belt and each pulley are different, the generated heat is prone to flow into the pulley or belt which can easily dissipate a large amount of heat. However, the different belt system operational speeds affect heat dissipation capacities of the belt and each pulley by changing the local heat transfer coefficient h , because it is highly related to the local airflow speed. At the same time, these speed changes also affect heat generation within the system. Thus, heat flow directions and flux within the belt are unique under each belt operation condition.

C. Thermal Prediction Model

A thermal prediction model of belt drive system is an algorithm to predict the temperature distribution of belt drive system. The thermal model has been developed based on thermal transfer and exchange. It is divided into three domains: pulley inner thermal analysis, belt inner thermal analysis and the pulley-belt heat exchange. For each component of inner thermal analysis, an equation describing the relationship between the heat inflow flux P and its contact surface temperature T is created. For each pulley-belt contact, an equation describing heat exchange flow due to the temperature gradient at pulley-belt engaged surfaces is established. By combining and solving these equations, the thermal distribution of the whole system can be acquired, and target temperatures such as pulley surface temperatures can be outputted. Figure 3 is the workflow of this thermal prediction model.

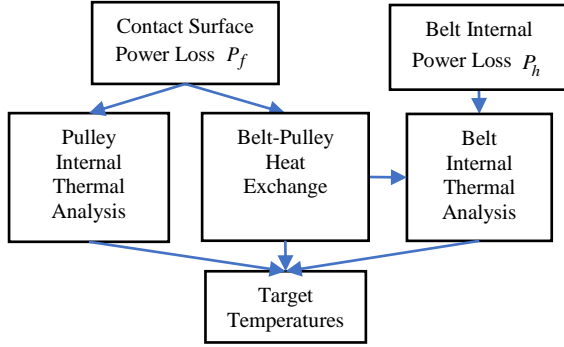


Figure 3. The flow chart of the methodology.

III. EXPERIMENTAL SETUP

The experiment is to measure temperatures at the outer surfaces of the belt and pulleys to validate the accuracy of thermal prediction algorithm by comparing temperature predictions under various operating conditions with corresponding experimental temperature distributions.

The preparation consists of four sections: the equipment and test sample preparation, the system layout and measured locations selection, as well as operating conditions determination.

A. Equipment and samples preparation

Some equipment is prepared to simulate the under-hood near engine environment. The engine simulator is designed to output the engine torque and speed as configured. A servomotor is used to generate a resistant load such as an alternator or a water pump. An enclosure chamber and thermal control system are also fabricated to simulate the ambient temperature in the engine under-hood environment. Three ambient temperature sensors and humidity sensors are to ensure the simulated environment is maintained and consistent. Moreover, a hub load sensor is to measure the belt tension. Three speedometers are to record speeds of crankshaft pulley (driver pulley) and test pulley (driven pulley), as well as the belt speed with the purpose of acquiring a belt slip data. Regarding temperature measurements, six infrared sensors are used to measure temperatures of important locations within the

belt system. A thermal image camera is used to measure temperatures of those locations as a second approach. Figure 4 shows a temperature measurement targeting on the outer surface of a test pulley through the thermal image camera in one operating condition.



Figure 4. The temperature on a thermal image camera

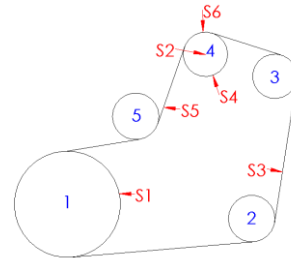
The sample selections prefer pulleys and belts widely used in the industry. The most prevalent 6K type of belt is selected in this test (figure 5). Pulley made of FRP A has been selected to cover the majority of used plastic pulleys in the industry.



Figure 5. The Belt cross-section (Courtesy of Litens Automotive Group).

B. System Layout and Measured Locations Selection

Figure 6 shows the belt drive system layout and its installation. Its design has to consider the following factors. Pulley 1 functions as a driver pulley to input loads into the system. Pulley 2 connects to the hub-load sensor to measure the belt tension. The speed of pulley 3 represents the belt speed because there is no load and belt slip on pulley 3. Pulley 4 acts as a resistant load pulley to output the transmission power. Pulley 5 connects to a tensioner to maintain the stability of the belt drive system. Figure 5 shows all six sensors (S1 to S6) and their target measured locations. The temperatures on these selected locations exhibit heat flows and their directions within the system, enabling us to optimize and validate the thermal prediction model. Furthermore, the belt drive layout also considers the sizes and locations of all six sensors to ensure all sensors can be mounted in the designed locations.



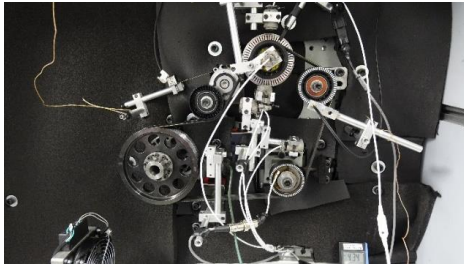


Figure 6. The Belt layout schematic and installation (Courtesy of Litens Automotive Group).

C. Operating Conditions Determination

The engine operations vary dramatically from idle to max revolution/torque conditions. It is necessary to select several acquisition points regarding pulley speed and revolution and ambient temperature to fully consider significant variance among operating conditions. Table II shows selected test points in three aspects and there are a total of 36 test configurations as each test is performed three times to reduce the measurement error. Similar to the speed and torque selections, the ambient temperature B is twice as high as temperature A to consider the effects of ambient temperature on several factors such as COF and material thermal conductivities.

TABLE II. OPERATION CONDITIONS ON PULLEY 4

Speed (RPM)	1400/2100/2800
Torque (N·m)	12/17
Ambient Temperature (°C)	30/60

IV. RESULTS AND COMPARISON

The experiments measure temperatures at six locations while tracking the operating condition parameters. The thermal model uses the operating inputs such as speeds and loads to calculate the power loss and then the temperature distributions. The validation is to compare the calculated temperatures with the ones acquired from the experiments. The following exhibits the analysis of two critical locations, the FRP pulley outer surface (location 4) and belt temperature (location 5). Location 4 has the highest temperature of the FRP pulley, and it determines whether the pulley will be below its thermal failure temperature. Location 5 has the inner belt surfaces and it determines whether the belt will be below its thermal failure temperature because the belt has a lower safety temperature than the pulley.

Table III shows the temperature rise of the FRP pulley under the ambient temperature A. The temperature rises along with the increase of pulley rotation speed and load.

TABLE III. TEMPERATURE RISE AT LOCATION 4 UNDER THE TEMPERATURE A

	Test1	Test2	Test3	Average
1400 RPM;12 N·m	30	27	25	27
2100 RPM;12 N·m	34	33	32	33

2800 RPM;12 N·m	41	41	38	40
1400 RPM;17 N·m	40	40	38	40
2100 RPM;17 N·m	51	48	46	48
2800 RPM;17 N·m	64	58	54	58

Table IV shows the temperature rise under a different ambient temperature B. The result shows the ambient temperature has a minor effect on the temperature rise in the system.

TABLE IV. TEMPERATURE RISE AT LOCATION 4 UNDER THE TEMPERATURE B

	Test1	Test2	Test3	Average
1400 RPM;12 N·m	25	26	22	24
2100 RPM;12 N·m	30	30	29	30
2800 RPM;12 N·m	38	38	37	38
1400 RPM;17 N·m	40	37	37	38
2100 RPM;17 N·m	49	45	44	46
2800 RPM;17 N·m	59	57	56	57

Table V shows the belt temperature rise under a different ambient temperature A. Comparing with the pulley temperatures, the belt temperature displays a lower temperature. It indicates that the frictional heat flux generated at pulley-belt engaged surfaces flows more into the belt than into the pulley, which is in accordance with the fact that the belt has higher thermal conductivity than the pulley. Considering that there is also a heat generation within the belt, the thermal dissipation ability of the belt is higher than of the pulleys. It is in accordance with the fact that the belt outer surfaces have higher air flow velocity than the pulley. The local air flows on all belt surfaces are uniform, while air flows on pulley surfaces are reduced with the decrease of pulley radius and its thermal dissipation ability is reduced as well.

TABLE V. TEMPERATURE RISE AT BELT UNDER THE TEMPERATURE A

	Test1	Test2	Test3	Average
1400 RPM;12 N·m	25	21	18	21
2100 RPM;12 N·m	30	27	25	28
2800 RPM;12 N·m	35	30	32	33
1400 RPM;17 N·m	33	30	28	30
2100 RPM;17 N·m	43	38	35	39
2800 RPM;17 N·m	53	50	45	49

Table VI shows the belt temperature rises under a different ambient temperature B. The result shows the ambient temperature does not affect the temperature rise.

TABLE VI. TEMPERATURE RISE AT BELT UNDER THE TEMPERATURE B

	Test1	Test2	Test3	Average
1400 RPM;12 N·m	22	23	19	21
2100 RPM;12 N·m	28	27	25	27
2800 RPM;12 N·m	35	35	34	34
1400 RPM;17 N·m	34	33	32	33
2100 RPM;17 N·m	42	42	40	42
2800 RPM;17 N·m	54	52	52	53

The temperature data from the experiment is reliable. It is averaged to reduce the measurement error. Therefore, it is eligible to provide a reference for the validation of the accuracy of the thermal prediction model.

Figure 7 shows the thermal prediction model gradually simulates the trend of temperature rise with the increase of speed and torque load from 1400RPM;12N·m to 2800RPM;17N·m, although there is some deviation in some conditions.

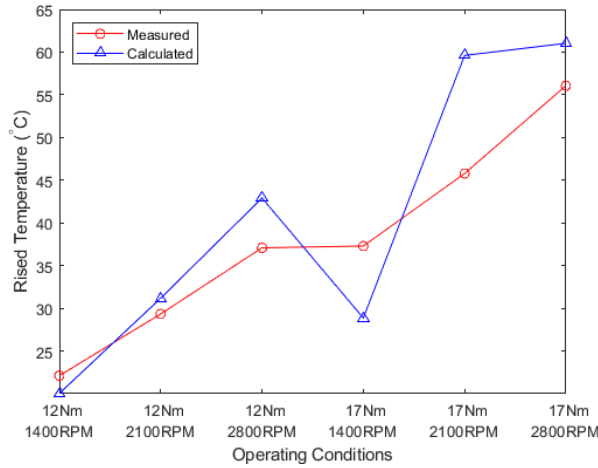


Figure 7. Relative temperature rise comparison of test pulley under the elevated ambient temperature A

Figure 8 shows the comparison between the calculated and measured relative temperature rise at the pulley inner surface under ambient temperature B. The calculated result can follow the trend of temperature rises, demonstrating that the thermal prediction model can work in a wide range of operating conditions.

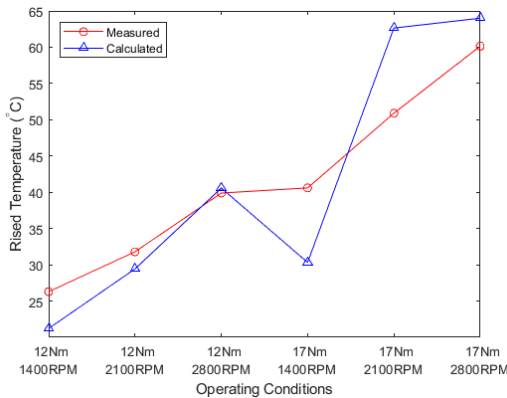


Figure 8. Relative temperature rise comparison of test pulley under the elevated ambient temperature B

Figure 9 shows the comparison between the calculated and measured relative temperature rises at the belt inner surface under ambient temperature A. The lower temperature predictions on the belt outer surfaces shows that it can predict flow directions and flux in the system which verifies that the thermal model could effectively calculate the partition of frictional heat at the pulley-belt engaged surface.

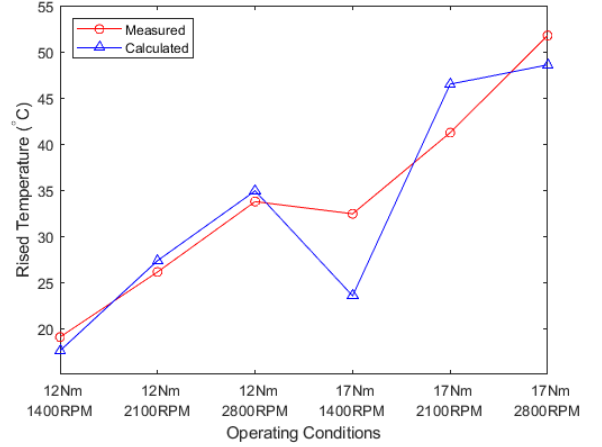


Figure 9. Relative temperature rise comparison of the belt under the elevated ambient temperature A

Figure 10 shows the comparison between the calculated and measured relative temperature rises at the belt inner surface under ambient temperature B. Like the pulley temperature prediction, this belt temperature prediction also is not influenced significantly by ambient temperature.

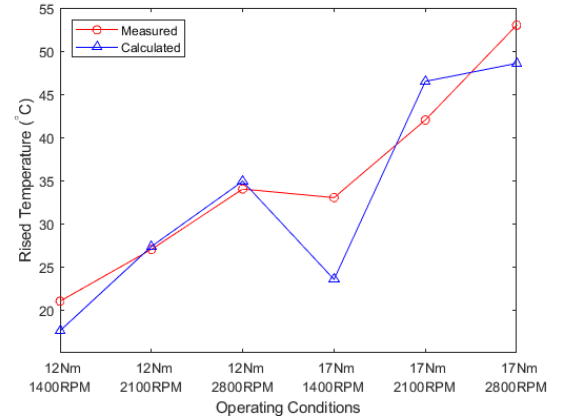


Figure 10. Relative temperature rise comparison of belt under the elevated ambient temperature B

The experimental results provide plenty of information for the thermal analysis of the belt drive system. The temperature rises of belt drive system are along with the increase of speed and torque in the operating conditions, which shows that the increase of heat generation is much more than the increase of heat dissipation at the condition of speed and torque growth. Although ambient temperatures affect the COF and thermal conductivities, it plays no significant role in the belt drive system.

The temperature result data is also used for the validation of the thermal prediction model. It proves that predicted temperatures generally match the experimental data, although there is some inaccuracy under some operating conditions. This is likely due to the simplification of the algorithm in the thermal model and further optimization will be performed to achieve accurate results.

V. CONCLUSION

The experiment has successfully measured the temperatures at various locations and expressed the thermal flows flux and directions by temperature difference at belt-pulley engaged surfaces. The work presents the theories and model of heat generation, transfer and exchanges in the system to describe the heat flow and the temperature distribution. In order to prove this model, this work has established a typical belt layout and designated operating conditions. It selects interesting locations of measuring temperature for the validation of thermal prediction model. It also includes procedures to select equipment to fully simulate the under-hood near engine environment. Last, the measured temperature data is exhibited, and comparisons between the experimental result and predicted values are performed. It proves the thermal prediction model can produce reasonable results. Future work will focus on the improvement of the accuracy of the thermal prediction model.

NOMENCLATURE

B	The belt width.
E_z	The rib axial stiffness.
H_i	The thickness of the rubber layer i .
R_j	The radius of the pulley j .
$P_{h,b}$	The power loss of bending.
$P_{h,rc}$	The power loss of radial compression.
P_h	The belt inner power loss.
P_f	The frictional power loss at bell-pulley engaged surfaces.
a_i, b_i	The experimental values for determination of the loss modulus.

i	The index number of rubber layer with total three layers.
j	The index number of pulley in a belt system.
p_v	The radial pressure.
x	The frictional power loss at bell-pulley engaged surfaces.
w_j	The angular rotational speed of pulley j with the unit of rad/s.
c	The belt longitudinal stiffness.
k_b	The radial compression stiffness of the belt.
V	The belt velocity.

ACKNOWLEDGMENT

The funding for this research from Mitacs Canada and Litens Automotive Group is gratefully acknowledged.

REFERENCES

- [1] H. Oka, Y. Tabuchi and H. Yazawa, "Development of plastic pulley for automotive air conditioner compressor," SAE Technical Paper, No. 2002-01-0603, 2002.
- [2] L. Manin, X. Liang and C. Lorenzon, "Power losses prediction in poly-v belt transmissions: application to front engine accessory drives," in International Gear Conference 2014, Lyon, 2014.
- [3] G. Gerbert, Traction Belt Mechanics. Göteborg: Chalmers Univ. of Technology, 1999.
- [4] C. Silva, L. Manin, R. Rinaldi, D. Remond, E. Besnier, M. Andrianoely, "Modeling of power losses in poly-V belt transmissions: Hysteresis phenomena (enhanced analysis)," in Mechanism and Machine Theory, vol. 121, pp. 373-397, 2018.
- [5] S. Abe, M. Tokoro, T. Yaegashi and K. Ogawa, "Thermal Analysis of Timing Belt," SAE Technical Paper, vol. 891988, 1989.
- [6] S. Merghache and M. Ghernaout, "Experimental and numerical study of heat transfer through a synchronous belt transmission type AT10," Applied Thermal Engineering, vol. 127, pp. 705-717, 2017.
- [7] Z. Qin, D. Cui, S. Yan and F. Chu, "Hysteresis modeling of clamp band joint with macro-slip," Mechanical Systems and Signal Processing, Vols. 66-67, pp. 89-110, 2016.
- [8] G. Gerbert, Heat in V-belt drives. Lund: Lund Technical University, 1981.
- [9] SimDrive 3D User Guide, Berlin: CONTECS engineering services GmbH, 2005.

EXPERIMENTAL AND NUMERICAL MODELLING OF A ONE-DEGREE OF FREEDOM NON-SMOOTH MECHANICAL SYSTEM

Solène Kojtych, Yann Colaïtis, Elsa Piollet, Alain Batailly

Department of Mechanical Engineering

Polytechnique Montréal

Montréal, Québec, Canada

solene.kojtych@polymtl.ca

Abstract— The aim of this paper is to highlight some numerical challenges occurring in the analysis of non-smooth mechanical systems by focusing on a single degree of freedom system with unilateral contact interface. An experimental setup is developed, its vibration behaviour is analyzed both with and without contact interface by means of forward and backward sweep tests. The parameters of the associated numerical model are carefully identified based on experimental measurements and observations. Numerical investigations are carried out assuming the system's response is periodic using the harmonic balance method. Numerical results are then confronted to experimental observations: frequency response curves and time responses are superimposed in order to assess the accuracy of the numerical model. Overall, a good agreement is obtained between numerical predictions and experimental measurements. Additionally, the sensitivity of accelerations computed with the harmonic balance method to the Gibbs phenomenon is highlighted.

Keywords- *non-smooth system; unilateral contact; harmonic balance method; experimental setup; Gibbs phenomenon*

I. INTRODUCTION

The modelling of non-smooth mechanical systems, particularly those featuring contact interfaces, is a challenge both from an experimental and a numerical standpoint. In a context where the design of sophisticated engineering applications now requires to account for inherent nonlinearities associated to contact [1, 2] or friction [3], the understanding of the physical phenomena at play is critical. Non-smooth mechanical systems are nonlinear systems characterized by the fact that their speed and acceleration fields are discontinuous. Because there is no equivalent to modal analysis in a non-smooth context, acquiring the intrinsic signature of a sophisticated mechanical non-smooth system is a very active field of research [4, 5]. Engineers and designers rely on several types of numerical methods [6, 7, 8] that may provide accurate results but that are also prone to

numerical sensitivity or instability. The development of a robust numerical strategy thus requires that attention be paid to both physical and numerical aspects [1, 9].

In this context, and as a first step towards the development of a robust numerical strategy dedicated to the analysis of non-smooth mechanical systems, this paper presents a numerical/experimental confrontation on an academic non-smooth system: a one-degree of freedom setup featuring a contact interface. Based on the assumption that the system's response is periodic, numerical investigations are carried out using the Harmonic Balance Method (HBM). Though fairly simple, the mechanical system of interest exhibits a rich dynamic behaviour calling for a careful analysis of both experimental observations and numerical results. The numerical sensitivity of the HBM is underlined, particularly when looking at acceleration fields. Nonetheless, the HBM advantageously provides a qualitative view of the system's dynamics that time integration techniques—also considered in this paper for the sake of validation—fail to provide.

The second section of the article describes the experimental setup. An analytic model for the underlying smooth system is proposed before its key mechanical parameters are identified. Frequency sweep tests are conducted for two distinct experimental configurations: with and without the contact interface. The third section of this paper briefly describes the theory behind the HBM. A comparison between numerical and experimental results for the two experimental configurations is presented and the differences are discussed. The influence of the number of harmonics taken into account in the HBM on the accuracy of the predicted solutions is investigated.

II. EXPERIMENTAL SYSTEM

A. Setup description

The experimental setup depicted in Fig. 1 aims at modelling a one-degree of freedom mass-spring system with contact-induced nonlinearity. The system consists of two main bodies: a sliding mass, excited along the x axis, and an impactor fixed on the bench structure and constraining the displacement

This research was supported by the Fonds de Recherche du Québec Nature et Technologies (FRQ-NT) and the Natural Sciences and Engineering Research Council of Canada (NSERC).

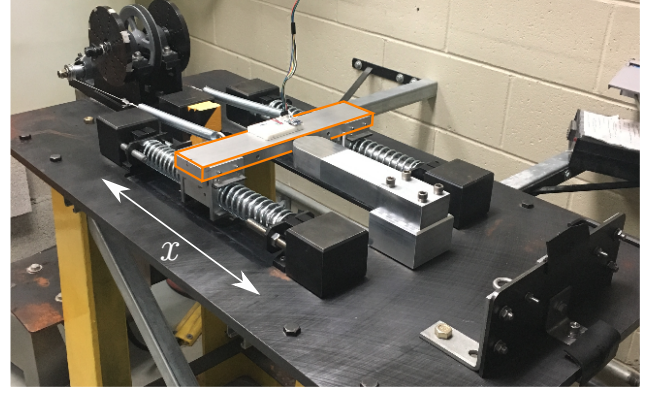
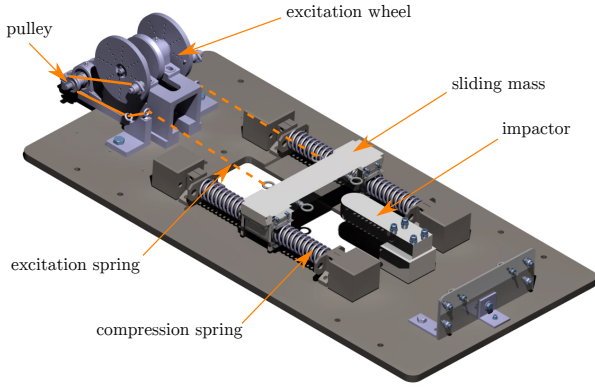


Figure 1. CAD and photo of the experimental setup.

of the mass. The impactor may be removed, thus making it possible to identify parameters of the system without contact interface. The impactor is made of steel, the gap between its extremity and the sliding mass at rest is $L = 10$ mm. The semi-cylindrical extremity of the impactor yields a line-to-surface contact interface. The sliding system is designed to meet specific criteria while the bench support and excitation system are adapted from [10].

Sliding system The sliding mass is made of steel, it is guided along the x direction by two parallel shafts fixed on the bench support. Sliding is made possible by two linear self-aligning bearings which also reduce friction with the shafts. Four compression springs of stiffness k_c placed along the guiding shafts allow the mass to oscillate. An initial compression of these springs ensures permanent contact with the mass in operation.

Excitation system A sinusoidal excitation is applied on the sliding system. A motor (not visible in Fig. 1) is fixed below the bench support, it provides a rotary motion which is transmitted to an excitation wheel. Two eccentric pins on this wheel are linked with two extension springs by means of a cable-pulley system. Finally, the sliding system undergoes a translation motion initiated by the springs. The magnitude of the excitation force depends on the stiffness k_e of the excitation springs and the excitation frequency can be chosen by the control panel of the motor and goes from about $0.85f_0$ to $1.15f_0$, where f_0 is the system's natural frequency. It is assumed that the global stiffness of the system is not significantly impacted by the excitation springs stiffness, because their contribution are very low compared to compression springs.

Measurement system A ± 16 g triple-axis accelerometer fixed on the mass is used to acquire accelerations and for transmission to a data acquisition system (DAQ). For cost-efficient reasons as well as academic purposes, the use of an ADXL326 accelerometer from Analog Devices and an Arduino UNO microcontroller (revision 3) as DAQ is considered in this study. The DAQ behavior is implemented through codes in C programming language, pre- and post-processing operations are carried out automatically by means of a dedicated Python 2 interface

developed specially for the experiment. Both the accelerometer and the DAQ were calibrated prior to the experiments. All codes are open source and available online [11]. The resolution of the measurement system is 0.05 g and the sampling frequency is set to 1000 Hz.

B. Model and parameters identification

Assuming the investigated mechanical system behaves as a one-degree of freedom system, its equation of motion is given by:

$$m\ddot{x} + d(\dot{x}) + kx = F(\omega, x, t) \quad (1)$$

where m and k are respectively the mass and the global stiffness of the system, $d(\dot{x})$ is a damping function to be determined, F stands for the external force, ω is the pulsation of the excitation and t is the time.

The aim of this section is to introduce and quantify sources of uncertainty for each of the system parameters m , $d(\dot{x})$, k and F in order to identify a suitable value for each of them. A precise estimation of F is particularly arduous to obtain as a non-negligible transverse motion of the extension springs is observed experimentally. For this reason, the characterization of the system follows a two-pronged approach: (1) values of m , d and k are first obtained without any excitation ($F = 0$) then, (2) F is identified.

Stiffness identification An uniaxial load frame is used to obtain an accurate value of the static stiffnesses of both compression and extension springs. The measurements confirmed the linear behavior of the springs. Averaged stiffnesses are summed up in Tab. I and the global stiffness of the system is given by:

$$k = 4k_c + 2k_e. \quad (2)$$

Due to the fairly low excitation frequency $f_e \approx 13$ Hz, a characterization of dynamical stiffnesses is assumed unnecessary.

Mass identification Although the mass of each component is known, a precise estimate of the mass of the sliding system

TABLE I. Parameters identified for the numerical model.

quantity	expression	value
compression spring stiffness	k_c	6 278 N/m
excitation spring stiffness	k_e	138 N/m
viscous damping coefficient	ξ	0.014
dry friction coefficient	μ	0.135
mass of sliding system	m	4.262 kg

is difficult to assess due to the unknown contribution of compression springs to the global sliding system's mass m . Indeed, the mass of compression springs is distributed between the steel mass and the guiding shaft, thus only bounds can be given for m :

$$3.718 \text{ kg} \leq m \leq 4.612 \text{ kg} \quad (3)$$

A more accurate estimate of the system's mass calls for a thorough analysis of the free response of the system.

Damping analysis Looking at the test bench depicted in Fig. 1, it is assumed that structural damping may essentially result from friction in linear bearings and from losses within the springs. Free vibration decay is thus analyzed: the sliding mass is moved to $x = 10 \text{ mm}$, it is then released and accelerations are recorded. This test is conducted three times for the sake of repeatability.

Three damping models are considered, see Tab. II. Raw acceleration data depicted in Fig. 2 are properly filtered and converted into displacements in order to identify the best parameters for each model (μ and/or ξ). The best matching is observed with the mixed damping model (viscous damping ξ and dry friction μ) which yields very good agreement with experimental data as shown in Fig. 3. Averaged values of the best fit parameters μ and ξ for the three tests are given in Tab. I.

The low value of ξ underlines that viscous damping is weak. For this reason, free-decay tests may be used to approximate the system's natural frequency $f_0 = 12.28 \text{ Hz}$ and thus determine its mass $m = 4.262 \text{ kg}$. Values of m , k_c , k_e , ξ and μ —which define the left term of (1)—used for the numerical model are reported in Tab. I.

TABLE II. Damping model equations ($g = 9.81 \text{ m/s}^2$ gravitational constant, $\text{sgn}()$ sign function, ξ viscous damping coefficient, μ dry friction coefficient, $\omega_0 = 2\pi f_0$ natural pulsation of the system).

damping model	equation
viscous damping	$d(\dot{x}) = 2m\xi\omega_0\dot{x}$
dry friction	$d(\dot{x}) = \mu mg \text{sgn}(\dot{x})$
viscous damping and dry friction	$d(\dot{x}) = 2m\xi\omega_0\dot{x} + \mu mg \text{sgn}(\dot{x})$

Frequency analysis A frequency sweep is conducted to determine the frequency response of the system as well as the

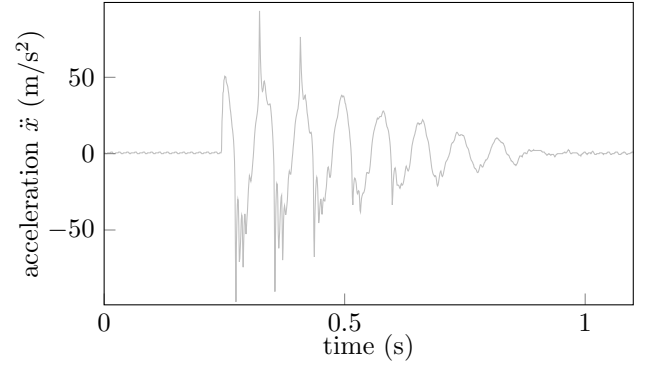


Figure 2. Free decay experimental data.

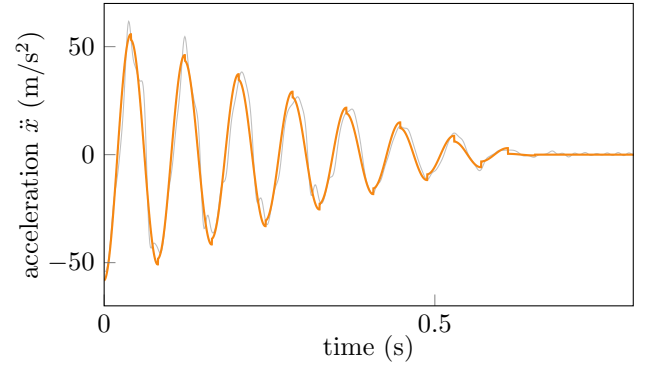


Figure 3. Simulation of damping model with viscous damping and dry friction (—) and experimental acceleration data (—).

expression of the right term of (1). The excitation frequency is updated step-by-step; for each increment the amplitude of the system's steady state response is recorded. Acquisitions last for 20 s and are recorded at a frequency rate of 1000 Hz.

In order to improve the experimental measurements accuracy, the frequency step is decreased to 0.05 Hz in the vicinity of the resonance peak. For each frequency excitation value, the maximal magnitude is plotted to draw an experimental response curve, see Fig. 4. Tests were made both for forward and backward frequency sweeps: the good match between the two responses underlines that the system without impactor does not exhibit any unwanted nonlinearity.

High-resolution video-tracking used during the sweep tests shows that the excitation springs undergo large frequency-dependent transverse displacements in operation as depicted in Fig. 5. As a consequence, it is not possible to establish any straightforward relation between the excitation springs stiffness and the magnitude of the force F . Due to the fact that the excitation is driven by the rotary motion of the motor, it is assumed that the right term of (1) may be written as:

$$F(\omega, x, t) = f_{\text{ext}}(\omega, t) = F(\omega) \sin(\omega t) \quad (4)$$

where $\omega = 2\pi f_e$ is the pulsation of the excitation and $F(\omega)$ the unknown force magnitude.

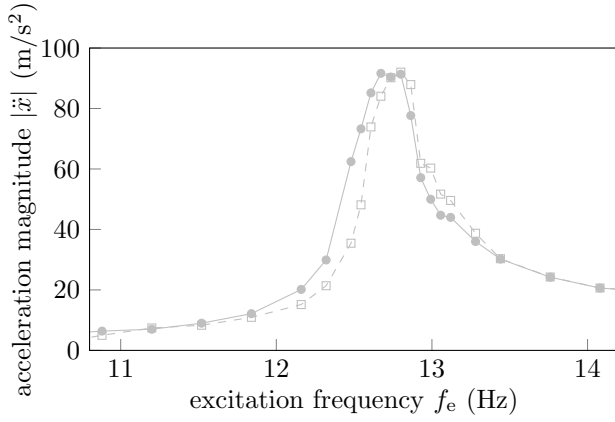


Figure 4. Experimental frequency response for the system without impactor: forward (●) and backward (□) frequency sweeps.

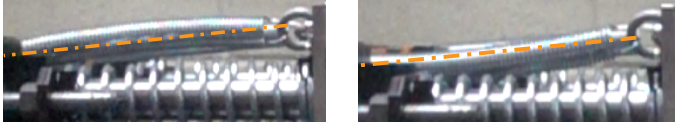


Figure 5. Extreme positions of an excitation spring undergoing transverse displacement in operation : (---) theoretical baseline.

Through a computation of the system's forced response using the previously identified parameters m , $d(\dot{x})$ and k , the values of $F(\omega)$ that corresponds to the experimentally measured amplitudes are retrieved for each considered excitation frequency. $F(\omega)$ related to both forward (●) and backward (□) frequency sweeps are pictured in Fig. 6. The similarity of the curves attest that the calculated excitation force can be considered as independent from the type of sweep.

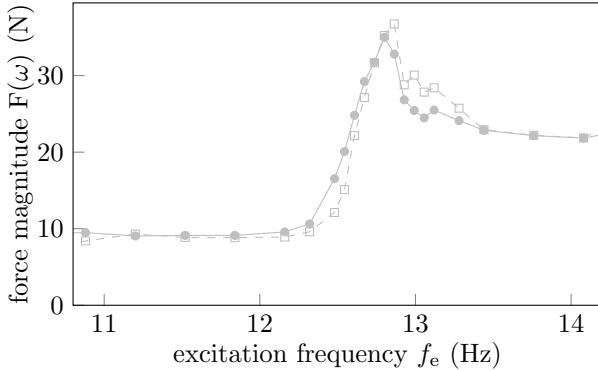


Figure 6. Magnitude of the excitation force obtained for forward (●) and backward (□) frequency sweeps.

C. Experimental results for system with impactor

Forward and backward sweep tests are conducted with impactor. Results in acceleration are depicted in Fig 7 and su-

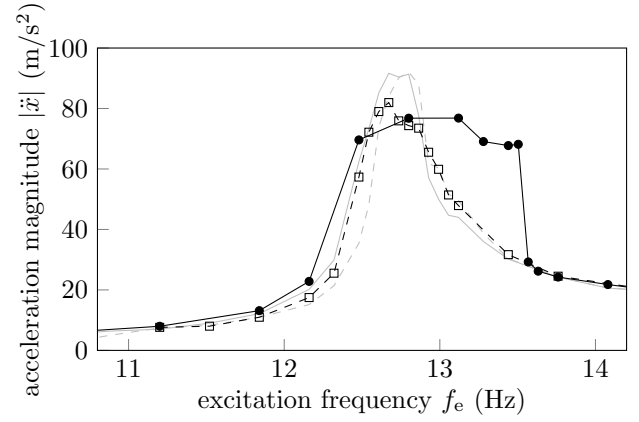


Figure 7. Experimental frequency response for the system with impactor: forward (●) and backward (□) frequency sweeps.

perimposed with experimental results without impactor. Curves are significantly different: for the forward sweep, the highest magnitude is located at a higher frequency than for the backward sweep and a jump in magnitude is visible at about 13.5 Hz. These behaviours are typical of the contact stiffening phenomenon and attest that a strong nonlinearity is induced by the contact interface. The slight decrease in magnitude for the three excitation points located immediately before the jump is assumed to result from a slight rotary motion of the sliding mass, due to unavoidable clearances in the linear bearings. A more in-depth analysis of the influence of this potential second degree of freedom goes beyond the scope of this study.

III. NUMERICAL MODEL

Several numerical strategies may be employed to solve (1). In particular, one may distinguish two classes of methods: (1) time integration and (2) frequency methods. Time integration is costly but advantageously does not rely on any assumption with respect to the system's behaviour; it is still commonly used today in the industry. To the contrary, frequency methods typically rely on an assumption of periodicity but may provide fast and reliable results. More importantly, frequency methods such as the HBM, when coupled to continuation algorithms, are able to provide a qualitative view of the system's dynamics which is key for understanding the underlying physical phenomena.

A. Theoretical presentation of the harmonic balance method

The HBM is widely used for the computation of nonlinear mechanical problems' periodic solutions [12, 13]. It relies on the assumption that the solution of a nonlinear system (1) undergoing a harmonic forcing (4) may be approximated by a truncated Fourier series up to the H -th harmonic:

$$x(t) \simeq \frac{a_0}{\sqrt{2}} + \sum_{j=1}^H [a_j \cos(j\omega t) + b_j \sin(j\omega t)] \quad (5)$$

where a_j and b_j are the unknowns Fourier coefficients related to cosine and sine terms. The same decomposition may also be used for the nonlinear forces f_{nl} (e.g. contact or friction) and the periodic external force f_{ext} :

$$\begin{aligned} f_{nl}(x, \dot{x}) &\simeq \frac{a_0^{nl}}{\sqrt{2}} + \sum_{j=1}^H [a_j^{nl} \cos(j\omega t) + b_j^{nl} \sin(j\omega t)] \\ f_{ext}(\omega, t) &\simeq \frac{a_0^{ext}}{\sqrt{2}} + \sum_{j=1}^H [a_j^{ext} \cos(j\omega t) + b_j^{ext} \sin(j\omega t)] \end{aligned} \quad (6)$$

From this point, all the Fourier coefficients may be gathered into the $(2H + 1)$ -dimensional vectors:

$$\begin{aligned} \tilde{\mathbf{x}} &= [a_0 \ a_1 \ b_1 \ \dots \ a_H \ b_H]^T \\ \tilde{\mathbf{f}}_{nl} &= [a_0^{nl} \ a_1^{nl} \ b_1^{nl} \ \dots \ a_H^{nl} \ b_H^{nl}]^T \\ \tilde{\mathbf{f}}_{ext} &= [a_0^{ext} \ a_1^{ext} \ b_1^{ext} \ \dots \ a_H^{ext} \ b_H^{ext}]^T. \end{aligned} \quad (7)$$

where the symbol (\sim) refers to frequency domain variables. By defining the Fourier basis vector \mathbf{T}_H related to (5):

$$\mathbf{T}_H = \left[\frac{1}{\sqrt{2}} \cos(\omega t) \ \sin(\omega t) \ \dots \ \cos(H\omega t) \ \sin(H\omega t) \right], \quad (8)$$

(5) and (6) are now read as:

$$\begin{aligned} x(t) &= \mathbf{T}_H \tilde{\mathbf{x}} \\ f_{nl}(x, \dot{x}) &= \mathbf{T}_H \tilde{\mathbf{f}}_{nl} \\ f_{ext}(\omega, t) &= \mathbf{T}_H \tilde{\mathbf{f}}_{ext}. \end{aligned} \quad (9)$$

Also, velocities and accelerations may be written as:

$$\begin{aligned} \dot{x}(t) &= \dot{\mathbf{T}}_H \tilde{\mathbf{x}} = \omega(\mathbf{T}_H \nabla) \tilde{\mathbf{x}} \\ \ddot{x}(t) &= \ddot{\mathbf{T}}_H \tilde{\mathbf{x}} = \omega^2(\mathbf{T}_H \nabla^2) \tilde{\mathbf{x}}. \end{aligned} \quad (10)$$

where ∇ is a derivative operator defined as:

$$\nabla = \begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & \nabla_j & \\ & & & \ddots \\ & & & & \nabla_H \end{bmatrix} \quad \text{and} \quad \nabla^2 = \nabla \nabla \quad (11)$$

with

$$\nabla_j = j \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{for } j = 1, \dots, H. \quad (12)$$

Using (9) and (10), the equation of motion (1) becomes:

$$m\omega^2(\mathbf{T}_H \nabla^2) \tilde{\mathbf{x}} + c\omega(\mathbf{T}_H \nabla) \tilde{\mathbf{x}} + k\mathbf{T}_H \tilde{\mathbf{x}} + \mathbf{T}_H \tilde{\mathbf{f}}_{nl} \simeq \mathbf{T}_H \tilde{\mathbf{f}}_{ext} \quad (13)$$

where $c = 2\xi\omega_0 m$ accounts for viscous damping. The difference between the left and right terms of (13) is related to the truncation of $x(t)$, see (5), it is called the residual $\mathbf{r}(\tilde{\mathbf{x}}, t)$ and can be expressed as:

$$\begin{aligned} \mathbf{r}(\tilde{\mathbf{x}}, t) &= m\omega^2(\mathbf{T}_H \nabla^2) \tilde{\mathbf{x}} + c\omega(\mathbf{T}_H \nabla) \tilde{\mathbf{x}} \\ &\quad + k\mathbf{T}_H \tilde{\mathbf{x}} + \mathbf{T}_H \tilde{\mathbf{f}}_{nl}(\tilde{\mathbf{x}}) - \mathbf{T}_H \tilde{\mathbf{f}}_{ext}(\omega, t). \end{aligned} \quad (14)$$

A Galerkin projection [6] on the Fourier basis \mathbf{T}_H is then applied to remove the time dependency of $\mathbf{r}(\tilde{\mathbf{x}}, t)$ so as to obtain a relation between the unknowns, a_j and b_j . This leads to the following set of nonlinear algebraic equations:

$$\begin{aligned} \mathbf{R}(\tilde{\mathbf{x}}, \omega) &= m\omega^2 \nabla^2 \tilde{\mathbf{x}} + c\omega \nabla \tilde{\mathbf{x}} + k\mathbf{I}_{2H+1} \tilde{\mathbf{x}} \\ &\quad + \mathbf{I}_{2H+1} \tilde{\mathbf{f}}_{nl}(\tilde{\mathbf{x}}) - \mathbf{I}_{2H+1} \tilde{\mathbf{f}}_{ext}(\omega) \end{aligned} \quad (15)$$

with \mathbf{I}_{2H+1} the identity matrix of size $(2H + 1)$. Equation (15) may be written in a more compact form:

$$\mathbf{R}(\tilde{\mathbf{x}}, \omega) = \mathbf{Z}(\omega) \tilde{\mathbf{x}} + \tilde{\mathbf{f}}_{nl}(\tilde{\mathbf{x}}) - \tilde{\mathbf{f}}_{ext}(\omega) = 0 \quad (16)$$

where $\mathbf{Z}(\omega)$ is the square linear dynamic stiffness matrix of size $(2H + 1)$ defined by:

$$\mathbf{Z}(\omega) = m\omega^2 \nabla^2 + c\omega \nabla + k\mathbf{I}_{2H+1}. \quad (17)$$

The nonlinear algebraic system (16) may be solved iteratively, for instance with a Newton-type or jacobian-free algorithms [12].

Nonlinear forces In general, the expression of the external forcing $\tilde{\mathbf{f}}_{ext}$ is known. Nonlinear forces $\tilde{\mathbf{f}}_{nl}$ however depend on the displacement and velocity amplitudes and are not known *a priori*. In order to address this issue, an Alternating Frequency/Time (AFT) procedure is adopted [14, 13]. It consists in the use of direct and inverse Discrete Fourier Transform (DFT) to evaluate the expression of nonlinear forces $f_{nl}(x, \dot{x})$ as well as their derivatives $\frac{\partial f_{nl}}{\partial x}$, $\frac{\partial f_{nl}}{\partial \dot{x}}$ in the time domain, then to obtain the frequency domain representations of $\tilde{\mathbf{f}}_{nl}$ and $\frac{\partial \tilde{\mathbf{f}}_{nl}}{\partial \tilde{\mathbf{x}}}$ (the latter is only required for a Newton nonlinear solver). The AFT procedure is illustrated in Fig. 8.

$$\tilde{\mathbf{x}} \xrightarrow{\text{DFT}^{-1}} x(t), \dot{x}(t) \longrightarrow f_{nl}(x(t), \dot{x}(t)) \xrightarrow{\text{DFT}} \tilde{\mathbf{f}}_{nl}(\tilde{\mathbf{x}})$$

Figure 8. Evaluation of nonlinear forces with AFT.

In this study a distinction is made between the normal f_n and tangential f_t components of the nonlinear forces, respectively related to contact and dry friction. An exponential penalty law is employed to model unilateral contact constraints:

$$f_n = \begin{cases} 0 & \text{if } \delta \leq -c_0 \\ \frac{f_0}{e^{(\delta/c_0)} - 1} \left[\left(\frac{\delta}{c_0} + 1 \right) \left(e^{(\frac{\delta}{c_0} + 1)} - 1 \right) \right] & \text{if } \delta > -c_0 \end{cases} \quad (18)$$

where δ is the penetration between solids, and (c_0, f_0) are contact regularization parameters. With respect to dry friction, the computation of f_t relies on a regularized Coulomb law:

$$f_t = \mu mg \operatorname{sgn}(\dot{x}) \simeq \mu mg \tanh(\gamma \dot{x}) \quad \text{with} \quad \gamma \gg 1 \quad (19)$$

where γ is the sign function regularization parameter.

The choice of parameters c_0 , f_0 and γ is based on a compromise between an accurate modelling of the physical problem and the need to ensure numerical stability. In this study, $c_0 = 5 \times 10^{-4}$ m, $f_0 = 10$ N and $\gamma = 10$ are used.

Path following : continuation Nonlinear dynamics systems often exhibit complex behaviours, such as coexistence of multiple solutions. The possibility to obtain some of these distinct solutions is a key asset of the HBM in comparison with time integration techniques. To this end, an arc-length continuation procedure is combined with the HBM/AFT algorithm. It consists in a prediction-correction approach to build the frequency response curve where ω is unknown and constrained by a curvilinear abscissa parametrization of the curve. For the sake of brevity, details of this algorithm are not given in this paper, the reader may refer to [15] for more details.

B. Numerical simulations without contact interface

Without contact interface, it is assumed that dry friction is the only source of nonlinearity:

$$f_{nl}(x, \dot{x}) = f_t. \quad (20)$$

Experimental data and numerical frequency response curves for the identified set of parameters given in Tab. I are presented in Fig. 9. Both forward and backward frequency sweeps are considered numerically and experimentally. A good agreement between HBM calculations and experimental values is evidenced as numerical predictions are almost perfectly superimposed with experimental measurements.

When looking at the response of the system in the time domain at the resonance peak, see Fig. 10, it appears that the numerical prediction matches experimental observation. In addition to experimental data and HBM results, the solution obtained using time integration (an embedded Runge-Kutta RK5 (4) time integration scheme is used) is also plotted in order to ensure the validity of numerical calculations. To give an idea, HBM computation time for one solution is about less than one second, and for time integration is about 1-2 minutes. As mentioned above, it is assumed that the minor discrepancy between numerical results and experimental observations in the vicinity of the highest amplitudes in Fig. 10 is related to a minor rotation of the mass due to bearing clearances.

C. Numerical simulations with contact interface

Accounting for the contact interface, the nonlinear forces may be written as:

$$f_{nl}(x, \dot{x}) = f_n + f_t. \quad (21)$$

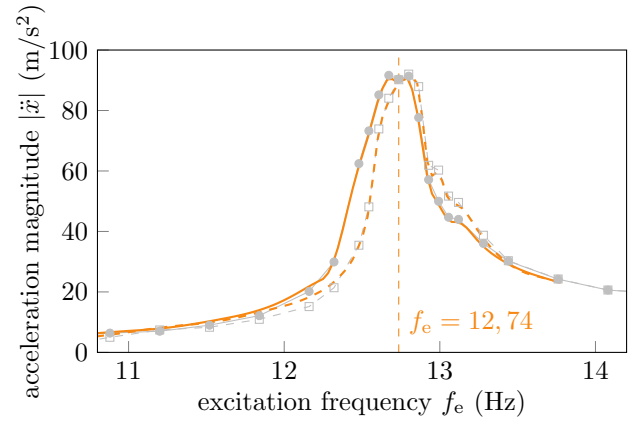


Figure 9. Comparison of experimental frequency response curves (forward (●) and backward (□) frequency sweeps) to the HBM model with $H = 8$ (forward (—) and backward (---) frequency sweeps).

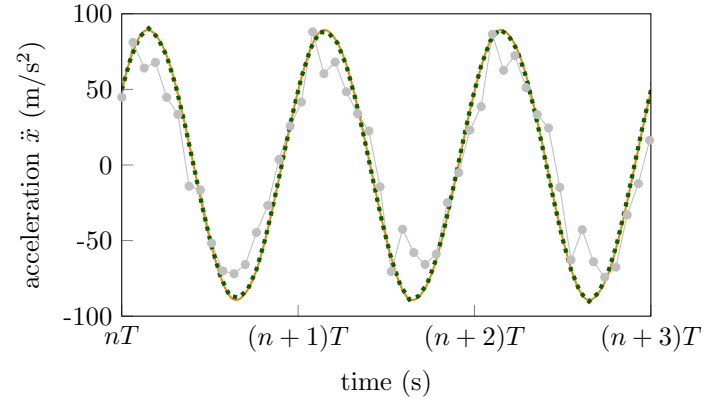


Figure 10. Comparison between experimental acceleration evolution (forward frequency sweep (●)), HBM (—) with $H = 8$, and RK5 (4) time integration (-.-) at $f_e = 12.74$ Hz.

The superimposition of results obtained with the HBM considering $H = 20$ harmonics and experimental measurements is given in Fig. 11. Significant differences of amplitudes are found between the numerical prediction (—) and experimental measurements (—). It is noticeable that the discrepancy between the two plots increases as the excitation frequency increases. In order to better understand the root cause of such discrepancy, time responses predicted numerically and observed experimentally are superimposed for two distinct excitation frequencies: $f_{e1}, f_{e2} = 12.8, 13.5$ Hz. The comparison between experimental and numerical accelerations are presented in Figs. 12 and 13. Interestingly, numerical predictions and experimental measurements essentially differ over a very narrow time frame, corresponding to the exact instant over which contact with the impactor occurs. During the contact phase, HBM results present a large peak of amplitude which is responsible for the error in amplitude evidenced in Fig. 11. The direct comparison of Figs. 12 and 13 underlines that the amplitude of this peak

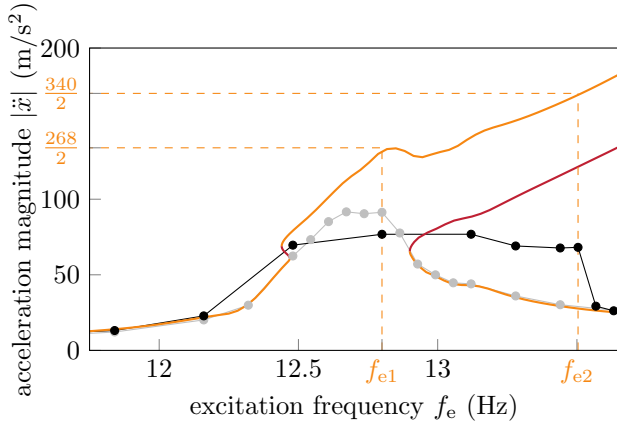


Figure 11. Comparison between experimental frequency response for the system : without impactor (●), with impactor (forward frequency sweep (●)) and HBM (stable (—), unstable (—) periodic solutions) with $H = 20$.

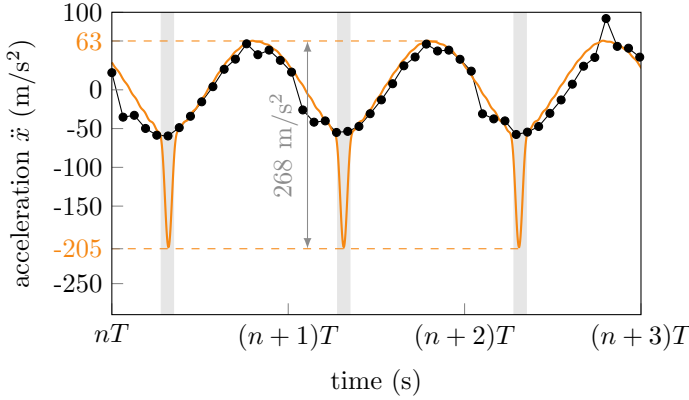


Figure 12. Comparison between experimental acceleration evolution (forward frequency sweep (●)) and HBM (—) with $H = 20$ at $f_{e1} = 12.8$ Hz. Time interval over which contacts occur is highlighted in grey.

increases with the excitation frequency.

The solutions computed by HBM at $f_e = 12.8$ Hz for different values of the number of harmonics H —not shown in this paper due to a lack of space—underline that the amplitude of the peak is strongly influenced by the number of harmonics H thus indicating it is essentially a numerical artifact. The well-known Gibbs phenomenon is responsible for this peak: in the vicinity of the time over which contacts occur, speeds and accelerations are non-differentiable and the Fourier basis used in (5) becomes ill-suited to accurately represent time responses.

There exists strategies to mitigate the Gibbs phenomenon, including the use of different bases of functions, see for instance [8]. One may note that the definition of the acceleration given in (10) in the frequency domain is proportional to ω^2 . For this reason, accelerations are particularly sensitive to the Gibbs phenomenon since the excitation frequency acts as a quadratic error amplification factor. To the contrary, results in terms of

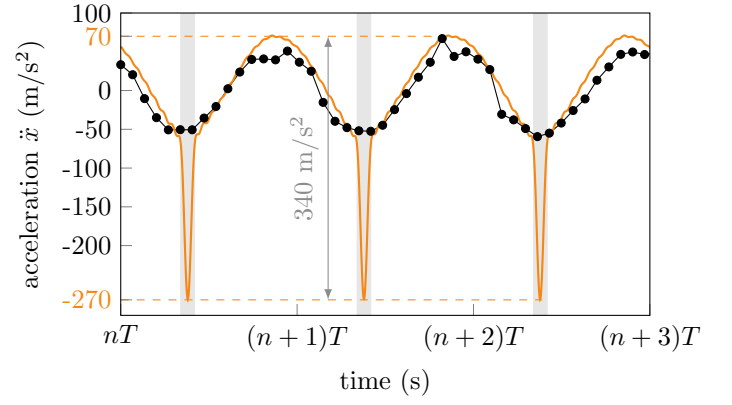


Figure 13. Comparison between experimental acceleration evolution (forward frequency sweep (●)) and HBM (—) with $H = 20$ at $f_{e2} = 13.5$ Hz. Time interval over which contacts occur is highlighted in grey.

displacements, which are not displayed for the sake of brevity do not feature such discrepancy with experimental results.

One way to mitigate the influence of the Gibbs phenomenon on the numerical results is to make a truncation of the predicted numerical solution computed for $H = 20$ to a lower number of harmonics; such truncations are depicted in Fig. 14. It is noticeable that the truncation of the solution to its first harmonic yields an acceptable approximation of the experimental observation.

Finally, the amplitude of all numerical solutions computed with $H = 20$ and truncated to their first harmonic (—) are plotted in Fig. 15. There is a better agreement between the numerically predicted frequency response of the system and experimental observations throughout the frequency range of interest. One should note that these solutions are distinct from the solutions directly computed by HBM with $H = 1$ (---), also depicted in Fig. 15.

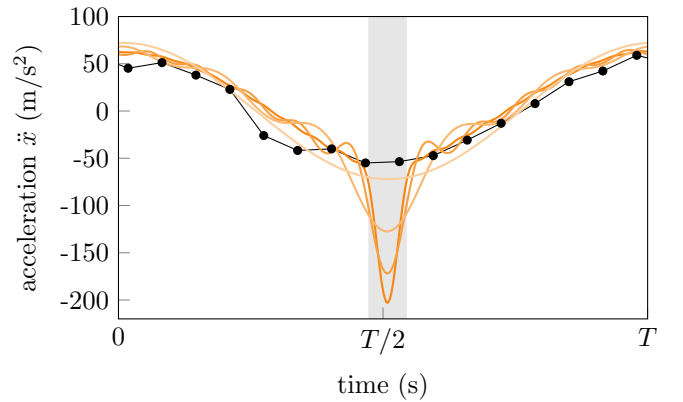


Figure 14. HBM results at $f_{e1} = 12.8$ Hz : full solution with $H = 20$ (—) and truncations to the first: 10 harmonics (—), 5 harmonics (—) and 1 harmonic (—).

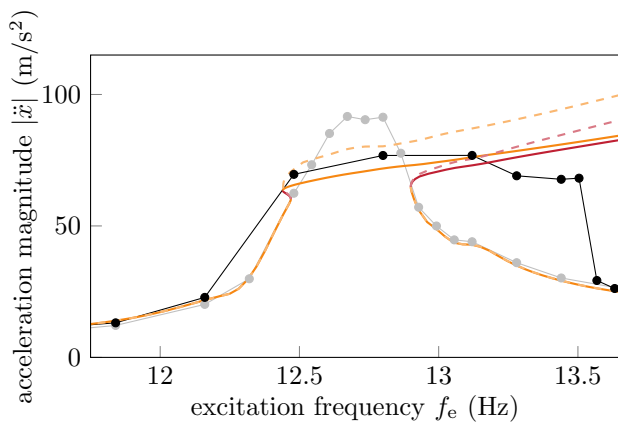


Figure 15. Frequency responses: experiment without impactor (\bullet), with impactor (\bullet), HBM (—) with $H = 20$ truncated to the first harmonic, and HBM (---) with $H = 1$.

IV. CONCLUSION

The numerical and experimental modelling of a single degree of freedom mechanical system with contact interface is investigated in this study. Based on a newly designed experimental setup, the mechanical parameters of the associated numerical model are identified. Frequency sweep tests are conducted with and without the contact interface. Assuming the periodicity of the system's response, the HBM is used to predict vibrations of the system. A good agreement is found between numerical prediction and experimental results in both time and frequency domains. Moreover, this study highlights the sensitivity of numerical results to the Gibbs phenomenon, particularly when looking at the accelerations.

This article underlines challenges in design and analysis of non-smooth one-degree of freedom mechanical system and attests to the importance of numerical/experimental confrontation. Future work will focus on an in-depth analysis of experimental data in displacements as well as investigations on the stability of experimental solutions prior to the jump discontinuity in forward frequency sweep. From a numerical standpoint, work is in progress for including within the HBM a numerical procedure of wear management in order to simulate complex industrial systems featuring wear and fretting.

ACKNOWLEDGMENT

The authors would like to thank Charles Audet (Polytechnique Montréal) for comments and careful proofreading that helped improve the manuscript.

REFERENCES

[1] A. Batailly, M. Legrand, A. Millicamps, and F. Garcin, "Numerical-Experimental Comparison in the Simulation of Rotor/Stator Interaction Through Blade-Tip/Abradable Coating Contact," *Journal of Engineering*

for Gas Turbines and Power, vol. 134 (8), Jun. 2012, pp. 082504–082504–11.

[2] G. Jacquet-Richardet, M. Torkhani, P. Cartraud, F. Thouverez, T. N. Baranger, M. Herran, C. Gibert, S. Baguet, P. Almeida, and L. Peletan, "Rotor to stator contacts in turbomachines. Review and application," *Mechanical Systems and Signal Processing*, vol. 40 (2), Nov. 2013, pp. 401–420.

[3] M. Krack, L. Salles, and F. Thouverez, "Vibration Prediction of Bladed Disks Coupled by Friction Joints," *Archives of Computational Methods in Engineering*, vol. 24 (3), Jul. 2017, pp. 589–636.

[4] J. Nol and G. Kerschen, "Nonlinear system identification in structural dynamics: 10 more years of progress," *Mechanical Systems and Signal Processing*, vol. 83, Jan. 2017, pp. 2–35.

[5] D. Laxalde and F. Thouverez, "Complex non-linear modal analysis for mechanical systems: Application to turbomachinery bladings with friction interfaces," *Journal of Sound and Vibration*, vol. 322 (4-5), May 2009, pp. 1009–1025.

[6] W.-J. Kim and N. Perkins, "Harmonic balance/Galerkin method for non-smooth dynamic systems," *Journal of Sound and Vibration*, vol. 261 (2), Mar. 2003, pp. 213–224.

[7] A. Thorin, P. Delezoide, and M. Legrand, "Non-smooth modal analysis of piecewise-linear impact oscillators," *SIAM Journal on Applied Dynamical Systems*, vol. 16 (3), 2017, pp. 1710–1747.

[8] S. Jones and M. Legrand, "Forced vibrations of a turbine blade undergoing regularized unilateral contact conditions through the wavelet balance method," *International Journal for Numerical Methods in Engineering*, vol. 101 (5), Oct. 2014, pp. 351–374.

[9] C. Issanchou, S. Bilbao, J.-L. L. Carrou, C. Touz, and O. Doar, "A modal-based approach to the nonlinear vibration of strings against a unilateral obstacle: Simulations and experiments in the pointwise case," *Journal of Sound and Vibration*, vol. 393, 2017, pp. 229 – 251.

[10] C. Charreton, C. Béguin, A. Ross, S. Étienne, and M. Pettigrew, "Two-phase damping for internal flow: Physical mechanism and effect of excitation parameters," *Journal of Fluids and Structures*, vol. 56, Jul. 2015, pp. 56–74.

[11] S. Kojtych, "Mesures d'accélération et système d'acquisition par microcontrôleur Arduino UNO," *Tech. rep.*, Polytechnique Montréal, Feb. 2019.

[12] G. Von Groll and D. J. Ewins, "The harmonic balance method with arc-length continuation in rotor/stator contact problems," *Journal of Sound and Vibration*, vol. 241 (2), 2001, pp. 223–233, doi : 10.1006/jsvi.2000.3298.

[13] T. Detroux, L. Renson, L. Masset, and G. Kerschen, "The harmonic balance method for bifurcation analysis of large-scale nonlinear mechanical systems," *Computer Methods in Applied Mechanics and Engineering*, vol. 296, 2015, pp. 18–38, doi : 10.1016/j.cma.2015.07.017.

[14] T. Cameron and J. Griffin, "An alternating frequency/time domain method for calculating the steady-state response of nonlinear dynamic systems," *Journal of Applied Mechanics*, vol. 56 (1), 1989, pp. 149–154, doi : 10.1115/1.3176036.

[15] R. Seydel, *Practical bifurcation and stability analysis*, vol. 5, Springer Science & Business Media, 2009, ISBN : 978-1-4419-1739-3.

Review and data consolidation of Ranque Hilsch Vortex Tube

Mark Parker* and Anthony G. Straatman
Department of Mechanical & Materials Engineering
Western University
London, Canada
*mparke46@uwo.ca

Abstract — A review and consolidation of experimental data was performed for the Ranque-Hilsch vortex tube. The objective was to determine optimal values for primary design aspects using the consolidated data. The review also served to determine the possibility of using literature on experimental studies to validate numerical simulations of a vortex tube.

Keywords: *Ranque-Hilsch vortex tube; design review; experimental data consolidation; temperature drop; boundary conditions; validate;*

I. INTRODUCTION

A Ranque-Hilsch vortex tube (RHVT) is a device that is capable of separating a compressed gas into two streams, one being heated and the other being cooled. It does this without any moving parts, electrical input, or chemical reactions. Compressed gas is injected tangentially into a tube that allows the peripheral flow (that has been heated) to exit at one end and the axial flow (that has been cooled) to exit at the other end. The device is composed of 4 parts: the vortex generator that utilizes nozzles to create the vortex, the main tube that contains the vortex, an orifice through which the cooled gas exits, and an adjustable plug through which the heated gas exits. The mechanism that drives the temperature separation has been studied for more than 80 years, but is yet not fully understood.

A literature review has been conducted with the first intention being to determine and evaluate the various design parameters of a RHVT, and the second to determine appropriate boundary conditions for numerical simulation of a RHVT and to assess the potential to use reported experimental data to validate the numerical results. As such there is an emphasis on experimental studies of the RHVT, with studies comparing numerical models to experimental setups also being of interest.

II. LITERATURE REVIEW

Markal et al. [1] examined the effect of the tip angle of the plug on the temperature separation for various aspect ratios. It was found that a tip angle of 30-degrees was optimal, but the effect of the tip angle was negligible for larger aspect ratios. Dincer et al. [2] also examined the effect of the plug tip angle, in addition to testing the diameter of the plug and the number of inlet nozzles. It was again shown that a 30-degree plug tip

resulted in an increased drop in temperature between the inlet and cold exit. Additionally, by having the ratio between the plug's diameter and the main tube's diameter being 0.55, and having 4 nozzles further increased the temperature drop. Kirmaci [3] studied the effect of the number of nozzles as well as the effect of using oxygen as the working fluid compared to air. It was determined that using oxygen achieves a greater temperature drop. However, it was also found that 2 inlet nozzles achieved a more substantial temperature drop compared to 3, 4, 5, and 6 inlet nozzles. Dincer et al. [4] further examined the effect of the number of inlet nozzles and the aspect ratio; however, they compared the temperature difference between the hot exit and the cold exit. As such, they found that an aspect ratio (ratio between the length and diameter of the main tube) of 15 resulted in a larger temperature separation compared to an aspect ratio of 10. Furthermore, 4 and 6 inlet nozzle configurations also increased the temperature separation compared to 2 inlet nozzles. Aydin et al. [5] proposed the use of a helical nozzle and tested different lengths of the helical nozzle with varying aspect ratios, where the parameter being compared was the temperature drop between the inlet and the cold exit. The resulting tests showed that a shorter helical out-performed a larger helical and that an aspect ratio of 30 out-performed aspect ratios of 10, 20, and 40. Eiamsa-ard [6] also proposed a snail nozzle as a new nozzle design. This nozzle design was tested with different numbers of injection ports, as well as different orifice diameters. The snail nozzle with 4 injection ports resulted in the largest temperature drop, out-performing snail nozzles with fewer injection ports, as well as a more typical tangential nozzle with 4 inlets. Additionally, a ratio between the orifice's diameter and the diameter of the main tube of 0.5 resulted in the most significant temperature drop when compared to larger and smaller ratios. Celik et al. [7] and Kandil & Abdelghany [8] further tested the effect of the orifice diameter and the aspect ratio on the drop in temperature between the inlet and the outlet. However, Celik et al. [7] showed that a ratio between the diameter of the orifice and main tube of 0.64 gave the most significant temperature drop, whereas Kandil & Abdelghany showed a ratio of 0.46 was optimal. Lastly, Thakare & Parekh [9] and Liu & Liu [10] compared numerical models to experimental data. Thakare & Parekh [9] used a 2D axis-symmetric approach where the inlet boundary condition was treated as a pressure inlet, and both the cold and the hot exits were treated as pressure outlets. Liu & Liu [10] used a 3D with

similar outlet boundary conditions; however, the inlet was treated as a mass flow inlet.

III. RESULTS AND DISCUSSION

A. Experimental Data Consolidation

The literature review shows that not only are there numerous design and operation parameters that affect the performance of a RHVT, but there are also several ways to quantify the performance. Furthermore, it was shown that there were discrepancies as to what values would achieve improved performance. As such the data from 29 papers [1]–[29] (resulting in 4896 data points) was consolidated with all available information for each point being stored. This resulted in some data points containing information on specific design, operation, and performance parameters while having a lack of information on other parameters. For example, a data point might contain design information pertaining to the main tube, vortex generator, and orifice, but will not have any information on the plug.

A custom MATLAB plotter was created to visualize the consolidated data. This graphical user interface plotter allowed for the creation of both 2D and 3D plots, with the addition of colour mapping, to display the data. This plotter was designed to be able to select any of the design, operation, and performance parameters to be used as the axis/colour map, allowing for easy comparison of the parameters. Furthermore, any point on the plot could be selected, and a readout of all the information for that point would be shown.

To accurately compare the data gathered from the literature, only the temperature drop between the inlet and the cold exit was considered for the performance criteria of interest. The operating conditions that had a significant effect on the temperature drop were the inlet pressure and the mass flow split. Thus, the custom plotter's 'Y' axis was set to the mass flow split, and the 'Z' axis was set to the temperature drop, with the data points being colour mapped based on the inlet pressure. This left the 'X' axis available to be used for the various design parameters. It was determined that there were 5 key design parameters, they were as follows; the ratio between the main tube's length to its diameter (aspect ratio), the ratio between the orifice diameter and the internal diameter of the main tube (orifice diameter ratio), the number of injection nozzles (number of nozzles), the ratio between the diameter of the plug to that of the internal diameter of the main tube (plug diameter ratio), and the tip angle of the plug (plug tip angle).

B. Aspect Ratio

Figure 1 shows the effect of the aspect ratio at the various cold mass splits and inlet pressures. When focusing on the inlet pressure range of 0.4 [MPa] to 0.6 [MPa], it was determined that an aspect ratio between 20 and 30 would produce a temperature drop, with 25 providing the most significant temperature drop. Figure 1 also showed that for an aspect ratio of 25, the temperature drop between the inlet and cold exit could be significantly increased by increasing the inlet pressure. It was

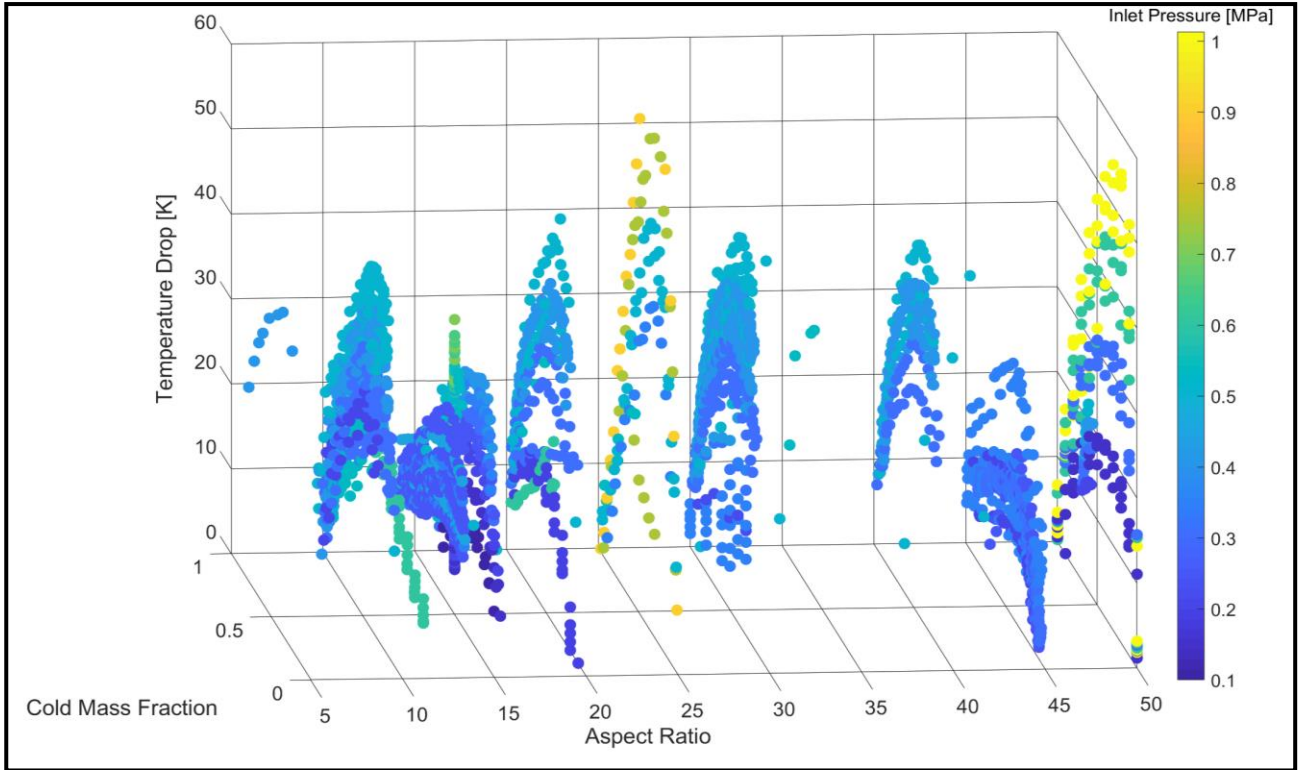


Figure 1. Aspect Ratio

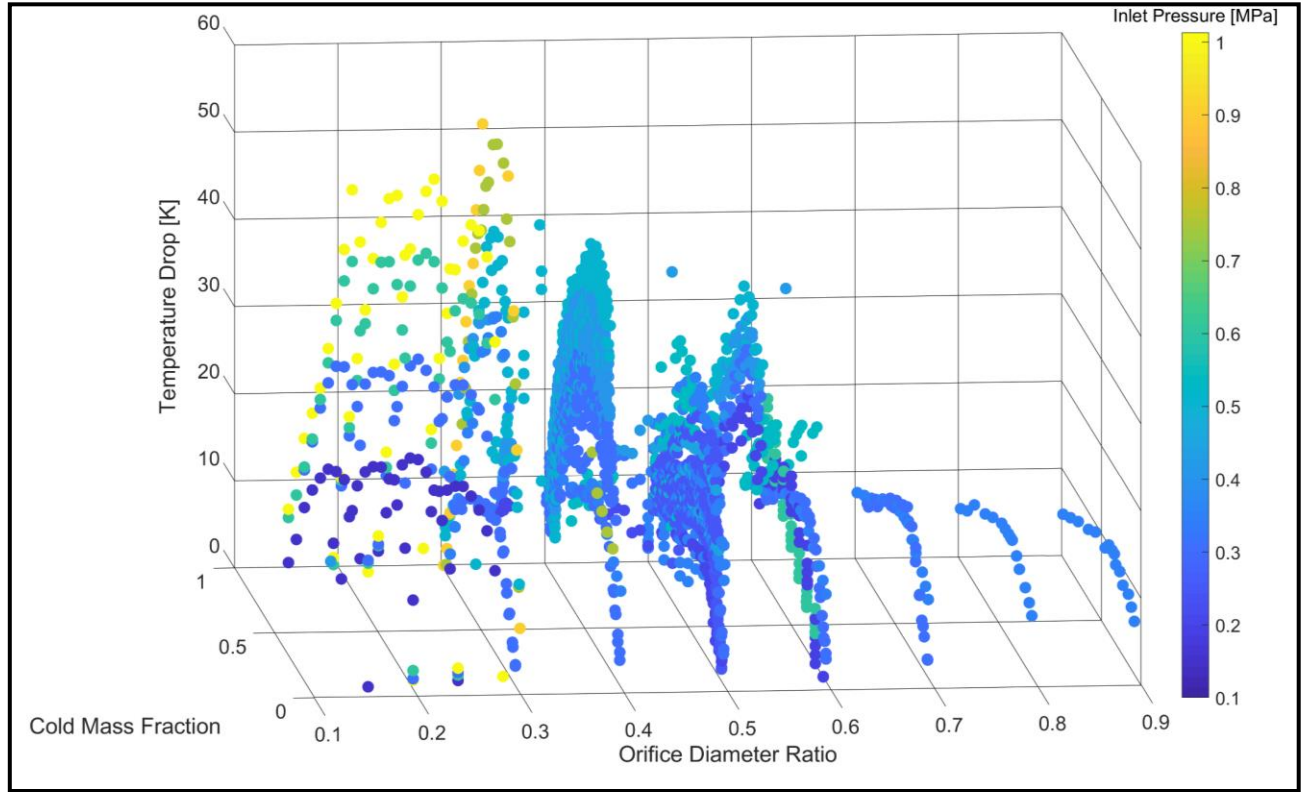


Figure 2. Orifice Diameter Ratio

also noticed that there were 2 clusters of data, one being for an aspect ratio approximately 15 and the other for an aspect ratio of 45. By changing the parameter controlling the colour mapping, it was determined that other parameters were also affecting these data points.

C. Orifice Diameter

Figure 2 shows the effect that the ratio between the cold exit orifice diameter to the diameter of the main tube had on the temperature drop. This parameter was of particular interest as from the literature review it was found that there was a discrepancy as to whether a larger or smaller ratio resulted in higher performance. When comparing to similar inlet pressures, Fig. 2 showed that ratios between 0.3 and 0.4 resulted in the most significant temperature drop.

D. Number of Injection Nozzles

Next, the number of injection nozzles was examined, as shown in Fig. 3. This plot is misleading, as even when accounting for the inlet pressure, the plot indicates that a single injection nozzle produced the most significant temperature drop. However, when changing the colour mapping parameter, it was found that other parameters were significantly different from that of the rest of the data shown in Fig. 3. Thus, looking at the data for a pressure range of 0.2 [MPa] to 0.5 [MPa] (which have comparable values for other design parameters), it was determined that an increase in the number of nozzles aids in increasing the temperature drop. However, this effect starts to diminish when increasing beyond 4 nozzles, resulting in 6

nozzles only providing a slight improvement over the 4 nozzle configurations.

E. Plug Shape and Size

Figures 4 and 5 pertain to the plug size and shape respectively. Most design configurations of a RHVT have a plug diameter ratio of 0.5 or 1.0 with a few configurations using a ratio less than 0.5. From Fig. 4 it was determined that the plug diameter ratio of approximately 0.3 showed slightly increased temperature separation compared to ratios of 0.5 and 1.0. However, the plug diameter ratio of 0.1 had the lowest temperature drop, even when comparing to similar inlet pressures. When reviewing the angle of the plug tip (shown in Fig. 5), it was determined that a flat plug (angle of 0 degrees) resulted in the smallest temperature drop. Furthermore, it was found that a 30-degree tip angle produced a slightly higher drop in temperature compared to plug tips with a 60-degree angle. This suggests that a smaller plug diameter ratio with a sharp angle would produce the most significant temperature drop.

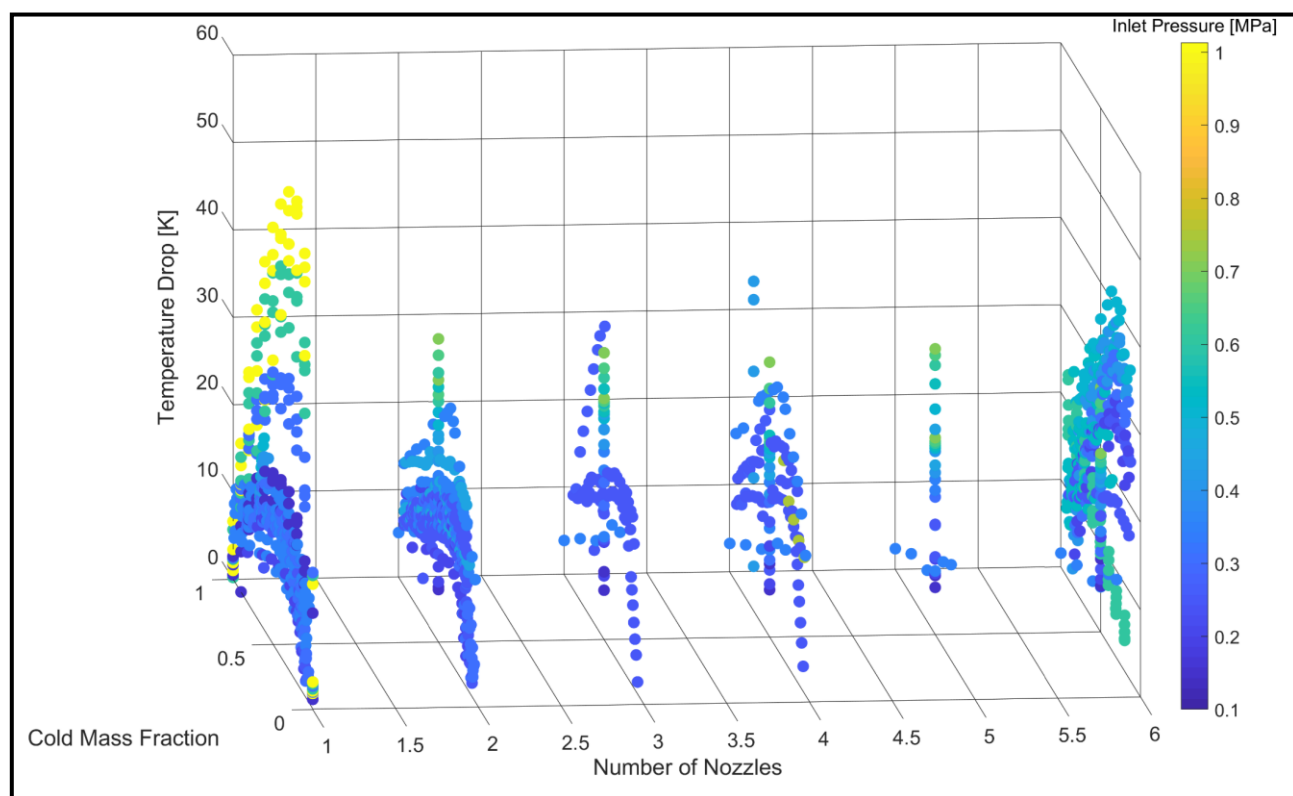


Figure 4. Number of Nozzles

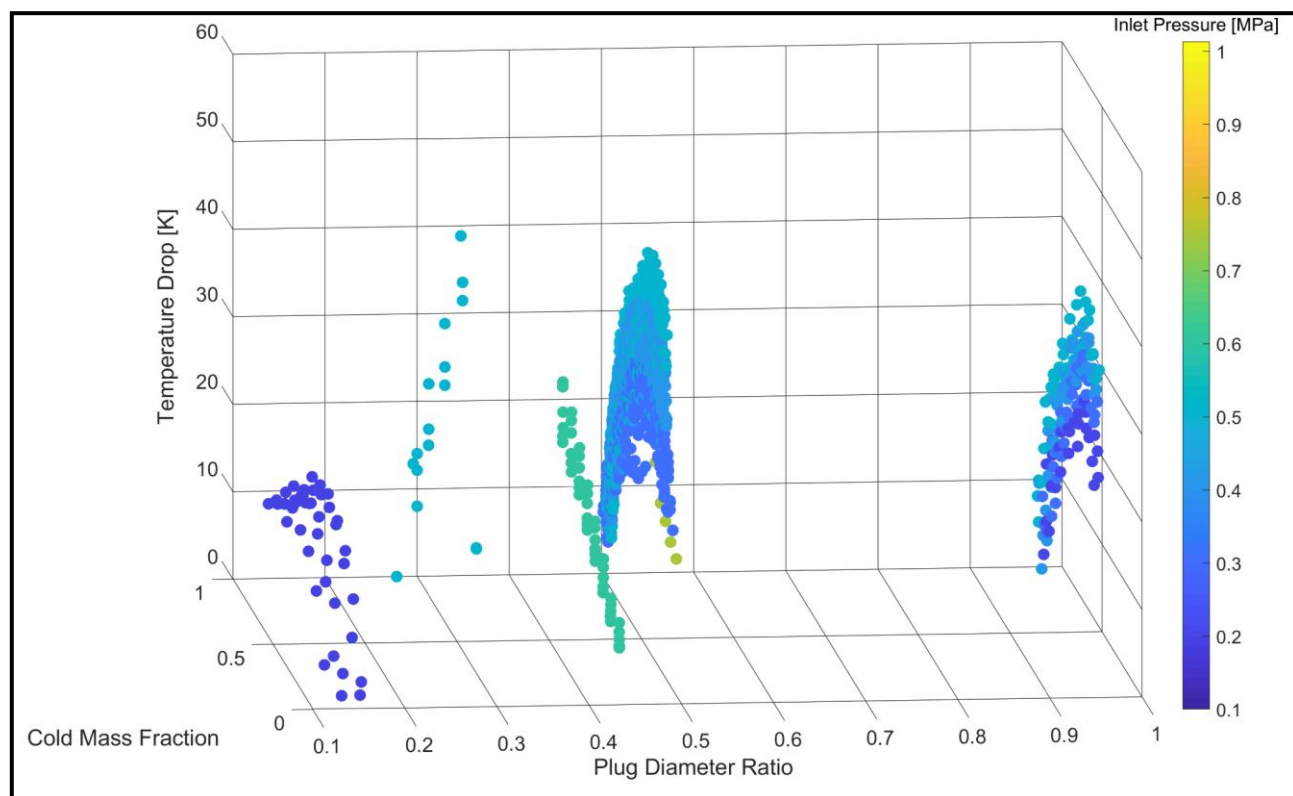


Figure 3. Plug Diameter Ratio

F. CFD Boundary Conditions and Validation using Experimental Study Literature

There are two options for setting the inlet boundary condition, the first being a pressure inlet and the second being a mass flow condition. However, when looking at the available literature with experimental results, the inlet pressure is primarily reported with no mention of the mass flow rate. Of the 29 papers reviewed for design purposes, only Nimbalkar & Muller [13] reported the inlet mass flow rates used during their experiment. Thus, validating a numerical simulation with a mass flow boundary condition from the available RHVT experimental literature would be difficult.

When looking at the cold and hot exits, numerical simulations primarily use only pressure outlet boundary conditions. By using different pressure values on the two exits of the RHVT simulation, the cold mass split can be adjusted. Again, when looking at the available literature, the pressure at the outlets were not reported.

Aside from difficulties related to using literature of experimental studies for boundary condition information, further difficulties arise with modeling the RHVT. As stated before, it was found that most papers did not report on all design parameters of the RHVT used in the experiments. This was evident when comparing the number of data points that contained an aspect ratio value (shown in Fig. 1) being 4879 points compared to 2691 data points with a value for the plug diameter ratio (shown in Fig. 4). Thus, accurately modeling the geometry of an available experimental study becomes impractical, and would require the model to assume dimensions for some parameters.

IV. CONCLUSION

After reviewing and collecting experimental data from the available literature, recommended design values were observed for primary design aspects. It was determined that to produce a significant temperature drop between the inlet and the cold exit the design parameters should be as follows: an aspect ratio of 25; a ratio between the orifice diameter and the main tube diameter of 0.35; 4 or more injection nozzles; and a plug with a sharp tip angle with a diameter ratio around 0.3.

However, there was also a trend of lack of information being reported in published articles. Whether it be information on the design parameters, (i.e., only reporting the design parameter values of what was being tested without reporting other design parameter values) or on the fluid properties (mass flow, pressure, temperature) for the inlet and outlets. This lack of design, operation and performance information results in an inability to generate a model that would be geometrically accurate or have clearly defined boundary conditions imposed on the inlet and outlets. Therefore, validating a numerical simulation of a RHVT using the available experimental literature on RHVT would be difficult and of questionable quality.

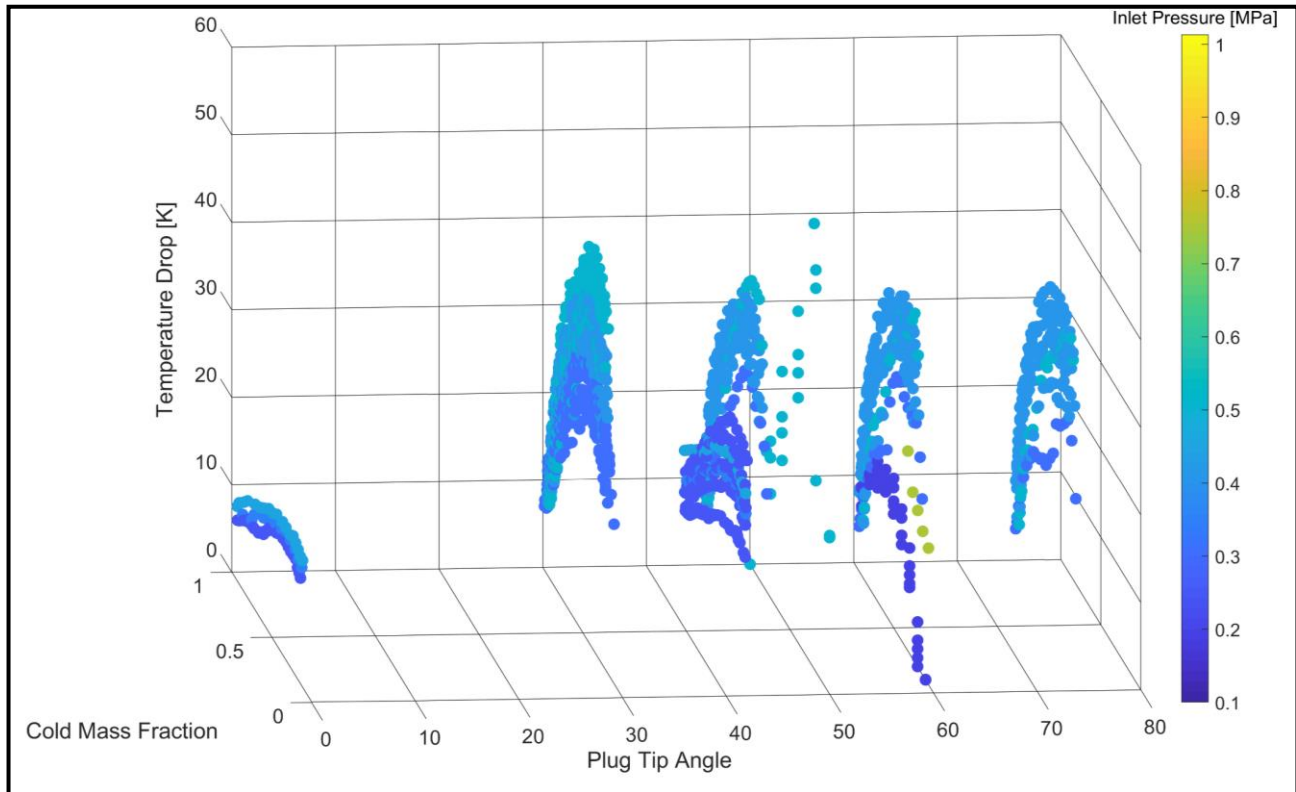


Figure 5. Plug Tip Angle

REFERENCES

- [1] B. Markal, O. Aydin, and M. Avci, "An experimental study on the effect of the valve angle of counter-flow Ranque-Hilsch vortex tubes on thermal energy separation," *Exp. Therm. Fluid Sci.*, vol. 34, no. 7, pp. 966–971, 2010.
- [2] K. Dincer, S. Baskaya, B. Z. Uysal, and I. Ucgul, "Experimental investigation of the performance of a Ranque-Hilsch vortex tube with regard to a plug located at the hot outlet," *Int. J. Refrig.*, vol. 32, no. 1, pp. 87–94, 2009.
- [3] V. Kirmaci, "Exergy analysis and performance of a counter flow Ranque-Hilsch vortex tube having various nozzle numbers at different inlet pressures of oxygen and air," *Int. J. Refrig.*, vol. 32, no. 7, pp. 1626–1633, 2009.
- [4] K. Dincer, S. Tasdemir, S. Baskaya, and B. Z. Uysal, "Experimental investigation of the effects of length to diameter ratio and nozzle number on the performance of counter flow Ranque-Hilsch vortex tubes," *Appl. Therm. Eng.*, vol. 28, no. 17–18, pp. 2380–2390, 2008.
- [5] O. Aydin, B. Markal, and M. Avci, "A new vortex generator geometry for a counter-flow Ranque-Hilsch vortex tube," *Appl. Therm. Eng.*, vol. 30, no. 16, pp. 2505–2511, 2010.
- [6] S. Eiamsa-ard, "Experimental investigation of energy separation in a counter-flow Ranque-Hilsch vortex tube with multiple inlet snail entries," *Int. Commun. Heat Mass Transf.*, vol. 37, no. 6, pp. 637–643, 2010.
- [7] A. Celik, M. Yilmaz, M. Kaya, and S. Karagoz, "The experimental investigation and thermodynamic analysis of vortex tubes," *Heat Mass Transf.*, vol. 53, no. 2, pp. 395–405, 2016.
- [8] H. A. Kandil and S. T. Abdelghany, "Computational investigation of different effects on the performance of the Ranque-Hilsch vortex tube," *Energy*, vol. 84, pp. 207–218, 2015.
- [9] H. R. Thakare and A. D. Parekh, "Experimental investigation & CFD analysis of Ranque-Hilsch vortex tube," *Energy*, vol. 133, pp. 284–298, 2017.
- [10] X. Liu and Z. Liu, "Investigation of the energy separation effect and flow mechanism inside a vortex tube," *Appl. Therm. Eng.*, vol. 67, no. 1–2, pp. 494–506, 2014.
- [11] R. Hilsch, "The use of the expansion of gases in a centrifugal field as cooling process," *Rev. Sci. Instrum.*, vol. 18, no. 2, pp. 108–113, 1947.
- [12] Y. Xue and M. Arjomandi, "The effect of vortex angle on the efficiency of the Ranque-Hilsch vortex tube," *Exp. Therm. Fluid Sci.*, vol. 33, no. 1, pp. 54–57, 2008.
- [13] S. U. Nimbalkar and M. R. Muller, "An experimental investigation of the optimum geometry for the cold end orifice of a vortex tube," *Appl. Therm. Eng.*, vol. 29, no. 2–3, pp. 509–514, 2009.
- [14] M. H. Saidi and M. S. Valipour, "Experimental modeling of vortex tube refrigerator," *Appl. Therm. Eng.*, vol. 23, no. 15, pp. 1971–1980, 2003.
- [15] K. D. Devade and A. T. Pise, "Effect of Mach number, valve angle and length to diameter ratio on thermal performance in flow of air through Ranque Hilsch vortex tube," *Heat Mass Transf.*, vol. 53, no. 1, pp. 161–168, 2017.
- [16] S. E. Rafiee and M. M. Sadeghiazad, "Experimental and 3D CFD investigation on heat transfer and energy separation inside a counter flow vortex tube using different shapes of hot control valves," *Appl. Therm. Eng.*, vol. 110, pp. 648–664, 2017.
- [17] O. Aydin and M. Baki, "An experimental study on the design parameters of a counterflow vortex tube," *Energy*, vol. 31, no. 14, pp. 2427–2436, 2006.
- [18] C.-H. S. C.-S. Kim, "Experimental and Numerical Studies in a Vortex Tube," *J. Mech. Sci. Technology*, vol. 20, no. 3, pp. 418–425, 2006.
- [19] P. Promvong and S. Eiamsa-ard, "Investigation on the Vortex Thermal Separation in a Vortex Tube Refrigerator," *ScienceAsia*, vol. 31, no. 3, p. 215, 2005.
- [20] S. Eiamsa-ard, K. Wongcharee, and P. Promvong, "Experimental investigation on energy separation in a counter-flow Ranque-Hilsch vortex tube: Effect of cooling a hot tube," *Int. Commun. Heat Mass Transf.*, vol. 37, no. 2, pp. 156–162, 2010.
- [21] A. Kumar, Vivekanand, and S. Subudhi, "Cooling and dehumidification using vortex tube," *Appl. Therm. Eng.*, vol. 122, pp. 181–193, 2017.
- [22] B. A. Shannak, "Temperature separation and friction losses in vortex tube," *Heat Mass Transf.*, vol. 40, no. 10, pp. 779–785, 2004.
- [23] S. Mohammadi and F. Farhadi, "Experimental analysis of a Ranque-Hilsch vortex tube for optimizing nozzle numbers and diameter," *Appl. Therm. Eng.*, vol. 61, no. 2, pp. 500–506, 2013.
- [24] O. Tlili El May, I. Mokni, H. Mhiri, and P. Bournot, "CFD investigation of a vortex tube: Effect of the cold end orifice in the temperature separation mechanism," *Sci. Acad. Trans. Renew. Energy Syst. Eng. Technol.*, vol. 1, no. 3, pp. 84–89, 2011.
- [25] U. Behera *et al.*, "CFD analysis and experimental investigations towards optimizing the parameters of Ranque-Hilsch vortex tube," *Int. J. Heat Mass Transf.*, vol. 48, no. 10, pp. 1961–1973, 2005.
- [26] T. Farouk and B. Farouk, "Large eddy simulations of the flow field and temperature separation in the Ranque-Hilsch vortex tube," *Int. J. Heat Mass Transf.*, vol. 50, pp. 4724–4735, 2007.
- [27] T. Dutta, K. P. Sinhamahapatra, and S. S. Bandyopdhyay, "Comparison of different turbulence models in predicting the temperature separation in a Ranque-Hilsch vortex tube," *Int. J. Refrig.*, vol. 33, no. 4, pp. 783–792, 2010.
- [28] S. Akheshmeh, N. Pourmahmou, and H. Sedgi, "Numerical Study of the Temperature Separation in the Ranque-Hilsch Vortex Tube," *Am. J. Eng. Appl. Sci.*, vol. 1, no. 3, pp. 181–187, 2008.
- [29] S. Y. Im and S. S. Yu, "Effects of geometric parameters on the separated air flow temperature of a vortex tube for design optimization," *Energy*, vol. 37, no. 1, pp. 154–160, 2012.

Environmental Engineering

Novel use of Fluent to Simulate Oxides of Nitrogen (NO_x) Disperion Downwind of an Air Pollution Barrier

Wood, David*; Aliabadi, Amir A.; Van Heyst, Bill

School of Engineering
University of Guelph
Guelph, Canada
woodd@uoguelph.ca

ABSTRACT

Atmospheric levels of oxides of nitrogen (NO_x; including nitric oxide (NO) and nitrogen dioxide (NO₂)) have increased in proportion to the increases in vehicle ownership and commute length. Concentrations are typically highest at the roadside and decrease with increasing distance from the road. The rate of decrease is dependent on the distance travelled, the meteorological conditions, and local topography. There is a growing body of evidence which suggests that the risk of adverse health effects from traffic related air pollution is increasing as a result of populations spending a larger amount of time near major roadways. It has been estimated that in densely populated North American cities, nearly 45% of the population live in the areas that are most affected by traffic emissions. It is this portion of the population that is at the greatest risk of the adverse health effects related to traffic emissions. An estimated 58,000 premature deaths in the US have been linked directly to poor air quality from the transportation industry.

There have been many proposals in North America and Europe to develop solutions that can be implemented at the roadside to protect nearby residents. A roadside air pollution barrier has been a common theme among the proposed solutions. Previous designs have attempted to create a physical barrier to prevent air pollution from moving to the residential side and others have included photocatalytic technologies, all to limited degrees of success. To that effect, this study presents the downwind (0 - 100 m) dispersion profiles, at inhalation height (1.5 m), of NO released on the highway side under three scenarios; no barrier, a standard barrier, and a novel air pollution barrier. The novel air pollution barrier uses a design that channels air, contaminated from vehicle emissions, into a double walled system where the interior is coated with a sunlight activated photocatalytic coating. Results were obtained using ANSYS Fluent and a transient detached eddy simulation (DES) in a domain that was 500 m x 75 m x 150 m. Boundary layer development and turbulence were achieved using both the synthetic and precursor methods as outlined in Aliabadi et al. (2017)¹. A 1/7th wind profile power law was used to develop the inlet conditions, with a wind speed of 4 ms⁻¹ at a height of 10 m. All cases were run using a neutral atmospheric stability. Results indicate that the presence of a barrier will improve downwind dispersion when compared to having no barrier present. Elimination of NO and the subsequent dispersion downwind of the novel air pollution barrier further reduces the concentrations of NO at inhalation height beyond that of a conventional barrier.

-
1. Aliabadi Amir A., Krayenhoff ES, Nazarian N, et al. Effects of Roof-Edge Roughness on Air Temperature and Pollutant Concentration in Urban Canyons. *Boundary-Layer Meteorol.* 2017;164(2):249-279. doi:10.1007/s10546-017-0246-1

SEA LETTUCE TO BIOCARBON AND BIOGAS WITH RECOVERY OF NUTRIENTS

Ankita Shrestha

Faculty of Sustainable Design Engineering
University of Prince Edward Island
Charlottetown, Canada
ashrestha3@upei.ca

Bishnu Acharya

Faculty of Sustainable Design Engineering
University of Prince Edward Island
Charlottetown, Canada
bacharya@upei.ca

ABSTRACT

Human activity in Prince Edward Island contributes to over 90% of nitrate in island waterways. Increase in nitrate level is responsible for excessive growth of sea lettuce which blocks waterways causes anoxic events, harms aquatic life and its decay process creates unpleasant smell. A sustainable way to mitigate the problem of overgrowing sea lettuce is needed. An integrated process of hydrothermal carbonization with anaerobic digestion is used for turning sea lettuce into hydrochar and biogas. While previous works on hydrothermal carbonization is limited to batch type reactor, a semi-continuous reactor with cooling chamber has been developed for this work. The carbonization process was carried out in a batch type reactor as well as in semi-continuous reactor at different temperature and residence time and the effect of these operating parameters on the biochar and the process water was studied. The severity factor increased with temperature and time and the heating value of biochar increased with severity factor and was in the range of 14 to 20 MJ/kg. This research aims to solve an environmental problem with the proper conversion technology while creating products that can find applications in energy and soil.

Yana Nec
Department of Mathematics and Statistics
Thompson Rivers University
Kamloops, BC
ynec@tru.ca

Greg Huculak
GNH Consulting Ltd.
Delta, BC

ABSTRACT

A horizontal well is a perforated pipe placed in the landfill waste layer and connected to a vacuum inducing device. This system comprises two flow fields: flow through porous media characterised by peculiarities unique to the landfill application, and flow along the pipe. We discuss the solutions to respective flow fields and their coupling, and the parameters essential for effective control of this system. Analytical and numerical approaches are combined to provide insight into the engineering design of such a collection mechanism, including the determination of radius of influence (and ensuing spacing between adjacent wells) and surface flux, the two questions for whose answers the industry has fumbled for many decades. The study culminates in an approximate solution to a non-axisymmetric landfill flow field when gravity is insignificant, and identifies situations, where gravity cannot be disregarded. This discovery makes the landfill flow field one of the very few gas flows, where gravity cannot be neglected. The conclusions tie well with the phenomena observed in the field that have baffled successful management of this collection mechanism, and offer design and control strategies. From an engineering perspective we classify the geometric design parameters (such as pipe diameter, perforation frequency etc.) and outcome quantities of interest (e.g. production rate, radius of influence, surface flux) into 3 groups: easily computable (closed form solutions), computable with moderate effort (iterative process) and requiring high end computational means (numerical solution of partial differential equations intractable analytically).

EGGSHELL-DERIVED SORBENT FOR BIOMASS GASIFICATION: CALCINATION AND CARBONATION KINETIC STUDIES

Shakirudeen A. Salaudeen, Animesh Dutta

School of Engineering
University of Guelph
Guelph, Canada
ssalaude@uoguelph.ca
adutta@uoguelph.ca

Bishnu Acharya

School of Sustainable Design Engineering
University of Prince Edward Island
Charlottetown, Canada
bacharya@upei.ca

ABSTRACT

This study presents an investigation into the calcination and carbonation reaction kinetics of eggshell. To analyze the calcination and carbonation processes, experiments were performed in a thermogravimetric analyzer. Non-isothermal (dynamic) and multiple heating rates were used to analyze the calcination process under nitrogen environment. The carbonation process, on the other hand, was studied at isothermal condition and at multiple temperatures in an atmosphere of carbon dioxide. The calcination kinetic parameters were evaluated with several model-based and isoconversional kinetic methods. The methods include the Friedman, Coats-Redfern, modified Coats-Redfern, Kissinger, Flynn-Wall-Ozawa and Kissinger-Akahira-Sunose methods. Furthermore, an analytical solution method was developed to evaluate the kinetic parameters and to predict the calcination conversion. The carbonation reaction was modelled with a modified form of the shrinking core model. Both the rapid surface reaction-controlled and the slow diffusion-limited stages of carbonation were analyzed. Results showed that the kinetic parameters obtained with the various methods are in good agreement, an activation energy of around 215 kJ/mol was obtained for the calcination reaction. It is also observed that the activation energy varies with the extent of conversion, suggesting that the mechanism is not a single-step type. In addition, results showed that the carbonation reaction mechanism of eggshell is controlled by the combination of surface reaction and product layer diffusion. An activation energy of 49.56 kJ/mol was obtained for the chemical reaction stage and 72.52 kJ/mol for the diffusion-limited stage for carbonation temperatures from 500°C – 700°C.

BIO-CARBON PRODUCTION BY OXIDATION PROCESS OF PULP AND PAPER BYPRODUCT

Zainab AL-Kaabi, Naresh Thevathasan
School of Environmental Science,
University of Guelph,
Guelph, Canada
zzwayyer@uoguelph.ca
nthevath@uoguelph.ca

Animesh Dutta, Ranjan Pradhan
School of Engineering,
University of Guelph,
Guelph, Canada
adutta@uoguelph.ca
rpradhan@uoguelph.ca

ABSTRACT

Bio-carbon production from black liquor that is an industrial waste, also generally called spent liquor is a promising way to add value to the industrial waste stream. A novel oxidative procedure using hydrogen peroxide as a green liquid oxidant at room temperature for processing of paper recycling neutral sulphite semi-chemical spent liquor(NSSCSL) with ash content of $23.27 \pm 0.8\%$, softwood (pine and spruce) kraft spent liquor (PSKSL) with ash content of $35.93 \pm 0.50\%$, and hardwood (aspen and balsam poplar) kraft spent liquor (APKSL) with ash content of $1.98 \pm 0.05\%$ was established and evaluations were carried out for its efficiency. The oxidation procedure was performed at room temperature using 5% of H_2O_2 followed by washing in diluted sulfuric acid (0.15 N). Bio-carbon produced from the three different spent liquors were similar in terms of ash percentage and ranged from 1.25 ± 0.05 to $1.48 \pm 0.05\%$, 1.82 ± 0.05 to $2.73 \pm 0.08\%$, 0.71 ± 0.07 to $1.98 \pm 0.05\%$. The carbon contents ranged from 60.54 to 61.50%, 58.89 to 58.89%, and 59.67 to 62.91% while higher heating values ranged from 25.32 to 26.11 MJ/kg, 24.11 to 24.74 MJ/kg, and from 24.58 to 26.89 MJ/kg for NSSCSL, PSKSL, APKSL respectively. The mass yield ranged from 29.98 to 31.88%, 46.27 to 45.79%, and from 41.33 to 43.93% respectively. These highly improved bio-carbon enhance their potential as green substitutes solid fuel, and bio-chemicals.

SEPARATION OF CO₂ FROM BIOGAS BY ADSORPTION ON ACTIVATED CARBON

Eniola Adewakun, Animesh Dutta

School of Engineering,

University of Guelph,

Guelph, Canada

eadewaku@uoguelph.ca

adutta@uoguelph.ca

ABSTRACT

Carbon dioxide (CO₂) has become a major concern due to its sudden increase in the atmosphere since the dawn of the industrial era. Alternative fuel use and CO₂ capture using bio-adsorbents are methods widely discussed and implemented. Biogas, which is considered an alternative fuel, consists of a significant amount of CO₂, which must be captured to achieve cleaner fuel. This study provides an in-depth understanding of hydrochar produced from switchgrass with different solvents and their applications as bio-adsorbents. In this study, switchgrass was hydrothermally carbonized at 200 and 240°C and at residence times of 30 and 60 mins by using deionized water and 0.01M of KOH solution. The mass yield, proximate analysis, ultimate analysis and functional group analysis were carried out to determine the most suitable hydrochar for the production of bio-adsorbents. The most suitable hydrochars were chosen and activated using KOH as an activating agent at 700°C for an hour to produce mesoporous and microporous bio-activated carbon. The Brunauer–Emmett–Teller (BET) analysis was carried out to determine the pore volume and surface area of the bio-adsorbent. A lab-scale fixed bed was filled with activated carbon and bio-adsorbents. The biogas consisting of 60% methane and 40% carbon dioxide was passed through the adsorption column at atmospheric temperature and varying pressures at different residence times. The effects of the various pre-treatments of the bio-adsorbents on the adsorption capacity were then determined.

A Novel Option for City Water Recycle with Carbon Capture

Bishnu Acharya
Faculty of Sustainable Design Engineering,
University of Prince Edward Island,
Charlottetown, Canada
bacharya@upei.ca

Prabir Basu & Akash Kulashreshtha
Mechanical Engineering Department
Dalhousie University
Halifax, Canada
Prabir.basu@dal.ca

ABSTRACT

Rapid urbanization is putting great pressure on the water supply of many large cities around the world. Major cities such as Cape Town, Sao Paulo, Bangalore, Beijing, Jakarta, Cairo are facing severe water crisis. Water available per inhabitant fell much below the World Bank standard of 1000 cubic meter per head per year. While no additional water source is coming up, population is increasing in these cities due to migration. The only way out is substantial recycling of water in use with minimum make up water requirement.

The sewage treatment plant of Charlottetown city treats approximately 7 million cubic meters of wastewater annually producing roughly 3700 tons of biosolid. The sewage sludge is pre-pasteurized, anaerobically digested and dewatered producing biogas and class-A biosolid. The biosolid is then taken away with a tipping fee of \$35/ton. Earlier work of the authors found a value added option for this biosolid that was subjected to hydrothermal carbonization. It produced hydrochar with 50% higher heating value and process water, which is subsequently subjected to anaerobic digestion to produce further biogas and relatively clean water. This water was though free from many impurities still contains nitrates that makes it unacceptable for drinking. Current research proposes leading this water to an aerated lagoon where macro algae would be grown, which will in turn absorb the harmful nitrates from the water and carbon dioxide from the atmosphere. The algae can be used for bio-oil production. Present work gives preliminary results of this process with mass and energy balance.

Kinetics of oil uptake using plastic waste

Junaid Saleem, and Gordon McKay

Division of Sustainable Development, College of Science & Engineering, Hamad Bin Khalifa University, Qatar Foundation,
Doha, Qatar

Abstract— Plastic waste constitutes around 15% of total municipal solid waste produced all over the world and a large volume of it is leaked into the ocean and ingested by seabirds and fish. In this work, we have up-cycled existing waste bottles made up of high-density polyethylene (HDPE) into value-added super-fast oil sorbent and have performed a kinetic study to explain saturation uptake, and retention capacity. The practical pick-up value of as-prepared sample was found to be 99 g/g which is considerably higher than existing commercial sorbent. The results reported in this work provides a useful basis to select as prepared sorbent for oil-spilled sites.

Keywords- *Sorption Kinetics; Saturation Uptake; Oil Spill*

I. INTRODUCTION (HEADING 1)

It is essential to control an oil spill at an initial stage, otherwise it may undergo various processes simultaneously that include emulsification, spreading, evaporation, dispersion etc. [1] Any of these processes further complicate the cleanup procedures, for example, oil which is evaporated, or dissolved or dispersed as droplets into the water literally escaped the extent of the treatment process. Hence, it is critically important to mitigate the spill as early as possible.

Another environmental problem which is more prevalent than oil spill is the generation of huge quantity of plastic waste. Plastics form a considerable portion 5–15% of MSW (municipal solid waste) by weight, which equals to 20–30% volumetric proportion [2]. It was reported that total MSW generation was 250 million tons in the US in 2011. Thus it can be estimated that the PSW (plastic solid waste) disposed is around 25 million in weight [3]. Hence, there is a need to convert this PSW into value added products and one such alternative is to use plastic waste as oil sorbent. To date, only few studies have shown the application of plastic waste as oil sorbent [4]–[6]. Recently, our research group has introduced a method for producing oil sorbent using plastic waste with high oil uptake capacity (99 g/g) which was much higher than previously reported values in the literature [7]. In this study, we

provide oil uptake saturation kinetics of the as prepared sorbent.

II. EXPERIMENTAL

Waste HDPE bottles were supplied by Nestle Hong Kong Ltd. Kaydol White Mineral oil and White Protopet 1S Petrolatum were supplied by Sonneborn. Irganox® 1010 was supplied by BASF and used as an antioxidant. Marine diesel oil was provided by Peak Biomass Energy Co. Ltd Its standard specifications state that the viscosity (at 50°C) and density (at 15°C) do not exceed 380 mm²/s and 991 kg/m³ respectively. UHMWPE (ultra-high molecular weight polyethylene) was supplied by Ticona.

Development of polymer film:

Initially sorbent film using only HDPE bottles was targeted; however HDPE, with relatively smaller molecular chains and molecular weight, could not be swollen significantly with the oil and could not be transformed into a pseudo-gel state; hence most of the oil did not mix well with HDPE and remained unmixed resulting in an inhomogeneous gel sheet. From the literature [8]–[10], it was observed that in order to meet this criteria, the polymer's molecular weight should be sufficiently higher with values ranging from 300,000 – 6,000,000 g/mol. Hence, it is in this context, 25 wt. % of higher molecular weight polymer UHMWPE (ultra-high molecular weight polyethylene) was added to help the blend undergo swelling by the solvent to get a pseudo gel state followed by extraction of solvent to create pores.

Polymer-Oil ratio:

The polymer-oil wt. ratio of 10:90 was used in which HDPE and UHMWPE constitute 7.5 and 2.5 wt. % respectively. The oil was later on removed leaving polymer that comprises of 75% HDPE and 25% UHMWPE. The sample was thus named HDP-75.

The waste HDPE bottle was chopped into small pieces and then washed with detergent. It was further cleaned with acetone and then dried in an oven for several hours. The blends were prepared using melt mixing in a static-mixer equipped with roller rotors. A similar procedure for material preparation was employed in our previous work [7], [11], [12].

Sorbent film of size 9cm² was initially weighed. Sorbent film of size 9 cm² was initially weighed and was then placed on the surface of marine diesel oil and allowed to float on the surface of oil for a certain period of time. Different contact times with the oil, namely 0.5, 1, 2, and 5 minutes were tested in order to find out the time required for the films to reach saturation.

The sorbent film was then removed from one of the four corners with the help of a forceps and allowed to hang in the air and drain for a specific duration. The oil laden sorbent film was then weighed to give the total weight of the sorbent and sorbate.

III. RESULTS AND DISCUSSION

Fig. 1 presents the SEM image of HDP-75. The pores of sorbent film as well as void spaces facilitate in oil sorption. Although few pores were in the range of 200 nm due to the stretching of fibrils, the pore size of the sorbent ranges from few nanometers to 100 nm.

Fig. 2 shows the saturation uptake kinetics of HDP-75 sorbent film. This experiment reveals that how much time our sorbent film is taken to reach its maximum capacity. It is a useful test as it will help us to understand how quickly saturation is reached. This information may prove to be significant especially to the sites where initial few minutes are crucial and any delay will cause the spreading of oil which will then be difficult to remove. It can be seen that sorbent film has reached its saturation time within 2 minutes. The experiment was further validated using different dripping time of 2 minutes, 5 minutes, and 45 minutes. Although, the oil uptake capacity is reduced for longer dripping durations, the saturation capacity is reached within 2 minutes of contact with oil. A set of 12 experiments were performed to obtain this plot and each experiment was repeated three times to obtain consistency.

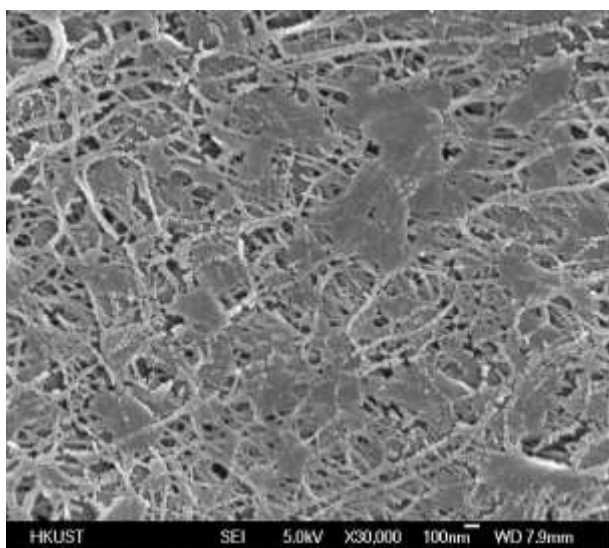


Figure 1. Porous structure of HDP-75 sorbent film

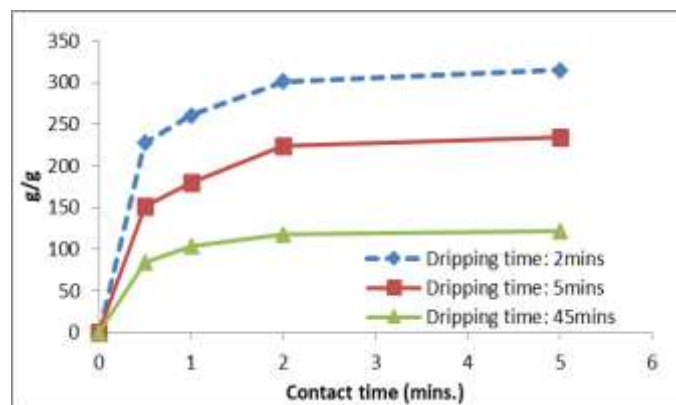


Figure 2. Sorbent-sorbate saturation study

ACKNOWLEDGMENT

Authors acknowledge the financial support from Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar.

REFERENCES

- [1] A. A. Al-Majed, A. R. Adebayo, and M. E. Hossain, "A sustainable approach to controlling oil spills," *J. Environ. Manage.*, vol. 113, pp. 213–27, Dec. 2012.
- [2] A. Bazargan, C. Hui, and G. McKay, "Porous Carbons from Plastic Waste," *Adv. Polym. Sci.* Springer, pp. 1–25, 2013.
- [3] J. Gug, D. Cacciola, and M. J. Sobkowicz, "Processing and properties of a solid energy fuel from municipal solid waste (MSW) and recycled plastics," *Waste Manag.*, vol. 35, pp. 283–292, Nov. 2015.
- [4] A. Aboul-Gheit, "Adsorption of spilled oil from seawater by waste plastic," *Oil Gas Sci. Technol.*, vol. 61, no. 2, pp. 259–268, Mar. 2006.
- [5] V. A. Tanobe, "Evaluation of flexible postconsumed polyurethane foams modified by polystyrene grafting as sorbent material for oil spills," *J. Appl. Polym. Sci.*, vol. 111, pp. 1842–1849, 2009.
- [6] A. M. Atta et al., "Porous polyurethane foams based on recycled poly(ethylene terephthalate) for oil sorption," *Polym. Int.*, vol. 62, no. 1, pp. 116–126, Jan. 2013.
- [7] J. Saleem, C. Ning, J. Barford, and G. McKay, "Combating oil spill problem using plastic waste," *Waste Manag.*, vol. 44, pp. 34–38, Jul. 2015.
- [8] A. Zachariades and P. Shukla, "Process for obtaining ultra-high modulus line products with enhanced mechanical properties," *US Pat.* 5,407,623, 1995.
- [9] L. Young-Keun, R. Jang-Weon, S. Jung-Moon, and J. Byoung-Cheon, "Microporous High Density Polyethylene film and Preparing method thereof," *US 2007/0218271 A1*, 2007.
- [10] X. Zhang and G. Rumierz, "Biaxially oriented porous membranes, composites, and methods of manufacture and use," *US 2011/0223486 A1*, 2011.
- [11] J. Saleem, A. Bazargan, J. Barford, and G. McKay, "Super-fast oil uptake using porous ultra-high molecular weight polyethylene sheets," *Polym. Adv. Technol.*, vol. 25, no. 10, pp. 1181–1185, Oct. 2014.
- [12] J. Saleem, A. Bazargan, J. Barford, and G. McKay, "Application of Strong Porous Polymer Sheets for Superior Oil Spill Recovery," *Chem. Eng. Technol.*, no. 3, pp. 482–488, 2015.

WRF SIMULATIONS OF URBAN CLIMATE: DO WE NEED LOW OR HIGH TOPOGRAPHIC RESOLUTION?

Md. Rafsan Nahian¹, Amir Nazem², Mohsen Moradi³, Shohel Mahmud⁴, William D. Lubitz⁵, Amir A. Aliabadi*

School of Engineering
University of Guelph
Guelph, Canada

¹mnahian@uoguelph.ca, ²anazem@uoguelph.ca, ³moradim@uoguelph.ca, ⁴smahmud@uoguelph.ca, ⁵wlubitz@uoguelph.ca,
*Corresponding Author: aliabadi@uoguelph.ca

Abstract— The performance of the Weather Research and Forecasting (WRF) model is evaluated with low and high topographic resolutions to observe the urban climate both in the surface level and higher altitudes. Overall, the model has presented a better performance in predicting the heat related properties compared to the momentum related properties. It has been observed that high wind bias is one of the main limitations of the model and the bias is higher at surface level compared to higher altitudes. High topographic resolution WRF simulations show better agreements with experimental observations in terms of predicting horizontal wind velocity while low topographic resolution WRF simulations deliver better results for reanalysis of the vertical wind velocity. Low topographic resolution WRF simulations predict relative humidity closer to experimental measurements while high topographic resolution WRF simulations predict surface level potential temperature more accurately. Overall, it is found that WRF under predicts the vertical wind velocity, potential temperature, and relative humidity while it over predicts the horizontal wind velocity both in surface level and higher elevations.

Keywords-component; WRF; mesoscale model; urban climate; GEOTPO 30 s; SRTM 1 s

I. INTRODUCTION

The Weather Research and Forecasting (WRF) model is a numerical weather prediction (NWP) mesoscale model designed for both research and operational atmospheric applications. It is suitable for an extensive span of applications across scales ranging from large eddy to global simulations. Such applications include real time NWP, data assimilation development and studies, parameterized physics research, regional climate simulations, air quality modeling, atmosphere-ocean-coupled simulations, and idealized simulations. In WRF, it is possible to combine and match the dynamical cores and physics packages of various models to optimize performance and this feature is particularly advantageous for intermodal comparisons and sensitivity studies [1].

The Weather Research and Forecasting (WRF) model has presented a high wind speed bias over land since the early versions of the model [2]. It has been identified that the bias still

exists in the latest versions which represents a limitation for the high demand of accurate wind estimations by different sectors [3]. Since current atmospheric models present an extensive spectrum of configuration options and parameters, selecting the optimum configuration among these alternatives has its own inherent challenges [4].

Researchers have emphasized the importance of the sensitivity of a model to change in its configuration settings. Various model configurations and parameter settings along with different initialization fields have been evaluated in this study [5]. WRF can be used for short term forecasting, reanalysis, assimilating in situ LIDAR observations and selecting the appropriate planetary boundary layer (PBL) scheme for a complex terrain [6]. It is shown that error minimization in the WRF simulation can be achieved by testing and choosing a suitable numerical and physical configuration for the region of interest. Increasing the horizontal and vertical grid resolution may lead to better reproduction of fine scale meteorological processes but this may not necessarily be true due to uncertainties in the overall performance of the various physical parameterizations and their responses to grid resolution [7].

Several physical parameterization schemes are available for microphysics, radiation and clouds, planetary boundary layer (PBL) schemes, surface layer (SL), and the land surface model (LSM). Such schemes have nonlinear interactions with each other and with the dynamical core of the model. Therefore, it becomes challenging to optimize the model due to these complex relationships [1]. Therefore, the effect of such nonlinearities on the simulation results cannot be predicted without performing the simulation. Also, such nonlinearities may cause model divergence. As a result, the effects must be investigated by parametric simulations. Model errors consist of the dependence on different numerical solvers, domain sizes, site location, initialization, boundary conditions, terrain and vegetation characteristics, and grid resolution both in horizontal and vertical directions [8]. Another important factor is the topography which has a great influence on the climate of a region. The substantial orographic features greatly influence the local and global weather and climate by changing the dynamics of the atmospheric circulation and interactions between the atmosphere and the land surface [9].

In most of the existing literature, the low topographic resolution (Global 30 Arc-Second) data-set has been applied to run WRF simulations and the results are compared with surface level observations. It is still a challenge for researchers to initialize high topographic resolution datasets for WRF simulations as well as observing the atmospheric properties in higher altitudes due to the complexity and unavailability of expensive weather sensing instruments. Hence, the main objectives of this paper are to observe the urban atmospheric boundary layer both near surface level and higher elevations to test the capability of WRF model using low and high topographic resolutions, and to observe the diurnal variations of different momentum and heat related properties. Subsequently, the numerical outcomes are compared with experimental observations of different meteorological instruments to compute the bias and error of the WRF model.

II. METHODOLOGY

A. Domain configurations

A particular distribution of WRF titled the Unified Environmental Modelling System (UEMS) version 18.8.1 has been tested with various grid configurations including the way of nesting, grid resolution, time invariant data resolution and geographical shifts in the domain center. The minimum grid distance of the nested domains should be at least four grid cells from parent boundary [10]. For computational stability, the magnitude of time steps (in seconds) are prescribed as the maximum 6 times the magnitude of the coarsest grid distance. This rule is associated with the coarsest grid and the time step is divided by 3 in each nest to maintain a ratio of 1/3 [1].

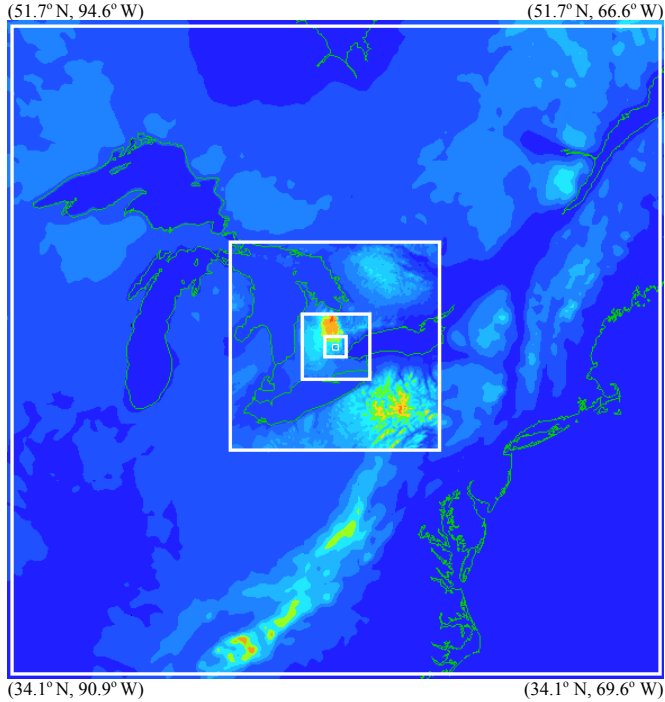


Figure 1. Size of the nested domains with Latitude and Longitude centered at Guelph, Ontario, Canada

Fig. 1 and Table I illustrate the size and configuration of the domains, respectively, where domain 1 is the largest domain and domain 5 is the smallest one. 45 vertical levels are used as opposed to more levels to gain computational speed.

TABLE I. CONFIGURATIONS OF THE NESTED DOMAINS

Parameter	Domain number				
	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5
Domain size (km × km)	2000	633	198	61.4	18.4
Grid size (km)	10	3.33	1.11	0.37	0.12
Number of grid elements	200	190	178	166	154
Top of the domain (km)	25	25	25	25	25
Vertical levels	45	45	45	45	45
Output frequency (minute)	60	60	60	60	60
GTOPO 30s topographical resolution (m)	900	900	900	900	900
SRTM 1s topographical resolution (m)	900	900	900	30	30

B. Initialization

The WRF simulations are run for three different days on August 03, 04, and 13 in 2018, for 36 hours including a spin up time of 12 hours. The Global Forecasting System (GFS) data-set has been used in the simulation that provides initial and time varying boundary conditions to WRF every three hours. WRF provides the option for obtaining time varying meteorological initialization fields from different sources. These sources of initialization fields provide data at different topographic resolutions and these predicted atmospheric properties are the critical aspect of this study.

The Global 30 Arc-Second (GTOPO 30s) and the Shuttle Radar Topography Mission 1 Arc-Second (SRTM 1s) datasets are used to modify the resolution of topography in domain 4 and domain 5. GTOPO 30s is a global dataset covering the full extent of latitude from 90 degrees south to 90 degrees north, and the full extent of longitude from 180 degrees west to 180 degrees east. As a low topographic initialization, the GTOPO 30s has geographic resolution around 900 m × 900 m. On the other hand, the Shuttle Radar Topography Mission (SRTM) is an international research effort that obtained digital elevation models on a near-global scale to generate the most complete high-resolution digital topographic database of the Earth. The SRTM 1s data-set has topographic resolution of about 30 m × 30 m, which is 30 times higher than the GTOPO 30s dataset [11]. Moreover, time invariant data including land water masks, land use, land cover classification and albedo have been obtained from the NCAR and MODIS database at all available resolutions. The surface height of domain 5 from sea level for both GTOPO 30 s and SRTM 1s datasets are shown in Fig. 2 and Fig. 3 respectively.

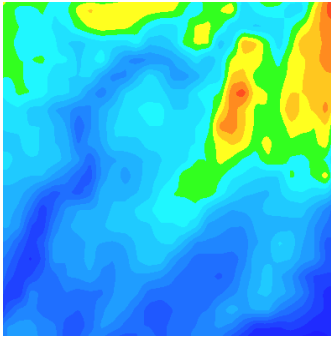


Figure 2. Surface height from sea level (GEOTPO 30s)

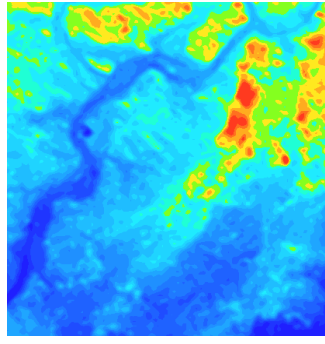


Figure 3. Surface height from sea level (SRTM 1s)

C. Physical options

The planetary boundary layer (PBL) scheme implemented in the model plays a decisive role on the accuracy of reanalyzed state and flow within the PBL because the wind varies according to the stability and baroclinic instability of the PBL. Some researchers have studied performing a sensitivity test of the WRF model and found that the Yonsei University scheme (YSU) shows improvement over the other PBL schemes of WRF [7, 12]. Hence, YSU scheme is applied in the simulation which uses identical profile functions for momentum and heat assuming turbulent Prandtl number as a constant.

D. Validation

The results of WRF model should be verified because of the uncertainty in model implementation for any specific location. Hence, experimental datasets from a Mini SODAR 4000 series by Atmospheric Systems Corporation located in the Turfgrass area at Guelph are used to validate the WRF simulations in higher elevation wind for every 4 hours while Guelph Turfgrass weather station's hourly dataset is used for surface level. The Guelph Turfgrass Institute station is identified with World Meteorological Organization (WMO) identification number 71833. This data is accessible with an hourly time resolution. Temperature data is collected at 2 m while wind data is collected at 10 m above the ground [13].

A quantitative comparison between the experimental observations and WRF model is performed by determining the Bias, Fractional Bias (FB) and the Normalized Mean Square Error (NMSE) [14] defined by

$$\text{BIAS} = \frac{\sum_{i=1}^n O_i - \sum_{i=1}^n M_i}{n}, \quad (1)$$

$$\text{FB} = \frac{\sum_{i=1}^n O_i - \sum_{i=1}^n M_i}{0.5 (\sum_{i=1}^n O_i + \sum_{i=1}^n M_i)}, \quad (2)$$

$$\text{NMSE} = \frac{\sum_{i=1}^n (O_i - M_i)^2}{(\sum_{i=1}^n O_i)(\sum_{i=1}^n M_i)}. \quad (3)$$

Here, O and M represent the experimental observation and WRF model, respectively, while n represents the number of

sample. Furthermore, for wind direction, the bias is calculated differently because wind direction is a circular variable. For each model (M) and observation (O) comparison, the bias is calculated using

$$\text{BIAS} = (|O-M|) \% 180. \quad (4)$$

Here all angles are in degrees and the remainder of a positive difference is taken by dividing this difference with 180 degrees. This ensures that the difference is always between 0 and 180 degrees. Of course, for multiple comparisons an average or median can be taken for this bias.

III. RESULTS AND DISCUSSION

The comparison between WRF simulations in domain 5 and the experimental observations on August 3, 4 and 13, 2018 is evaluated in this discussion. In each time interval, the median value of every property is considered to perform statistical analysis because the mean value may not be a fair representation of the data due to the mean value being easily influenced by the outliers in the data.

A. Horizontal wind velocity

Horizontal wind speed is a fundamental atmospheric field property caused by air moving from high to low pressure. The dispersion process of pollutants in any region is highly affected by wind speed where diffusion takes place in the direction of plume transport [15]. Figs. 4-6 show the fractional bias of horizontal wind speed at 10 m (near surface level), 70 m (Mini SODAR) and 170 m (Mini SODAR) respectively. These altitudes are pressure heights¹. In these figures the error bars represent the Normalized Mean Square Error (NMSE).

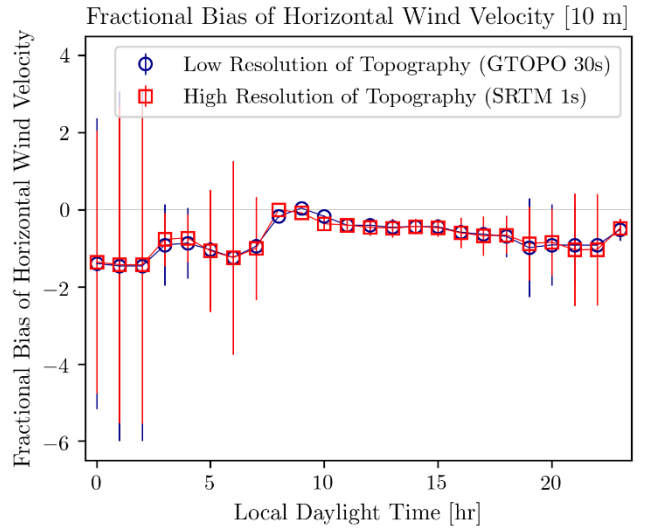


Figure 4. Fractional bias of horizontal wind velocity at 10 m

¹ www.weather.gov/epz/wxcalc_pressurealtitude

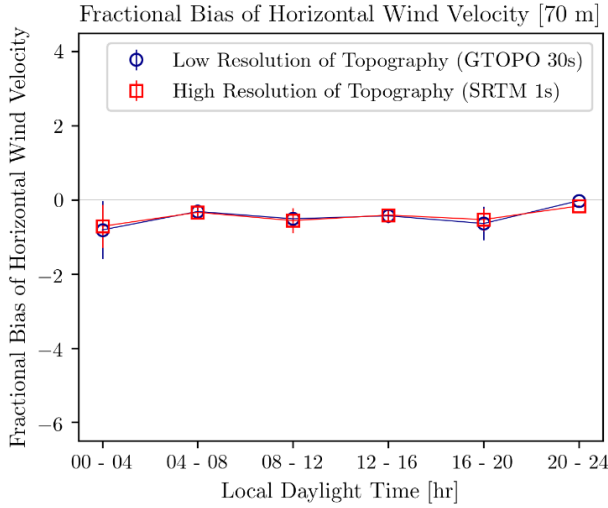


Figure 5. Fractional bias of horizontal wind velocity at 70 m

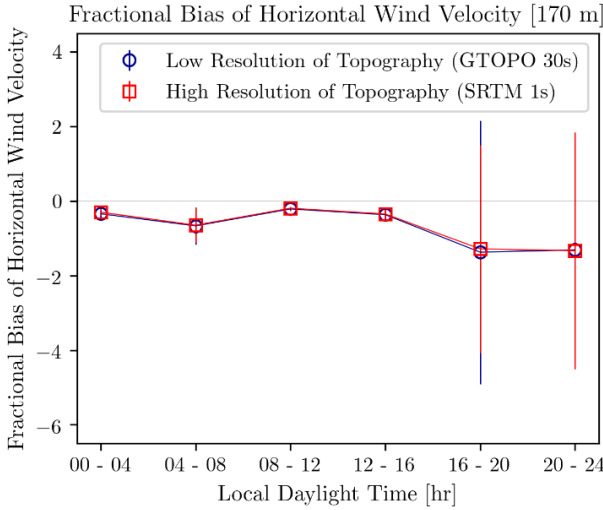


Figure 6. Fractional bias of horizontal wind velocity at 170 m

It has been observed that in convective hours when the atmosphere is thermally unstable, WRF shows less fractional bias and error than the thermally stable conditions at late night and early morning. There are no significant differences between low and high topographic resolution WRF simulations although at 10 m elevation high topographic resolution WRF simulations sometimes show marginally better agreement than the low topographic resolution WRF simulations. Low altitude wind speeds at surface level are influenced by surface roughness and atmospheric stability. Low topographic resolution WRF simulations do not account for realistic surface roughness in the topography and therefore exhibit unwanted bias in wind speed compared to high topographic resolution WRF simulations.

B. Vertical wind speed

Vertical velocity is a key factor for cloud development, precipitation, and development of weather systems. Figs. 7-8 show the fractional bias of vertical wind speed at 70 m (Mini SODAR) and 170 m (Mini SODAR), respectively. It is also

noted that the 10 m observation does not include vertical wind speed. In these figures the error bars represent the Normalized Mean Square Error (NMSE).

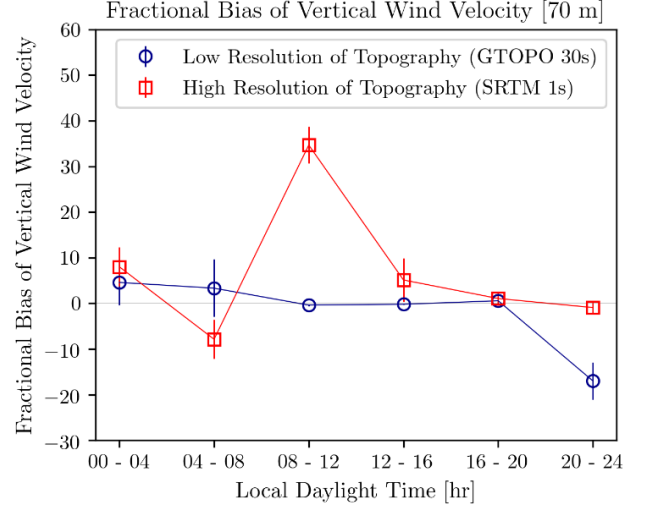


Figure 7. Fractional bias of vertical wind velocity at 70 m

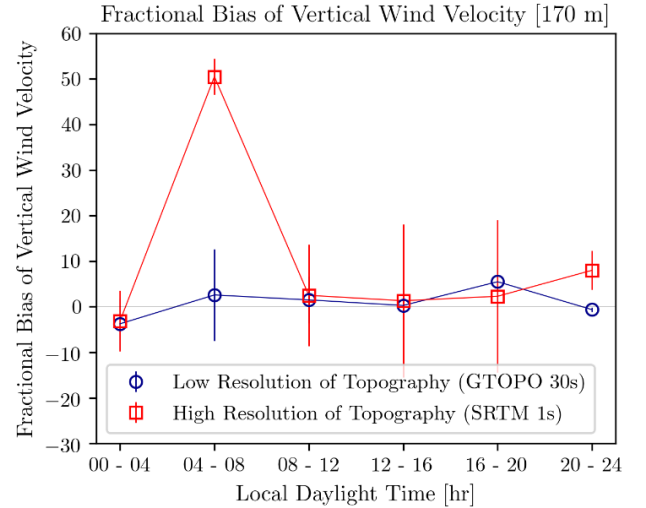


Figure 8. Fractional bias of vertical wind velocity at 170 m

The present results of vertical velocity profile show better agreements with Mini SODAR at thermally unstable condition compared to stable condition. However, there are large differences between low and high topographic resolution WRF simulations at stable conditions and with low topographic resolution WRF simulations predicting lower bias and error. Nevertheless, it should be borne in mind that, as a sonic instrument, the Mini SODAR has limitations in measuring profiles of vertical wind accurately in higher altitudes which may affect the experimental measurements.

C. Wind direction

Wind usually flows from high pressure to low pressure regions of the atmosphere. The median values of wind direction bias at 10 m (near surface level), 70 m (Mini SODAR) and 170 m (Mini SODAR) are shown in Fig. 9-11.

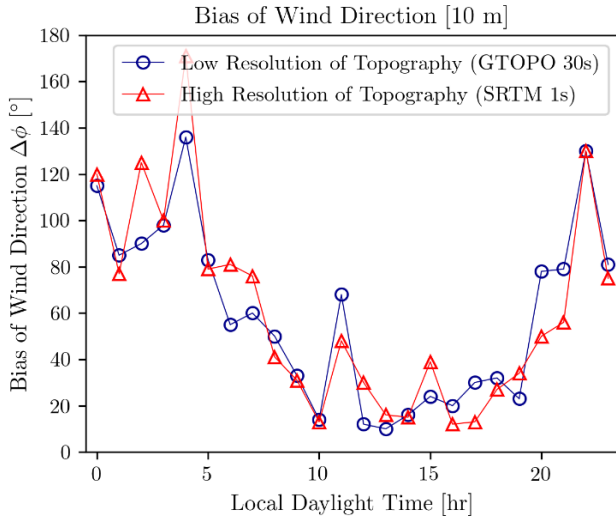


Figure 9. Median of wind direction bias at 10 m

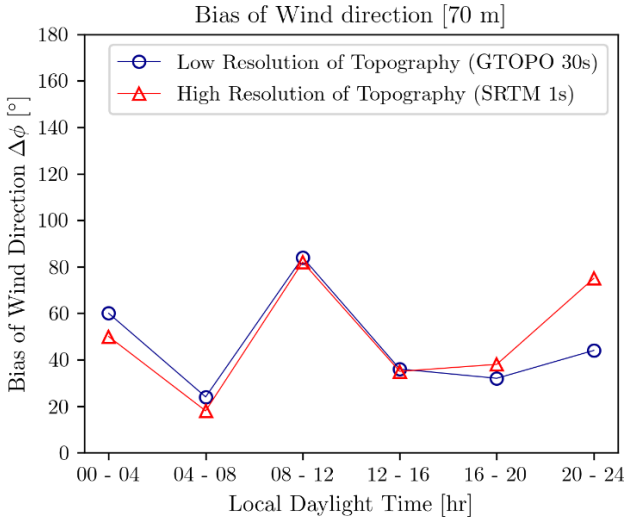


Figure 10. Median of wind direction bias at 70 m

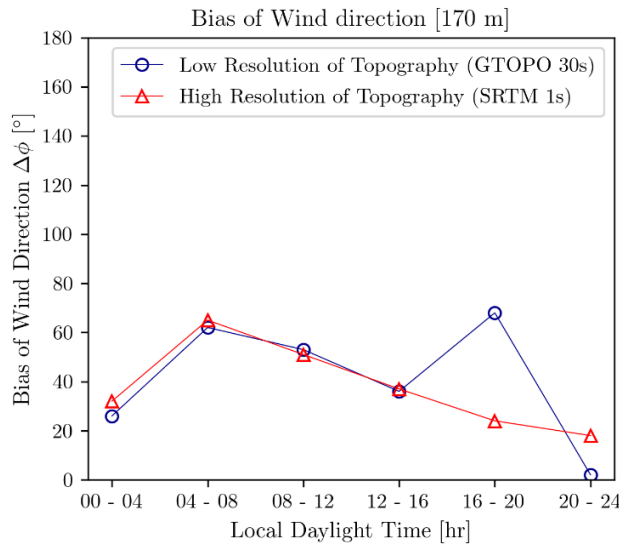


Figure 11. Median of wind direction bias at 170 m

It has been certainly identified that there is an extensive bias of wind direction in thermally stable conditions while the bias is less in unstable condition. In stable condition, the wind flows very slowly which causes more complexity to determine the exact wind direction for a specific time. Moreover, the bias of wind direction is more near surface level than the higher altitudes. This is due to the surface roughness, and shear stress which results in frequent changes of wind directions. However, there is no significant difference of wind direction between low and high topographic resolution WRF simulations during most of the times.

D. Potential temperature

Potential temperature is a more dynamically important quantity than the actual temperature as potential temperature is not affected by the physical lifting or sinking associated with flow over obstacles or large-scale atmospheric turbulence [16]. Fig. 12 shows the fractional bias of potential temperature at 2 m (near surface level) in domain 5. Figs. 13-14 show the spatial distributions of potential temperature at stable condition (0200 LDT) while Figs. 15-16 illustrate the potential temperature at unstable condition (1400 LDT) in domain 5 for low and high topographic resolutions, respectively.

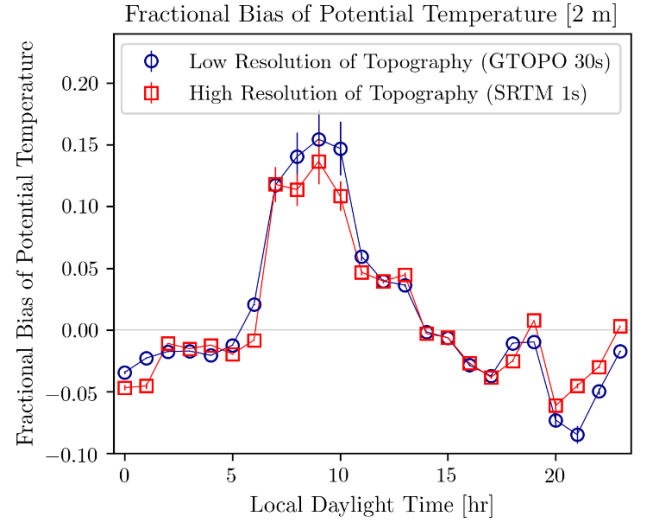


Figure 12. Fractional bias of potential temperature at 2 m

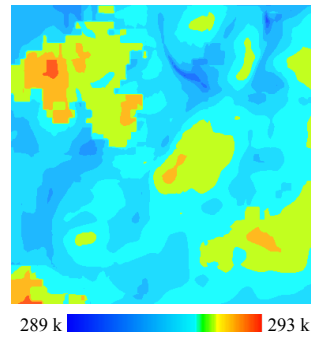


Figure 13. Potential temperature at 0200 LDT (GTOPO 30s)

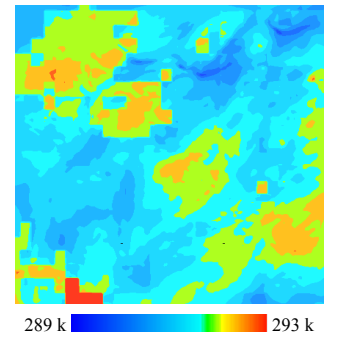


Figure 14. Potential temperature at 0200 LDT (SRTM 1s)

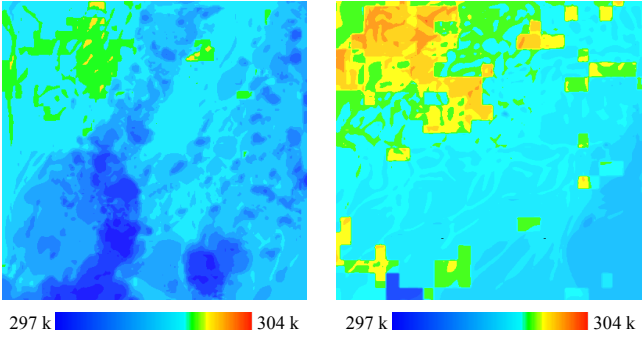


Figure 15. Potential temperature at 1400 LDT (GTOPO 30s)

Figure 16. Potential temperature at 1400 LDT (SRTM 1s)

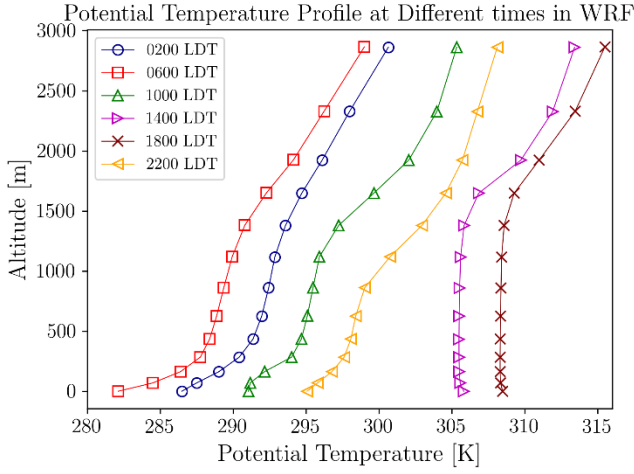


Figure 17. Potential temperature profile in WRF (SRTM)

In terms of predicting the potential temperature WRF shows better agreements at stable condition. Moreover, it has been observed that the bias is negative in stable condition which means WRF overpredicts potential temperature during this time. In contrast, WRF underestimates the potential temperature at unstable condition as the fractional bias shows positive trend during convective hours. However, the bias is almost the same in most of the times for low and high topographic resolution WRF outputs. Besides, the spatial distribution of potential temperature shows similar magnitude at stable conditions although there are much deviations of potential temperature between low and high topographic resolutions at unstable condition. It appears that the surface level temperature is better well-mixed for high topographical resolution during convective hours. This can explain the difference in the results during convective hours because surface roughness variations can possibly be responsible for a higher degree of surface level mixing. Fig. 17 shows the potential temperature profile of WRF in thermally stable and unstable conditions. When the atmosphere is stable at night and early morning, there is positive potential temperature gradient, while the potential temperature gradient is negative at lower altitudes in thermally unstable condition, especially during the afternoon hours.

E. Relative humidity

Besides water content, relative humidity depends on temperature and the pressure of the system of interest. The same amount of water vapor results in higher relative humidity in cool air than warm air. Fig. 18 shows the fractional bias of relative humidity at 2 m (near surface level) in domain 5. Moreover, the spatial distributions of relative humidity at stable condition (0200 LDT) are shown in Figs. 19-20 while the relative humidity at unstable condition (1400 LDT) in domain 5 are illustrated in Fig. 21-22 for low and high topographic resolutions.

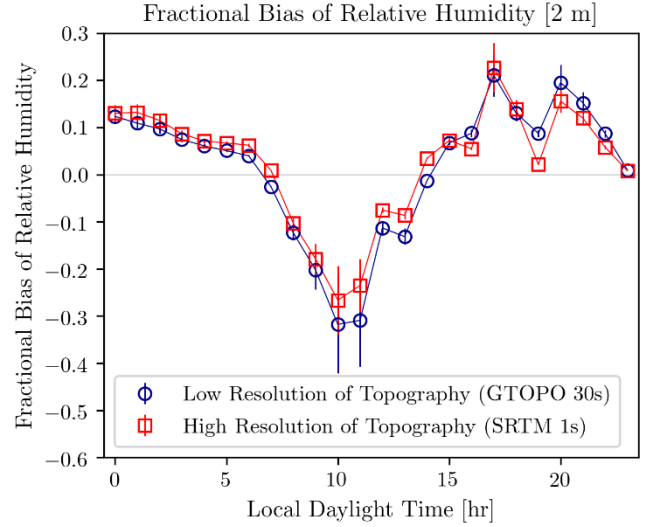


Figure 18. Fractional bias of relative humidity at 2 m

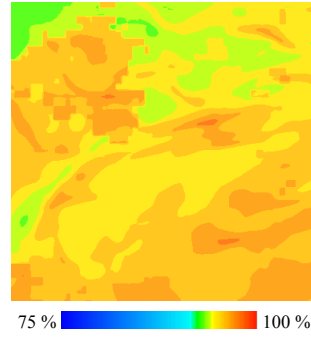


Figure 19. Relative humidity at 0200 LDT (GTOPO 30s)

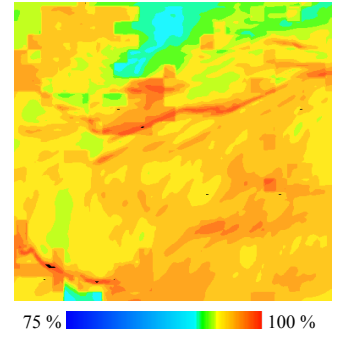


Figure 20. Relative humidity at 0200 LDT (SRTM 1s)

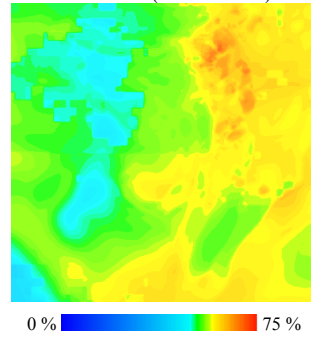


Figure 21. Relative humidity at 1400 LDT (GTOPO 30s)

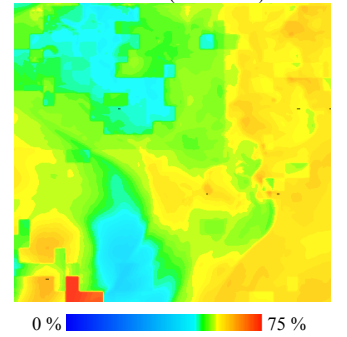


Figure 22. Relative humidity at 1400 LDT (SRTM 1s)

WRF shows a negative co-relation for fractional bias of relative humidity and temperature. WRF under predicts the relative humidity at stable condition as the bias is positive while it overestimates the relative humidity at unstable condition. However, there is no significant difference of relative humidity between low and high topographic resolution WRF outputs in both stable and unstable conditions.

F. Turbulent vertical heat flux

The turbulent vertical heat flux involves the eddy induced fluxes of heat. In vertical direction, the kinematic heat flux is denoted by $\overline{w\theta}$, where w is the standard deviation of vertical velocity and θ is the fluctuation of potential temperature [17]. At surface level, the small-scale turbulence determines the vertical heat flux while the large-scale turbulence heat flux is zero. With the increase of height, the large-scale turbulence heat flux increases and reaches to a maximum value [18].

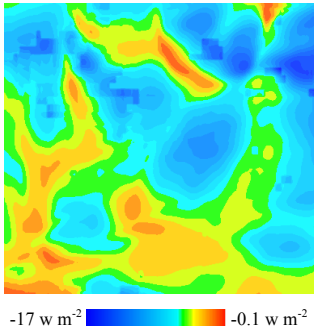


Figure 23. Vertical turbulent heat flux at 0200 LDT (GTOPO 30s)

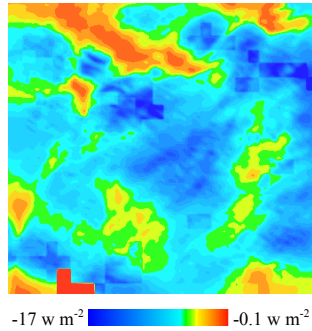


Figure 24. Vertical turbulent heat flux at 0200 LDT (SRTM 1s)

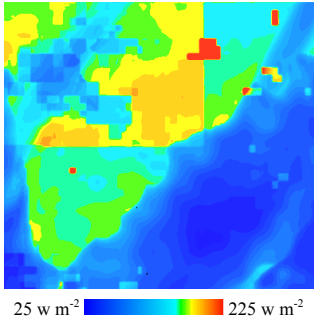


Figure 25. Vertical turbulent heat flux at 1400 LDT (GTOPO 30s)

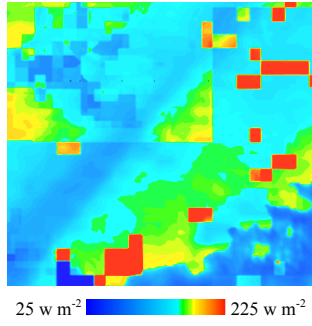


Figure 26. Vertical turbulent heat flux at 1400 LDT (SRTM 1s)

WRF shows less variations of vertical turbulent heat flux in both low and high topographic resolution simulations in stable condition. However, the differences are more at unstable condition when the heat flux changes rapidly with time. From Figs. 23-26, it has been observed that high topographic resolution WRF simulation predicts higher heat flux than the low topographic resolution WRF outputs. During convective hours, heat flux seems to exhibit a different spatial pattern between low and high topographic resolution simulations.

G. Error analysis

The results of WRF model are verified because of inherent uncertainty in model implementation at a specific location. The

experimental data from the Mini SODAR and Guelph Turfgrass weather station are used to validate the WRF simulations for those specific locations in domain 5. A quantitative comparison between the experimental observations and WRF model is also performed by determining the Fractional Bias (FB) and the Normalized Mean Square Error (NMSE). Table II shows the FB and NMSE of different parameters to compare WRF outputs with experimental observations while Table III shows the optimum topographic resolutions of WRF for different properties based on the results in Table II.

TABLE II. FRACTIONAL BIAS AND NMSE OF WRF

Parameter	Fractional bias		NMSE	
	Low resolution	High resolution	Low resolution	High resolution
Horizontal Velocity (higher elevation)	-0.363	-0.345	0.137	0.123
Vertical Velocity (higher elevation)	0.071	0.289	0.005	0.086
Horizontal Velocity (10 m)	-0.661	-0.662	0.491	0.492
Potential temperature (2 m)	0.012	0.011	0.00014	0.00011
Relative humidity (2 m)	0.017	0.029	0.00029	0.00089

TABLE III. OPTIMUM TOPOGRAPHIC RESOLUTIONS OF WRF

Parameter	Optimum topography
Horizontal Velocity (higher elevation)	Low topographic resolution
Vertical Velocity (higher elevation)	Low topographic resolution
Horizontal Velocity (10 m)	High topographic resolution
Potential temperature (2 m)	High topographic resolution
Relative humidity (2 m)	Low topographic resolution

It is clearly identified that although there are small differences in predicting horizontal wind velocity at surface level and higher altitudes, the differences between low and high topographic resolution WRF simulations are significant in predicting the vertical wind velocity. Low topographic resolution simulations realize better results than high topographic resolution simulations as far as vertical wind is concerned. WRF over predicts the horizontal wind velocity both at surface level and higher elevation while it under predicts the vertical velocity at the measured elevations. Moreover, the bias of horizontal wind velocity is also higher at surface level compared to higher altitudes. WRF shows a correlation in fractional bias between potential temperature and relative humidity. High topographic resolution simulations produce slightly better results than low topographic resolution simulations in predicting the potential temperature while low topographic resolution simulations show better agreements with experimental observation in predicting the relative humidity.

IV. CONCLUSION

The sensitivity tests of low and high topographic resolution WRF simulations are performed to observe the urban climate both in the surface level and higher altitudes. The numerical results are compared to the experimental observations of a Mini

SODAR and Guelph Turfgrass weather station to validate the WRF outputs. Overall, the model has presented a better performance in predicting heat related properties such as temperature and relative humidity compared to the momentum related properties. It is also determined that low topographic resolution simulations show excellent agreement with observed relative humidity while high topographic resolution simulations predict the potential temperature more accurately at surface level.

On the other hand, the error of horizontal wind velocity is less with high topographic resolution WRF simulations while low topographic resolution WRF simulations deliver much better results in terms of predicting the vertical wind velocity. It is also found that predicting the surface level wind velocity is one of the main limitations of WRF as the fractional bias is almost double compared to the bias of wind velocity at higher elevations. As far as wind direction is concerned, it is also found that predictions of surface level wind direction are less reliable, specially at stable conditions, compared to wind directions at higher altitudes. However, as the atmosphere is much calm during stable hours, the wind moves very slowly which may also cause inaccurate measurement of wind directions by the anemometers and result in a large bias of wind direction in thermally stable condition.

ACKNOWLEDGMENT

The Mini SODAR was implemented with the help of Benjamin Dyer and Manoj K. Nambiar. In-kind technical support for this work was provided by Rowan Williams Davies and Irwin Inc. (RWDI). This work was supported by the Discovery Grant program (401231) from the Natural Sciences and Engineering Research Council (NSERC) of Canada; Government of Ontario through the Ontario Centres of Excellence (OCE) under the Alberta-Ontario Innovation Program (AOIP) (053450); and Emission Reduction Alberta (ERA) (053498). OCE is a member of the Ontario Network of Entrepreneurs (ONE).

REFERENCES

- [1] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, and J. G. Powers, "A description of the advanced research WRF Version 3", NCAR technical note, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA, 2008.
- [2] W. Y. Cheng, and W. J. Steenburgh, "Evaluation of surface sensible weather forecasts by the WRF and the Eta models over the western United States", *Weather and Forecasting*, vol. 20(5), pp. 812-821, 2005.
- [3] G. Roux, Y. Liu, L. D. Monache, R. S. Sheu, and T. T. Warner, "Verification of high resolution WRF-RTFDDA surface forecasts over mountains and plains", In 10th WRF users' workshop, pp. 20-23, 2009.
- [4] J. Nossent, P. Elsen, and W. Bauwens, "Sobol'sensitivity analysis of a complex environmental model", *Environmental Modelling & Software*, vol. 26(12), pp. 1515-1525, 2011.
- [5] S. Hirabayashi, C. N. Kroll, and D. J. Nowak, "Component-based development and sensitivity analyses of an air pollutant dry deposition model", *Environmental Modelling & Software*, vol. 26(6), pp. 804-816, 2011.
- [6] R. L. Carpenter, B. L. Shaw, M. Margulis, K. S. Barr, T. Baynard, D. Yates, and J. Sharp, "Short-term numerical forecasts using WindTracer LIDAR data", In Fourth Conference on Weather, Climate, and the New Energy Economy, American Meteorological Society, Austin, TX., 2013.
- [7] D. Carvalho, A. Rocha, M. Gómez-Gesteira, and C. Santos, "A sensitivity study of the WRF model in wind simulation for an area of high wind energy", *Environmental Modelling & Software*, vol. 33, pp. 23-34, 2012.
- [8] N. K. Awan, H. Truhetz, and A. Gobiet, "Parameterization-induced error characteristics of MM5 and WRF operated in climate mode over the Alpine region: an ensemble-based analysis", *Journal of Climate*, vol. 24(12), pp. 3107-3123, 2011.
- [9] S. V. Kumar, C. D. Peters-Lidard, J. L. Eastman, and W. K. Tao, "An integrated high-resolution hydrometeorological modeling testbed using LIS and WRF", *Environmental Modelling & Software*, vol. 23(2), pp. 169-181, 2008.
- [10] W. Wang, C. Bruyere, M. Duda, J. Dudhia, D. Gill, H. C. Lin, and J. Mandel, "ARW version 3 modeling system user's guide", Mesoscale & Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA, 2015.
- [11] T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, and D. Seal, 2007, "The shuttle radar topography mission", *Reviews of geophysics*, vol. 45(2), 2007.
- [12] V. S. Challa, J. Indracanti, M. K. Rabarison, J. Young, C. Patrick, J. M. Baham and A. Yerramilli, "Numerical experiments on the sensitivity of WRF-CMAQ simulations of air quality in the Mississippi Gulf coastal region to PBL and Land surface models", In The 6th Annual CMAS Conference, Chapel Hill, NC, October 13, 2007.
- [13] A. A. Aliabadi, M. Moradi, D. Clement, W. D. Lubitz, and B. Gharabaghi, "Flow and temperature dynamics in an urban canyon under a comprehensive set of wind directions, wind speeds, and thermal stability conditions", *Environmental Fluid Mechanics*, pp. 1-29, 2018.
- [14] A. A. Aliabadi, E. S. Krayenhoff, N. Nazarian, L. W. Chew, P. R. Armstrong, A. Afshari, and L. K. Norford, "Effects of roof-edge roughness on air temperature and pollutant concentration in urban canyons", *Boundary-Layer Meteorology*, vol. 164(2), pp. 249-279, 2017.
- [15] D. A. Vallero, "Fundamentals of air pollution", Academic press, 2014.
- [16] J. T. Moore, "Isentropic Analysis Techniques: Basic Concepts", COMET COMAP, Retrieved on January 11, 2018.
- [17] A. A. Aliabadi, "Theory and Applications of Turbulence: A Fundamental Approach for Scientists and Engineers", Amir Abbas Aliabadi Publications, Guelph, 2018.
- [18] E. Palmén, "On the mechanism of the vertical heat flux and generation of kinetic energy in the atmosphere", *Tellus*, vol. 18(4), pp. 838-845, 1966

A VERTICAL DIFFUSION MODEL TO PREDICT PROFILES OF TEMPERATURE WITHIN THE LOWER ATMOSPHERIC SURFACE LAYER: SIMPLE OR COMPLICATED?

Mohsen Moradi¹, Manoj K. Nambiar, Amir Nazem, Md. Rafsan Nahian, William D. Lubitz, Amir A. Aliabadi*

School of Engineering
 University of Guelph

¹moradim@uoguelph.ca,

*Corresponding Author:
 aliabadi@uoguelph.ca

Benjamin Dyer
 Department of Physics and
 Astronomy
 McMaster University
 Hamilton, Canada

E. Scott Krayenhoff
 School of Environmental Science
 University of Guelph
 Guelph, Canada

Abstract—Knowledge of urban and rural climates is of crucial importance in urban climate control and wind farm development. A vertical diffusion model that predicts vertical profiles of temperature in a rural area was developed. This model has the capability to predict the thermally stable and unstable conditions as they vary diurnally. A different parameterization, which is independent of the well-known Obukhov length, was introduced to calculate the turbulent diffusion coefficient. The source/sink term in the diffusion equation was also parameterized based on heat fluxes at surface level. The net heat flux was composed of shortwave/longwave radiation fluxes and release of sensible/latent heat fluxes from vegetation/ground. As a result, the model can transition between thermally stable and unstable conditions over the diurnal cycle. The assumption of a constant specific humidity profile in this model was assessed by comparing vapour and saturation vapour pressures up to low altitudes, well within the lower portion of the atmospheric surface layer. In addition, a Doppler miniSoDAR system was operated to sample the vertical wind speed profiles in a typical rural area. The diurnal wind speed profile was compared to the logarithmic law. From an urban climate point of view, this model can be employed on its own or it can be coupled to existing urban canopy models to predict the urban microclimate.

Keywords—component; length scale, temperature profile, vapour pressure, vertical diffusion, wind profile

I. INTRODUCTION

Modeling the vertical profile of meteorological quantities in the lowest range of altitudes within the atmospheric surface layer is becoming important since this layer drives the microclimate environments. For example, accurate wind speed profiles are required to perform wind energy feasibility studies and to predict shear loads on the wind turbine blades. In the context of urban planning, vertical profiles of wind speed and temperature in both urban and rural areas are of great importance to predict and control the urban environment to meet human comfort.

It has been suggested that turbulent transport in rural areas, which is characterized by horizontally uniform distribution of temperature and wind, is only significant in the vertical direction. The well-known K – theory, which is based on the down-gradient transport hypothesis, is generally accepted to approximate turbulent transport in the vertical direction, where K_t is turbulent diffusivity governing the strength of transport [1]. This hypothesis indicates that, for example, heat flows down from warm regions to cold regions at a rate proportional to K_t multiplied by gradients of mean temperature. In this theory, vertical motion is accounted for small scales of turbulent mixing based on mixing length scales [2].

During the last few decades, numerous studies have focused on modeling vertical profiles of wind speed and temperature based on turbulent mixing length models [2-4]. The very first version of the length scale parameterization suggested a linear relationship between mixing length and height for near neutral conditions but $(-1/2)$ and $(-1/4)$ power laws for unstable conditions [3]. It has been suggested that wind speed and temperature profiles, specifically turbulent diffusion coefficient within the surface layer, is largely controlled by friction velocity, surface roughness, and stability condition [2]. The majority of existing models parameterize the mixing length scale using the Obukhov length, which is defined as

$$L = \frac{-u_{*0}^3}{\kappa(g/T)w'T'} \quad (1)$$

where u_{*0} is friction velocity near the ground, κ is the von Kármán constant, (g/T) is buoyancy parameter, and $\overline{w'T'}$ is the kinematic turbulent vertical heat flux at the surface. It has been suggested that the sign of (z/L) , where z is the height above the surface, implies the stability condition: negative represents the thermally unstable atmosphere and positive represents the thermally stable atmosphere [5]. Despite extensive use of L to relate momentum and heat fluxes to bulk meteorological parameters, serious limitations have been found with the use of L particularly for near neutral and stable atmospheric conditions [6]. During neutral condition, vertical heat flux approaches close to zero leading to large values (positive or negative) for

Obukhov length. Hogstrom [7] also showed substantial scatter between experiments for wind and temperature gradient as a function of L for neutral conditions. Efforts have been made to parameterize the turbulent mixing length scales in the lowest few hundred meters of the atmospheric boundary layer using alternative methods [8]. For example, it has been suggested that turbulent mixing length scales can be divided into three regions. In the first region mixing length increases linearly with height within the surface layer considering stability correction; in the next region just above the surface mixing length depends only on stability condition; and in the last region at the top of the surface layer the mixing length is negligible compared to the other two regions [8].

II. OBJECTIVE

Many studies in the literature parameterize turbulent diffusivity based on mixing length scales, which are still suffering from Obukhov length limitations. To cope with these problems, a different parameterization for turbulent diffusivity is introduced, which is mainly inspired by explicit formulation of the turbulent length scale [9] and systematic scaling analysis of turbulent flow [10]. The model will distinguish between stable and unstable conditions based on direct solar radiation in a rural area. The study also investigates how well the rural wind speed profile can be formulated using field measurements in a typical area in Guelph, Canada.

III. OBSERVATION SITE

In this study, data was collected from a rural field campaign held in the Guelph Turfgrass Institute from July 14, 2018 to September 4, 2018. The rural area is far away from the nearby urban areas by 1-2 km (Fig. 1).

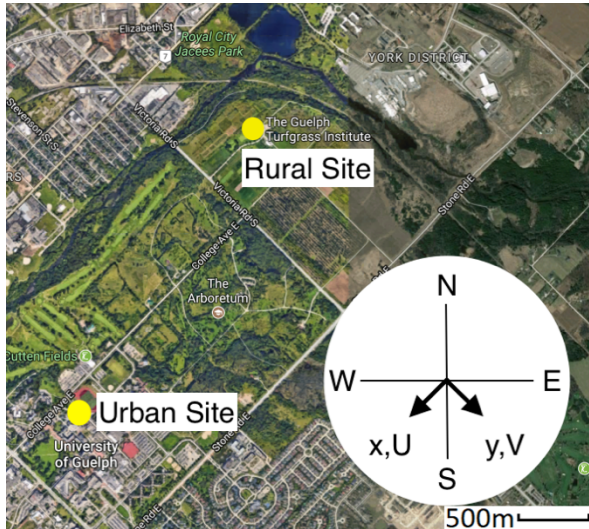


Figure 1. Top view of rural area which is located at the Guelph Turfgrass Institute at 43.5473°N and 80.2149°W.

In the rural site, a Doppler miniSoDAR was operated to measure wind speed and wind direction from 30 m to 200 m at 10 m vertical resolution (Fig. 2). It outputs average data every 30 minutes. In the same location, Environment and Climate Change Canada is operating a meteorological station by

measuring temperature and humidity at 2 m elevation and wind speed and direction at 10 m elevation.



Figure 2. The Doppler miniSoDAR instrument operated in the rural area to measure the wind speed and direction profiles

IV. METHODOLOGY

In this study, wind speed profile is calculated using the logarithmic equation

$$\bar{S} = \frac{u_*}{\kappa} \ln \left(\frac{z-d}{z_0} \right), \quad (2)$$

where d is rural displacement height and z_0 is rural aerodynamic roughness length scale. Raupach et al. [11] suggested that z_0 is 0.2 for trees. u_* is friction velocity. Aliabadi et al. [10] performed a systematic scaling analysis of turbulence parameters using data collected from a microclimate field campaign in Guelph, Canada, to parameterize u_* . They found a high linear correlation coefficient between u_* and mean horizontal velocity and suggested the following equation for the rural cite

$$u_* = 0.07\bar{S} + 0.12. \quad (3)$$

Fig. 3 shows measurements from the miniSoDAR from 30 meters up to five times of average building height (around 80 m) in the surrounding areas compared to logarithmic wind speed profile. The logarithmic model with a new equation for friction velocity predicts wind speed profile in the rural area reasonably well. However, it very often overestimates wind speed up to the height of 50 m under thermally stable conditions.

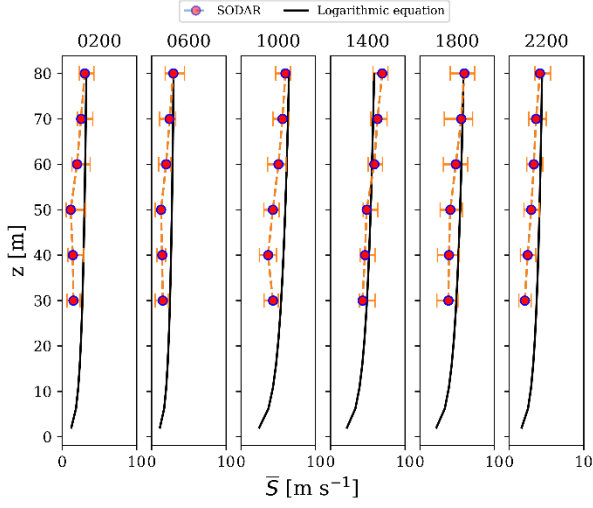


Figure 3. Comparison between logarithmic wind speed profile and miniSoDAR measurements; times in Local Standard Time (LST)

In atmospheric modeling, it is a common practice to characterize turbulence using vertical diffusion of an atmospheric quantity of interest, such as momentum, heat, and humidity [12]. This approach follows the Reynolds averaging procedure in which the variable of interest (here temperature) is decomposed into mean and fluctuating parts. The heat equation in the horizontally-homogeneous region is reduced to

$$\frac{\partial \bar{\theta}}{\partial t} + \bar{W} \frac{\partial \bar{\theta}}{\partial z} = \frac{\partial \overline{w'\theta'}}{\partial z} + \gamma, \quad (4)$$

where $\bar{\theta}$ is the potential temperature, \bar{W} is the mean velocity component in the z or vertical direction, $\overline{w'\theta'}$ is the kinematic turbulent vertical heat flux, and γ is the heat sink/source term also known as cooling or heating rate. Assuming steady state conditions and zero mean vertical velocity the left-hand side of equation (4) is forced to zero. From K -theory, turbulent heat flux is proportional to the gradient of mean potential temperature. Thus, equation (4) can be simplified to

$$0 = \frac{\partial}{\partial z} \left(-\frac{K_t}{Pr_t} \frac{\partial \bar{\theta}}{\partial z} \right) + \gamma, \quad (5)$$

where K_t is turbulent diffusivity and Pr_t is the turbulent Prandtl number (assumed to be one here [13] but can be changed). It has been suggested that K_t is a function of mixing length scale, which needs to be parameterized given the stability condition [8].

A. Parameterization of Mixing-Length

Prandtl [14] proposed that turbulent diffusivity is proportional to wind shear and it can be formulated by length scale (ℓ), which inherently depends on atmospheric stability condition

$$K_t = \ell^2 \frac{\partial \bar{S}}{\partial z}, \quad (6)$$

where \bar{S} is the mean horizontal wind speed. The approach to parameterize length scale, as opposed to previous models, is not a function of Obukhov length. As mentioned before, the diurnal variation of Obukhov length exhibits very large values in neutral

or weakly stable conditions, where the vertical heat flux is negligible (see Fig. 4 in local standard time (LST) as obtained from a climate model). According to previous formulations for mixing length [8], it can cause unexpected values for mixing length and consequently turbulent diffusivity.

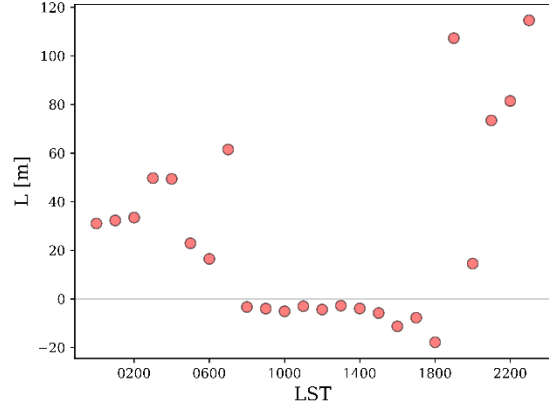


Figure 4. Diurnal variation of Obukhov length; times in Local Standard Time (LST)

Inspired by Gryning et al. [8], the mixing length scale can be parameterized as

$$\frac{1}{\ell} = \frac{1}{C_{cr}} \left(\frac{1}{\kappa z} + \frac{1}{C_* u_*} \right), \quad (7)$$

where the first term on the right-hand side indicates a linear relationship with height near the surface, while the second term restricts the value of length scale in the upper part of the atmospheric surface layer, and C_* is a correction factor for the second term. C_{cr} is a scaling correction factor which can be optimized to 2 during unstable conditions and 1.5 during stable conditions. The most appropriate and practical indicator for thermally unstable conditions is presence of incoming solar radiation on the surface while the absence of incoming solar radiation on the surface indicates stable conditions. This approach is mostly applicable for not-so-much humid air at short vertical scales.

B. Parameterization of Heat Sink/Source Term

We still need to parameterize heat sink/source term (γ) to close equation (5). Cooling or heating of the atmospheric surface layer is tightly coupled to heat transfer with the earth surface. When a positive upward heat flux from the surface is injected into the atmosphere, unstable conditions are created with the consequence that the column of air warms above the surface. On the other hand, a negative upward heat flux from the surface, i.e. heat sink into the earth surface, will create stable conditions while cooling the column of air above the surface. Such heat fluxes can be estimated using sensible and latent heat components.

Parameterization of γ , which is composed of sensible and latent heat fluxes from the rural site, requires meteorological information measured at the rural area. For example, a weather file typically contains information about direct and diffusive

solar radiations, temperature at the height of 2 m and wind speed at the height of 10 m on an hourly basis. In this study, we assumed the following equations to formulate γ

$$\gamma = C_\gamma \left(\frac{Q_{net}}{\rho C_p} \right) \frac{1}{H_{bl}}, \quad (8)$$

$$Q_{net} = \underbrace{(Q_{Hveg} + h_{conv}(T_0 - T_{air}))}_{\text{sensible heat flux}} + Q_{rad} + \underbrace{(Q_{Lveg} + Q_{Lsoil})}_{\text{latent heat flux}}, \quad (9)$$

where Q_{net} is the net heat flux (positive upward from the surface into the atmosphere at the rural site), ρ is air density near the rural surface, C_p is air specific heat capacity, C_γ is a scaling factor for heat sink/source term (optimized to be 10) and H_{bl} is the diurnally-averaged boundary layer height taken to be 2000 m. In equation (9), Q_{Hveg} is the sensible heat flux from vegetation, h_{conv} is the convection heat transfer coefficient at the rural surface, T_0 is the rural surface temperature, T_{air} is the air temperature at the height of 2 m, Q_{rad} is the long wave and shortwave radiation absorbed by the rural surface, Q_{Lveg} is the latent heat flux from vegetation, and Q_{Lsoil} is the latent heat flux from soil. Bueno et al. [15] provided more details about calculation of these heat fluxes. Palyvos [14] suggested that the convection heat transfer coefficient is a linear function of wind speed

$$h_{conv} = 3.7\bar{S}_0 + 5.8, \quad (10)$$

where \bar{S}_0 is mean wind speed obtained from the rural weather station.

V. RESULTS AND DISCUSSION

From the above parameterization of mixing length and heat source/sink term we can now determine the diurnal variation of temperature and wind speed profiles in the rural area. Note that the model is only applied to rural climate up to the height of 5 times of average building height in the surrounding urban area, which is 80 m for the top of the domain for the vertical diffusion model. The boundary conditions for potential temperature are fixed value at the bottom of the domain, set to temperature at 2m elevation, and zero gradient on the top of the domain. Wind shear gradient is estimated from the logarithmic law at any elevation. However, this assumption is more accurate during thermally-neutral conditions.

As shown in Fig. 5, vertical profile of turbulent diffusivity exhibits negligible value very near the surface under all stability conditions. It is also known that the turbulent diffusivity very far from the surface close to the top of the planetary boundary layer is also negligible (not shown here). K_t increases with height at different rates depending on the stability condition. Under unstable conditions, atmospheric turbulence is mostly driven by thermal convection and shear production, as opposed to stable condition. Thus, as shown in Fig. 5, turbulent mixing at 1000 LST, 1400 LST and 1800 LST increases at higher rates.

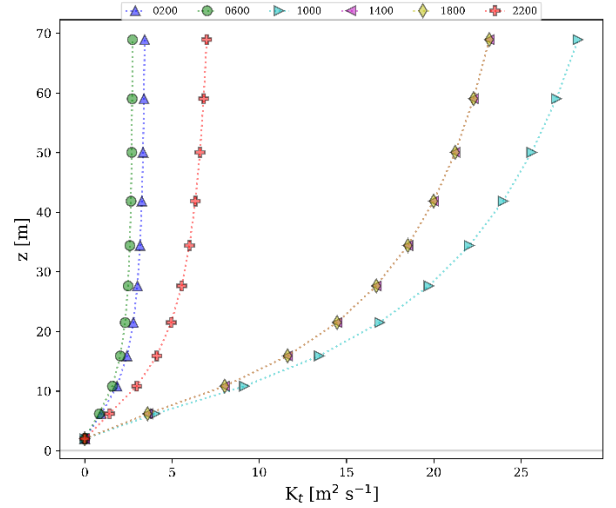


Figure 5. Diurnal variation of turbulent heat diffusivity profiles in the rural area; times in Local Solar Time (LST)

Fig. 6 shows the potential temperature profiles in the rural area over the course of a day. The atmosphere goes through a cycle starting from strongly stable conditions in the early morning (0200 LST) to unstable condition in the midafternoon (1400 LST). Under stable conditions, potential temperature increases with height and the atmospheric surface layer is weakly turbulent. As potential temperature profiles evolve with time and the sun rises, the rural surface begins warming up and the atmosphere becomes more turbulent. At 0200 LST the rural surface layer retains the heat from the previous day, explaining higher potential temperatures aloft, but the atmosphere is calm and stable near the surface where cooling occurs. As time evolves into the night and close to the morning, the surface layer potential temperature further drops at all altitudes. During the day, higher surface temperatures cause upward heat flux and the atmosphere becomes more turbulent. A clear transition between unstable to stable condition is evident at 1800 LST, where the potential temperature profile does not exhibit any significant change with height near the surface, representing the neutral condition.

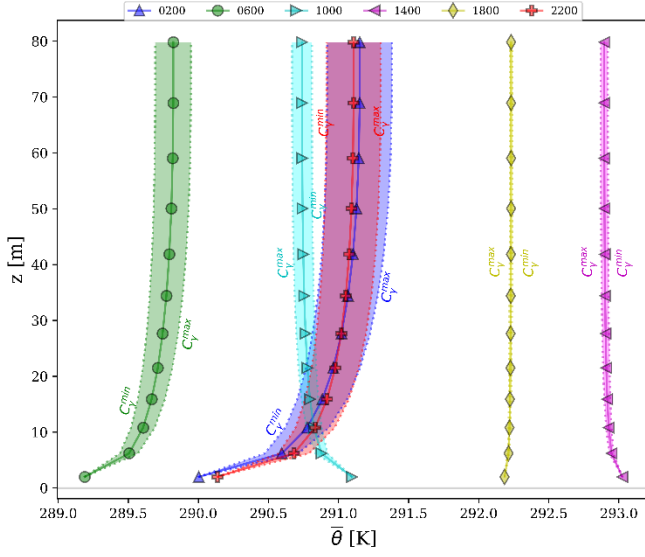


Figure 6. Diurnal variation of temperature profile in the rural area; times in Local Solar Time (LST)

An important parameter which can significantly affect parameterization of cooling/heating rate is C_γ . The effect of this coefficient on γ and the potential temperature profile is investigated by performing a sensitivity analysis. C_γ is varied in the range 8, 10, 12. As shown in Fig. 6, temperature profiles during stable atmospheric condition shift more under the influence of C_γ . In addition, diurnal variation of the heat sink/source term is provided in Fig. 7. This figure clearly shows that the rural heat flux varies between negative to positive values under a cycle of stable-neutral-unstable-neutral conditions. Sensitivity of the sink/source term to C_γ is also shown in Fig. 7.

Another sensitivity analysis is conducted to understand the effect of scaling correction factor C_{cr} for the mixing length scale on the temperature profile. For this purpose, C_{cr} is set to vary in the range 1.2, 1.5, 1.8 under stable condition and 1, 2, 3 under unstable condition. As shown in Fig. 8, this scaling factor has a large effect on the potential temperature profiles particularly during the stable condition.

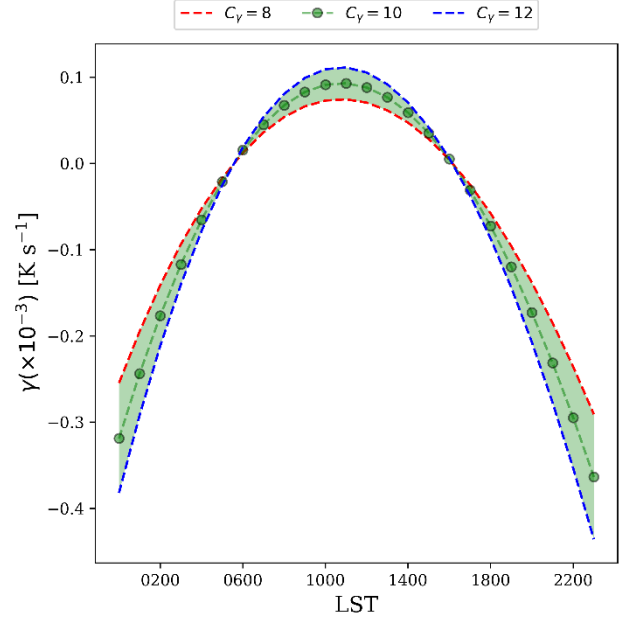


Figure 7. Diurnal variation of cooling or heating rate for different C_γ ; times in Local Solar Time (LST)

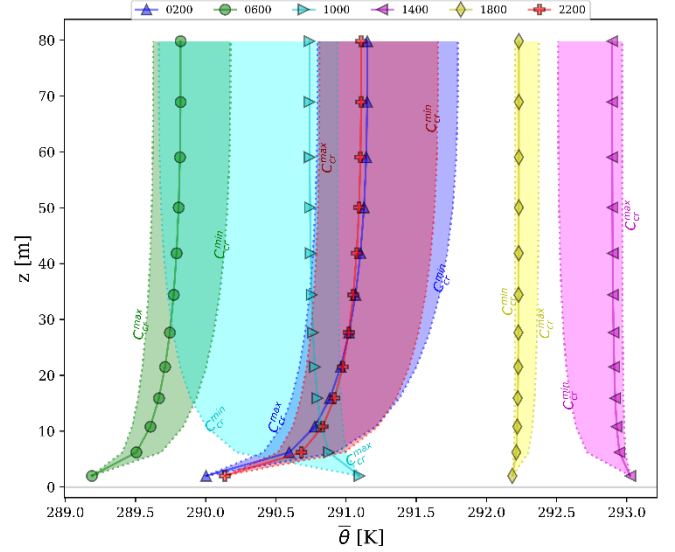


Figure 8. Effect of scaling factor on diurnal variation of temperature profile; times in Local Solar Time (LST)

Another assumption made in this rural model is to consider that the specific humidity is constant, i.e. invariant with height. This assumption is valid until vapour pressure is less than the saturation vapour pressure for a given altitude. This condition must be checked to validate the adequacy of this assumption. From the specific humidity equation, we have

$$q = \frac{\rho_v}{\rho_a} = \frac{\rho_v}{\rho_v + \rho_d}, \quad (11)$$

where ρ_v is the water vapour density, ρ_a is the air density and ρ_d is dry air density. This can be simplified more using the ideal gas law

$$q = 0.622 \frac{P_v(z,T)}{P_a(z,T)}, \quad (12)$$

where P_v and P_a are the water vapour pressure and the air pressure, respectively, as a function of height and real temperature (T) (not potential temperature). Using the Clausius-Clapeyron equation, we can determine the saturation pressure (P_{sat}) for water using

$$P_{sat}(T) = 0.61094 \exp\left(\frac{17.625T}{T+243.04}\right), \quad (13)$$

where P_{sat} is in kPa and T is in $^{\circ}\text{C}$. Note that the temperature in the diffusion equation is the potential temperature, so, a density profile is required to convert potential temperature to real temperature. Using density (ρ_0), real temperature (T_0) and pressure (P_0) at the surface level from the weather station at 2 m elevation, and considering a lapse rate of $-0.000133 \text{ kg m}^{-3} \text{ m}^{-1}$ for density within the surface layer, the density profile can be simplistically parameterized by

$$\rho = \rho_0 - 0.00133(z - z_0). \quad (14)$$

Now, we can calculate the vapour pressure and saturation vapour pressure at the top of the domain using equations (12) and (13), respectively. Fig. 7 shows the time series of P_{sat} and P_v on the domain top over the course of a day. Vapour pressure is almost always less than saturation pressure, particularly during the day. However, vapour pressure tends to saturation pressure under strongly stable condition from 0000 LST to 0600 LST. Assuming constant specific humidity up to the height of five times of average building height in the surroundings, i.e. 80 m, does not significantly violate the requirement for vapour pressure to remain below saturation vapour pressure on the top of the domain. So the assumption for constancy of specific humidity within the lowest range of altitudes in the surface layer is a practical assumption. Note, however, that this assumption does not apply to high altitudes where condensation may occur due to vapour pressure exceeding saturation vapour pressure.

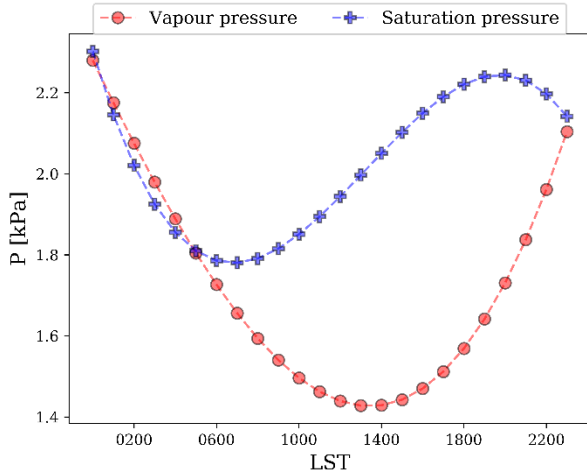


Figure 9. Diurnal variation of vapour pressure and saturation vapour pressure on the top of the vertical diffusion model domain; times in Local Solar Time (LST)

VI. CONCLUSION

A different approach was used to parameterize meteorological quantities, specifically vertical profiles of the potential temperature, in the lowest range of altitudes for the atmospheric surface layer of a typical rural area. As opposed to previous models suffering from limitations on Obukhov length, the model was developed independent of Obukhov length to formulate turbulent heat diffusivity. A vertical diffusion equation was used for temperature using K -theory with a new parameterization for the heat sink/source term, and subsequently the turbulent heat diffusivity. A relation between turbulent diffusivity and mixing length was established based on the availability of direct incoming solar radiation reaching the rural surface. In other words, a solar-basis correction factor for mixing length was introduced to consider effect of atmospheric stability, which is optimized as 2 under unstable condition (incoming solar radiation) and 1.5 under stable condition (lack of incoming solar radiation). Diurnal variation of turbulent heat diffusivity profile showed higher values during the day than over night, when the surface layer is less turbulent. Longwave and shortwave radiation, sensible and latent heat fluxes from soil and vegetation, and convection heat flux from the ground were considered as contributing factors in parameterization of heat sink/source term in the energy equation. As a result, vertical profile of the potential temperature clearly showed a cycle through stable, neutral, unstable, and neutral conditions. A logarithmic equation was used to determine the wind speed profile, which was more accurate during neutral and unstable conditions. The diurnal variation of the logarithmic wind profile was compared to measurements from a miniSoDAR in the rural area. An important assumption in this rural climate model was to consider specific humidity to be constant, i.e. invariant with height. This was assessed by comparing the vapour pressure and the saturation vapour pressure at the top of the vertical diffusion model domain (five times of building height in the nearby urban areas), and it was shown that almost always the vapour pressure is less than the saturation vapour pressure. However, vapour pressure approached saturation pressure over night. Future work shall involve integration of this rural climate model into an urban climate model to predict the vertically-resolved microclimate of urban areas. The vertical diffusion model, however, may have many other potentials for use such as wind farm assessment as well.

ACKNOWLEDGMENT

The study was supported by the Undergraduate Student Research Awards (USRA) and the Discovery Grant program from Natural Sciences and Engineering Research Council (NSERC) of Canada (Grant No. 401231). Special thanks go to Steve Nyman, Chris Duiker, Peter Purvis, Manuela Racki, and Jeffrey Defoe at the University of Guelph, who helped with the rural campaign logistics. The computational platforms were set up with the assistance of Jeff Madge, Joel Best, and Matthew Kent.

REFERENCES

- [1] A. A. Aliabadi, "Theory and applications of turbulence: a fundamental approach for scientists and engineers," Amir A. Aliabadi Publications, Guelph, 2018.

- [2] S. E. Gryning, E. Batchvarova, B. Brummer, H. Jorgensen, S. Larsen, "On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer," *Bound.-Layer Meteor.*, 2007, vol. 124, pp. 251-268.
- [3] A. J. Dyer, "A review of flux-profile relationships," *Bound.-Layer Meteor.*, 1974, vol. 7, pp. 363-372.
- [4] G. D. Hess, J. R. Garratt, "Evaluating models of the neutral, barotropic planetary boundary layer using integral measures: Part II. modelling observed conditions," *Bound.-Layer Meteor.*, 2002, vol. 104, pp. 359-369.
- [5] R. B. Stull, "An introduction to boundary layer meteorology," Kluwer Academic Publishers, Dordrecht, 2003.
- [6] A. A. Aliabadi, R. M. Staebler, J. de Grandpre, A. Zadra and P. A. Vaillancourt, "Comparison of Estimated Atmospheric Boundary Layer Mixing Height in the Arctic and Southern Great Plains under Statically Stable Conditions: Experimental and Numerical Aspects," *Atmos. Ocean*, 2016, vol. 54, pp. 60-74.
- [7] U. Hogstrom, "Non-dimensional wind and temperature profiles in the atmospheric surface layer: A re-evaluation," *Bound.-Layer Meteor.*, 1988, vol. 42, pp. 55-78.
- [8] S. E. Gryning, E. Batchvarova, B. Brummer, H. Jorgensen, S. Larsen, "On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer," *Bound.-Layer Meteor.*, 2007, vol. 124, pp. 251-268.
- [9] A. Pena, S. E. Gryning and J. Mann, "On the length scale of the wind profile," *Q. J. R. Meteorol. Soc.*, 2010, vol. 136, pp. 2119-2131
- [10] A. A. Aliabadi, M. Moradi, D. Clement, W. D. Lubitz, B. Gharabaghi, "Flow and temperature dynamics in an urban canyon under a comprehensive set of wind directions, wind speeds, and thermal stability conditions," *Environ. Fluid Mech.*, 2019, vol. 19 (1), pp. 81-109.
- [11] M. R. Raupach, R. A. Antonia, S. Rajagopalan, "Rough-wall turbulent boundary layers," *Appl. Mech. Rev.*, 1991, vol. 44, pp. 1-25.
- [12] J. D. Wilson, "An Alternative Eddy-Viscosity Model for the Horizontally Uniform Atmospheric Boundary Layer," *Bound.-Layer Meteor.*, 2012, vol. 145, pp. 165-184.
- [13] A. A. Aliabadi, R. M. Staebler, M. Liu and A. Herber, "Characterization and Parametrization of Reynolds Stress and Turbulent Heat Flux in the Stably-Stratified Lower Arctic Troposphere Using Aircraft Measurements," *Bound.-Layer Meteor.*, 2016, vol. 161, pp.99-126.
- [14] L. Prandtl, "Meteorological application of fluid mechanics," *Beitr. Phys. Atmos.*, 1932, pp. 188-202
- [15] B. Bueno, L. Norford, J. Hidalgo, G. Pigeon, "The urban weather generator," *J. Build. Perform. Simu.*, 2013, vol. 6, pp. 269-281.

An Airborne Thermal Imaging Methodology for Mapping Land Surface Temperature (LST) with a High Spatiotemporal Resolution

Ryan Byerlay¹, Mohammad Biglarbegan², Amir A. Aliabadi*

School of Engineering, University of Guelph, Guelph, Canada

rbyerlay@uoguelph.ca, mbiglarb@uoguelph.ca, [*aliabadi@uoguelph.ca](mailto:aliabadi@uoguelph.ca)

Abstract—A novel airborne thermal imaging technique is developed to map Land Surface Temperature (LST) with a high spatiotemporal resolution. A thermal camera is utilized as flown from a platform levitated by a tethered balloon. The developed method is suitable for near field observations with oblique view angles of the surrounding surface. In comparison to satellite observations, our approach results in less than 6% relative errors with a median relative error of 1.1% in predicting LST as compared to LST from images captured with the Moderate Resolution Imaging Spectroradiometer (MODIS).

Keywords—component; Georeferencing; Land Surface Temperature (LST); Airborne Thermal Imaging;

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) and other airborne imaging platforms have become commonly used tools for remote sensing [1]. A remote sensing parameter of significant interest for atmospheric modeling, meteorology, and climatology is the spatial distribution of Land Surface Temperature (LST) [2]. Sources of LST commonly include sensors located on satellites such as Landsat 8, which carries the Thermal Infrared Sensor (TIRS) and Terra¹ and Aqua² which carry the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor [3]. However, satellite LST data can be missing due to a variety of factors including cloud cover and sensor failure [4]. Furthermore, high spatial resolution LST data from satellites, such as Landsat 8³, is available at a low temporal resolution [5, 6]. Conversely, high temporal resolution LST data from satellite sensors, such as the Advanced Baseline Imager (ABI) on the Geostationary Operational Environmental Satellites R series of satellites (GOES-R)⁴, is available at a low spatial resolution [6]. That is, satellite-based sensors either offer high temporal – low spatial or low temporal – high spatial resolutions of LST. Recent advancements in both UAV and thermal imaging technologies have created opportunities for environmental LST to be accurately measured at a high spatiotemporal resolution. On-board UAV imaging systems, Global Positioning System (GPS) and Inertial Measurement Unit (IMU) data can be correlated to directly georeference image pixels to GPS coordinates without the use of Ground Control Points (GCPs) [7]. Coupling thermal imagery with common meteorological instruments to monitor environmental turbulent statistics on a UAV makes it possible to evaluate vertical heat fluxes among other meteorological data. Quantitative measurements of thermal imagery commonly utilize proprietary

software packages including PhotoScan Professional [8, 9], Pix4Dmapper [10] and MATLAB. Open source thermal image processing software such as Thermimage was developed for thermal image analysis⁵. However, open source thermal image processing coupled with integrated direct georeferencing of pixels is not widely distributed.

A Python based thermal image processing methodology was developed by the authors to calculate LST and to directly georeference images collected during a May 2018 field campaign for a remote northern mining location in Canada. The developed image processing methodology calculates georeferenced LST in decimal degrees of latitude and longitude. In this paper, we compare the results from the May 2018 campaign with respect to MODIS daytime LST at 1-kilometer horizontal spatial resolution. Percent relative error of LST was calculated between the method and the median of the MOD11A1⁶ data product, recorded on-board the Terra satellite, for each day from May 2018. The rest of the paper is organized as follows: Section II presents the data collection methodology and Python program development, Section III presents the LST results from the field campaign derived by the method and relative comparison to MODIS, Section IV discusses LST spatial patterns and relative errors of LST, and finally Section V concludes the paper.

II. MATERIALS AND METHODS

A. Field Campaign

During May 2018, the Tethered And Navigated Air Blimp 2 (TANAB2) completed multiple near surface profile measurements of the atmosphere. Thermal images were also collected using an uncooled FLIR Zenmuse XT 19-millimeter lens thermal camera. The camera was controlled by a DJI N3 Flight Controller and a DJI Lightbridge2 from either an Android or iOS device attached to the Lightbridge2. Wind velocities in the x, y and z directions, air pressure, and air temperature were also recorded from an onboard TriSonica Mini Ultrasonic Anemometer at 10Hz. All these components were attached to an aluminum structure and frame referred to as the gondola of the TANAB2.

One surface profile included the launching of the TANAB2 from grade level and a controlled release of line attaching the blimp to a fixed location at the surface of the Earth. Depending on environmental wind conditions, multiple mooring lines can be used and controlled by personnel. Utilizing three mooring lines, the TANAB2 was deployed in conditions with a maximum

¹ <https://terra.nasa.gov/about>

² <https://aqua.nasa.gov/>

³ <https://landsat.usgs.gov/what-are-band-designations-landsat-satellites>

⁴ <https://www.goes-r.gov/spacesegment/abi.html>

⁵ <https://github.com/gtatters/Thermimage>

⁶ https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod11a1_v006

wind speed of 10 meters per second. During ascent, images were collected by utilizing the Android or iOS device in conjunction with the Lightbridge2 to pan the thermal camera vertically and horizontally. Images were recorded up to a maximum of approximately 150 meters above grade level. Between May 5, 2018 and May 31, 2018, 11682 thermal images, each with a resolution of 640 pixels by 512 pixels were recorded.

B. Methodology

An image processing program utilizing Python, and associated open source image processing software, was developed on Ubuntu 16.04. Two additional open source software tools were used to extract information associated with the image metadata and camera signal values. These programs are ExifTool⁷ and ImageMagick respectively. Both programs were executed through the Linux terminal from the Python script and the outputs of each respective program were saved to variables in the Python program.

ExifTool was used to extract physical camera constants and image specific information. Data recorded by the integrated GPS and IMU in the DJI N3 flight controller was saved in the metadata of each image. ExifTool uses image tags to extract data. The tags vary by the manufacturer of the camera. Image specific information extracted from each picture included the latitude and longitude of the gondola when the image was recorded, the camera gimbal roll degree, the camera gimbal yaw degree, the camera gimbal pitch degree, the gondola roll degree, the gondola pitch degree and the altitude of the gondola. The raw signal recorded by individual pixels were extracted through using both ExifTool and ImageMagick. ImageMagick specified individual pixels to be extracted from the raw signal data while the ExifTool extracted the raw thermal pixel data.

All images have a gimbal pitch angle greater than -2° and less than or equal to -76° , where the horizontal plane is assumed to be zero degrees with positive angles upward. The gimbal pitch represents the angle for the center of an image. As per manufacturer specifications, the camera's mechanical pitch range is 45° and -135° . Images with a gimbal pixel angle greater than -30° may introduce error to the LST calculation⁸. Many images were recorded with oblique gimbal pitch angles between -2° and -30° . Excluding images within this range would have significantly reduced the spatial distribution of LST, such that spatial temperature gradients within the mining facility would have been omitted. Any top image pixels with a pitch angle greater than -1° at the top of each image (gimbal pitch angle plus half of the vertical field of view (VFOV) (26°)) were disregarded. This filtering was needed so that only pixels pointing to the land surface are included in the LST calculations. With this assumption, the pixel row for the top of each image was calculated through a mathematical relationship discussed below and visualized in Fig. 1 and Fig. 2. Images with a pitch angle of less than -76° were omitted for georeferencing simplicity as the pitch angle for the bottom of each image would be equivalent to less than -90° otherwise.

The altitude of the TANAB2 gondola (camera) was calculated through deriving a mathematical relationship between the atmospheric pressures recorded by the TriSonica relative to the atmospheric pressure at the start of each launch. The land

surface elevations in the eight cardinal directions (north, north-east, east, south-east, south, south-west, west and north-west) 10 kilometers away from each TANAB2 launch location were determined with the Geocontext-Profiler⁹ and exported to separate text files. In Python, a polynomial was fitted to the elevation data for each direction. A linear relationship representing the line of sight for the center of the camera was derived. The intersection(s) of these two equations was calculated and the smallest real positive solution was used as the surface level elevation and horizontal distance away from the TANAB2.

The geographic coordinates for the top, center, and bottom midpoints as well as corners of each image were calculated through a variation of the Haversine formula as per Equations 1 and 2.

$$Lat_2 = \arcsin[\sin(Lat_1) \cos\left(\frac{HDist}{R}\right) + \cos(Lat_1) \sin\left(\frac{HDist}{R}\right) \cos(Yaw)] \quad (1)$$

$$Lon_2 = Lon_1 + \text{atan}^2[\sin(Yaw) \sin\left(\frac{HDist}{R}\right) \cos(Lat_1), \cos\left(\frac{HDist}{R}\right) - \sin(Lat_1) \sin(Lat_2)] \quad (2)$$

Where Lat_2 and Lon_2 represent the calculated geographic coordinates, Lat_1 and Lon_1 represent the geographic coordinates recorded by the N3 when each image was captured (i.e. gondola's GPS coordinates), $HDist$ represents the surface level horizontal distance away from the TANAB2, R represents the equatorial radius of the Earth in kilometers¹⁰ and Yaw represents the heading of the camera gimbal in degrees from north positive clockwise. All angles and geographic coordinates were converted to radians before calculating the new geographic coordinates. When deriving the geographic coordinates for pixels on the edges of the images, the horizontal distance away from the TANAB2 was determined through considering half of the horizontal field of view (HFOV) (32°) with the known horizontal distance away for the top, center and bottom of each image, respectively. Using simple trigonometric relationships, the $HDist$ was determined. If geographic coordinates of pixels in the middle of the image were to be calculated, the $HDist_{Vertical}$ was required. Conversely, if geographic coordinates of pixels on either the left or right edges of each image were to be calculated, the $HDist_{Edge}$ was required. To calculate the geographic location of pixels within the image, a relation between geographic distance and image pixels was identified.

A mathematical relationship was derived to determine the new top pixel row and the geographic distance away from the TANAB2 with respect to image pixels. This relation is illustrated in Figs. 1 and 2. Angles based on the vertical and horizontal fields of view properties of the camera are used to georeference nine locations for each image recorded including the corners and midpoints for the top, center and bottom of each image, respectively.

⁷ <http://u88.n24.queensu.ca/exiftool/forum/index.php?topic=7116.0>

⁸ https://dl.djicdn.com/downloads/zenmuse_xt/en/sUAS_Radiometry_Technical_Note.pdf

⁹ <http://www.geocontext.org/publ/2010/04/profiler/en/>

¹⁰ <https://nssdc.gsfc.nasa.gov/planetary/factsheet/earthfact.html>

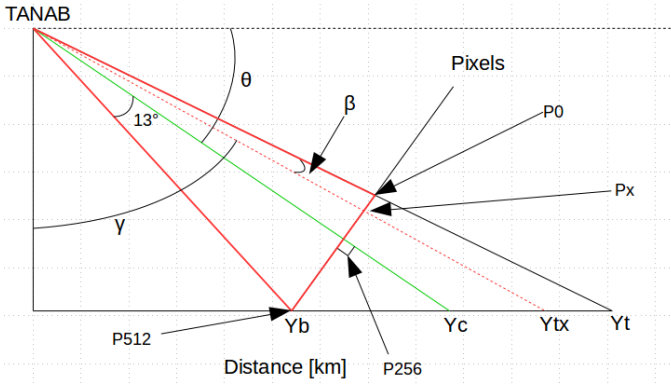


Figure 1. Relation between image pixels and horizontal distances.

Where Yb , Yc , Ytx and Yt represent the horizontal distance away from the TANAB2 for the bottom, center, new top and top of each image; $P0$, Px , $P256$ and $P512$ represent the top pixel row, the new top pixel row, the center pixel row, and the bottom pixel row for each image, 13° is half of the VFOV, θ represents the pitch angle, γ and β represent angles used in equations that follow to correlate pixels to distances. The red triangle in Fig. 1 is displayed in more detail in Fig. 2.

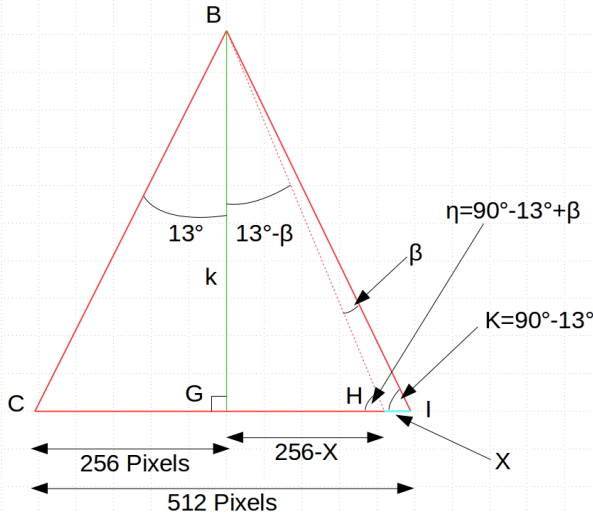


Figure 2. Relationship between VFOV and image pixels.

Where X represents the number of pixel rows omitted between the top of the image and another pixel row which itself is a function of the newly assigned top pitch angle. A geometric step function was used to determine the pixel steps for each row, where the majority of the pixel rows were located at the top of each image. In addition, every 64th pixel across each column was selected and a corresponding geographic coordinate pair was calculated for each pixel column and row pair. At most each image would yield a maximum of 80 coordinate pairs. Most images yielded less as many images were assigned a new top pixel row. This pixel row (X) was calculated as a function of angles γ , β , κ and η as represented in the equations below

$$\gamma_{Dist-Pix} = \arctan\left(\frac{HDist_{vertical}}{Altitude}\right), \quad (3)$$

$$\beta_{Dist-Pix} = 90^\circ - \text{abs}(\text{Gimbal Pitch}) + 0.5 \cdot VFOV - \gamma_{Dist-Pix}, \quad (4)$$

$$\kappa = 90^\circ - 0.5 \cdot VFOV, \quad (5)$$

$$\eta = 90^\circ - 0.5 \cdot VFOV + \beta_{Dist-Pix}, \quad (6)$$

$$X_{Dist-Pix} = \frac{[0.5 \cdot VPR \cdot \sin(\beta)]}{[\sin(VFOV \cdot 0.5) \cdot \sin(180^\circ - \eta)]}, \quad (7)$$

where VPR is the Vertical Pixel Range and is 512 as per the camera specifications. For coordinates on the left and right edges of each image, the $HDist_{Edge}$ from the TANAB2 was calculated using Eq. 8

$$HDist_{Edge, Dist-Pix} = \frac{HDist_{vertical}}{\cos(0.5 \cdot HFOV)} \quad (8)$$

Using the same relationship as illustrated in Figs. 1 and 2, a set of equations were derived to relate numerical pixel locations to geographic distances. The relations described below were used when iterating through the pixel coordinates of an image. Rearranging Eq. 9, $\beta_{Pix-Dist}$ was solved for as shown in Eq. 10

$$\frac{\sin(0.5 \cdot VFOV)}{\sin(\kappa) \cdot [0.5 \cdot VPR]} = \frac{\sin(0.5 \cdot VFOV - \beta_{Pix-Dist})}{\sin(\eta) \cdot [0.5 \cdot VPR - j]}, \quad (9)$$

$$\beta_{Pix-Dist} = -\arctan\left(\frac{[0.5 \cdot VPR - j] \cdot \sin(0.5 \cdot VFOV)}{0.5 \cdot VPR \cdot \sin(\kappa)}\right) + 0.5 \cdot VFOV, \quad (10)$$

$$\gamma_{Pix-Dist} = 90^\circ - \text{abs}(\text{Gimbal Pitch}) + 0.5 \cdot VFOV - \beta_{Pix-Dist}, \quad (11)$$

$$Slope_{Pix-Dist} = \frac{-1}{\tan(\gamma_{Pix-Dist})}, \quad (12)$$

where j represents the vertical pixel row and $Slope_{Pix-Dist}$ represents the slope for the line of sight from the camera for each corresponding pixel location. The horizontal distance is derived using the same method described before using the data from Geocontext-Profiler. If georeferencing pixels between the edge and the center of an image, an angular offset based on the HFOV must be used when calculating the horizontal distance. This angular offset is directly related to the location of the pixel column with respect to the center of the image. Equations 1 and 2 were used to calculate the geographic coordinates for each pixel in an image. The angular offset was added or subtracted to the Yaw value depending on the location of the pixel column.

Using the raw signal value extracted from each pixel with ExifTool and ImageMagick in conjunction with heat transfer formulas based on Planck's Law, the LST was calculated. Martiny et al. [11] developed a partially empirical relationship

between Planck's Law (as described by Çengel and Ghajar [12] in Eq. 13) and the infrared pixel value recorded by a thermal camera (Eq. 14)

$$E_{b\lambda} = \frac{C_1}{\lambda^5 \cdot [\exp(\frac{C_2}{\lambda T}) - 1]}, \quad (13)$$

$$I = \frac{R}{\exp(\frac{B}{T}) - 1}, \quad (14)$$

where $E_{b\lambda}$ represents the emissive radiative power for a spectral black body, λ represents the wavelength of radiation released by the black body, T represents the surface temperature of the black body and C_1 and C_2 represent constants [12]. In Eq. 14, I represents the thermal radiation emitted from the imaged surface and T represents the temperature of the imaged surface. The remaining parameters represent constants which are directly influenced by the thermal camera and are determined during the calibration process by the camera manufacturer [11].

The total radiative energy recorded by a thermal camera as described by Usamentiaga et al. [13] and FLIR Systems [14] is a function of three radiative energy sources as per Eq. 15

$$U_{tot} = \varepsilon \tau U_{obj} + (1 - \varepsilon) \tau U_{refl} + (1 - \tau) U_{atm}, \quad (15)$$

where U_{tot} represents the total radiative energy recorded by a thermal camera, ε represents the emissivity of the imaged surface, τ represents the transmissivity of the atmosphere between the surface and the camera, U_{obj} represents the fraction of radiative energy emitted from the imaged surface, U_{refl} represents the theoretical fraction of radiative energy that is reflected from the imaged object based on an assumed reflective temperature and U_{atm} represents the radiative energy theoretically emitted from the atmosphere based on an assumed atmospheric temperature [14]. Usamentiaga et al. [13] noted that the transmissivity of the atmosphere is usually close to one, as a result, the atmospheric transmissivity was assumed to be equivalent to one. From the camera manufacturer, the apparent reflective temperature used to calculate the U_{refl} was 295.15 K. The U_{refl} term was calculated using the following equation

$$U_{refl} = \frac{R_1}{R_2 \cdot [\exp(\frac{B}{T_{refl}}) - F]} - O, \quad (16)$$

where R_1 , R_2 , B , F and O are all constants determined by the camera manufacturer. These constants in addition to the apparent reflective temperature were extracted from the metadata of each image with ExifTool.

The emissivity of the land surface was derived from a remote sensing satellite. MODIS imaged the land surface twice daily over the entire field campaign. Using the MOD11B3¹¹ data product, recorded from the Terra satellite, the average land surface emissivity at a spatial resolution of 6 kilometers for three specific spectral ranges over the entire study area was calculated. Wang et al. [15] developed a relationship to determine the Broadband Emissivity (BBE) as a function of the three spectral

bands recorded by MODIS. The BBE relation (Eq. 17) was used to calculate the land surface temperature from the thermal images.

$$BBE = a\varepsilon_{29} + b\varepsilon_{31} + c\varepsilon_{32}, \quad (17)$$

where ε_{29} , ε_{31} and ε_{32} are the spectral emissivity bands from MODIS (Bands 29, 31 and 32, respectively) and a , b and c are constants determined to be acceptable for soil, vegetation and anthropogenic land surfaces [15].

The radiative energy signal from the imaged object recorded by the thermal camera was calculated using Eq. 18 and the corresponding object surface temperature was calculated with Eq. 19

$$U_{obj} = \frac{U_{tot} - [(1 - \varepsilon)U_{refl}]}{\varepsilon}, \quad (18)$$

$$T_{obj} = \frac{B}{\ln[\frac{R_1}{R_2(U_{obj} + O)} + F]}, \quad (19)$$

where ε is the BBE of a specific geographic location and T_{obj} is the LST of a specific geographic location.

III. RESULTS

The median LST values from the entire field campaign were calculated at a spatial resolution of 1-kilometer for six four-hour intervals in Local Daylight Time (e.g. 00:00-04:00 LDT, 04:00-08:00 LDT, 08:00-12:00 LDT, 12:00-16:00 LDT, 16:00-20:00 LDT, 20:00-24:00 LDT). Spatial distribution of LST with respect to important land surface features are included below. Geographically important land features include the perimeter of the facility (black), the perimeter of a tailings pond (blue), the perimeter of a mine (red), and the TANAB2 launch locations (white dots in Figs. 3 and 5). The temperature distribution of LST of both the tailings pond and mine for the corresponding four-hour intervals are represented as boxplots in Fig. 4.

A comparison of the median LST at 1-kilometer spatial resolution collected during May 24th was also completed with MODIS LST data from the MOD11A1 data product. The TANAB2 was deployed within the mine perimeter on this day, as denoted by the white dot within the mine perimeter in Fig. 5. The spatial distribution of percentage relative error between the two methods with respect to key land features is also presented.

IV. DISCUSSION

The spatial distribution and variation of LST varies diurnally as shown in Figs. 3 and 4. In general, the mine and areas around the mine are warmer as compared to other areas of the mining facility. The spatial distribution of LST within the mine and pond varies the least between 04:00 and 08:00 LDT. Between 00:00 and 04:00 there is a noted temperature gradient between the pond and the mine. This gradient is apparent for all other time periods except for the 04:00 to 08:00 LDT time period.

¹¹ https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod11b3_v006

The LST derived from images recorded from the May 24th, 2018 field campaign compared well with the MODIS LST as shown in Fig. 5. LST in and around the mine appear to have a percentage error of approximately 6% or less. Furthermore, the overall median relative error was determined to be 1.1%. Areas west and north-west of the mine have a higher percentage error which may be attributed to higher surface elevations as

compared to the mine itself and areas east of the mine. Temperatures at these pixels require very oblique view angles, i.e. pitch angles closer to the horizontal, therefore reduced accuracy due to solar reflection and a thick atmospheric boundary layer, through which the image quality may be disrupted.

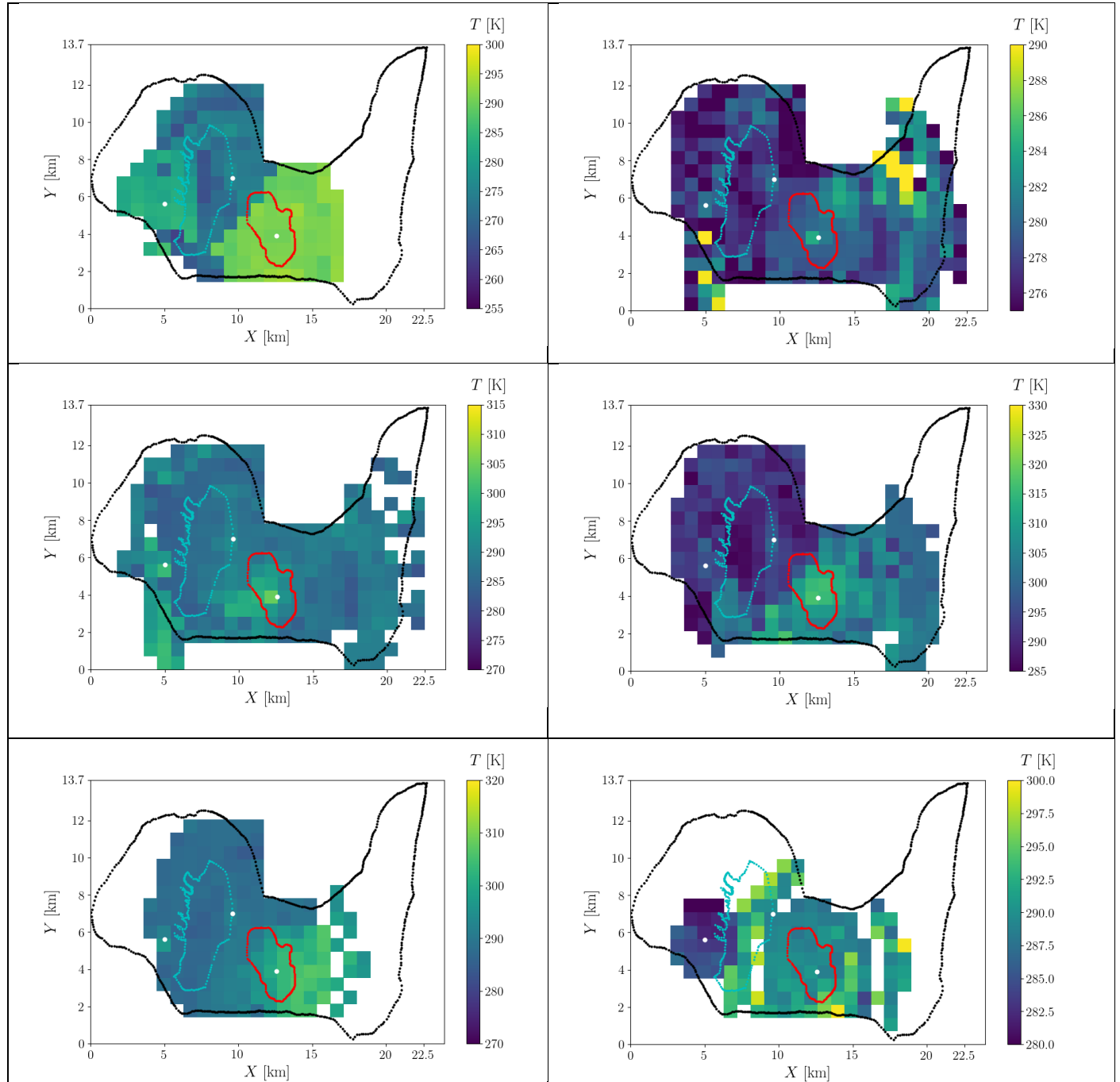


Figure 3. Median temperatures at 00:00 – 04:00 LDT (top left), 04:00–08:00 LDT (top right), 08:00–12:00 LDT (middle left), 12:00–16:00 LDT (middle right), 16:00–20:00 LDT (bottom left), 20:00–24:00 LDT (bottom right); at 1-km by 1-km horizontal spatial resolution.

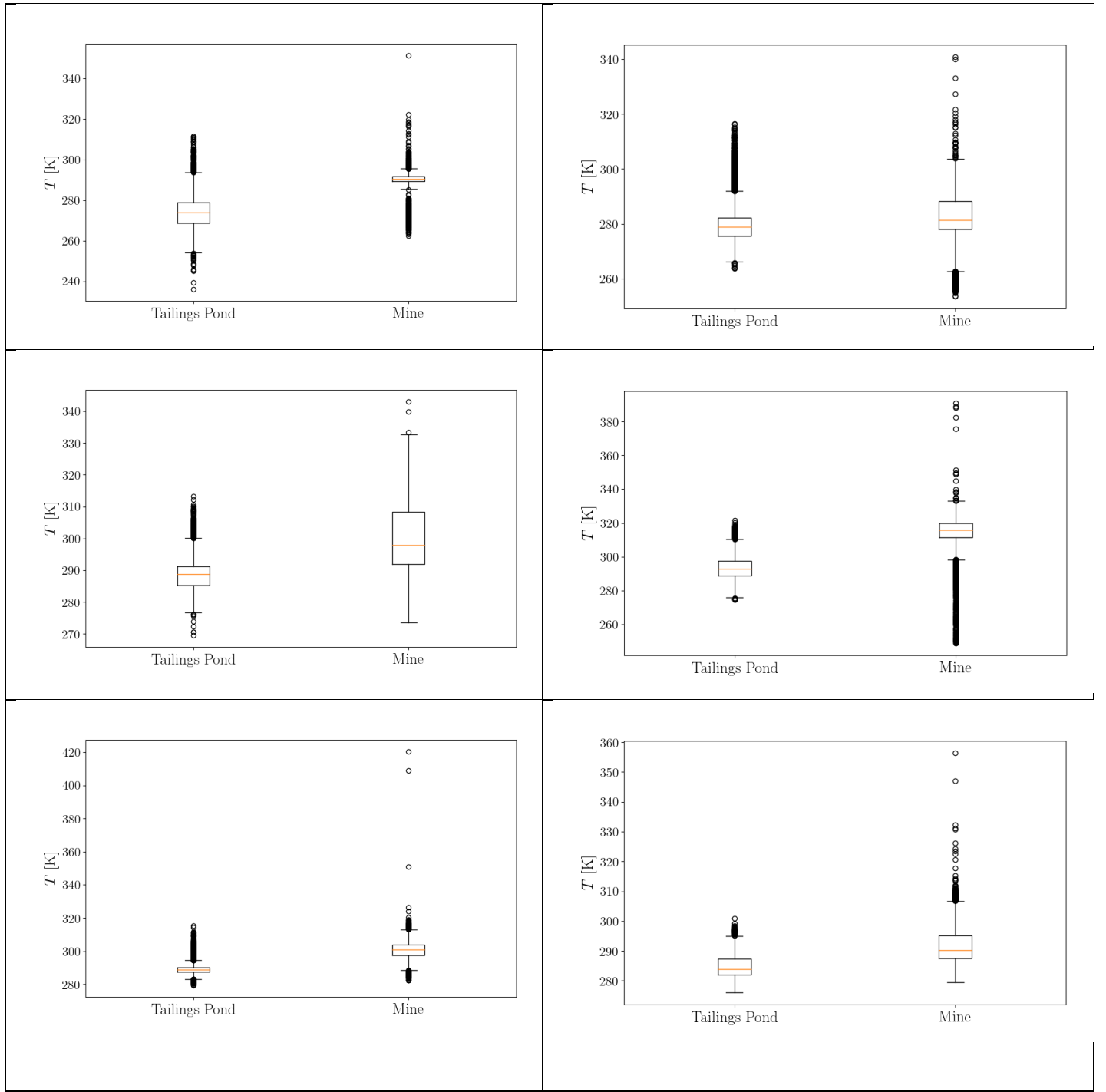


Figure 4. Boxplot of temperatures at 00:00-04:00 LDT (top left), 04:00-08:00 LDT (top right), 08:00-12:00 LDT (middle left), 12:00-16:00 LDT (middle right), 16:00-20:00 (bottom left), 20:00-24:00 LDT (bottom right); for tailings pond and mine.

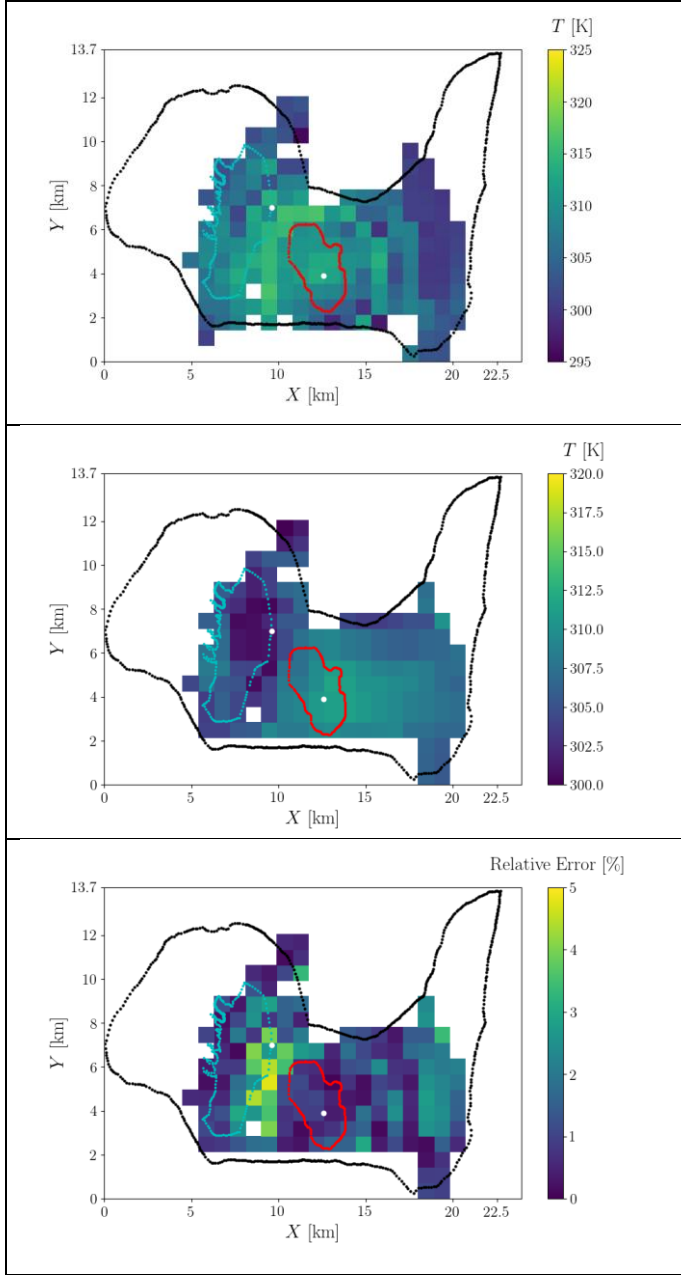


Figure 5. Median temperatures from May 24, 2018 12:00-14:00 LDT recorded from images captured by the thermal camera (top), daytime temperatures from the early afternoon on May 24, 2018 derived from the MODIS MOD11A1 data product (middle), percentage error between thermal camera and MODIS LST on May 24 (bottom).

In total, 98% of the total 11 682 images were processed with the method and LST from individual pixels were calculated. The number of images or pixels within an image processed with the method can be changed depending on the data set size and preferences of the user. The size of the image data set is much larger than any singular satellite image from Landsat 8 or GOES. However, with an image dataset, parameters within the methodology, can be adjusted to spatially represent LST with a very high spatial resolution over any specified time period.

Using satellite image datasets, these advantages cannot be realized due to inherent limitations of satellite sensors.

V. CONCLUSION

A novel airborne thermal imaging technique is developed to measure Land Surface Temperature (LST) with a high spatiotemporal resolution. A thermal camera is deployed on a tethered balloon in a remote mine field in northern Canada in May 2018. The terrain variability and camera's GPS position, altitude, pitch, and yaw angles were used to georeference each recorded pixel in the thermal image by tagging a latitude and longitude to it. The developed method was suitable for near field observations with moderately oblique view angles of the surrounding surface. In comparison to MODIS satellite observations, our method results in less than 6% relative errors in predicting LST close to the launch. This error increases with pitch angles closer to the horizontal. Our approach provides high spatial and temporal resolution of LST measurements simultaneously and therefore overcome limitations of satellites that cannot achieve both high spatial and temporal resolutions at the same time.

ACKNOWLEDGMENT

The airborne thermal imaging system was developed by the assistance of Denis Clement, Jason Dorssers, Katharine McNair, James Stock, Darian Vyriotes, Amanda Pinto, and Phillip Labarge. The authors thank the assistance of the following individuals in operating the thermal imaging system for the May 2018 campaign: Amir Nazem, Md. Rafsan Nahian, and Manoj Kizhakkeniyil. In-kind technical support for this work was provided by Rowan Williams Davies and Irwin Inc. (RWDI). This work was supported by the Discovery Grant program (401231) from the Natural Sciences and Engineering Research Council (NSERC) of Canada; Government of Ontario through the Ontario Centres of Excellence (OCE) under the Alberta-Ontario Innovation Program (AOIP) (053450); and Emission Reduction Alberta (ERA) (053498). OCE is a member of the Ontario Network of Entrepreneurs (ONE).

REFERENCES

- [1] Jose A. J. Berni, Pablo J. Zarco-Tejada, Lola Suárez, and Elias Fereres, "Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring From an Unmanned Aerial Vehicle," *IEEE Trans. Geosci. Remote Sens.*, 2009, 47(3):722-738.
- [2] Charlie J. Tomlinson, Lee Chapman, John E. Thornes, and Christopher Baker, "Remote sensing land surface temperature for meteorology and climatology: a review," *Meteorol. Appl.*, 2011, 18(3):296-306.
- [3] Estoque, Ronald C. and Murayama, Yuji, "Classification and change of built-up lands from Landsat-7 ETM+ and Landsat-8 OLI/TIRS imageries: A comparative assessment of various spectral indices," *Ecol. Indic.*, 2015, 56:205-217.
- [4] Hamid Reza Ghafarian Malamiri, Iman Roustas, Haraldur Olafsson, Hadi Zare, and Hao Zhang, "Gap-Filling of MODIS Time Series Land Surface Temperature (LST) Products Using Singular Spectrum Analysis (SSA)," *Atmosphere-Basel*, 2018, 9(9):334.
- [5] Panagiotis Sismanidis, Iphigenia Keramitsoglou, and Chris T. Kiranoudis, "Diurnal Analysis of Surface Urban Heat Island Using Spatially Enhanced Satellite derived LST Data," *Joint Urban Remote Sensing Event*, June 2015, Lausanne.

- [6] Klemen Zakšek and Krištof Oštir, "Downscaling land surface temperature for urban heat island diurnal cycle analysis," *Remote Sens. Environ.*, 2012, 117:114-124.
- [7] Darren Turner, Arko Lucieer, and Luke Wallace, "Direct Georeferencing of Ultrahigh-Resolution UAV Imagery," *IEEE Trans. Geosci. Remote Sens.*, 2014, 52(5):2738-2745
- [8] Yoann Malbêteau, Stephen Parkes, Bruno Aragon, Jorge Rosas, and Matthew F. McCabe, "Capturing the Diurnal Cycle of Land Surface Temperature Using an Unmanned Aerial Vehicle," *Remote Sens.*, 2018, 10(9):1407.
- [9] H. Hoffmann, H. Nieto, R. Jensen, R. Guzinski, P. Zarco-Tejada, and T. Friborg, "Estimating evaporation with thermal UAV data and two-source energy balance models," *Hydrol. Earth Syst. Sci.*, 2016, 20(2): 697-713.
- [10] Abdul Nishar, Steve Richards, Dan Breen, John Robertson, and Barbara Breen, "Thermal infrared imaging of geothermal environments and by an unmanned aerial vehicle (UAV): A case study of the Wairakei- Tauhara geothermal field, Taupo, New Zealand," *Renew. Energ.*, 2016, 86:1256-1264.
- [11] M. Martiny, R. Schiele, M. Gritsch, A. Schulz, and S. Wittig, "In Situ Calibration for Quantitative Infrared Thermography," *Quant. Infr. Therm. J.*, 1996.
- [12] Yunus A. Çengel and Afshin J. Ghajar, *Heat and Mass Transfer: Fundamentals and Applications*, 5th Edition. New York, NY: McGraw-Hill Education, 2015.
- [13] Rubén Usamentiaga, Pablo Venegas, Jon Guerediaga, Laura Vega, Julio Molleda, and Francisco G. Bulnes, "Infrared Thermography for Temperature Measurement and Non-Destructive Testing," *Sensors-Basel*, 2014, 14(7):12305-12348.
- [14] FLIR Systems, *Installation Manual: FLIR A3XX series FLIR A6XX series*, 2010.
- [15] Kaicun Wang, Zhengming Wan, Pucui Wang, Michael Sparrow, Jingmao Liu, Xiuji Zhou, and Shigenori Haginoya, "Estimation of surface long wave radiation and broadband emissivity using Moderate Resolution Spectroradiometer (MODIS) land surface temperature/emissivity products," *J. Geophys. Res.*, 2005, 110(D11).

AIRFLOWS IN NARROW STREET CANYONS: SINGLE OR DOUBLE VORTEX?

Lup Wai Chew
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA
lupwai@mit.edu

Amir A. Aliabadi
Environmental Engineering Program
University of Guelph
Guelph, ON, Canada
aaliabad@uoguelph.ca

Leslie K. Norford
Department of Architecture
Massachusetts Institute of Technology
Cambridge, MA, USA
lnorford@mit.edu

Abstract—Airflows in narrow urban street canyons (canyons with height-to-width aspect ratio of 2.0 or higher) are generally understood to induce two counter-rotating vortices. We show that the double-vortex regime only exists at low Reynolds numbers, Re ($\sim 10^4$). At high Re ($\sim 10^5$ or higher), only one vortex is observed, consistent with full-scale field measurement at $Re \sim 10^6$. The change from double-vortex regime at relatively low Re to the single-vortex regime at high Re suggests that the widely adopted critical Re (where $Re \geq 10,000$ is sufficiently high to ensure Re -independent flow) is not applicable for narrow street canyons.

Keywords - *CFD simulations; high aspect ratio canyon; Reynolds number independence; skimming flow*

I. INTRODUCTION

Airflows across urban areas have been studied extensively, as they have direct impacts on many aspects of built environments. For example, architectural features such as roof shape and building porosity can channel more winds into urban areas to enhance pollutant dispersion and improve the air quality [1–4]. Britter and Hanna [5] categorized the study of urban airflows into four scales: regional (~ 100 km), city (~ 10 km), neighborhood (~ 1 km), and street (~ 0.1 km). This paper focuses on the street scale, where we can resolve the flow features in individual urban street canyons. Urban street canyons (“canyons” hereafter) are outdoor spaces between buildings. When the street length is much larger than the building height, a two-dimensional (2D) canyon is formed. Under perpendicular winds, 2D canyons with a height-to-width aspect ratio (H/W) larger than 0.7 exhibit the skimming flow pattern [6], which is the most severe in terms of pollutant or heat trapping inside canyons [5,7]. Therefore, many studies have investigated the skimming flow regime in 2D canyons.

TABLE I. SUMMARY OF STUDIES ARRANGED BY CANYON ASPECT RATIO (H/W) AND REYNOLDS NUMBER (Re).

No.	Authors	H (m)	W (m)	H/W	Re	No. of vortices
1	Caton et al. [8]	0.07	0.07	1.0	3,000	1
2	Gerdes & Olivari [9]	0.03	0.03	1.0	4,000	1
3	Baik et al. [10]	0.24	0.24	1.0	10,000	1
4	Li et al. [11]	0.1	0.1	1.0	12,000	1
5	Kovar-Panskus et al. [12]	0.29	0.29	1.0	19,000	1
6	Meroney et al. [13]	0.06	0.06	1.0	20,000	1
7	Brown et al. [14]	0.15	0.15	1.0	30,000	1
8	Nakamura & Oke [15]	17.0	16.0	1.06	2.1×10^6	1
9	Baik et al. [10]	0.24	0.12	2.0	10,000	2
10	Li et al. [11]	0.1	0.05	2.0	12,000	2
11	Eliasson et al. [16]	15.0	7.1	2.1	1.9×10^6	1
12	Santamouris et al. [17]	20.0	8.5	2.35	2.5×10^6	1

Table I lists selected experiments and field studies with H/W between 1.0 and 2.4. Note that the height of canyons could be as small as 0.03 m (in reduced-scale experiments) and as large as 37 m (in full-scale field measurements). The Reynolds number, $Re = HU_{ref}/\nu$, is a dimensionless parameter to compare the scale, where U_{ref} is a reference wind speed (often taken to be the freestream wind speed), H is building height, and ν is the kinematic viscosity. Reduced-scale

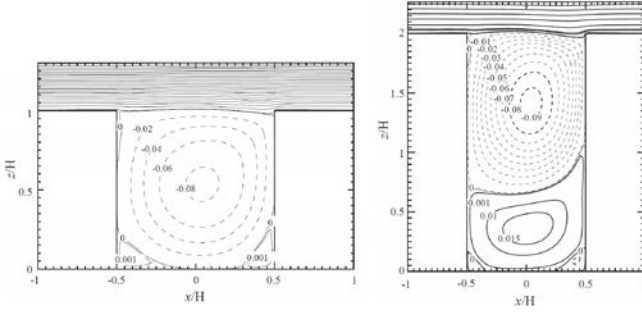


Fig. 1. Contours of stream-functions at Reynolds number 12,000. Left: flow across a canyon with $H/W = 1.0$, where a single vortex is observed (excluding the minor corner vortices). Right: flow across a canyon with $H/W = 2.0$, where two counter-rotating vortices are observed. Images from Li et al. [18].

experiments have Re on the order of 10^3 to 10^4 , while full-scale field measurements have Re two orders of magnitude higher at 10^6 . At $Re \sim 10^4$, a single quasi-steady vortex is induced in a canyon with $H/W = 1.0$. In a narrow canyon, where $H/W = 2.0$, two vortices could be induced, as shown in Fig. 1 [18].

The Re required for Re -independent flows is often taken to be 10,000 [10,11,13,19]. This means that the flow pattern does not change with increasing Re when Re exceeds 10,000. Studies 1-8 in Table I show that increasing the Re from 3,000 to 30,000, and further to 10^6 , does not change the number of vortices in canyons with $H/W = 1.0$ (all studies reported one vortex). This means that the effect of Re is negligible, or the flows are Re -independent. However, this is not true for canyons with $H/W = 2.0$. Studies 9 and 10 in Table I reported two vortices at $Re \sim 10,000$, but studies 11 and 12 reported only one vortex at $Re \sim 10^6$. Clearly, the flows are not Re -independent in this range of Re , since the flow pattern changes when the Re is increased above 10,000. Our previous experimental work has shown evidence that Re -independence is achieved at $Re \sim 70,000$ for canyons with $H/W = 2.0$ [20]. However, only the velocity profiles were measured and no flow visualization is performed. This paper extends the study numerically by conducting computational fluid dynamics (CFD) simulations to visualize the overall flow fields in the canyons at different Re .

II. NUMERICAL MODEL

A. Description of Simulation Setup

Fig. 2 shows the 2D CFD domain. The dimension of the buildings and canyon follow those in our water channel experiments [20], which are used for model validation. The canyon height is $H = 0.2$ m while the canyon width is $0.5H$. Both the upstream and downstream buildings have a width of $0.6H$. Since a periodic boundary is employed on the left and right boundaries, the building width is halved to $0.3H$ in the CFD model. The top of the surface is $2.75H$ from the ground to match the water depth in the experiments. The top surface is a free-slip wall. The roof, walls, and ground have a no-slip boundary condition. The left and right boundaries have a periodic boundary condition, meaning that the domain repeats itself in the streamwise direction, forming an array with an infinite number of canyons. The initial conditions are zero for

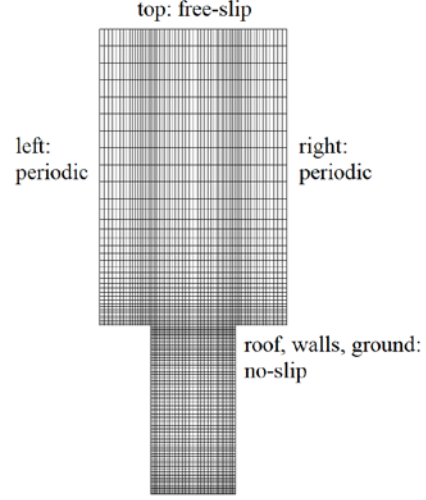


Fig. 2. Mesh and boundary conditions of the simulation domain

all parameters, except for the streamwise velocity, where a freestream velocity is prescribed based on the Re in each simulation.

The model is built and meshed in the ANSYS workbench package version R17.2. All grids are perfectly orthogonal (hexahedral). The grids near the walls are refined. The maximum grid expansion ratio is 1.2. The simulations are run with ANSYS FLUENT. A steady Reynolds-Averaged Navier-Stokes (RANS) solver with the standard k-epsilon turbulence closure scheme is used. We repeat the simulations with realizable and RNG k-epsilon schemes and obtain identical results. The enhanced wall treatment is selected based on the recommendation for wall-resolving flows [21]. The SIMPLEC (Semi-Implicit Method for Pressure-Linked Equations-Consistent) algorithm is used for pressure-velocity coupling. For discretization, the Least Squares Cell Based method is used for gradients and the second order scheme is used for all other parameters. The convergence criterion is set at 10^{-5} for all variables, which is achieved after 10,000 iterations. To test for convergence, the simulation is continued for another 10,000 iterations. The results do not change, confirming that a converged solution has been obtained.

B. Mesh Independence Study

Three models with coarse, normal, and fine mesh resolutions are built for mesh sensitivity study. The numbers of grids are 4,358, 17,572, and 39,432, respectively. For brevity, only the coarse mesh model is shown in Fig. 2. Fig. 3 compares the normalized streamwise velocity profiles at the middle of the canyon. The vertical distance from the ground, z , is normalized by the canyon height, H , while the streamwise velocity, u , is normalized by the freestream velocity, U_{ref} . The Re is 105,000, and the initial and boundary conditions are identical for all three models. The profiles from all three models have negligible difference above $z/H = 0.2$. Near the ground level (z/H between 0 and 0.2), the coarse mesh model underpredicts the u/U_{ref} magnitude. The normal mesh model produces a profile identical to that of the fine mesh model, confirming that the normal mesh resolution is sufficiently fine. However, to

achieve a dimensionless wall distance (y^+) on the order of 1, the fine mesh model is used for all subsequent simulations.

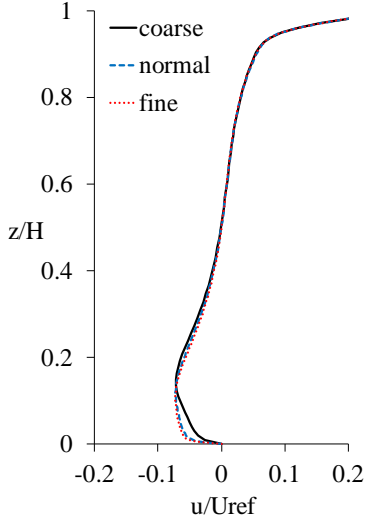


Fig. 3. Comparison of the normalized streamwise velocity profiles from the three models with different mesh resolutions.

C. Model Validation

The water channel experiments in [20] are used for CFD model validation. The detailed experimental setup and methodology is available in [22]. The Re range is between 12,000 to 105,000 in the experiments. We conduct CFD simulations for each Re and compare the simulation results to the experiments. Fig. 4 shows that the simulated u/U_{ref} profiles agree well with the experiments. At the lowest Re (12,000), the CFD model correctly predicts the low velocity region at the bottom half of the canyon ($z/H < 0.5$). At the highest Re (105,000), the CFD model also correctly predicts the negative velocity near the ground level. The fractional bias, $FB = 2 (\overline{u_s} - \overline{u_e}) / (\overline{u_s} + \overline{u_e})$, and the normalized mean-square error, $NMSE = \overline{(u_s - u_e)^2} / (\overline{u_s} \times \overline{u_e})$ are used to quantify the errors [23]. The over-bar represents average (the spatial average of the line profile in this study), u is the velocity, subscript s represents simulation, and subscript e represents experiment. The FB for each case (Re of 12,000, 28,000, 57,000, 87,000, and 105,000) is -0.033, 0.017, 0.058, 0.050, and 0.055, respectively. The NMSE is 0.016, 0.019, 0.025, 0.025, and 0.031, respectively. Since the FB and NMSE for all cases are small, the CFD models are considered validated.

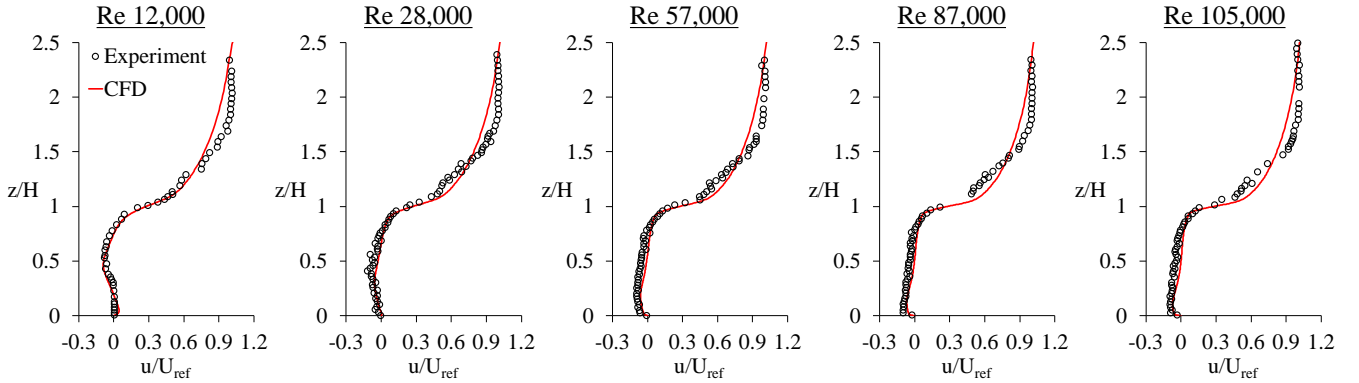


Fig. 4. Comparison of simulated streamwise velocity profiles to experiments at different Reynolds number. All profiles taken at the middle of the canyon.

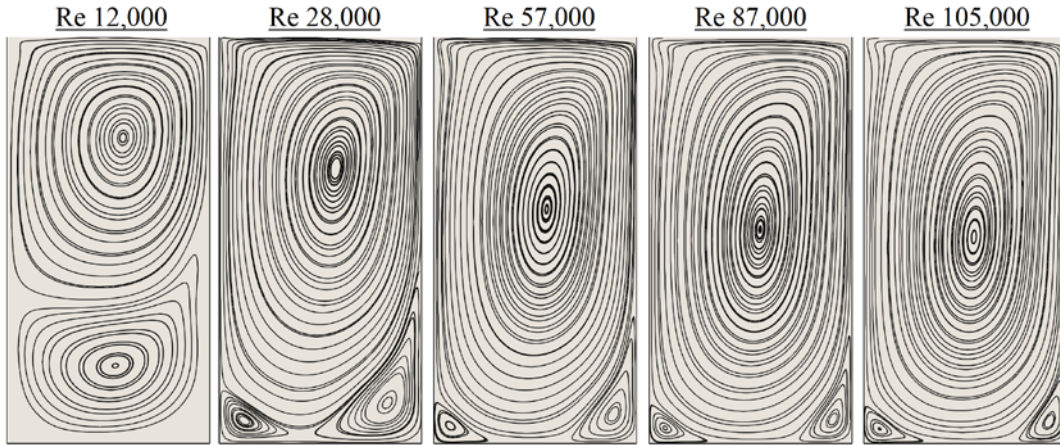


Fig. 5. Streamlines at different Reynolds number. When the Reynolds number is increased, the top vortex grows and the bottom vortex shrinks.

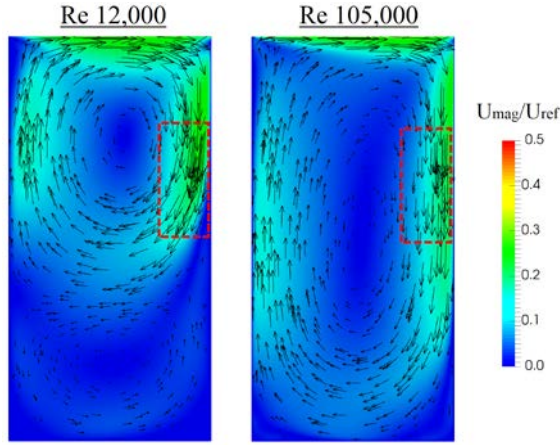


Fig. 6. Contours and vector plots of normalized velocity magnitude. At Re 12,000, the downward flow along the windward wall loses its vertical momentum near the mid-canyon height, as indicated by the red box. In contrast, at Re 105,000, the downward flow continues to flow along the windward wall.

III. RESULTS AND DISCUSSION

The experiments provide only the line velocity profiles but not the full flow fields. The (validated) CFD models complement this drawback by providing visualization of the whole flow fields in the canyons. Fig. 5 compares the streamlines at different Re . At Re of 12,000, the double-vortex flow regime is obtained, similar to the results in the literature [10,11]. When the Re is increased to 28,000, the top vortex grows and occupies a larger region in the canyon, while the bottom vortex shrinks. The top vortex grows even larger when the Re is further increased to 57,000, while the bottom vortex breaks into two corner vortices. Increasing the Re from 57,000 to 87,000 and further to 105,000 does not result in significant changes in the flow pattern.

Why does Re affect the number of vortices in a canyon? Referring to Fig. 6, the shear layer at the building height induces a downward velocity near the windward wall of the building (the wall facing the incoming wind). This downward velocity indicated in the red box is about $0.3U_{ref}$. In other words, the local Re is about 30% of the global Re , i.e., the local Re is 3,600 and 31,500 for the cases with Re 12,000 and Re 105,000, respectively. For the case with a higher local Re , this downward velocity penetrates deeper into the canyon and creates only one major vortex inside the canyon. With higher Re , the entire regime (energy-containing, inertial subrange, and dissipation range) in the turbulence kinetic energy cascade shifts. It is speculated that the double-vortex pattern, which is caused by viscous dissipation at relatively low Re , does not form at high Re . In fact, at high Re , the inertial subrange, which is driven only by inertia-dominated dissipation, is extended to cover scales of motion close to the canyon length scale. In this range, viscous dissipation is negligible compared to the inertia-dominated dissipation of turbulence kinetic energy [24]. This explains why only one vortex forms at relatively high Re .

To summarize, Fig. 5 shows that the commonly adopted critical $Re = 10,000$ for Re -independent flows does not hold in canyons with aspect ratio 2.0. Many reduced-scale studies have reported the double-vortex flow pattern in narrow street canyons and generalized the results to full-scale canyons. Consequently, applications derived based on the double-vortex flow regime (e.g., pollutant dispersion in narrow canyons [25,26]) may not be applicable at full-scale canyons. Therefore, we should revisit the applicability of previous studies on narrow canyons and be cautious when generalizing the results from reduced-scale experiments to full-scale built environments.

IV. CONCLUSIONS

The critical Reynolds number (Re) for Re -independent flows in urban street canyon is often taken to be 10,000. Using validated CFD models, we show that this critical Re is not applicable for narrow canyons with height-to-width aspect ratio of 2.0. In addition, narrow canyons are often understood to exhibit the double-vortex flow regime. We show that this regime only exists at relatively low Re ($< 28,000$). At sufficiently high Re ($> 57,000$), only one major vortex is observed, consistent with full-scale field measurement [19]. This finding changes the common understanding of flows in narrow canyons, where multiple vortices are expected.

REFERENCES

- [1] C. Yuan, E. Ng, L.K. Norford, "Improving air quality in high-density cities by understanding the relationship between air pollutant dispersion and urban morphologies," *Build. Environ.* 71 (2014), pp. 245–258.
- [2] L.W. Chew, L.K. Norford, "Pedestrian-level wind speed enhancement in urban street canyons with void decks," *Build. Environ.* 146 (2018), pp. 64–76.
- [3] P. Kastner-Klein, R. Berkowicz, R. Britter, "The influence of street architecture on flow and dispersion in street canyons," *Meteorol. Atmospheric Phys.* 87 (2004), pp. 121–131.
- [4] A.A. Aliabadi, E.S. Krayenhoff, N. Nazarian, L.W. Chew, P.R. Armstrong, A. Afshari, L.K. Norford, "Effects of Roof-Edge Roughness on Air Temperature and Pollutant Concentration in Urban Canyons," *Bound.-Layer Meteorol.* 164 (2017), pp. 1–31.
- [5] R. Britter, S. Hanna, "Flow and dispersion in urban areas," *Annu. Rev. Fluid Mech.* 35 (2003), pp. 469–496.
- [6] T. Oke, *Boundary layer climates*, 2nd edition, Methuen, London and New York, 1987, pp. 267.
- [7] X.-X. Li, C.-H. Liu, D.Y. Leung, K. Lam, "Recent progress in CFD modelling of wind field and pollutant transport in street canyons," *Atmos. Environ.* 40 (2006), pp. 5640–5658.
- [8] F. Caton, R.E. Britter, S. Dalziel, "Dispersion mechanisms in a street canyon," *Atmos. Environ.* 37 (2003), pp. 693–702.
- [9] F. Gerdes, D. Olivari, "Analysis of pollutant dispersion in an urban street canyon," *J. Wind Eng. Ind. Aerodyn.* 82 (1999), pp. 105–124.
- [10] J.-J. Baik, R.-S. Park, H.-Y. Chun, J.-J. Kim, "A laboratory model of urban street-canyon flows," *J. Appl. Meteorol.* 39 (2000), pp. 1592–1600.
- [11] X.-X. Li, D.Y. Leung, C.-H. Liu, K. Lam, "Physical modeling of flow field inside urban street canyons," *J. Appl. Meteorol. Climatol.* 47 (2008), pp. 2058–2067.
- [12] A. Kovar-Panskus, L. Moulinneuf, E. Savory, A. Abdelqari, J.-F. Sini, J.-M. Rosant, A. Robins, N. Toy, "A wind tunnel investigation of the influence of solar-induced wall-heating on the flow regime within a simulated urban street canyon," *Water Air Soil Pollut. Focus.* 2 (2002), pp. 555–571.
- [13] R.N. Meroney, M. Pavageau, S. Rafailidis, M. Schatzmann, "Study of line source characteristics for 2-D physical modelling of pollutant

- dispersion in street canyons,” *J. Wind Eng. Ind. Aerodyn.* 62 (1996), pp. 37–56.
- [14] M.J. Brown, R.E. Lawson, D.S. DeCroix, R.L. Lee, “Mean flow and turbulence measurements around a 2-D array of buildings in a wind tunnel,” in: 11th Jt. AMSAWMA Conf. Appl. Air Pollut. Meteorol. Long Beach CA, 2000.
- [15] Y. Nakamura, T.R. Oke, “Wind, temperature and stability conditions in an east-west oriented urban canyon,” *Atmospheric Environ.* 1967. 22 (1988), pp. 2691–2700.
- [16] I. Eliasson, B. Offerle, C. Grimmond, S. Lindqvist, “Wind fields and turbulence statistics in an urban street canyon,” *Atmos. Environ.* 40 (2006), pp. 1–16.
- [17] M. Santamouris, N. Papanikolaou, I. Koronakis, I. Livada, D. Asimakopoulos, “Thermal and air flow characteristics in a deep pedestrian canyon under hot weather conditions,” *Atmos. Environ.* 33 (1999), pp. 4503–4521.
- [18] X.-X. Li, C.-H. Liu, D.Y. Leung, “Development of a k - ϵ model for the determination of air exchange rates for street canyons,” *Atmos. Environ.* 39 (2005), pp. 7285–7296.
- [19] W.H. Snyder, *Guideline for fluid modeling of atmospheric diffusion*, Environmental Protection Agency, Research Triangle Park, NC (USA), 1981.
- [20] L.W. Chew, A.A. Aliabadi, L.K. Norford, “Flows across high aspect ratio street canyons: Reynolds number independence revisited,” *Environ. Fluid Mech.* 18 (2018), pp. 1275–1291.
- [21] ANSYS Inc, *ANSYS FLUENT User’s Guide Release 14.0*, (2011).
- [22] L.W. Chew, N. Nazarian, L. Norford, “Pedestrian-level urban wind flow enhancement with wind catchers,” *Atmosphere*. 8 (2017), pp. 159.
- [23] S. Hanna, J. Chang, “Acceptance criteria for urban dispersion model evaluation,” *Meteorol. Atmospheric Phys.* 116 (2012), pp. 133–146.
- [24] A.A. Aliabadi, *Theory and applications of turbulence: A fundamental approach for scientists and engineers*, Amir A. Aliabadi Publications, Guelph, Ontario, Canada, 2018, pp. 168.
- [25] J. Zhong, X.-M. Cai, W.J. Bloss, “Large eddy simulation of reactive pollutants in a deep urban street canyon: Coupling dynamics with O₃-NO_x-VOC chemistry,” *Environ. Pollut.* 224 (2017), pp. 171–184.
- [26] X.-X. Li, C.-H. Liu, D.Y. Leung, “Large-eddy simulation of flow and pollutant dispersion in high-aspect-ratio urban street canyons with wall model,” *Bound.-Layer Meteorol.* 129 (2008), pp. 249–268.

Fluid Mechanics

MECHANICS OF ACTIVE FLUIDS: FLUID MECHANICS REFASHIONED BY MICROORGANISMS

Hassan Peerhossaini

Professor & Western Research Chair in Urban Resilience and Sustainability

University of Western Ontario
London, Canada

and

AstroParticles and Cosmology Laboratory
University of Paris
Paris, France

The mechanics behind *active fluids* is truly impressive. Particularly fascinating is the recorded manifestations of the fluid mechanics employed and refashioned by microorganisms for survival, growth, and habitat colonization.

Active fluid refers to the fluids of self-propelled particles such as bacteria or algae. *Active fluid* properties fundamentally differ from passive fluids and exhibit remarkable physical phenomena over a wide range of scales, from microscopic diffusion and motility of microbial suspensions to the large-scale colonization of aqueous spaces. To proliferate, bacteria usually form aggregates as a precursor to biofilms. This is an important virulence mechanism on solid surfaces. For example, it provides a barrier against drugs. At the water-air interface, blooming is also detrimental from an ecological and economical point-of-view since it causes oxygen-depleted areas where aquatic life cannot develop.

After a short introduction we explore the dynamics of bacterial colonization by studying the time and space variation of bacterial diffusion in a thin film of fluid close to the solid surfaces of different topographies, elasticity, and geometry.

Acknowledgements: This lecture is partially constructed based on some experimental and theoretical results of my research team. I especially acknowledge the results of H. Fadlallah, M. Jarrahi, J. Leopoldes and T.Vourc'h.

MULTILAYER VISCOPLASTIC FLUID FLOWS IN AXIALLY ROTATING PIPES

Seyed Mohammad Taghavi
Département de Génie Chimique
Université Laval
Quebec, Canada
Seyed-Mohammad.Taghavi@gch.ulaval.ca

Shan Lyu
Département de Génie Chimique
Université Laval
Quebec, Canada
shan.lyu.1@ulaval.ca

Oil and gas well primary cementing operations involve pumping a sequence of fluids into the well, for example, cement along a circular pipe (casing) to remove (displace) in situ drilling mud. As the fluids used have different densities, multilayer flows are expected to occur. A casing rotation is generally believed to improve the displacement process, but without solid laboratory experiments to prove that such rotation is indeed effective. In fact, rotation is not widely used to enhance the displacement that occurs within the casing, although this would be feasible as there are various types of casing heads and special adaptors in the industry to enable a casing axial rotation.

We perform laboratory fluid mechanics experiments to help develop a crucial understanding about enhancing the displacement process using a pipe rotation. These experiments conducted in a long, inclined pipe reveal in detail how the dynamics of multilayer flows are significantly affected by a pipe rotation, resulting in improving the displacement process in certain conditions. Increasing the pipe rotation velocity enhances the transverse mixing between the fluids involved and, above a critical transition, eventually leads to a removal of the displaced fluid. The dynamics of the flow can be described by the dimensionless groups that govern it, such as the Rossby number, the Froude number, etc.

Experimental Study on the Effects of Surface Hydrophobicity on Drag Reduction in a Taylor Couette Flow

Ahmed Mohamed, Xili Duan, Yuri Muzychka

Department of Mechanical Engineering

Memorial University of Newfoundland

St. John's, NL, Canada

afamohamed@mun.ca

ABSTRACT

In recent years, superhydrophobic coatings have been applied to achieve non-wetting surfaces for various thermal fluidic applications. Air packets trapped in micro/nano-meter scales near the coated solid surface provide a lubrication to the fluid flow which leads to a slip boundary condition along the wall. The literature shows that flow slipperiness helps in reducing drag, but there are inconsistencies in the magnitude of observed drag reduction (DR) between different studies, and its dependence on the surface characteristics and Reynolds number in turbulent flow remains unclear. In this study, a MCR301 rheometer is used to investigate drag reduction in a Taylor Couette flow. Coatings are applied on an aluminum surface to achieve three different wettabilities, with water static contact angles of 171° , 150° , and 116° respectively. Disposal concentric measuring cups are coated with a dipping-and-spanning technique. The contact angles of water on the coated surfaces are measured using an OCA 15EC contact angle instrument. Constant and variable shear rates are studied to observe the coating degradation effects for all samples. Measurement results show that the DR is more than 10%. This study paves the way for further study of the drag reduction techniques in pipe flow.

Relationship between POD Modes and Physical Mechanisms in Tornado-Like Vortex

Mohammad Karami

The Wind Engineering, Energy and Environment
(WindEEE) Research Institute, Western University,
London, Ontario, Canada N6A 5B9
line 4: mkarami3@uwo.ca

Horia Hangan

The Wind Engineering, Energy and Environment
(WindEEE) Research Institute, Western University,
London, Ontario, Canada N6A 5B9

Luigi Carassale

Department of Mechanical Engineering, University of
Genova, Italy

Hassan Peerhossaini

The Wind Engineering, Energy and Environment
(WindEEE) Research Institute, Western University,
London, Ontario, Canada N6A 5B9

ABSTRACT

The use of proper orthogonal decomposition (POD) as a feature-recognition tool to identify typical or recurrent pressure patterns, possibly associated to some specific physical phenomenon, has been experimented by several authors. Sometimes the results of the analyses appeared encouraging, but the general opinion that POD modes is unable to identify physically-meaningful coherent structures was derived by several authors. A major reason of this failure is believed to derive from the orthogonality constraint of the modal shapes that is intrinsically present in POD and has no physical justification.

Here, we investigated the relationship between the physics of the vortex flow and POD modes. For this purpose, we simulated a bivariate Gaussian synthetic vortex whose center is moving in a circumference, as shown in Figure 1. Then, we applied POD algorithm on fluctuating field of the synthetic vortex.

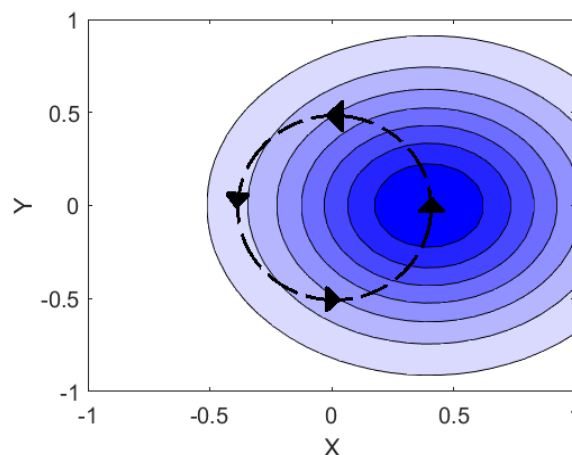


Figure 1. Bi-variate Gaussian synthetic vortex, moving around a circumference.

DEVELOPING MORE ACCURATE MODELS OF TORNADOS THROUGH COMPUTATIONAL ANALYSIS OF THE TORNADO WIND FIELDS PRODUCED BY SIMULATED SUPERCELL EVENTS

Eric Savory, Niall Bannigan

Department of Materials and Mechanical Engineering
University of Western Ontario
London, Canada

Leigh Orf

Space Science and Engineering Centre
University of Wisconsin–Madison
Madison, United States of America

ABSTRACT

Tornados are a major hazard and can form at any time; to understand them and the dangers that they pose is an important matter. Currently, the Rankine vortex model is used as the basis for theoretical tornado analysis and various small scale wind facilities are used for experimental analyses, see [1, 2], but these both assume axial symmetry, impose symmetric boundary conditions on the entire domain, and do not incorporate the changes in intensity of wind and shape of the vortex like an actual tornado. To more accurately model tornados, simulated supercells created on the Blue Waters supercomputer in the USA at large scales were analysed. After the extraction of the wind fields at all time steps, the data were circumferentially averaged and statistically analysed with computational methods for peak tangential wind speeds, anomalies in the vortex characteristics, and movement of the tornado centre. These methods provide more accurate models of tornados, which may be used to predict the formation of tornados, how much damage that can occur for a given tornado size, and over what area is the tornado considered dangerous.

REFERENCES

- [1] H. L. Kuo, "Axisymmetric Flows in the Boundary Layer of a Maintained Vortex," *Journal of Atmospheric Sciences*, vol. 28, pp. 20-41, 1970.
- [2] M. Refan and H. Hangan, "Characterization of Tornado-Like Flow Fields in a New Model Scale Wind Testing Chamber," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 151, pp. 107-121, 2015.

Numerical study on gas–solids circulating fluidized bed downer

Zeneng Sun

Dept. Chemical and Biochemical Engineering
University of Western Ontario
London, Canada
Zsun82@uwo.ca

Chao Zhang

Dept. Mechanical and Materials Engineering
University of Western Ontario
London, Canada
czhang@eng.uwo.ca

ABSTRACT

Gas-solids flow structures inside a circulating downward fluidized bed reactor (CFB downer) are numerically studied via CFD approach. An Eulerian-Eulerian two-fluid model coupling with k - ϵ turbulence model for both phases and kinetic theory of granular flow for the solids phase is setup for simulations. CFB downer under various operating conditions from low-density to high-density has been simulated and the results agree well with experimental data. Axial and radial profiles of solids concentration show that a more homogeneous flow structure is achieved in the downer reactor comparing with the circulating upward fluidized bed reactor (riser). Velocity profiles also show that a more uniform velocity distribution takes place of both the gas and solids phases inside a downer. A much thinner wall region is found in the downer reactor instead of the denser annular region in the CFB riser. More details of the velocity profiles at different heights along the downer indicate that the particles go through a short accelerating stage and reach a uniform motion stage quickly with the help of gravity after entering into the downer. The scale-up effects of the CFB downer reactor is also studied by CFD approach between two downers with inner diameters of 0.0762m and 0.2m respectively. CFD results show that the wall effects are more obvious in the wider downer.

MECHANICS OF ACTIVE FLUIDS AT A SOLID LIQUID INTERFACE: DIFFUSION OF SYNECHOCYSTIS SP.PCC 6803 IN DILUTE REGIME

Zahra Habibi

Department of Civil and Environmental Engineering
 The University of Western Ontario
 Mechanics of Active Fluids and Bacterial Physics Lab
 London, Canada

Thomas Vourc'h

Institut Curie Physico-Chimie Curie
 Paris, France

Julien Léopoldès

Institut Langevin, Université Paris-Est
 Marne-La-Vallée
 Paris, France

Malihe Mehdizadeh Allaf

Department of Civil and Environmental
 Engineering
 The University of Western Ontario
 Mechanics of Active Fluids and Bacterial
 Physics Lab
 London, Canada

Hassan Peerhossaini

Department of Civil and Environmental
 Engineering
 The University of Western Ontario
 Mechanics of Active Fluids and Bacterial
 Physics Lab
 London, Canada and Laboratoire
 AstroParticules et Cosmologie Université
 Paris Diderot
 Paris, France

ABSTRACT

Active fluid refers to a fluid whose constituent elements are self-propelled; they transform chemical energy of nutrients to mechanical energy to assure their propulsion. Examples include suspensions of microorganisms like bacteria, algae, and sperm cells. They display properties that differ fundamentally from passive fluids. Contrary to conventional fluid flows in which one needs gradients of pressure, velocity and temperature to break equilibrium and drive the flow, in active fluids, cells as the microstructural elements of the fluid convert the chemical energy of nutrients into mechanical energy for driving the flow. Therefore, by using their molecular motors to activate their flagella (or pili), microorganisms in active fluids can develop complex spontaneous fluid motions in the absence of external gradients.

From a biological and a physical point-of-view, it is important to investigate active fluids such as hydrodynamics of microorganisms due to their direct and indirect effects on human lives, the environment, and also its applications in industrial, medical, food and energy fields. In this paper, we focus on the diffusion phenomenon in active fluids, which is directly related to the motility of the active fluid particles. Motility of the model cyanobacterium *Synechocystis* sp.PCC 6803 at a solid liquid interface in the dilute regime has been investigated. Cyanobacteria — also called “blue-green algae” — are a phylum of bacteria

that obtain their energy through photosynthesis transformation of CO_2 and are the only oxygenic phototroph prokaryotes. To proliferate and survive against environmental stresses, they attach irreversibly and diffuse on solid surfaces by forming biofilms (i.e. masses of adherent cells embedded in slimy extracellular matrices). Their motion in biofilm is an intermittent “run” and “tumble” motility composed of alternating immobile “tumble” and mobile “run” periods. We have recently shown that the diffusion of *Synechocystis* sp. PCC 6803 tends to slow down due to the recognition of traces of secreted extracellular matrix left by the bacteria

on a solid surface. Thereafter, the bacteria continue to diffuse slowly and frequent interactions between them start to occur due to their high surface density. We have also experimentally measured the rate of aggregates formation, and observed that it depends on the cellular motility as provided by surfaces of different hardness. In this work, MATLAB software is used for particle tracking analysis and investigation of the correlation between the motilities of neighboring bacteria. It is of special interest to learn whether close bacteria follow almost similar paths.



Figure 1. *Synechocystis*

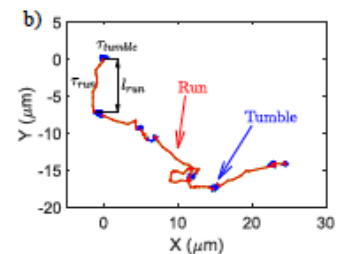


Figure 2. Example of a trajectory of *Synechocystis* (T. Vourc'h et al.)

* T. Vourc'h et al. PHYSICAL REVIEW E - Volume: 97 Issue: 3 Article Number: 032407

Influence of trailing edge flap pitching on dynamic stall in pitching airfoils

Farid Samara

Department Mechanical and Mechatronics Engineering
 University of Waterloo
 Waterloo, ON, Canada
 fsamara@uwaterloo.ca

David A. Johnson

Department Mechanical and Mechatronics Engineering
 University of Waterloo
 Waterloo, ON, Canada
 david.johnson@uwaterloo.ca

As wind turbines increase in size, predicting loading on the blades becomes more challenging and obstructs the development of larger and more efficient wind turbines. It is thus important to develop active control strategies to decrease the loading and mitigate the fluctuations experienced by the blades in normal environmental conditions. Wind turbine load reduction techniques could increase the power produced by the turbine, reduce the capital and maintenance cost while prolonging the lifespan of not just the blades but also the drive-train and tower¹. A trailing edge flap (TEF), a small movable control surface, has the potential to be an effective method to control blade loads as they can change the forces generated at a specific blade location¹.

In this experimental work the influence of a trailing edge flap (TEF) on dynamic stall effects is investigated on a pitching S833 airfoil at a Reynolds number of 1.8×10^5 and a reduced frequency (k) of 0.06 and 0.1 ($k = \pi * c * f / U_\infty$ where c is the airfoil chord, f is the pitching frequency and U_∞ is the freestream velocity). The TEF covers 20% of a chord of 178 mm. A 3D printed blade section was designed and built to house the actuation and sensing on the airfoil in real time. Coefficient of lift and moment were obtained from 54 surface pressure taps. The blade section was designed to be installed on a 3.4 m wind turbine but initial testing was done inside a 0.61 m square wind tunnel. More details about the experimental setup could be found in Samara and Johnson².

The static lift and moment coefficients will be presented for a range of pitch and flap angles well beyond what is already published. The lift and moment coefficients are presented for the dynamic cases for a mean pitch angle of 0° , 10° , and 20° to represent stall onset, dynamic stall and deep stall cases. The flap oscillating motion follows the same pitch motion but with different phase lags (ϕ). Phase lag is defined by the lag between the pitch and flap sinusoidal motion.

Figure 1 shows the coefficient of lift (C_L) and moment (C_M) for a reduced frequency of 0.06 and a mean angle of attack of 20° . In both plots the same trend could be seen across the different flap phases but magnitude and location of peaks change. The plots show the hysteresis cycle for an airfoil under dynamic stall. The steep spike when $\alpha = 19^\circ$ is due to the formation and convection of a leading edge vortex (LEV)³. The second peak that occurs at $\alpha = 20^\circ$ is due to a second LEV. After the LEV the airfoil is completely stalled and reattachment occurs at $\alpha = 8^\circ$.

This sharp increase in lift and moment forces could be detrimental to wind turbine operation. It was found, as one example result, that if the aim is to reduce the load fluctuations around the mean angle then it would be best to set the flap phase to $\pi/2$. On the other hand, if the aim is to reduce load fluctuations at the min and max then it would be best to set the flap phase to 0π .

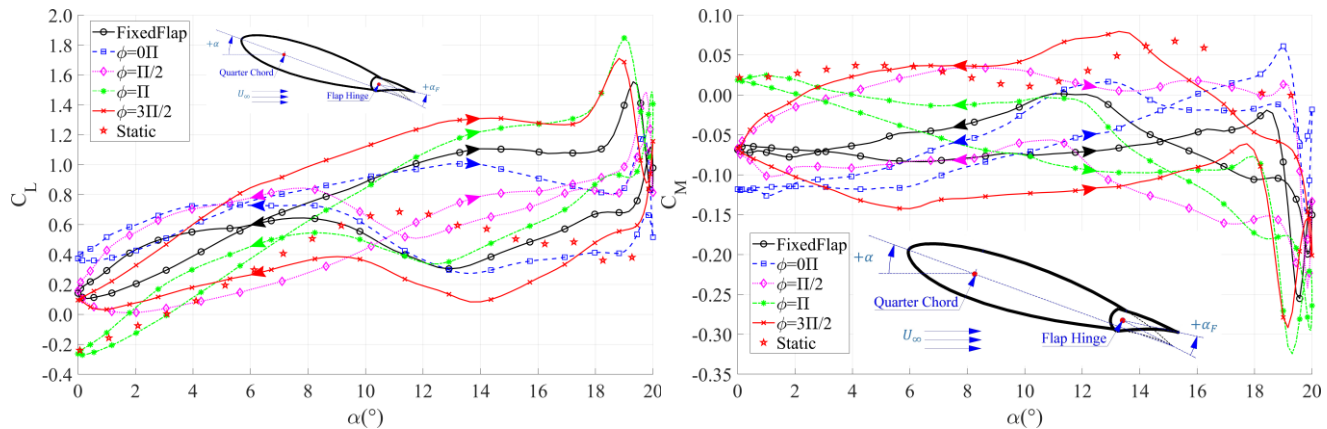


Figure 1: a) Coefficient of lift (C_L) and b) coefficient of moment (C_M) versus α for different flap phases (ϕ)

¹ Barlas and van Kuik, Progress in Aerospace Sciences, 46(1), (2010)

² Samara and Johnson, Journal of Physics: Conference Series, 1037, 052023, (2018)

³ McCroskey, Annual Review of Fluid Mechanics, 14:285-311, 285-311 (1982)

Bed Voidage Predictions for Inverse Liquid-solid Fluidized Beds

Saleh Srabet

Department of Mechanical and Materials Engineering
Western University
London, Ontario, Canada
ssrabet@uwo.ca

Dr. Jesse Zhu

Department of Chemical and Biochemical Engineering
Western University
London, Ontario, Canada
jzhu@uwo.ca

Dr. Dominic Pjontek

Department of Chemical and Biochemical Engineering
Western University
London, Ontario, Canada
dpjontek@uwo.ca

ABSTRACT

Inverse liquid-solid fluidization refers to a two-phase system where dispersed light solid particles, whose density is less than the continuous liquid density, are suspended by a downward fluid flow. Inverse liquid-solid fluidization systems present some advantages in practical applications when compared to conventional upward liquid-solid fluidized beds. For example, light solid particles are characterized by intensive random movement, improving the liquid-solid contact efficiency and enhancing the system's mass transfer [1]. Furthermore, biofilm thicknesses in biological processes can be efficiently controlled with this configuration [2]. The increasing demand of liquid-solid reactors for various applications, particularly wastewater treatment (i.e., particle-supported biofilm), has created considerable interest in improving the understanding of hydrodynamic characteristics for inverse liquid-solid fluidization bed configurations.

The steady-state bed voidage, which is related to the bed expansion, constitutes a significant parameter for the design, operation and scale-up of fluidized bed bioreactors. Many past studies have proposed empirical bed voidage correlations by modifying the predicted terminal particle velocities and/or the bed voidage index in the Richardson-Zaki equation [3] based on experimental measurements in small-scale systems (i.e., bed column ID less than 80 mm). This work experimentally measures the bed voidage of an inverse liquid-solid fluidized bed in a large scale system with downer column reactor of 200 mm internal diameter and a height of 4.5 m. Furthermore, this work includes solid particles whose densities are less than 300 kg/m³, which have unique characteristics as their behavior does not obey Newton's free settling law due to the low particle inertia and spiral motion trajectory during free-raising [4].

A correlation is proposed based on an analysis of forces (i.e., effective buoyant and drag force) which act on a single suspended particle. The correlation is developed in terms of the drag coefficient as well as the Archimedes and Reynolds numbers to predict the bed voidage of an inverse liquid-solid fluidized bed with using wide extensive solid particle densities. This correlation provides a good agreement with experimental data of this study and is compared to existing correlations from past studies.

References

- [1] Nikov, I., and Karamanev, D., 1991. "Liquid-solid mass transfer in inverse fluidized bed." *AIChE Journal*. 37, 5 (1991): 781–784.
- [2] Arun, N., Razack, A., and Sivasubram 2013. "Recent Progress in Hydrodynamics of Inverse Fluidized Bed Reactors: A Review." *Chemical Engineering Communications*. 200, 9, (2013): 1260–1277.
- [3] Richardson, J. E., and Zaki, W. N., "Sedimentation and Fluidization: Part I." *Transactions of the Institution of Chemical Engineers*. 32 (8), (1954): 35-53.
- [4] Karamanev, D., and Nikolov, L., "Free rising spheres do not obey newton's law for free settling." *AIChE Journal* 38.11 (1992): 1843-1846.

Minimum Fluidization Velocity of Inverse Liquid-solid Fluidized Bed With Using Light Solid Particles

Saleh Srabet

Department of Mechanical and Materials Engineering
Western University
London, Ontario, Canada
ssrabet@uwo.ca

Dr. Jesse Zhu

Department of Chemical and Biochemical Engineering
Western University
London, Ontario, Canada
jzhu@uwo.ca

Dr. Dominic Pjontek

Department of Chemical and Biochemical Engineering
Western University
London, Ontario, Canada
dpjontek@uwo.ca

ABSTRACT

Inverse liquid-solid fluidized beds have recently received more attention, particularly for wastewater treatment bioreactors (i.e., particle-supported biofilm) due to their advantages compared to conventional upward liquid-solid fluidized beds. The solid particles in an inverse configuration can be fluidized at low liquid velocities, reducing solid particle attrition and efficiently controlling biofilm thickness [1]. The flow behavior of free-rising light particles in inverse liquid-solid fluidization was initially thought to be similar to free-falling heavy particle in upward liquid-solid fluidization. Nevertheless, Karamanev and Nikolov [2] experimentally determined that light particles, with densities lower than 300 kg/m^3 , do not obey Newton's law related to free settling since these particles have a spiral trajectory during free-raising. These particles are particularly of interest because their drag coefficients deviate from the standard drag curve. The minimum fluidization velocity is an important hydrodynamic parameter for the design, operation and scale-up of a fluidized bed bioreactor. Fluidization occurs when the frictional pressure drop across the fixed bed equals the net buoyant force per unit area. The Ergun equation [3] is a well-known equation for presenting pressure drop through fixed bed which is applicable for a wide range of Reynolds number and for non-spherical particles. Wen and Yu [4] proposed a minimum fluidization velocity correlation based on a modified Ergun equation, in terms of the fluid-solid Archimedes number and two constant factors. Previous minimum fluidization velocity studies in inverse liquid-solid fluidized beds have focused on particle with densities approaching to the liquid density (typically water) and have concluded that their experimental data matched with Wen and Yu correlation [4].

In this study, the minimum fluidization velocity for three kinds of light and free-rising particles (28 , 122 and 300 kg/m^3) was experimentally investigated in a large-scale system (i.e., reactor column ID 200 mm and a height of 4.5 m) by two measurement methods: the frictional pressure drop and bed height methods. The experimental minimum fluidization velocities had large deviations when compared to the predictions from the Wen and Yu correlation. It is believed that the drag characteristics of the light particles, which has been previously demonstrated to be affected by the low particle inertia and spiral motion flow trajectory flow, are the cause for the observed deviations. As such, a modification to the Wen and Yu correlation is proposed to improve the predictions for the current experimental data as well as relevant experimental data from a prior study.

References

- [1] Arun, N., Razack, A., and Sivasubram 2013. "Recent Progress in Hydrodynamics of Inverse Fluidized Bed Reactors: A Review." *Chemical Engineering Communications*. 200, 9, (2013): 1260–1277.
- [2] Karamanev, D., and Nikolov, L., "Free rising spheres do not obey newton's law for free settling." *AIChE Journal* 38.11 (1992): 1843-1846.
- [3] Elgun, S., "Fluid Flow Through Packed Columns" *Chem. Eng. Prog.*, 48, 89 (1952).
- [4] Wen, C. Y and Yu, Y. H., "A Generalized Method for Predicting the Minimum Fluidization Velocities," *AIChE Journal*, 12, 610 (1966b).

CFD and CAA Analysis of the SD 7037 Airfoil for Low Reynolds Number Flow

Alison Zilstra

Department of Mechanical and Mechatronics Engineering
University of Waterloo
Ontario, Canada
alison.zilstra@uwaterloo.ca

David A. Johnson

Department of Mechanical and Mechatronics Engineering
University of Waterloo
Ontario, Canada
david.johnson@uwaterloo.ca

The prediction of aeroacoustic noise using computational methods is an important area of research that has seen large growth recently due to the improvements in computational software as well as computing power. There is significant interest in the use of these fully analytical methods to aid in the design of airfoils and blade shapes to reduce the noise of propellers, helicopters and wind turbines. The airfoil used for these simulations, SD 7037, is specifically designed for use on small scale wind turbines that operate at low Reynolds numbers. The Wind Energy Group at the University of Waterloo has conducted experimental PIV analysis and acoustic analysis of the SD 7037 airfoil at low Reynolds numbers for static and oscillating angles of attack (AOAs) [1, 2]. This large set of experimental data allows for the validation of both the flow parameters and acoustic measurements of the simulation and can shed light on the capabilities and inefficiencies of the prediction model.

This research uses a combination of computational fluid dynamics (CFD) and computational aeroacoustics (CAA) methods to predict the airfoil self-noise for the SD 7037 airfoil. The simulations are conducted to determine the ability of the current computational models to accurately predict the aeroacoustic noise generated by an airfoil. The geometry tested is a 2D blade segment of the SD 7037 airfoil surrounded by a structured C-mesh. The methods tested are Large Eddy Simulation (LES) to solve the flow parameters and the Ffowcs-Williams and Hawkins (FW-H) acoustic analogy to determine the sound generated by the airfoil's surface [3]. These methods were tested using the ANSYS Fluent software, which has these analysis methods built-in.

The static AOAs investigated are 0° and 1° at a Reynolds number of 4.25×10^4 , and Figure 1 shows the simulated acoustic spectra compared against the experimental measurements by Tam [2]. The results show good agreement but bring up valid questions about the simulation requirements to predict the strong tonal noise that is generated at 1° by the laminar boundary layer – vortex shedding (LBL-VS) airfoil self-noise mechanism. This is investigated by comparison of compressible and incompressible simulations and their impact on the simulation of the boundary layer and its interaction with the trailing edge (TE) of the airfoil. The 2D vortex shedding behaviour in the boundary layer and wake is directly related to the tonal noise generated at 4100Hz through a transient comparison of vorticity magnitude contours, lift coefficient, and acoustic pressure data. The 3400Hz tone was only measured clearly in the compressible 0° simulation (as seen in Figure 1), which calls into question the simulation requirements for predicting the feedback loop responsible for the more complex airfoil self-noise mechanism. The full paper will provide a thorough explanation of the findings, as well as aspects of the prediction model that require further investigation.

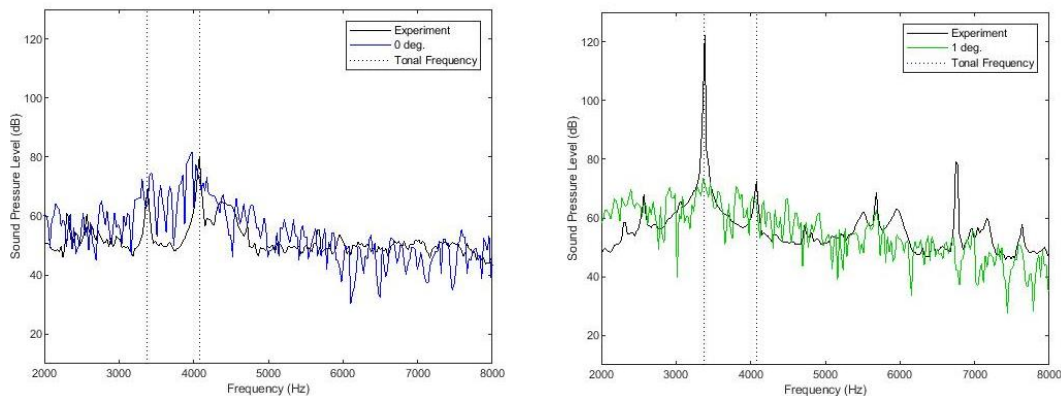


Figure 1: Acoustic spectra for SD 7037 at 0° (left) and 1° (right) compared with experimental data by Tam [2]

- [1] F. Ghorbanishohrat, *Study of a low Re airfoil considering laminar separation bubbles in static and pitching motion*. PhD thesis, University of Waterloo, PhD thesis, 2019
- [2] N. Tam, *An Aeroacoustic Study of Airfoil Self-Noise for Wind Turbine Applications*. University of Waterloo, MSc thesis, 2017.
- [3] J. Ffowcs Williams and D. Hawkins, "Sound Generation by Turbulence and Surfaces in Arbitrary Motion," *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 264, no. 1151, pp. 321-342, 1969.

Experimental study of controlling flow separation in channel expansions

Devi Ram Thapa
D.M. Wills Associates Limited
150 Jameson Drive Peterborough, Canada
dthapa@dmwills.com

ABSTRACT

Irrigation networks which require an expansion linking a narrow rectangular channel to a wider trapezoidal channel. This can cause a significant head loss and increase the risk of boundary erosion. A narrow rectangular channel connected to a wider trapezoidal channel by a wedge expansion which is relatively simple to construct. However, the flow in this expansion experiences a sudden change in flow direction as it crosses the diagonal connecting the two planes that form the side walls of the wedge expansion. For the first time, this problem was overcome by providing a narrow bridging device along the diagonal of the wedge expansion. This reduces the flow separation and hence limited the loss of energy. The characteristics of three-dimensional flow in both wedge and modified wedge systems were investigated by means of non-intrusive measurements of the flow field. Flow separation characteristics were recorded by tracing the separation zone and the points of reattachment based on velocity data, and were confirmed by visual observation involving dye tests. The results from this experimental study have a direct application to the design of efficient irrigation channel networks. The Laser Doppler Anemometry (LDA) measurements of flow velocity were obtained and used for the determination of three-dimensional (3D) flow parameters, including the longitudinal water surface profile, the expansion energy loss and the maximum velocity in the expansion exit flow. Details of flow separation, flow reattachment, turbulence intensity and secondary flow characteristics were studied.

The result from experimental study showed that the separation occurred on one side, and the flow was three-dimensional. Compared to the wedge expansion, the modified wedge expansion effectively guides the flow and prevents abrupt change in flow direction. The modified wedge expansion gives a much earlier recovery of the flow in the downstream channel, lower levels of turbulence, and a less strong secondary flow. It also achieves a small gain in the energy at the location of flow recovery and reduces the head loss coefficient, as a result of the reduced separation zone.

Keywords: Wedge transition; Modified wedge transition; Flow separation; Secondary flow; Head loss; Laser Doppler anemometry measurements; Open channel flow; Hydraulics; Hydraulic design; Experimentation.

PLANE COUETTE FLOW OF HEAVY WATER WITH VARIABLE VISCOSITY

Robert L. Varty

Consultant in Engineering Science

Edmonton, Alberta, Canada

robvarty@telus.net

ABSTRACT

The dependence of the viscosity of a Newtonian fluid on its temperature can be incorporated into calculations of the velocity distribution in a flow. The steady flow between a fixed plate and a moving plate (plane Couette flow) was considered. The fluid was heavy water, at the temperatures and the pressures that exist in a heavy-water-cooled nuclear reactor. The temperature of the fixed lower plate was less than the temperature of the moving upper plate. The temperature varied linearly with the position between the plates. The pressure variation between the plates was much smaller than the absolute pressure of the fluid.

The differential equations for viscous flow were solved numerically by a finite-difference method. The values of viscosity for various temperatures and pressures were calculated using functions of a form indicated by Henry Eyring's molecular theory of liquids; the coefficients were determined from information that had been published by The International Association for the Properties of Water and Steam. A computer program was developed; the program was written in the dialect of BASIC that was available in IBM Lotus Symphony and in Apache OpenOffice. This technique provided a graphical user interface and a spreadsheet-based display of the results.

The program was run for flows with subcritical Reynolds numbers and less-than-one Brinkman numbers, so that the assumptions of laminar flow and negligible viscous dissipation were valid. A comparison of the resulting velocity distributions with those for constant viscosity indicates the importance of calculations with temperature-dependent viscosity.

System Simulation of Engine Cooling Water Jacket

Elle Mistruzzi

Dept. of Mechanical, Automotive & Materials Engg.
University of Windsor
Windsor, On, Canada
mistruze@uwindsor.ca

Shankar Natarajan

System Simulation
FCA Canada Inc.
Windsor, On, Canada
shankar.natarajan@fcagroup.com

Ronald Barron

Department of Mathematics & Statistics
University of Windsor
Windsor, On, Canada
az3@uwindsor.ca

Ram Balachandar

Dept. of Civil & Environmental Engineering
University of Windsor
Windsor, On, Canada
rambala@uwindsor.ca

ABSTRACT

The engine cooling system in a vehicle ensures the engine runs at its most efficient temperature at different operating conditions. The system includes heat exchangers, a thermostat, pump, plumbing lines and a cooling water jacket. Each branch of the system and its different components need to receive adequate coolant flow. A 1D system level model of the coolant loop is generated with each component of the system characterized using geometry and or performance data. Accurately modeling and capturing the flow behaviour of the coolant through the water jacket poses a particular challenge. With a 1D system model, we are able to simulate a complete transient drive cycle. This presentation explores the use of an experimental flow bench and CFD results to convert the engine cooling water jacket into a robust 1D system model. Non-dimensional pressure loss and Reynolds number are calculated based on pressure drop and flow rate data, for a wide range of cold to hot conditions. GT-SUITE software is used as the system simulation modeling platform, and its built-in application GEM-3D is used to import the 3D CAD geometry. A detailed investigation is performed by carefully splitting the cylinder block and head into multiple individual flow components with at least one inlet and one outlet. Gasket holes are modeled as the inlet/outlet between the block and head subassemblies. Outcomes of this research will include an improved water jacket modeling process, and an overall reduction in cost and time to achieve accurate results.

Three-Dimensional Characterization of Bursting and Sweeping Phenomena in the Turbulent Boundary Layer

Kadeem Dennis

Department of Mechanical Engineering
University of Western Ontario
London, Ontario, Canada
kdennis5@uwo.ca

Kamran Siddiqui

Department of Mechanical Engineering
University of Western Ontario
London, Ontario, Canada
ksiddiq@uwo.ca

ABSTRACT

The boundary layer has a well-established role in governing several important engineering quantities. At large scales the atmospheric boundary layer governs dynamic loading on structures and at small scales the boundary layer in a heat exchanger controls the convective heat transfer rate. In a majority of engineering systems, the boundary layer is turbulent in nature and thus numerous turbulent phenomena characterized by three-dimensional non-linear processes are present whose dynamics strongly influence the transportation of mass, heat, and momentum. As it is desirable to manipulate parameters such as the drag force and heat transfer rate to improve the performance of engineering systems, it is of interest to advance our understanding of governing turbulent boundary layer dynamics.

The composition of the turbulent boundary layer features several regimes whose dynamics and interactions influence key transportation processes in the boundary layer. The inner region, which contains the viscous sublayer, buffer layer, and logarithm layer, are known to be home to several important turbulent phenomena. This group of phenomena features turbulent coherent structures and burst and sweep events. The dynamics of turbulent coherent structures in the boundary layer is a highly active field of research with numerous studies dedicated to characterizing their nature. Burst events are characterized by low momentum (e.g. relatively low streamwise velocity) fluid in the viscous sublayer ejected upward into the buffer and logarithmic layers where fluid has higher streamwise velocity. As the ejected fluid accelerates upward, the associated momentum, mass or heat are carried away from the wall. The strength of the bursting process has been demonstrated in prior research studies where a significant increase in local Reynolds stress was reported during bursting events. In contrast to the upward motion induced by burst events, sweep events are identified by high momentum (e.g. relatively high streamwise velocity) fluid found in the logarithmic and buffer layers rushing downward into the viscous sublayer. While substantial work has been focused on characterizing burst and sweep events individually, there has been little experimental research on describing the interaction and three-dimensional nature of these phenomena. The focus of this study is to experimentally investigate and characterize the interaction of burst and sweep events in the turbulent boundary layer from 3D aspects.

Experiments were performed in a closed loop, low-disturbance wind tunnel featuring a 1 m long test section. The turbulent boundary layer is formed over a smooth horizontal flat plate. The multi-plane particle image velocimetry (PIV) technique was performed to capture two-dimensional velocity fields of multiple planes with respect to the mean flow direction. Five free stream velocities were tested ranging from $U = 1.0$ m/s to 7.5 m/s corresponding to Re_θ from 280 to 1700. In all experiments at least 5,000 instantaneous velocity vector fields are captured at a rate of 15 Hz. From the captured velocity fields in each experiment, first order turbulent statistics are calculated to determine the instantaneous fluctuating velocity fields. Quadrant analysis is then performed, i.e. using conditional sampling to categorize each turbulent velocity vector into four unique quadrant events (e.g. Q1, Q2, Q3, Q4) where burst and sweep events are identified as Q2 and Q4 respectively. The relative strength of each quadrant event is determined and used as a basis to describe the frequency and temporal interaction of each quadrant event via wavelet analysis. This analysis is extended to 3D by a comparison of findings from each of the velocity measurement planes. Results are used to construct a temporal model describing the interaction of quadrant events, thus burst and sweep events, and their relative strengths, as important transport phenomena in the turbulent boundary layer.

Effect of background turbulence on the evolution of and entrainment into a turbulent jet

Rana Sahebjam

Department of Civil Engineering and Applied Mechanics
McGill University
Montreal, Canada
Rana.sahebjam@mail.mcgill.ca

Susan Gaskin

Department of Civil Engineering and Applied Mechanics
McGill University
Montreal, Canada
Susan.gaskin@mcgill.ca

Environmental jets are often subject to the turbulent fluctuations of the receiving ambient. The effect of background turbulence on the evolution and entrainment of a turbulent jet is studied. The background turbulence results in meandering of the jet, and, at any instant, it deviates the jet from its time-averaged properties. The present investigation builds on previous work undertaken in this laboratory studying the effect of background turbulence on the velocity and concentration fields of a turbulent jet.

A quasi-homogeneous, isotropic and zero-mean-flow turbulent background was generated by means of a random jet array. The laser-induced fluorescent (LIF) method was employed to obtain the scalar field of a high Schmidt number ($Sc \sim 2000$) turbulent jet at different cross sections ($20 < x/D < 60$). Mean and root mean square (RMS) concentrations demonstrate a faster decrease of concentration due to the background turbulence. Moreover, probability density functions (PDFs) and the intermittency factor are compared to the quiescent background case to discuss the effect of background turbulence on the structure of the jet and entrainment into the jet.

Drag reduction in conduits with vibrating walls

J. K. Rogenski

Department of Mechanical & Materials Engineering
Western University
London, Canada
jrogensk@uwo.ca

J. M. Floryan

Department of Mechanical & Materials Engineering
Western University
London, Canada
floryan@uwo.ca

ABSTRACT

The efficiency of transportation systems is strongly related to the magnitude of drag. Flow control strategies have been observed occurring both naturally in biological systems as well as developed technologically, with the drag reduction being their objective. As an example, drag reduction has been observed in the peristaltic movement in human bodies, e.g. transport of food through the gullet, lymph circulation, and ejaculation. The efficiency of such transport can be improved by using conduit wall vibrations with proper characteristics. Wall vibrations belong to the active flow control methods and therefore they require energy input. Proper selection of parameters, e.g. amplitude of vibrations and phase speed of deformation waves, as well interaction of different wave patterns, determines if the drag can be reduced and what the associated energy cost is. The optimal situation occurs if the energy cost of actuation is lower than the energy savings resulting from the drag reduction. The objective of this work is to better understand the effects of vibration on the flow in a controlled conduit and to identify relations between the system parameters. The model problem used to develop and test the relevant ideas involves a steady two-dimensional flow of a Newtonian fluid confined between two infinite parallel walls and driven by a pressure gradient. This flow is modified by imposing wall vibrations in the form of waves traveling either downstream or upstream. The effect of these waves on the flows is measured by determining the change in the pressure gradient required to drive the same mass flow rate through the smooth reference channel and through the controlled, vibrating channel. A decrease of this pressure gradient corresponds to drag reducing waves. Detailed analysis of surface forces shows the roles played by the pressure and by the shear effects. The model problem is solved numerically with the help of the Galileo transformation that converts the original unsteady problem into a steady problem in a moving frame of reference. The numerical solution is carried out using both the spectrally accurate immersed boundary conditions (IBC) method and a modified version of the Open source Field Operation And Manipulation (OpenFOAM) toolbox. Results show that it is possible to identify wave parameters resulting in the reduction of pressure losses leading to energy savings larger than the energy cost of the actuation.

Effect of Anchor Impeller Speed on Gas Dispersion in Highly Viscous non-Newtonian fluids using Coaxial Mixers: A Peculiar Phenomenon

Sinthuran Jegatheeswaran (PhD student)

Chemical Engineering Department
Ryerson University
Toronto, Canada
s2jegath@ryerson.ca

Farhad Ein-Mozaffari (Professor)

Chemical Engineering Department
Ryerson University
Toronto, Canada
fmozaffa@ryerson.ca

ABSTRACT

In a wastewater treatment plant, the aeration process contributes to the highest energy consumption, which is about 50-60% of the total operational cost, to achieve the desired oxygen concentration for the survival of micro-organisms [1,2]. Since gases are less dense than bulk liquid, the residence time of bubbles in the fluid is very low leading to an inefficient gas-liquid mass transfer and poor aeration efficiency. As a result, the aerobic biological processes are only suitable if the Chemical Oxygen Demand (COD) concentration is less than 1000 mg/L [3]. Despite a high degree of treatment efficiency, the aerobic biological processes are limited by the Chemical Oxygen Demand (COD) concentration of the wastewater and its aeration efficiency.

In our research study, the effect of anchor impeller hydrodynamics on gas dispersion in non-Newtonian fluid was studied both experimentally and numerically, for the first time, in a coaxial configuration since there is absolutely no information available in the literature. The coaxial mixer consists of a central impeller and a slowly moving anchor near the wall. At a fixed central impeller speed of 290 rpm, a significant drop in the local gas holdup was observed when the anchor impeller speed increased from 30 to 50 rpm. Through air velocity vector and volume fraction contours obtained via CFD, this study shed light on how the anchor impeller hydrodynamics affect the gas dispersion and flow regime. By operating at a low critical impellers speed ratio (i.e. central impeller rotational speed/anchor impeller rotational speed), the research findings open avenue for energy-efficient gas dispersion routes at low Reynolds numbers.

1. F. Liu, A. Ouedraogo, S. Manghee, A. Danilenko, A primer on energy efficiency for municipal water and wastewater utilities. Energy Sector Management Assistance Program (ESMAP). 2012.
2. D. Mamais, C. Noutsopoulos, A. Dimopoulou, A. Stasinakis, T. Lekkas, Wastewater treatment process impact on energy savings and greenhouse gas emissions. Wat. Sci. Tech. 71 (2015) 303-308.
3. Y.J. Chan, M.F. Chong, C. L. Law, D.G. Hassell, A review on anaerobic-aerobic treatment of industrial and municipal wastewater, Chem. Eng. J. 155 (2009) 1-18.

EXPERIMENTAL SIMULATION OF A DENSITY-DRIVEN DOWNBURST TRANSLATING WITHIN AN ATMOSPHERIC BOUNDARY LAYER

Shivani Jariwala
 Dept of Mechanical and Materials
 Engineering
 University of Western Ontario
 London, Canada
 sjariwal@uwo.ca

Kyle Graat
 Dept of Mechanical and Materials
 Engineering
 University of Western Ontario
 London, Canada
 kgraat@uwo.ca

Eric Savory
 Dept of Mechanical and Materials
 Engineering
 University of Western Ontario
 London, Canada
 esavory@uwo.ca

ABSTRACT

In a downburst a dense air mass containing precipitate matter rapidly descends from a cumulonimbus cloud, baroclinicity forming a vortex ring that impinges on the ground. Downbursts having strong intensity pose a threat to buildings, transmission lines, aircraft during take-off and landing, and to vegetation. Hence, understanding the wind field of a downburst before and after it impinges on ground at model scale is imperative for characterizing the zones where maximum impact of the downburst outflow is experienced. In the current study a density-driven downburst is modelled using aqueous solutions of potassium dihydrogen phosphate as the dense fluid and glycerol as the ambient fluid with density difference of 3.37%, with a cylindrical fluid release mechanism (diameter= 12cm, height of cylinder= 9.2cm, scale ratio= 1:16000) arrangement to replicate a single stationary downburst event with release height of 10.5cm within a sheared wind environment. Comparing the model scale dimensions to a full-scale downburst event, the cylindrical fluid release mechanism with diameter of 12cm represents a full-scale downburst of diameter equal to 700m and the release height of 10.5 cm represents the height of 1.68 km of the cloud base from ground. Turbulence generating passive devices such as fence, spires and a uniformly-placed array of hexagonal nuts are used to create a realistic, scaled boundary layer within a 4.90m long and 1.06m wide hydraulic flume with working flow depth of 28cm. The resulting boundary layer due to the turbulence generating passive devices for the freestream velocity of the ambient fluid ranging from 3-7cm/s is quantified using hot-film anemometry. For understanding the scalar field and the mixing characteristics of the dense and ambient fluid throughout the event the planar laser induced florescence (PLIF) method is adopted, whilst particle image velocimetry (PIV) is used to quantify the velocity vector field. The main objective of this research work is to understand the influence of ambient shear on the outflow and velocity characteristics of a density-driven downburst.

Keywords: density-driven downbursts, atmospheric boundary layer, hot-film anemometry, planar laser induced florescence, particle image velocimetry, hydraulic flume

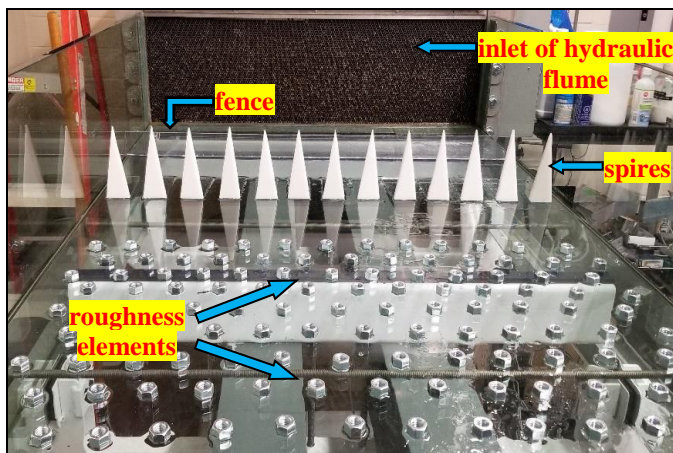


Figure 1 Experimental setup of turbulence generators over the bed of hydraulic flume

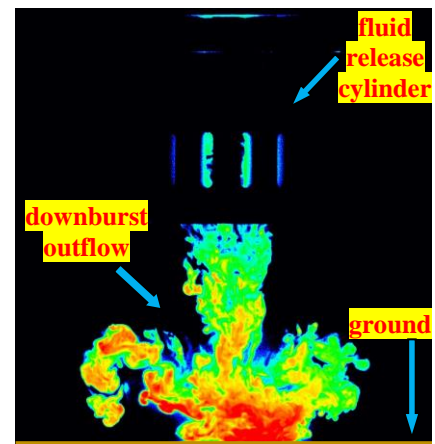


Figure 2 PLIF image of a downburst event at model scale (adapted from Babaei 2018)

AN EXACT SOLUTION TO THE 3D NAVIER-STOKES AND APPLICATIONS TO THE RANQUE-HILSCH VORTEX TUBE

Nolan J Dyck

Department of Mechanical and Materials Engineering
 The University of Western Ontario
 London, Ontario, Canada
 ndyck@uwo.ca

Anthony G Straatman

Department of Mechanical and Materials Engineering
 The University of Western Ontario
 London, Ontario, Canada
 astraatman@eng.uwo.ca

ABSTRACT

The Ranque-Hilsch Vortex Tube (RHVT) is a small tube which separates a supplied compressed air stream from a tangential inlet into hot and cold streams of air, emitted from a small hole at one end of the tube and an annular exit at the opposite end, respectively. Despite the vast literature available on RHVTs, no consensus has been reached on the source of the so-called “temperature separation” phenomenon occurring within the RHVT.

The objective of the present work is to study the energy movement in the RHVT. To this end, a new exact solution to the 3D incompressible Navier-Stokes equations describing a RHF-like flow has been developed. Starting from the incompressible Navier-Stokes equations, the streamfunction is introduced, and the vorticity equations have been derived. We have used an approach similar to that of Wang [1] and assumed the vorticity and circumferential velocity may be replaced with linear functions of the streamfunction plus known auxiliary functions. With this approach, the problem is reduced to three quasi-linear partial differential equations, which may be solved sequentially to yield a family of swirling, recirculating flows, such as the flow shown in Figure 1.

From the resulting family of solutions, we compute the “energy separation” of these flows by comparing the energies of the recirculating (cold exit) flow and the non-recirculating (hot exit) flow to the inlet flow.

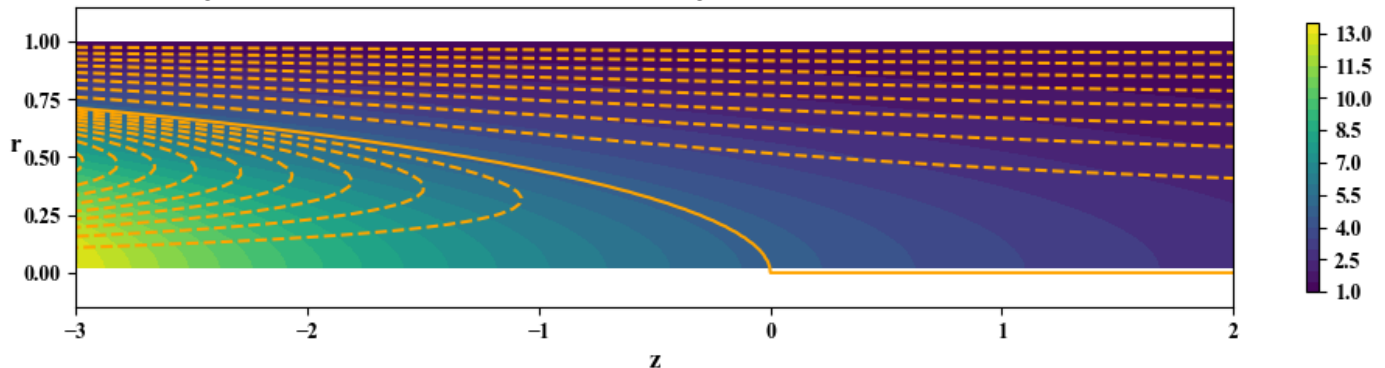


Figure 1: Swirling flow solution in a tube of radius 1. Streamlines are shown in Orange (the stagnation streamline is solid) and the colour contours show the variation in angular velocity about the tube axis. Fluid enters near the tube wall on the left hand side. Some fluid continues moving towards the right, while the remaining fraction recirculates along the axis.

References

- [1] Wang, C.-Y. (1966). On a Class of Exact Solutions of the Navier-Stokes Equations. *Journal of Applied Mechanics*, 33(3), 696–698. <http://doi.org/10.1115/1.3625151>

Analysis of Gas Dispersion in Non-Newtonian fluids with a Coaxial Mixer through Tomography and CFD and Response Surface Methodology

Maryam Jamshidzadeh (*Author*)

Dept. of Chemical Engineering
Ryerson University
Toronto, Canada

Maryam.jamshidzadeh@ryerson.ca

Dr. Farhad Ein-Mozaffari (*Author*)

Dept. of Chemical Engineering
Ryerson University
Toronto, Canada

fmozaffa@ryerson.ca

Dr. Ali Lohi (*Author*)

Dept. of Chemical Engineering
Ryerson University
Toronto, Canada

Maryam.jamshidzadeh@ryerson.ca

ABSTRACT

The gas dispersion in non-Newtonian fluids is a challenging task due to the complex rheology exhibited by this type of fluids. The main objective of this study was to analyze the performance of an aerated coaxial mixer comprising of two central impellers and an anchor impeller for the gas dispersion in the CMC solution, which is a pseudo-plastic fluid. To achieve this objective, electrical resistance tomography was utilized to measure the local and global gas holdup. The interfacial area, which has a pronounced effect on the mass transfer between the gas and liquid phases, is a function of the bubble size distribution inside the aerated system. Thus, the coupled dynamic gas disengagement-tomography technique was employed to measure the number of bubble classes and sizes. A computational fluid dynamics (CFD) model was developed to determine the optimal design parameters and operating conditions. The visualization of flow patterns, the shear distribution and its effect on local bubble size were also explored using CFD. In this study, the effects of different parameters such as speed ratio of impellers, pumping direction, and rotation mode on the power consumption, gas hold-up, and bubble size were investigated. The results showed that the combination of the downward pumping pitched-blade impeller and the anchor in the co-rotating mode enhanced the gas holdup and generated smaller bubble sizes.

Laminar boundary layer response to an impulse forcing by a spanwise array of plasma actuators

Hossein Khanjari

Department of Mechanical Engineering
Lassonde School of Engineering, York University
Toronto, Canada
khanjari@yorku.ca

Ronald E. Hanson

Department of Mechanical Engineering
Lassonde School of Engineering, York University
Toronto, Canada
hansonre@yorku.ca

Philippe Lavoie

Institute for Aerospace Studies
University of Toronto
Toronto, Canada

Plasma actuators are often used for actively controlling flows. For example, a spanwise array of dielectric barrier discharge (DBD) was used as vortex generators for applications relating to the delay of bypass transition [1] or separation control [2]. In the later example of [2], relatively large amplitude streaks are generated to enhance the region of attached flow over a ramp simulating airfoils at high angle of attack. Hanson *et al.* [1] used the actuator arrays to produce low-amplitude streaks typical of those preceding transition to turbulence following the bypass transition pathway. It was demonstrated that streaks undergoing transient growth could be attenuated using plasma actuator arrays, suggesting the feasibility of a bypass transition control system using this technology.

The experimental data available for DBDs is limited to steady or quasi-steady cases. Measurements of the streaks produced by plasma actuators in the work of Hanson *et al.* [1] in a laminar boundary layer are limited to only streamwise velocity, using hotwire anemometry. However, access to the wall-normal velocity (as implemented by particle image velocimetry of Jukes *et al.* [2]), would result in a more comprehensive understanding of the development of the streaks, which can support the development of large-bandwidth closed loop control systems.

In the present study, a 3-D laminar flow is simulated using the commercial CFD code Star-CCM+ using a 2nd order implicit unsteady numerical method. A rectangular flow domain of 150x200x1250 (*mm*) is considered. The boundary conditions are an inlet velocity of 5 m/s, symmetry along the side-walls, and pressure outlets along the top and downstream of the inlet section (the former is used to achieve a zero pressure gradient flow). The mesh is comprised of hexahedron cells in the core volume. Prism layers are applied throughout the boundary layer region. In total, approximately 1.5M cells are contained in the flow domain to adequately resolve the flow and regions near the actuators. In order to simulate the actuation conditions of Ref. [1], a momentum source term is used to model influence of the plasma actuators. First, a laminar boundary layer is established that follows the Blasius solution in absence of the momentum source term.

The effect of DBDs on the flow is simulated using the plasma actuator body force distribution model presented by Shyy *et al.* [3]. Since the momentum source term is unknown, for model validation, the source term was optimized to achieve the steady spatial distribution and magnitudes of disturbance shown in Ref. [1]. In the next step, an impulse forcing from the actuator is also considered in the model.

The simulations provide all three components of velocity which leads to a more detailed insight into the vortex formation process. Therefore, understanding of the spatial and temporal dynamics for the pulsed operation of the plasma actuator are also enhanced. Particular attention is directed to the downstream shear stress, as an important aspect of controller design, which has previously shown a non-minimum-phase response behavior. Practical control methods, restrict flow measurements to those feedbacks obtained at the wall. In the present work, it is shown that obtaining all three velocity components for the spanwise shear stress may provide predictive tool for streak formation. This can improve the control system bandwidth for attenuating the boundary layer streaks.

[1] Hanson. R, Bade. K, Belson. B, Lavoie. P, Naguib. A, & Rowley. C. 2014 Feedback control of slowly-varying transient growth by an array of plasma actuators. *Physics of Fluids* 26, 024102 (2014).

[2] Jukes, T. & Choi, K.-S. 2013 On the formation of streamwise vortices by plasma vortex generators. *J. Fluid Mech.* (2013), vol. 733, pp. 370–393.

[3] Shyy. W, Jayaraman. B, & Andersson. A. 2002 Modeling of glow discharge-induced fluid dynamics. *Journal of Applied Physics* 92, 6434 (2002).

MEASURING DYNAMIC CONTACT ANGLES OF A DROPLET ON INCLINED HYDROPHOBIC SURFACES WITH IRREGULAR ROUGHNESS

Kewei Shi, Xili Duan

Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, NL, A1B 3X5 Canada

Kristin M. Poduska

Department of Physics and Physical Oceanography
Memorial University of Newfoundland
St. Johns, NL A1B 3X7 Canada

ABSTRACT

Metallic superhydrophobic surfaces are being studied for corrosion and ice mitigation on offshore structures in harsh marine environments. A combination of surface roughening and low energy coating is often used to fabricate these non-wetting surfaces. For metallic materials, the surfaces are usually irregularly roughened, since regular patterned surfaces are hard to fabricate on an industrial scale. Quantification of the surface wettability often reduces to a measurement of the contact angles (static and dynamic) that a liquid droplet makes on the surfaces. This paper presents measurement results of dynamic contact angles of a water droplet on inclined (super)hydrophobic surfaces. Two sample surfaces are used in this study, both with zinc deposition on sandblasted surfaces by 250 μm and 100 μm Al_2O_3 particles respectively, and both are coated with stearic acid afterwards. A DataPhysics OCA 15EC contact angle instrument is used for the measurement of drops on an adjustable tilted plate. The results demonstrated special challenges encountered in measuring the dynamic wetting on randomly roughened hydrophobic surfaces. It is found that droplet pinning happens frequently and can severely affect the droplet shape and sliding dynamics and contact angles. These effects lead to reduced data quality if the traditional measurement of advancing and receding angles is used. With these results, we argue that the field of surface wettability should move to be less reliant on simple models and consider these realistic factors in characterizing the wettability of randomly roughened surfaces.

Numerical investigation of the effect of mid-section length on the aerodynamics of the Ahmed body

Naseeb A. Siddiqui* and Martin Agelin-Chaab
Faculty of engineering and applied sciences
University of Ontario Institute of Technology
Oshawa, Canada

* naseeb.siddiqui@uoit.net

ABSTRACT

The Ahmed body is a benchmark simplified bluff body for investigating the overall flow aerodynamics behaviour of a generic road vehicle. The literature reveals that it has been extensively studied from multiple perspectives using various techniques. However, the effect of the mid-section length on the overall drag is not understood. In fact, this was highlighted by (Ahmed, Ramm, & Faltin, 1984) that the interference between the rear end and the forebody is weak and the small drag at the forebody could be a consequence of the length of the mid-section. Also, the interference between upstream and downstream flow can be anticipated to generalize this result. There is also no reported work on this in the open literature, and hence this investigation fills that gap. In this study, the mid-section length of the Ahmed body is varied to examine the aerodynamic behavior of the bluff body under different conditions. Five aspect ratios (the length to height ratio) of the Ahmed body with 25° rear slant surface are considered. A Reynolds number of 2.8×10^6 based on the total length is used. The SST K-Omega model in Fluent is used to solve the governing equations. Preliminary results show that there is no significant change in the overall drag with the aspect ratios. However, the coefficient of drag at the forebody, midsection, and rear is compared to analyze the relationship between lengths and drag coefficient. This bifurcation of drag coefficients can easily be compared with the experimental result as expressed by (Ahmed et al., 1984) which is analyzed further. This comparison is expected to reveal the affected bifurcation of the total drag coefficient. In addition, the coherent structures and vorticity generation at the rear of the body is anticipated to be affected by different mid-section lengths. This investigation is expected to facilitate more understanding of the effect of the mid-section length on the flow around the Ahmed body in general.

References

Ahmed, S. R., Ramm, G., & Faltin, G. (1984). Some Salient Features Of The Time-Averaged Ground Vehicle Wake. *SAE Transactions*, 93(2), 473–503. <https://doi.org/10.4271/840300>

A Study on Vortex Dynamics in Laminar-Turbulent Transition by a Point Cloud Model

Kazuo Matsuura
 Graduate School of Science and Engineering
 Ehime University
 Matsuyama, Ehime, Japan
 matsuura.kazuo.mm@ehime-u.ac.jp

ABSTRACT

Different vortices appear in the process of laminar-turbulent transition in boundary layers. In this study, a new data mining method which reduces the degree of freedom of vortices based on a point cloud model, and can do list management of them by putting cluster IDs on the vortices is proposed for laminar-turbulent transition.

Two cases of computation on overall transition processes, i.e., Cases A and B, are conducted in this study. The scenario for each case is summarized in Table 1. Case A is a case without Free-stream turbulence (FST), and transition of typical K-regime is reproduced. Case B is a case with FST of intensity $Tu=0.5\%$, and the generation of hairpin vortices is influenced by the FST.

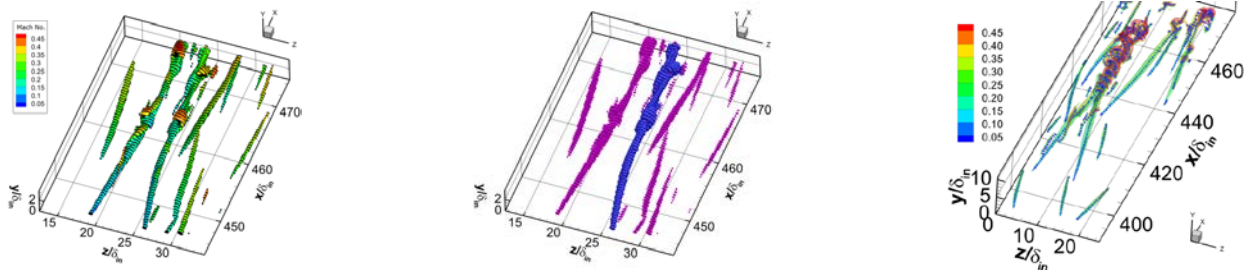
The governing equations are the unsteady three-dimensional fully compressible Navier-Stokes equations written in general coordinates for body-fitted mesh geometries. The system of equations is closed by the perfect gas law. The equations are solved by a sixth-order finite-difference method and the third-order explicit Runge-Kutta scheme.

Figure 1 shows the results of applying a newly devised algorithm judging interior points of a vortex, a clustering method and a skeltonization method to an instantaneous flow field of Cases A and B. Part (a) shows the vortical structures represented by the iso-surfaces of the second invariance of velocity gradient tensor (SIVGT), and also interior points enclosed by the iso-surface. Part (b) shows a cluster corresponding to an unstable hairpin leg. Part (c) shows hairpin and satellite vortices represented by the skelton points. These results show that the present algorithm works successfully, and the present clustering method can selectively pick up a connected vortex structure (a leg part in this example), which is located close to other longitudinal vortices.

Thus, the data mining method that reduces the degree of freedom of vortices based on a point cloud model, and can do list management of vortex structures by clustering the particles is proposed for the transition processes.

Table 1 Computational Cases, IC: initial condition, LBL: laminar boundary layer, SD: sinuous disturbances, FST: free-stream turbulence, IST: isotropic turbulence

Case	FST Intensity	IC	Inflow BC
A	0%	Steady LBL Profile	LBL+SD
B	0.5%	Case A	LBL+SD+FST(IST)



(a) Vortical structure visualized by SIVGT and its representation by particles (black points) enclosed by the iso-surfaces, The color in the legend is Mach No. (Case A)

(b) Extracted cluster corresponding to a unstable hairpin leg (blue points) (Case A)

(c) Hairpin and satellite vortices represented by the skelton points and the transparent iso-surfaces (Case B)

Fig. 1. The results of applying the newly devised algorithm and the clustering method to instantaneous flow fields.

A VERY LARGE-EDDY SIMULATION (VLES) MODEL FOR INVESTIGATION OF THE THERMALLY-STRATIFIED ATMOSPHERIC BOUNDARY LAYERS: REDUCTIONIST AND PRACTICAL

Amir A. Aliabadi

Environmental Engineering, University of Guelph
Guelph, Canada

Corresponding Author: aliabadi@uoguelph.ca

Gonalo Pedro

RWDI
Guelph, Canada

Abstract—A Very Large-Eddy Simulation (VLES) model was developed to simulate thermally-stratified atmospheric boundary layers. The model performance was validated against experimental wind tunnel observations. The effects of grid resolution and model input parameters were analyzed in the suitability of the model to predict the experimental observations. The model exhibited a short adaptation distance smaller than five boundary-layer heights. It produced most of the mean and turbulence statistics profiles reasonably well in agreement with the experiments. The VLES model shall be further tested at full scale to simulate thermally-stable atmospheric boundary layers. The model should also be tested on complex topography with differential surface temperature to ensure adequate performance in realistic applications.

Keywords—Atmospheric boundary layer; Computational Fluid Dynamics (CFD); Thermally-stratified flow; Very Large-Eddy Simulation (VLES);

I. INTRODUCTION

In the recent decades, Large-Eddy Simulation (LES) models have gained popularity among practical CFD models with a high degree of fidelity [4]. It is true that LES models do not resolve turbulent fluctuations of a flow at all scales of the energy cascade, but they resolve the energy-containing and a significant portion of the inertial subrange scales [3, 4, 6]. Such scales of fluctuations govern important aspects of turbulent flows such as pollutant dispersion and excitation of structures immersed in flows.

Despite such popularity, one of the main difficulties in implementing realistic LES models is the lack of the availability of simplistic yet adequate perturbation fields to be inserted at the model flow inlet. LES requires such perturbations because it is the nature of this type of modelling to simulate the evolution of perturbations in a numerical domain that eventually describe the turbulent fluctuations in the flow. From a theoretical stand point, the perfectly ideal inlet

fluctuations must meet several criteria: 1) they must be stochastically varying, on scales down to the spatial and temporal filter scales; 2) they must be compatible with the Navier-Stokes equations; 3) they must be composed of coherent eddies across a range of spatial scales down to the filter length; 4) they must allow easy specification of turbulent properties; and 5) they must be easy to implement [16].

In practice, however, no simple model exists that meets all such criteria for flow inlet perturbations. Models either employ simplistic approaches while missing important physical aspects of the flow [10, 17, 15, 7, 11, 18], or they employ fully realistic methods at the expense of complexity of implementation [1, 2, 8, 14].

A reductionist yet practical model was implemented by Aliabadi et al. [6] for the investigation of neutral atmospheric boundary layers. This model was known as a Very Large-Eddy Simulation (VLES) model and employed a synthetic method for inlet perturbations. The development aimed at minimizing the number of input parameters for the model while attempting to simulate mean flow, turbulence statistics, spectra, and anisotropy realistically. The input parameters were a single timescale and a single lengthscale specifying the turbulent inlet fluctuations, a parameter controlling the Sub-Grid Scale (SGS) transport, and an aerodynamic surface roughness lengthscale. The model was tested against experimental wind tunnel data and other LES models. It produced experimental profiles of mean velocity and turbulence velocity statistics reasonably well. It simulated spectra and anisotropy realistically. This model showed potential for use in industrial applications where it is impractical to perform high resolution simulations or implement complex synthetic inlet boundary conditions to match all flow properties beyond what is necessary for a practical application.

The objective of this study is to extend the VLES model capabilities in simulating thermally-stratified atmospheric boundary layers. The model performance is validated against experimental wind tunnel observations of Ohya [12]. The effect of grid resolution is analyzed. The model performance as a

function of inlet parameters is studied in a sensitivity investigation. Various thermal stability strengths are analyzed. In addition, the model performance is investigated with the use of momentum and thermal wall functions. The VLES model is implemented in OpenFOAM 4.0.

II. METHODOLOGY

A. Model Geometry

The model geometry is shown in Fig. 1. The tunnel height, width, and length are $Z=1.5\text{m}$, $Y=1.5\text{m}$, and $X=5\text{m}$, respectively. Airflow is in the x direction. Four vertical solution probes are considered for monitoring the simulation results. All results reported in this paper are obtained from the solutions monitored on probe 4. This choice ensures the flow is fully developed for statistical sampling.

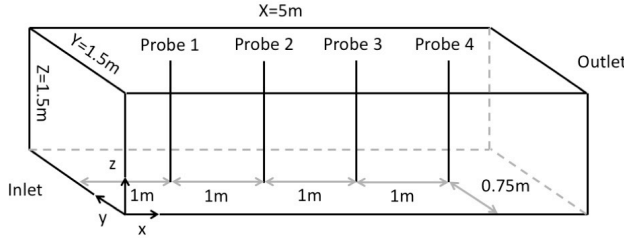


Figure 1. Model geometry.

B. The Synthetic Vortex Method

To generate turbulence at the inlet, a vortex method is used. This method was originally developed by Sergent [15] but has been continually improved until recently [18]. In this method, velocity fluctuations are inserted at the inlet in the form of synthetic two-dimensional eddies derived from mean statistical information about the flow as a function of space (height above ground) and time. This method does neither require a precursor method to generate fluctuations beforehand using another simulation, nor require periodic streamwise boundary conditions to recycle fluctuations from inside the domain. Therefore, it offers great simplicity. The velocity fluctuation field is given as

$$u(x) = \frac{1}{2\pi} \sum_{i=1}^N \Gamma_i \frac{(x_i - x) \times s}{|x_i - x|^2} \left(1 - e^{-\frac{|x_i - x|^2}{2(\sigma_i(x_i))^2}} \right) e^{-\frac{|x_i - x|^2}{2(\sigma_i(x_i))^2}}, \quad (1)$$

where $u(x)$ is velocity perturbation vector at the model inlet that is later superimposed on the mean inlet velocity, x is position vector on the inlet boundary, N is the number of vortices to be inserted at the inlet (in this study 200), i is the index for the current vortex, Γ_i is the circulation for the current vortex, x_i is the position vector for the centre of the current vortex, s is unit vector along the streamwise direction, and $\sigma_i(x_i)$ is characteristic length for the radius of current vortex. We assume that the wall-normal direction is $+z$ and that flow is in the $+x$ direction. A power-law profile is used for the mean inlet velocity

$$\bar{U}(z) = U_{ref} \left(\frac{z}{z_{ref}} \right)^\alpha, \quad (2)$$

where z_{ref} is a reference height (0.1m), U_{ref} is reference velocity, and α is an exponent parameterized as a function of aerodynamic surface roughness length of the surface z_0 given as

$$\alpha = \frac{1}{\ln\left(\frac{z_{ref}}{z_0}\right)}. \quad (3)$$

Next a turbulence intensity profile must be assumed. This is obtained from

$$I_u(z) = \frac{1}{\ln\left(\frac{z}{z_0}\right)}, \quad (4)$$

where $I_u(z)$ is limited by a maximum value given the fact that for atmospheric flows there is a limit to $I_u(z)$ of typically in the order of one. This allows parameterization of Turbulence Kinetic Energy (TKE) (k) such that

$$k(z) = 1.5(U(z)I_u(z))^2. \quad (5)$$

Calculation of the characteristic size for energy-containing eddies begins with the calculation of inlet boundary lengthscale

$$L = \frac{2L_z L_y}{L_z + L_y} \quad (6)$$

where inlet dimensions are used. It is reasonable to assume that the size of energy-containing eddies σ_{max} scales with L , using a constant a_σ , to be adjusted later, according to

$$\sigma_{max} = a_\sigma L \quad (7)$$

Meanwhile, the size of energy-containing vortices or eddies is a function of height and must decrease with height. The energy-containing vortex size is parameterized using the mixing length approach

$$\frac{1}{\sigma(z)} = \frac{1}{\sigma_{max}} + \frac{1}{\kappa(z + z_0)}, \quad (8)$$

where, $\kappa=0.41$ is the von Kármán constant. A characteristic time for the largest energy-containing eddies can be approximated using scaling. The characteristic velocity U_0 for the largest energy-containing eddies can be defined using the power law and the reference height so that $U_0 = \alpha z_{ref}^\alpha$. The lengthscale for such eddies can be defined as $\ell_0 = \sigma_{max}$. Calculation of these two scales enable the calculation of the largest energy-containing eddies Reynolds number $Re_{\ell_0} = U_0 \ell_0 / \nu$. We can subsequently calculate the Kolmogorov length scale $\eta = \ell_0 Re_{\ell_0}^{-3/4}$, the Kolmogorov velocity scale $u_\eta = U_0 Re_{\ell_0}^{-1/4}$, and the dissipation rate $\epsilon = \nu(u_\eta/\eta)^2$. Finally, we can calculate the characteristic lifetime for the largest energy-containing eddies

$$\tau_0(\ell_0) = \left(\frac{\ell_0^2}{\epsilon} \right)^{1/3}. \quad (9)$$

This timescale is not representative for all energy-containing eddies, but only the largest ones. For ease of implementation,

however, it is possible to define a representative timescale for all energy-containing eddies assuming a constant a_τ , to be adjusted later, with

$$\tau = a_\tau \tau_0 (\ell_0). \quad (10)$$

This timescale can be used to sample a new set of vortices at the inlet after every fixed number of timesteps, so that as soon as this timescale is elapsed new vortices will be sampled.

The circulation can be formulated for each vortex with the knowledge of computational face area S , in which the vortex center is located and the TKE (k) given for a height. The circulation sign is randomized as either positive or negative such that

$$\Gamma = \pm 4 \left(\frac{\pi S k}{3N(2 \ln 3 - 3 \ln 2)} \right)^{1/2}. \quad (11)$$

The inlet temperature profile is formulated using a power law with the same exponent found earlier such that

$$\Theta(z) = (\Theta_\infty - \Theta_s) \left(\frac{z}{z_{max}} \right)^\alpha + \Theta_s, \quad (12)$$

where Θ_∞ is the far field temperature on top of the boundary layer, Θ_s is surface temperature, and z_{max} is the height for the top of the domain, i.e. above which Θ_∞ is defined.

These formulations fully close the system of equations necessary for the synthetic eddy method.

C. The Sub-Grid Scale (SGS) Model

The Sub-Grid Scale (SGS) model is fully described by Aliabadi et al. [3]. This is based on a one-equation TKE (k) model, for which the sub-grid lengthscale is formulated as

$$l = C_\Delta (\Delta x \Delta y \Delta z)^{1/3}, \quad (13)$$

where C_Δ is a parameter to control l and therefore the SGS model.

D. Wall Functions

The VLES model requires two wall functions: one for momentum transport and the other for heat transport near the walls. The wall function for momentum transport is given by Raupach et al. [13]

$$U^+ = \frac{U}{u_\tau} = \frac{1}{\kappa} \ln \left(\frac{z+z_0}{z_0} \right), \quad (14)$$

where u_τ is friction velocity and z_0 is the characteristic aerodynamic roughness length of the surface. The wall function for temperature is given by Jayatilleke [9]

$$T^+ = \frac{(T_w - T) \rho C_p u_\tau}{q_w} = Pr_t (u^+ + Pr_f), \quad (15)$$

where $Pr_t=0.85$ is the turbulent Prandtl number. Here Pr_f is further parameterized as a function of Pr_t and the laminar Prandtl number $Pr=0.72$

$$Pr_f = 9.24 \left[\left(\frac{Pr}{Pr_t} \right)^{3/4} - 1 \right] \left[1 + 0.28 e^{(-0.007 \frac{Pr_t}{Pr})} \right]. \quad (16)$$

E. Numerical Schemes

The implementation of numerical grid, boundary conditions, finite volume schemes, finite volume solution control, and solution averaging are fully discussed in Aliabadi et al. [3] and Aliabadi et al. [6] and will not be provided here for brevity.

F. Numerical Grids

Four grid levels are chosen for the simulations that vary from very fine (level I) to very coarse (Level IV). These grids are listed in Table I with appropriate descriptions and specifications.

TABLE I. NUMERICAL GRID LEVELS

Grid Level	Grid Level Information		
	Description	$N_x-N_y-N_z$	N_{Total}
I	Very fine	100-100-100	1,000,000
II	Fine	100-75-75	562,500
III	Coarse	100-50-50	250,000
IV	Very Coarse	100-25-25	62,500

G. Validation Dataset

The validation dataset is obtained from wind tunnel experiments of Ohya [12]. For this wind tunnel a chain roughness was used with lengthscale $h=0.0055m$ or equivalently an aerodynamic roughness length of $z_0=0.00055m$ ($z_0 \sim 0.1h$). Table II shows the details of wind tunnel experimental cases. As can be seen, the experiments are mainly run to vary the thermal stability condition, given by the bulk Richardson number Ri_δ . In addition, the Reynolds number based on boundary layer height Re_δ and vertical temperature difference between far field and surface $\Delta\Theta = \Theta_\infty - \Theta_s$ are reported. These cases correspond to very weakly stable (Case 1) to very strongly stable (Case 4) conditions.

TABLE II. WIND TUNNEL EXPERIMENTAL CASES OF OHYA (2001)

Experimental Variables	Stability Case			
	Case 1	Case 2	Case 3	Case 4
$U_\infty [m s^{-1}]$	1.83	1.29	1.01	0.91
Re_δ	50,600	35,300	28,000	23,700
Ri_δ	0.12	0.24	0.40	0.74
$\Delta\Theta = \Theta_\infty - \Theta_s [K]$	27.4	27.4	28.7	43.3

III. RESULTS

A. Sensitivity to Grid Levels

First the VLES model has been run with a default set of synthetic eddy parameters on different grid resolutions. Fig. 2 shows the results of the analysis. It appears that most solution variables reasonably agree with the experimental wind tunnel observations for a grid level as coarse as grid level III. The choice of such a grid warrants a good agreement with experiments. However, results obtained on grid level IV

deviate from the experimental observations, suggesting that such a coarse level of a grid is not desirable. These simulations were wall resolving so that no wall function had to be used at the wall boundary. Here δ represents the boundary-layer height, which is used to normalize vertical distance from the wall.

B. Different Thermal Stability Cases

Next the VLES model was run with a default set of synthetic eddy parameters for different thermal stability cases according to Table II. Fig. 3 shows the results of the analysis. All stability cases were run on grid level III, which was shown to be resolved enough for this VLES model. The agreement is reasonably good. Although there are deviations from the wind tunnel observations, the VLES model predicts the same trends as the experiments for the turbulence statistics by indicating a suppression of the magnitude of such statistics with increasing thermal stability.

C. Sensitivity to a_σ

Next the VLES model has been run in a sensitivity investigation for the choice of the a_σ synthetic eddy parameter. This parameter controls the size of eddies fed at the inlet. Fig. 4 shows the results of the analysis. All cases were run on grid level III. The results show that mean quantities are not affected

by varying a_σ ; however, using larger values of a_σ results in larger magnitudes of turbulence statistics. This can be explained by the fact that larger eddy structures are more energetic and overall add to the turbulence levels in the simulation domain. The choice of $a_\sigma=3$ results in the best agreement in turbulence statistics with the wind tunnel experiments. Note that this choice is valid for wall-resolving simulations.

D. Sensitivity to a_τ

Next the VLES model has been run in a sensitivity investigation for the choice of the a_τ synthetic eddy parameter. This parameter controls how frequently new eddies are sampled at the inlet. In other words, this parameter controls the lifetime of the largest energy-containing eddies at the inlet. Fig. 5 shows the results of the analysis. All cases were run on grid level III. The results show that mean quantities are not affected significantly by varying a_τ ; however, using larger values of a_τ results in larger magnitudes of turbulence statistics. The choice of $a_\tau=0.01$ results in the best agreement in turbulence quantities. Note that this choice is valid for wall-resolving simulations.

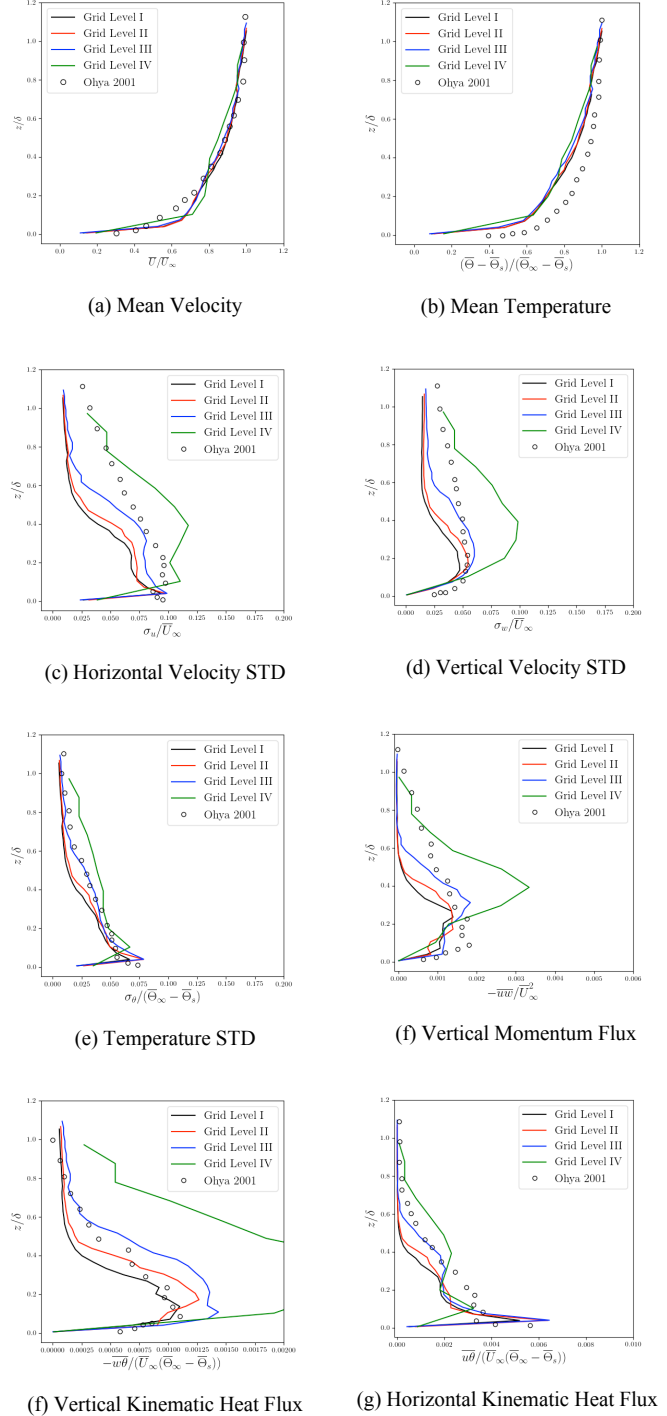


Figure 2. Sensitivity of the VLES model to grid level: very fine (Level I), fine (Level II), coarse (Level III), and very coarse (Level IV).

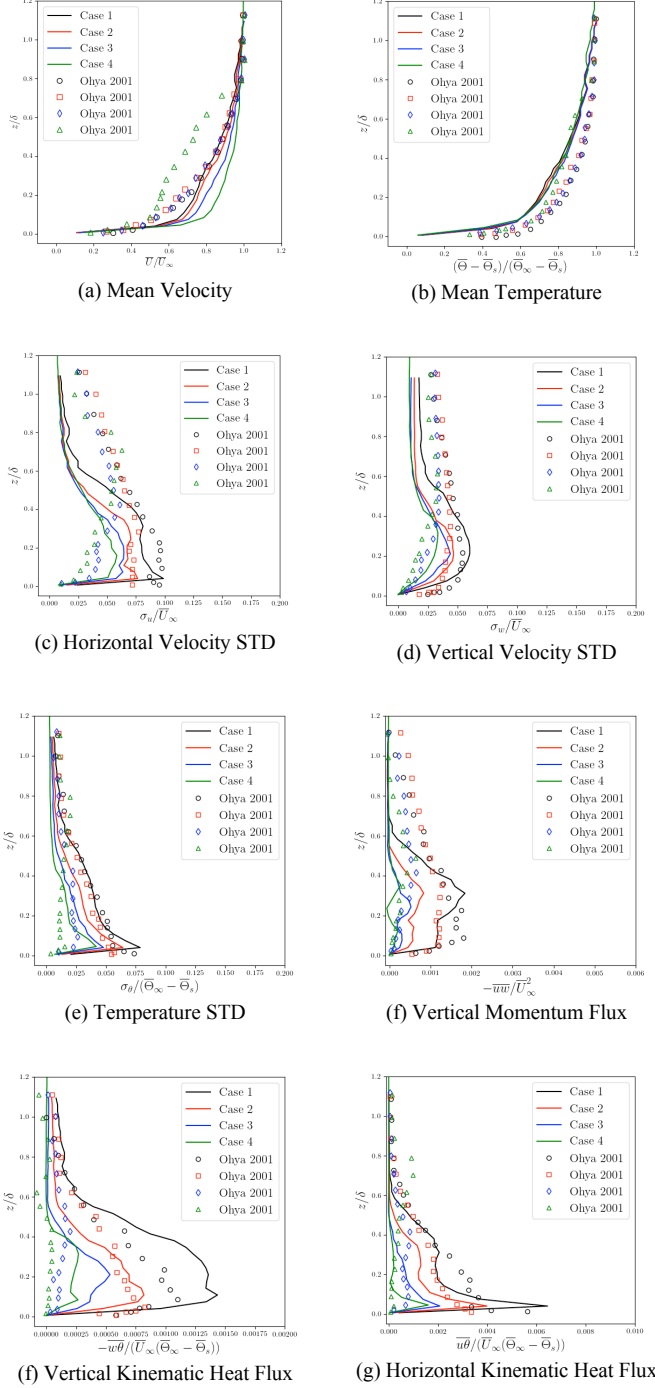


Figure 3. Response of the VLES model to different thermal stability conditions: very weakly stable (Case 1), stable (Case 2), strongly stable (Case 3), very strongly stable (Case 4).

E. Sensitivity to C_Δ

Next the VLES model has been run in a sensitivity investigation for the choice of the C_Δ parameter required for the SGS model. This parameter controls the transport phenomena at sub-grid scales by controlling the sub-grid mixing length.

Fig. 6 shows the results of the analysis. All cases were run on grid level III. The results show that mean quantities are not

affected significantly by varying C_Δ ; furthermore, varying this parameter has different effects on different turbulence statistics. For example, while the effect on turbulence variances is minimal, the effect on turbulent fluxes are greater. The choices

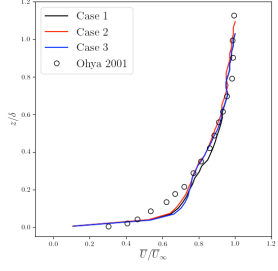
of C_Δ provide reasonable simulation results in good agreement with the wind tunnel experiments.

F. Use of Wall Functions

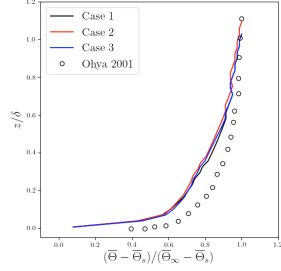
Finally, the VLES model has been run in a sensitivity investigation when using wall functions. Here the first layer of the grid adjacent to the wall is gradually coarsened, which would correspond to increasing values of z^+ in the simulation, while the quality of the solution is monitored. Fig. 7 shows the results of the analysis. All cases were run on grid level III. For these simulations, it was found that increasing the value of z^+ does not degrade the quality of the mean solution obtained. However, with increasing values of z^+ , turbulence statistics reduce and fluctuations are damped. It appears the z^+ values up to 120 still provide reasonable agreement with the wind tunnel experiments.

For use of wall functions, new choices of the synthetic eddy parameters must have been made for the simulations to

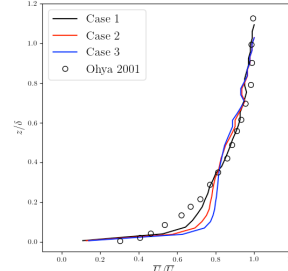
perform adequately. Here, while we still used $a_\sigma=3$, we found that $a_\tau=0.05$ would result in the best agreement. In other words, larger eddy lifetime must have been assumed for reasonable agreement with wind tunnel experiments. This can be explained by the fact that when wall functions are used, turbulence generation near the walls is modelled as opposed to resolved, in which case eddy formation at some distance away from the wall occurs with a larger time constant. This implies that TKE (k) transfer from the wall to the outer layer starts with larger time constants, and therefore, it necessitates more model timestep iterations before new eddies are sampled at the Inlet [6].



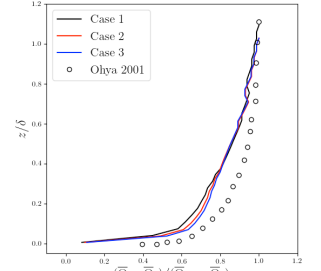
(a) Mean Velocity



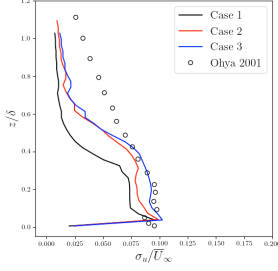
(b) Mean Temperature



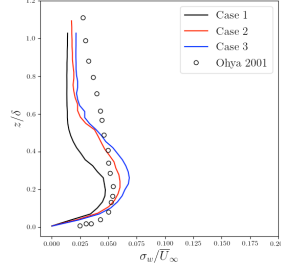
(a) Mean Velocity



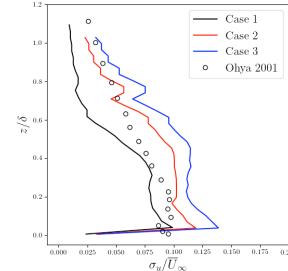
(b) Mean Temperature



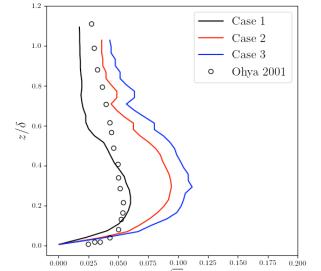
(c) Horizontal Velocity STD



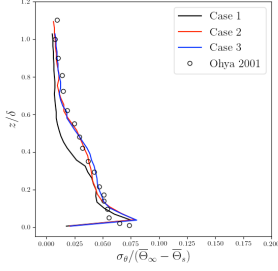
(d) Vertical Velocity STD



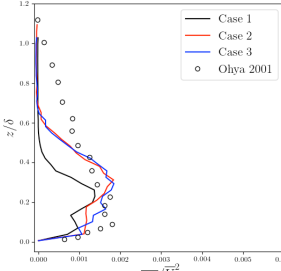
(c) Horizontal Velocity STD



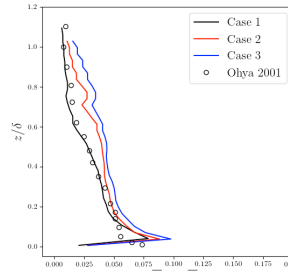
(d) Vertical Velocity STD



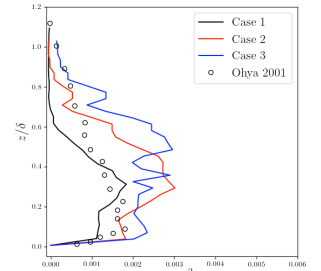
(e) Temperature STD



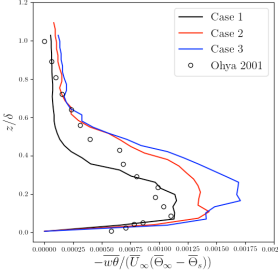
(f) Vertical Momentum Flux



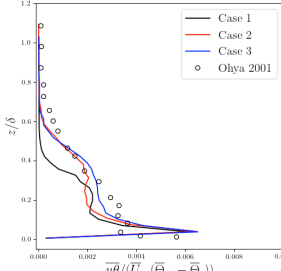
(e) Temperature STD



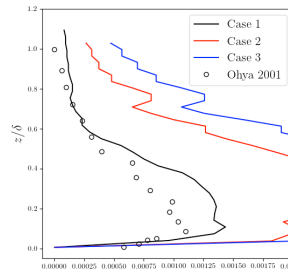
(f) Vertical Momentum Flux



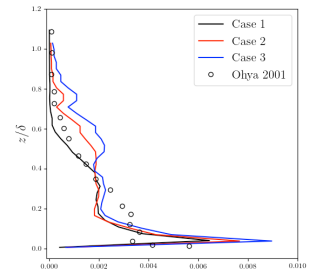
(f) Vertical Kinematic Heat Flux



(g) Horizontal Kinematic Heat Flux



(f) Vertical Kinematic Heat Flux



(g) Horizontal Kinematic Heat Flux

Figure 4. Sensitivity of the VLES model to different values of α_σ : Case 1 ($\alpha_\sigma=1$), Case 2 ($\alpha_\sigma=3$), and Case 3 ($\alpha_\sigma=5$).

Figure 5. Sensitivity of the VLES model to different values of α_τ : Case 1 ($\alpha_\tau=0.01$), Case 2 ($\alpha_\tau=0.05$), and Case 3 ($\alpha_\tau=0.1$).

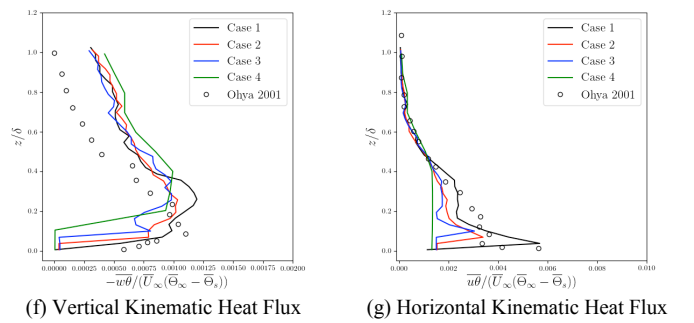
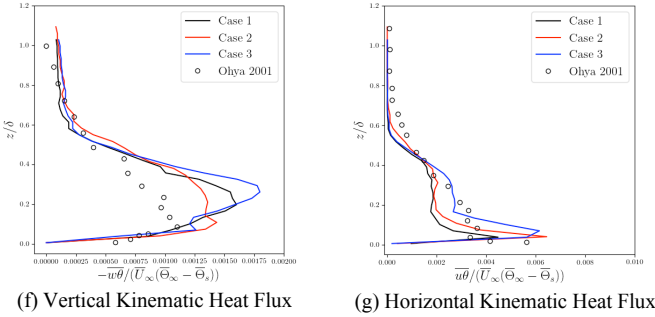
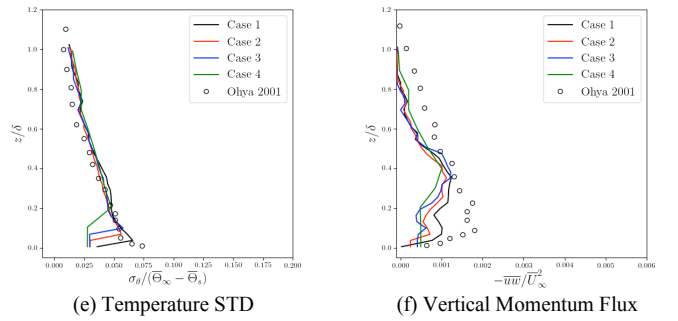
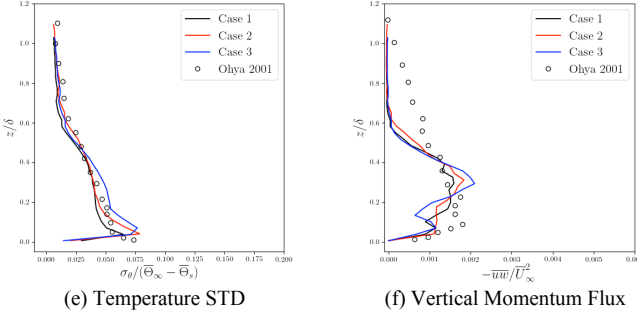
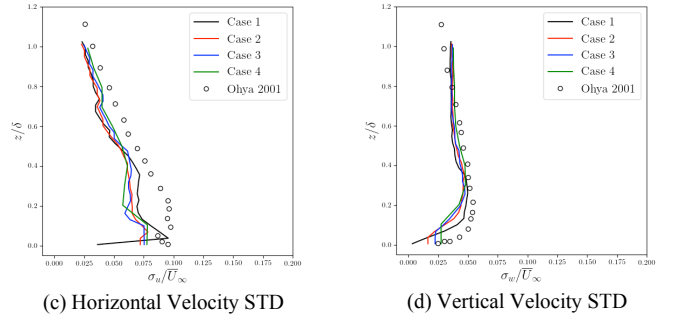
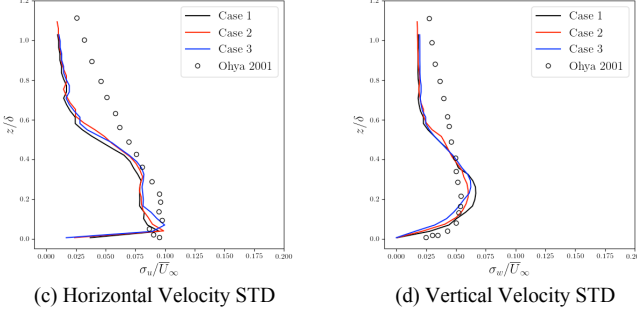
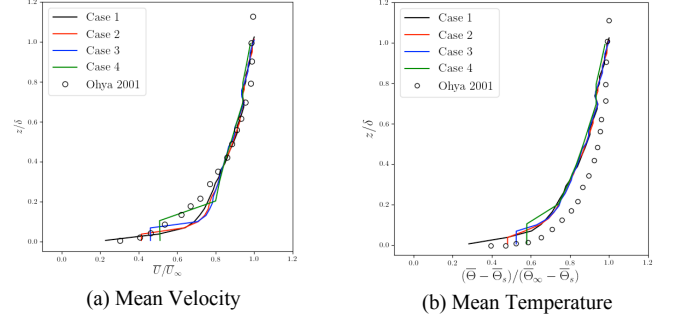
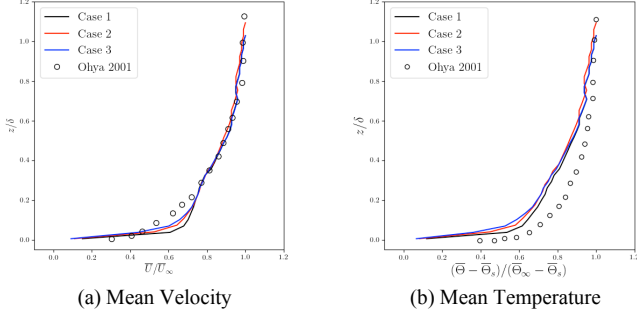


Figure 6. Sensitivity of the VLES model to different values of C_D : Case 1 ($C_D=0.5$), Case 2 ($C_D=1.0$), and Case 3 ($C_D=1.5$).

Figure 7. Sensitivity of the VLES model to the first layer grid height when using wall functions: Case 1 ($z^+ \sim 50$), Case 2 ($z^+ \sim 80$), Case 3 ($z^+ \sim 120$), and Case 4 ($z^+ \sim 150$).

IV. CONCLUSIONS AND FUTURE WORK

A Very Large-Eddy Simulation (VLES) model was developed to simulate thermally-stratified atmospheric boundary layers. The model performance was validated against experimental thermal wind tunnel observations. The effect of grid resolution was analyzed. The model performance as a function of inlet parameters was studied in a sensitivity investigation. Various thermal stability strengths were analyzed. In addition, the model performance was investigated with the use of momentum and thermal wall functions.

The model exhibited a short adaptation distance smaller than five boundary-layer heights. It produced most of the mean and turbulence statistics profiles reasonably well. While turbulence statistics showed sensitivity to the choice of model parameters, the mean profiles were not so sensitive to such parameters. The choice of model parameters allowed an optimization to reach close agreement between the model and wind tunnel experiments. While the choice of wall functions implied reduced computational cost, the implementation of wall functions resulted in suppression of turbulence. Such suppression can further be circumvented by optimization of model parameters.

The VLES model shall be further tested at full scale to simulate thermally-stable atmospheric boundary layers. The model should also be tested on complex topography with differential surface temperature to ensure adequate performance in realistic applications.

ACKNOWLEDGEMENTS

In-kind technical support for this work was provided by Rowan Williams Davies and Irwin Inc. (RWDI). This work was supported by the Discovery Grant program (401231) from the Natural Sciences and Engineering Research Council (NSERC) of Canada; Government of Ontario through the Ontario Centres of Excellence (OCE) under the Alberta-Ontario Innovation Program (AOIP) (053450); and Emission Reduction Alberta (ERA) (053498). OCE is a member of the Ontario Network of Entrepreneurs (ONE). The computational platform was installed with the assistance of the IT team: Joel Best, Jeff Madge, and Matthew Kent.

REFERENCES

- [1] H. Aboshosha, G. Bitsuamlak, and A. El Damatty, "LES of ABL flow in the built environment using roughness modeled by fractal surfaces," *Sustain. Cities Soc.*, vol. 19, pp. 40–60, 2015.
- [2] H. Aboshosha, A. Elshaer, G. T. Bitsuamlak, and A. El Damatty, "Consistent inflow turbulence generator for LES evaluation of wind-induced responses for tall buildings," *J. Wind Eng. Ind. Aerod.*, vol. 142, pp. 198–216, 2015.
- [3] A. A. Aliabadi, E. S. Krayenhoff, N. Nazarian, L. W. Chew, P. R. Armstrong, A. Afshari, and L. K. Norford, "Effects of roof-edge roughness on air temperature and pollutant concentration in urban canyons," *Boundary-Layer Meteorol.*, vol. 164, pp. 249–279, 2017.
- [4] A. A. Aliabadi, *Theory and Applications of Turbulence: a Fundamental Approach for Scientists and Engineers*. Amir A. Aliabadi Publications, Guelph, 2018, pp. 168.
- [5] A. A. Aliabadi, M. Moradi, D. Clement, W. D. Lubitz, and B. Gharabaghi, "Flow and temperature dynamics in an urban canyon under a comprehensive set of wind directions, wind speeds, and thermal stability conditions," *Environ. Fluid Mech.*, 2018 <https://doi.org/10.1007/s10652-018-9606-8>.
- [6] A. A. Aliabadi, N. Veriotes, and G. Pedro, "A Very Large-Eddy Simulation (VLES) model for the investigation of the neutral atmospheric boundary layer," *J. Wind Eng. Ind. Aerod.*, vol. 183, pp. 152–171, 2018.
- [7] S. Benhamadouche, N. Jarrin, Y. Addad, and D. Laurence, "Synthetic turbulent inflow conditions based on a vortex method for large-eddy simulation," *Prog. Comput. Fluid Dynam. Int. J.*, vol. 6, pp. 50–57, 2006.
- [8] H. G. Castro, R. R. Paz, J. L. Mroginski, and M. A. Storti, "Evaluation of the proper coherence representation in random flow generation based methods," *J. Wind Eng. Ind. Aerod.*, vol. 168, pp. 211–227, 2017.
- [9] C. L. V. Jayatilke, "The influence of Prandtl number and surface roughness on the resistance of the laminar sub-layer to momentum and heat transfer," pp. 271, 1966.
- [10] T. S. Lund, X. Wu, and K. D. Squires, "Generation of turbulent inflow data for spatially developing boundary layer simulations," *J. Comput. Phys.*, vol. 140, pp. 233–258, 1998.
- [11] F. Mathey, D. Cokljat, J. P. Bertoglio, and E. Sergent, "Assessment of the vortex method for large eddy simulation inlet conditions," *Prog. Comput. Fluid Dynam. Int. J.*, vol. 6, pp. 58–67, 2006.
- [12] Y. Ohya, "Wind-tunnel study of atmospheric stable boundary layers over a rough surface," *Boundary-Layer Meteorol.*, vol. 98, pp. 57–82, 2001.
- [13] M. R. Raupach, R. A. Antonia, and S. Rajagopalan, "Rough-wall turbulent boundary layers," *Appl. Mech. Rev.*, vol. 44, pp. 1–25, 1991.
- [14] M. Ricci, L. Patruno, and S. de Miranda, "Wind loads and structural response: benchmarking LES on a low-rise building," *Eng. Struct.*, vol. 144, pp. 26–42, 2017.
- [15] M. E. Sergent, *Vers une methodologie de couplage entre la simulation des grandes echelles et les modeles statistique*, pp. 198, 2002.
- [16] G. R. Tabor, and M. H. Baba-Ahmadi, "Inlet conditions for large eddy simulation: a review," *Comput. Fluids* vol. 39, pp. 553–567, 2010.
- [17] T. G. Thomas, and J. J. R. Williams, "Generating a wind environment for large eddy simulation of bluff body flows," *J. Wind Eng. Ind. Aerod.*, vol. 82, pp. 189–208, 1999.
- [18] B. Xie, "Improved Vortex Method for LES Inflow Generation and Applications to Channel and Flat-plate Flows," pp. 146, 2016.

CAN WE LIMIT THE PASSENGER EXPOSURE TO AIRBORNE CONTAMINANTS IN COMMERCIAL AIRCRAFT CABINS DURING THE CLIMB LEG?

Hossam A. Elmaghraby

Mechanical Engineering Program, School of Engineering
University of Guelph
Guelph, Ontario N1G 2W1, Canada
hmaghr@uoguelph.ca

Yi Wai Chiang, Amir A. Aliabadi

Environmental Engineering Program, School of Engineering
University of Guelph
Guelph, Ontario N1G 2W1, Canada
chiange@uoguelph.ca

*Corresponding Author: aliabadi@uoguelph.ca

Abstract—The climb leg is one of the most acceleration-intensive periods in a passenger aircraft flight. It was previously found that the passenger exposure to cough-released airborne contaminants during climb may reach 2.8 to 3.0 times when compared to other legs [1]. In the current study, airflow design and source control strategies are researched numerically for their ability to reduce cough-released airborne contaminant dispersion in the cabin of a Boeing 767-300 aircraft during climb. Sulfur Hexafluoride (SF₆) was used to mimic the contaminant, which mainly comprises of cough-released particles in the size range from 1.6 to 3 µm in diameter. The airflow design strategies involved altering the supply airflow direction, while the source control strategies involved moving the cougher to different locations in the cabin. Among all cases, the relocation of the cougher to the left-side, centre-row location and changing the airflow direction in two and three dimensions exhibited the highest reduction in passenger exposure to contaminant compared to the baseline climb case. The exposure reductions were 0.5-0.7 times for the first case, 0.5-0.7 times for the second case, and 0.4-0.9 times for the third case, respectively.

Keywords- Aircraft acceleration; Airborne contaminants; Air quality; Aircraft ventilation; Source control; CFD

I. INTRODUCTION

Air quality and disease transport aboard passenger aircraft has been an intensive research topic in the past few decades as inferred from a large number of studies in literature [2–13]. This is because the transmission of airborne viruses, such as influenza, tuberculosis, and SARS, is escalated in the closed cabin space through direct passenger-to-passenger exposure and/or from contaminated surfaces (European Centre for Disease Prevention and Control (ECDC) 2018; Mangili and Gendreau 2005). Examples for such transmissions are the in-flight SARS outbreak in China in 2003, and the outbreak of influenza A(H1N1) in 2009 whose introduction was caused by air travel [14]. In addition, the complex environment inside the passenger aircraft cabins due to the high occupant density and

the wide range of passenger activity provide suitable grounds for air quality deterioration and spread of airborne contaminants if no proper remedial measures are taken [15, 16].

Passenger aircraft perform several flight legs, which are ordered as: takeoff, climb, steady level flight (cruise), descent, and landing. During those legs, the aircraft move with high speeds and experience various accelerations [17]. With those accelerations occur body forces that can significantly affect the airflow patterns and airborne contaminants dispersion within aircraft cabins. To the authors' knowledge, no previous aircraft ventilation or air quality studies have investigated the effect of such body forces on in-cabin airflow patterns and contaminant dispersion behavior. Rather, studies in the literature always considered that aircraft are stationary or in cruise mode, for which the only applicable body force results from the gravitational acceleration.

However, Elmaghraby et al. 2019 [1] found in a recent study on a Boeing 767-300 aircraft model that among the steady level flight, climb, and descent legs, the climb leg exhibited the highest levels of contaminant surrogate (SF₆) exposures at two different monitoring locations in the cabin. Also, variations in airflow patterns and airflow circulation (Γ) from one flight leg to the other were noticed. This indicates that acceleration-induced body forces on aircraft have a significant influence on both airflow patterns and contaminant dispersion in the cabins and require further investigation especially in the form of parametric variations and finding possible means of mitigation [1].

In the current study, different airflow design and source control strategies are investigated as mitigation or reduction means for the increased cough-released contaminant exposure in a passenger aircraft cabin during the climb leg using numerical simulations. The case for the descent leg was not studied due to the evidence that contaminant exposure is not significantly influenced for those legs [1]. Airflow design strategies such as changing airflow supply direction are employed. On the other hand, the source control strategies considered involve moving the cougher to other locations in the cabin.

*Address all correspondence to this author.

II. METHODS

A. Cabin Model Geometry

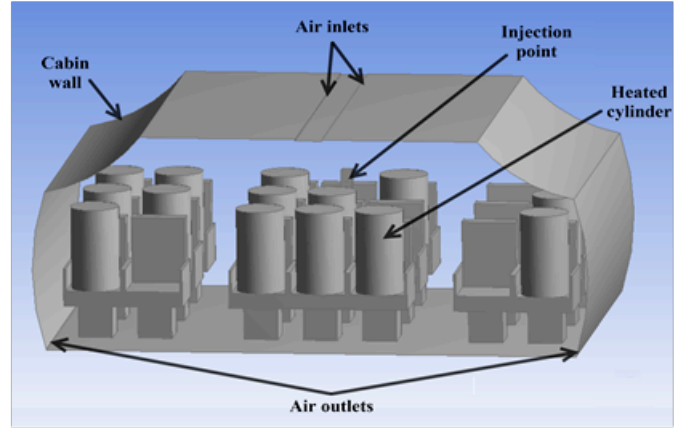
The aircraft cabin model designed and built in the previous study [1] was based on adopted measurements from two studies in literature: an experimental study by Sze To et al. 2009 [18] and a CFD simulation study by Wan et al. 2009 [19] that followed. The original cabin mock-up replicates a full-size sectional economy-class cabin of a Boeing 767-300 passenger aircraft which has 21 seats arranged in three rows. The dimensions of the cabin mock-up are $4.9 \text{ m} \times 3.2 \text{ m} \times 2.1 \text{ m}$ (W, L, H). The cabin mock-up is located at the International Centre for Indoor Environment and Energy, Technical University of Denmark, Lyngby, Denmark [20]. Also, detailed information about the experimental work performed on the dispersion and deposition of expiratory particles in the used aircraft cabin mock-up can be found in the studies [18] and [1]. Fig. 1 depicts the isometric view (DesignModeler software in the ANSYS 17.0 CFD package) and a plan view for the aircraft cabin model.

B. Boundary and Initial Conditions

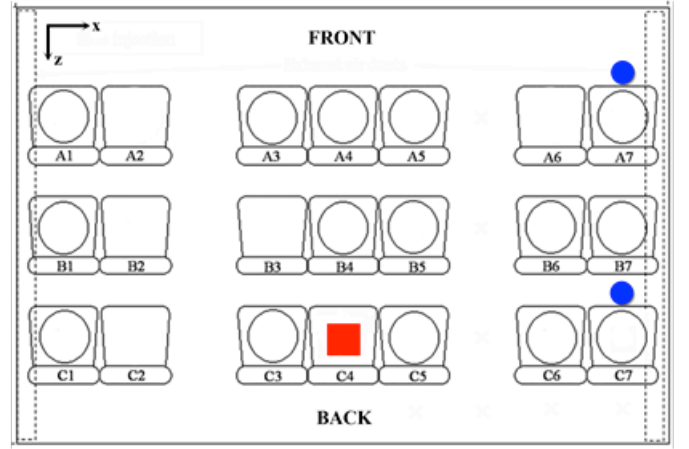
The model's boundary and initial conditions in the original studies [18, 19] were adopted and closely implemented in the numerical solver FLUENT 18.2, and later version 19.1, for the case of 200 L s^{-1} supply air flow rate through the conventional mixing ventilation system used. The Sulfur Hexafluoride (SF_6) gas was released in the cabin to mimic the injection and transport of the cough's smallest size droplets (typically 1.6 to $3.0 \mu\text{m}$), and which formed the largest number concentration of the injected droplet ensemble in the experiments. The SF_6 was introduced as a surrogate to the smallest size cough particles because of its high density and molecular weight (about 6.14 kg m^{-3} and $146.06 \text{ grams mol}^{-1}$, respectively), which make it capable of mimicking the flow behavior of those particles in the cabin [11]. This approach was also adopted to reduce the computational burden of simulating particle motion in the model considering that the current model adopts a reductionist approach. Table I shows the boundary and initial conditions for the current model.

The simulation was run in two parts. First, the airflow domain was completely solved in the steady mode, then the transient section of the simulation is initiated with the cough (SF_6) release for 1 s with a volume of 0.4 l . After this release is stopped, the transient simulation continues for a total time of 350 s .

The standard wall functions were used for near-wall flow treatment, and the SIMPLE solution algorithm was used for the pressure-velocity coupling. The least square cell-based method was employed as the spatial discretization scheme gradient, the second order method was used for solving the pressure, while the second order upwind method was used for solving all other equations (momentum, species transport, turbulence kinetic energy, dissipation rate, energy, etc.). For the temporal discretization, however, the first-order implicit method (implicit backward Euler method) was employed for the transient part of the simulation following a fixed time stepping procedure with a time step size



(a)



(b)

Figure 1. The Boeing 767-300 cabin model used in the current study; (a) isometric view of the geometry built in ANSYS, and (b) plan view for the seats with the cough/injector position (red square), and the contaminant concentration monitoring points (blue circles) [1]

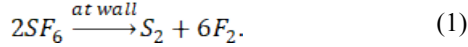
TABLE I. BOUNDARY AND INLET CONDITIONS FOR THE MODEL

Boundary/inlet Conditions	Value
Supply air temperature	24°C
Supply air flow rate	200 l s^{-1} (corresponds to a supply velocity of 2.61 m s^{-1})
Supply air absolute humidity	0.92 g kg^{-1} (corresponds to 5% RH at supply air temperature)
Cabin wall temperature	18°C
Heating cylinder heat release	60 W per cylinder (person)
SF_6 (cough) injection location	Seat C4
Air velocity at injection location	10.6 m s^{-1}

of 0.1 second while allowing 10 solution iterations per time step.

To accurately simulate cough particle deposition on walls and surfaces, the SF_6 gas was not allowed to bounce-off the walls and surfaces in the cabin model. A surface reaction

boundary condition at every wall and surface was set to dissociate SF_6 upon contact to its basic gaseous components: Sulfide (S_2) and Fluorine (F_2) gases according to the following reaction,



As the released SF_6 in the cabin is already at very low concentrations, the concentrations of the Sulfide and Fluorine gases produced from its dissociation are extremely low and do not affect the air composition, fluid properties, or the monitored SF_6 concentration in the cabin. This method allows a surrogate simulation of small particles that behave like gases in the aircraft cabin very computationally economically.

C. Model Validation and Error Estimation

The previous study [1] presents a thorough grid independence analysis of the current model. In addition, it presents model validation through error estimation of the calculated SF_6 concentration time series at the two monitoring seat locations (seats A7 and C7) using various Reynolds-averaged Navier-Stokes (RANS) turbulence models (Standard k-epsilon, RNG k-epsilon, Realizable k-epsilon, Standard k-omega, and SST k-omega). Thus, the grid independence study will not be repeated here again. However, a model validation of the normalized SF_6 concentration time series calculated using the RNG k-epsilon model, which was previously found to be the most accurate turbulence model, against experimental measurements is shown in Fig. 2.

The numerical solution was obtained on a fine grid (7,375,800 grid elements) with the RNG k-epsilon turbulence model that is capable of simulating buoyancy effects on the production and dissipation of turbulence kinetic energy (k).

In the preceding study [1], the error estimation for the model predictions were calculated as the fractional mean bias (FB) and normalized mean square error (NMSE) measures. The FB and NMSE are defined as follows,

$$FB = \frac{2(\overline{C_o} - \overline{C_p})}{(\overline{C_o} + \overline{C_p})}, \quad (2)$$

$$NMSE = \frac{((\overline{C_o} - \overline{C_p})^2)}{(\overline{C_o} * \overline{C_p})}, \quad (3)$$

where, C_o and C_p are the observed (experimental) and predicted (numerical) concentrations, respectively. While FB is a measure of the shift between the observed and predicted quantities, NMSE is a measure of the spread between observed and predicted quantities. For a perfect model, FB and NMSE are both equal to zero.

Table II gives the FB and NMSE values calculated for the air velocity and the normalized SF_6 concentration time series between the experimental measurements and the numerical predictions of the simulation using the RNG k-epsilon turbulence model.

From Table II it can be observed that the NMSE values for normalized SF_6 concentration time series at seat C7 are about 50% less than at seat A7, while FB values are almost identical

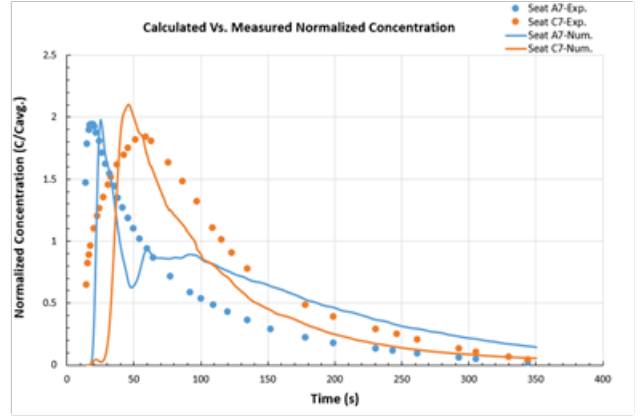


Figure 2. Comparison of the normalized SF_6 concentration time series between the experimental measurements and numerical calculations using the RNG k-epsilon turbulence model on the fine grid level

TABLE II. FB AND NMSE VALUES FOR THE NUMERICAL PREDICTIONS

Air Velocity		Normalized Concentration			
FB	NMSE	FB (Seat A7)	NMSE (Seat A7)	FB (Seat C7)	NMSE (Seat C7)
0.07977	0.02743	0.31867	0.71138	0.39909	0.37711

at both seats. This indicates that the shift between the observed and predicted concentration values is similar, but the spread of the predicted data with respect to the observed ones is two times higher at seat A7, which indicates less accurate predictions. Physically, this decrease in prediction accuracy can be attributed to the condition of airflow, and consequently that of the surrogate SF_6 gas, from the emission source (at seat C4) to each seat. From C4 to C7 the flow is mainly lateral, which is less susceptible to the bulk flow turbulence than the primarily longitudinal flow experienced from C4 to A7 (see Fig. 1).

D. Calculation of Aircraft Body Acceleration Components

The aircraft vertical acceleration (a_v) and horizontal acceleration (a_h) components were calculated during the climb leg using a basic approach adapted from different sources in aircraft dynamics literature.

The calculation procedure relies on applying Newton's second law ($\sum \vec{F} = m\vec{a}$) on two axes passing through the center of gravity of the aircraft; one is vertical and the other is horizontal. The forces in action are the lift (L), drag (D), the aircraft's weight (W), and the thrust of the jet engines (T). For example, the relative vertical acceleration on the aircraft cabin during climb is found to be 2.4 g, which is comprised of two parts; 1.4 g due to aircraft acceleration, and 1 g representing the gravitational component. More information on the calculation procedure followed during the climb leg can be found in [1].

III. RESULTS AND DISCUSSION

As highlighted in the introduction section, the SF_6 concentration was found to be the highest at the two monitoring locations considered during the climb leg as compared to the steady level flight and descent legs. The calculated SF_6 concentration time series during the three flight legs are graphically shown in Fig. 3. Also, it can be noticed from Fig. 3 that the SF_6 concentration time series during the descent leg is similar to that during the steady level flight leg with no significant difference. This can be attributed to the low speed of the passenger aircraft during descent yielding limited acceleration components. Although changing some model factors, such as the location of the cougher (contaminant injector), cough velocity, or airflow conditions may alter this resemblance between the steady level flight and the descent legs, the large relative difference in SF_6 concentration between the mentioned two legs and the climb leg favors the investigation of the climb leg. Therefore, the current study will only consider the climb leg, and different airflow design and source control strategies are to be investigated as remedial techniques to the increasing SF_6 concentration noticed during this leg.

In addition to the SF_6 concentration, the exposure of the passengers to SF_6 over time is used as another measure in this study to assess the effect of the acceleration-induced body forces on the contaminant dispersion in the cabin. The exposure is determined by calculating the area under the curve corresponding to each case using the following integral within the simulation time limits,

$$\text{Exposure} = \int_0^{350s} C_{\text{SF}_6}(t) dt. \quad (4)$$

Using this measure, it was found that the passenger exposure is always the highest during the climb leg when compared to the steady level flight and the descent legs. The highest exposure ratio was 3.0 to 1 calculated between climb and descent at seat C7. Nevertheless, the passenger at the same seat experienced a very similar exposure to the contaminant during the descent and steady level flight legs with a ratio of 0.9 to 1 [1].

A. Airflow Design Strategies

Airflow direction:

In the current study, the direction of the supply airflow was changed using two-dimensional and three-dimensional approaches. In 2-D, the airflow direction was tilted downwards from the cabin ceiling by an angle α . However, in 3-D, an angle β was added to direct the airflow either to the front or to the back of the cabin while still being tilted downwards with the angle α . Fig. 4 shows examples of using those angles in 2-D and 3-D views for the cabin model.

a) Changing angle α only:

First, only the change of the supply airflow direction angle α in 2-D is considered. Three airflow supply angles were investigated for their ability to reduce the SF_6 concentration exposure in the cabin during the climb leg: 20°, 30°, and 60°.

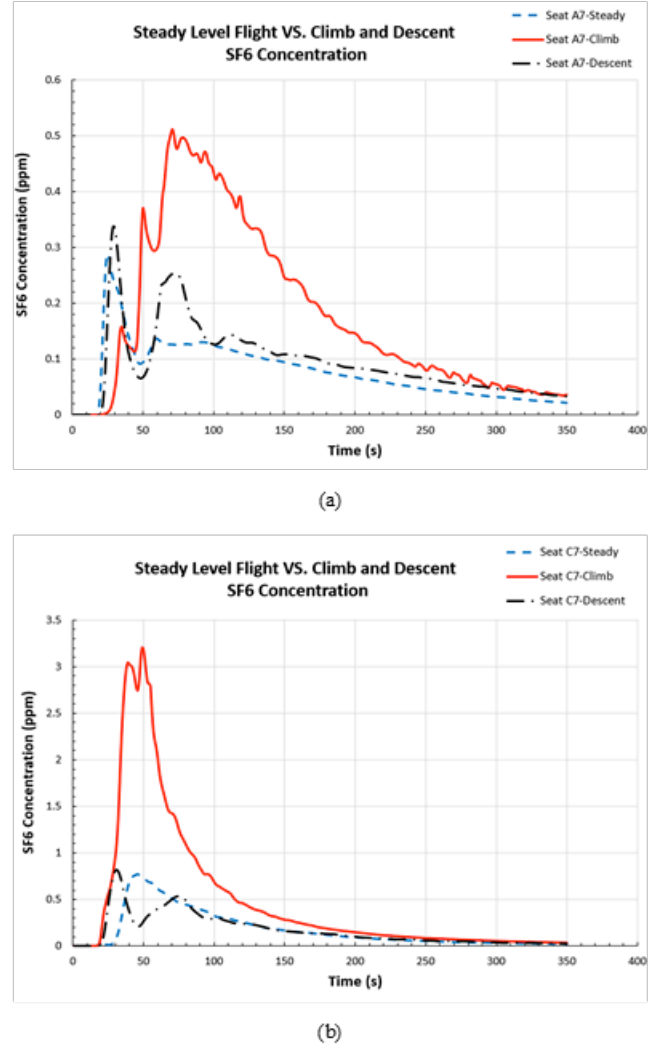
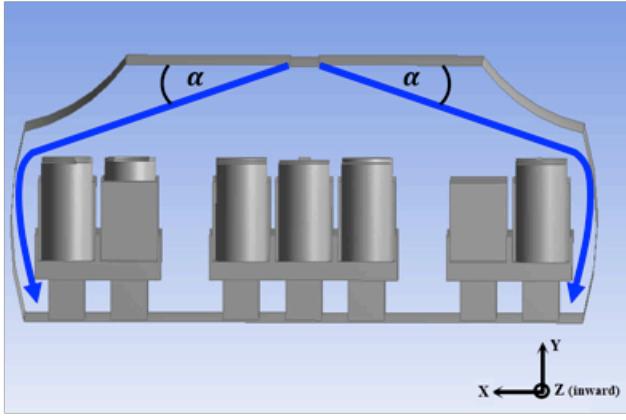


Figure 3. Comparison of the predicted SF_6 concentration time series among steady level flight, climb, and descent legs; (a) at seat A7, and (b) at seat C7 [1]

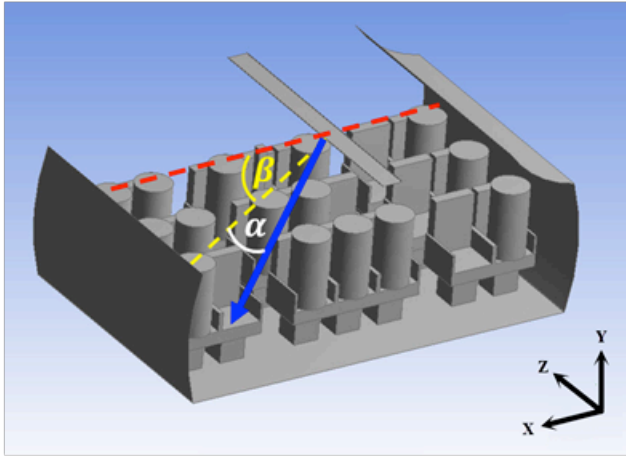
The SF_6 concentration time series for the 20° and 30° airflow supply scenarios during climb versus that for the standard climb and steady level flight from Elmaghraby et al. 2019 [1] at seats A7 and C7 are shown in Fig. 5. However, for the readability of the figures, the same comparison for the 60° supply case is shown separately in Fig. 6.

From Figs. 5 and 6 it can be noticed that there is a considerable difference in the calculated SF_6 concentration time series using each of the three airflow supply angles. The airflow supplied at 30° had the peak SF_6 concentration, and consequently the exposure of occupants, reduced to almost 50% of the original concentration during climb at the two monitoring locations. Conversely, the airflow supplied at 20° was not able to provide better air quality conditions at the two monitoring locations, and the exposure of passengers to the contaminant was almost the same as that for the original climb air supply scenario.

On the other hand, from Fig. 6, it can be seen that supply-



(a)



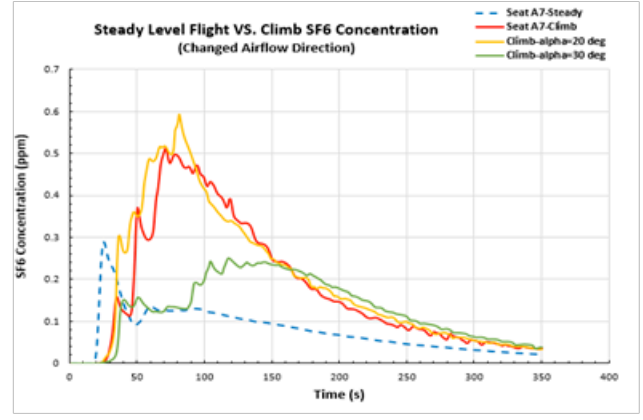
(b)

Figure 4. Redirecting the supplied airflow to the aircraft cabin; (a) in 2-D using angle α only, and (b) in 3-D using angles α (downwards) and β (front or back) together

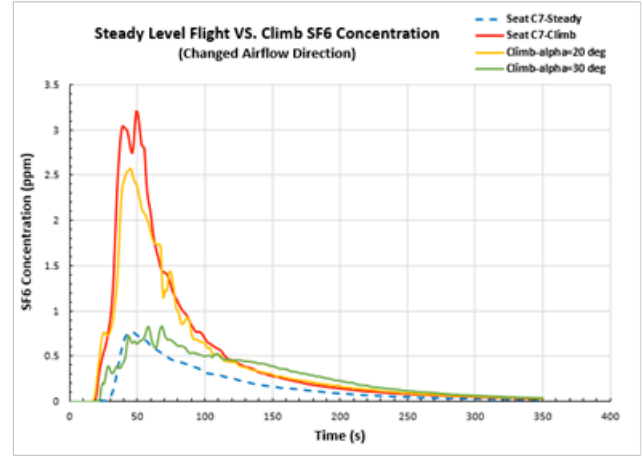
ing air at 60° to the cabin produced the worst air quality conditions at the two locations. This is backed by the very high passenger exposure to SF_6 under this air supply condition, especially at seat A7, where the exposure was around 400% of the original case.

To put this comparison in a more graphical way, SF_6 concentration contours are shown at the breathing level of the occupants during the 30° and 60° airflow supply cases in Fig. 7. Due to the transient nature of the simulations, the best representative time window was chosen to show the contours, which is at 350 s in this case.

As can be seen in the two contour plots, with $\alpha = 60^\circ$, the area covered by the supplied air is very limited at the cabin central area around the two rear seat rows. This leaves most of the seats on the two cabin sides exposed to the contaminant. Conversely, the airflow supplied at 30° efficiently reached the cabin sides and lead to reduced passenger exposure at most cabin seats. However, with $\alpha = 30^\circ$, a very minor increase in the SF_6 concentration is seen at the centre of the cabin due to the elevated mixing effects induced by the strong airflow eddies at this area.



(a)



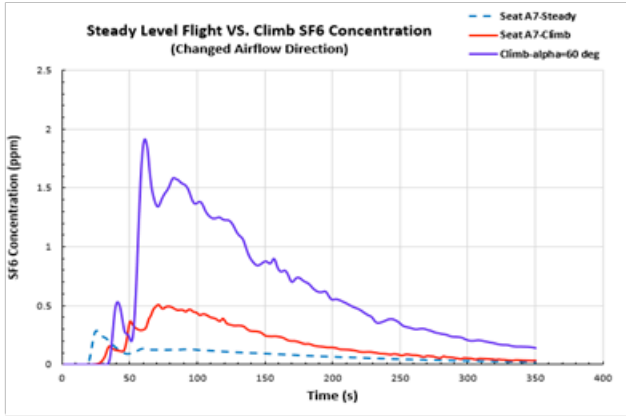
(b)

Figure 5. Comparison of the predicted SF_6 concentration time series between the steady level flight and climb legs using the default, 20° , and 30° airflow supply angles during the climb; (a) at seat A7, and (b) at seat C7

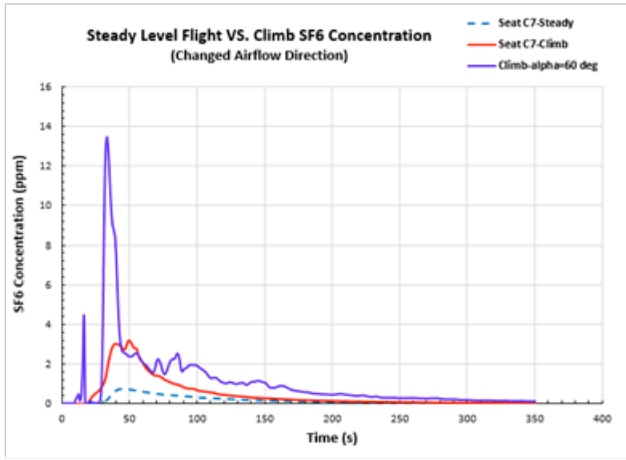
b) Changing angles α and β simultaneously:

In this alternative airflow redirection approach, angle β is simultaneously changed with angle α to add a three-dimensional perspective to this investigation. As $\alpha = 30^\circ$ provided the best cabin air quality relative to the other two airflow supply angles in 2-D, $\alpha = 30^\circ$ will be used again here with β also chosen to be equal to 30° with supply airflow either directed to the front or to the back of the cabin. This was performed to provide a clear comparison between those two scenarios while limiting the number of simulations required. Fig. 8 depicts the SF_6 concentration time series at the two monitoring locations using $\beta = 30^\circ$ to the front and to the back.

Comparing the SF_6 concentration time series at the two locations from Fig. 8, the $\beta = 30^\circ$ airflow supply to the back of the cabin could consistently reduce the time-integrated passenger exposure from the original climb case either at seat A7 or seat C7. More specifically, at seat C7, the exposure was reduced to a level close to that for the steady level flight condition. Conversely, in case of the $\beta = 30^\circ$ airflow supply to the front, the passenger exposure was higher



(a)

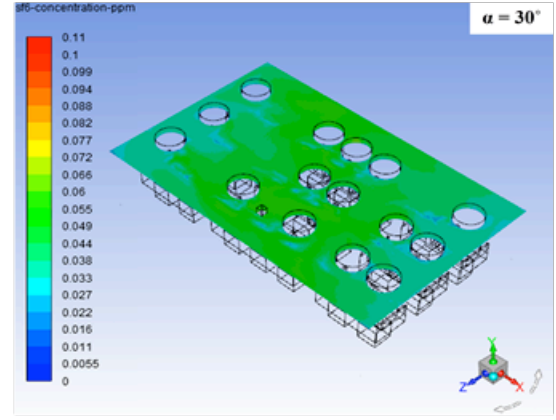


(b)

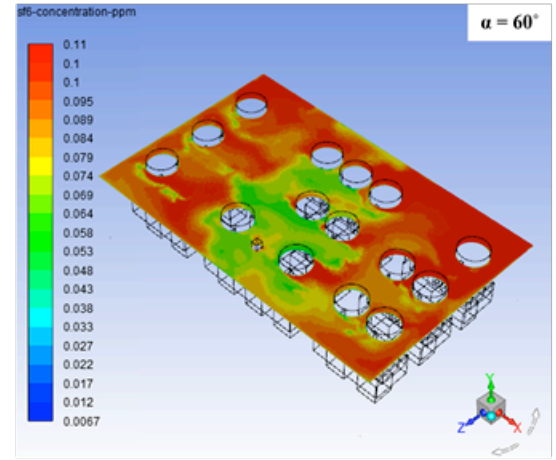
Figure 6. Comparison of the predicted SF_6 concentration time series between the steady level flight and climb legs using the default and 60° airflow supply angles during climb; (a) at seat A7, and (b) at seat C7

at seat A7 than the original case, but the exposure was almost halved for the same scenario at seat C7. Such complex response in concentration time series can be understood when the SF_6 contour plots for the mentioned two airflow supply cases at 350 s are compared as shown in Fig. 9.

As noticed from the figure, directing airflow to the front of the cabin pushes the contaminant eventually to the rear of the cabin at the end of simulation time (350 s). Before this happens, however, the air moves most of the contaminant to the frontal rows for a short period of time (70-80 s) after contaminant release in the cabin. This explains the very high contaminant concentration at seat A7 around this time. On the other hand, supplying airflow to the back of the cabin leads to steadily pushing the contaminant to the front of the cabin and providing appropriate dilution of it in the cabin air with no major dispersion patterns in the back rows. This can be attributed to the overall airflow direction being in the same direction of the cough flow in this case. The contaminant dilution provided by the back-directed airflow makes the overall passenger exposure to the contaminant consistently low at most cabin seats as previously seen in Fig. 9.



(a)



(b)

Figure 7. SF_6 concentration contour plots at passenger breathing level during climb at 350 s; (a) using airflow supply angle $\alpha = 30^\circ$, and (b) using airflow supply angle $\alpha = 60^\circ$

B. Source Control Strategies

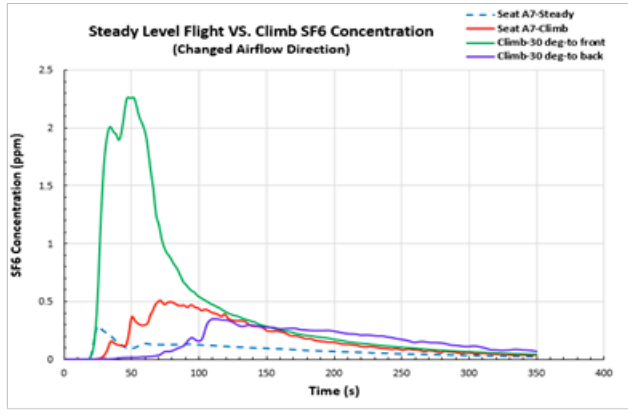
Cougher location in the cabin:

In this investigation, the location of the cougher was changed twice from the back row in the centre to the central row on the left side (LC), and to the front row on the right side (RF) (looking from the back of the cabin to the front). The new cougher locations with respect to the original case are shown in Fig. 10.

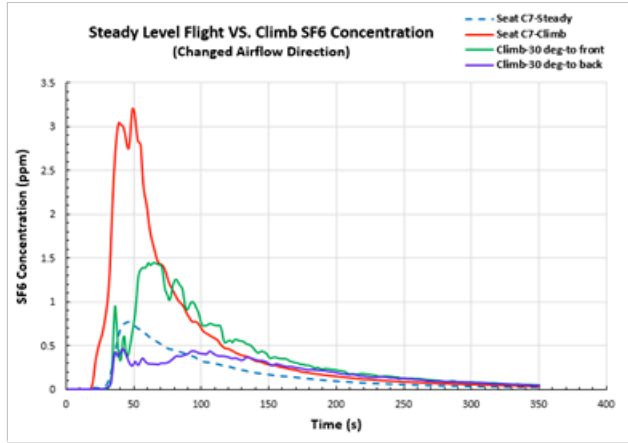
Figs. 11 and 12 depict the SF_6 concentration time series at the two seats A7 and C7 with the cougher positions at locations LC and RF, respectively. The series for the two cases were separated to enhance their readability.

Based on the concentration time series, the contaminant concentration at each monitoring point is significantly affected by the cougher location. For example, as the cougher moves to the LC location, the concentration at the two monitors for the full simulation time was significantly reduced to a nearly similar level. This is because the cougher at the LC location is almost equally distanced from the two monitoring points.

Conversely, as the cougher set at the RF location is very



(a)



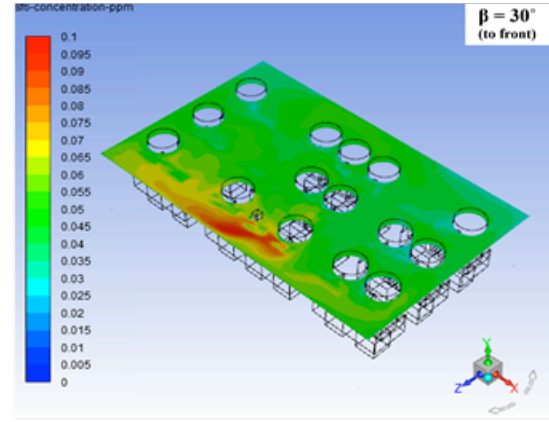
(b)

Figure 8. Comparison of the predicted SF₆ concentration time series between the steady level flight and climb legs using supply angles $\alpha = 30^\circ$ and $\beta = 30^\circ$ (to front and back) during climb; (a) at seat A7, and (b) at seat C7

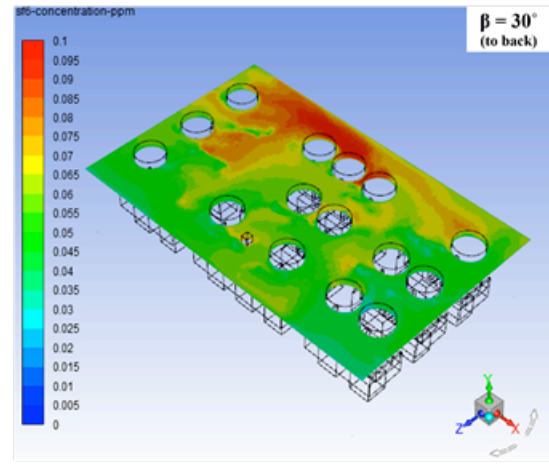
close to the monitoring point at seat A7, the contaminant concentration increased substantially to about eight-fold its maximum value during the baseline climb case as shown in Fig. 12a. On the other hand, at seat C7, the contaminant time-averaged concentration was much lower in the RF cougher location scenario than the baseline case because the cougher was moved further away from it.

In addition to the effect of cougher proximity, the ventilation airflow patterns in the cabin and/or the existence of walls or surfaces close to the cougher location have significant influence on the dispersion behavior of the released contaminant in the cabin by impacting or redirecting the cough. Such influence can be inferred from the SF₆ concentration contour plots for the two cougher location cases illustrated in Fig. 13.

Looking at the contour plot for the LC cougher location, the contaminant is noticed to reside at the back of the cabin at the end of simulation. This is impacted by the airflow in the cabin and the body forces during aircraft climb, which push the contaminant to the back rows from the front of the cabin on the left side where it initially dispersed after release. On the



(a)



(b)

Figure 9. SF₆ concentration contour plots at passenger breathing level during climb at 350 s; (a) using airflow supply angle $\beta = 30^\circ$ to the front, and (b) using airflow supply angle $\beta = 30^\circ$ to the back

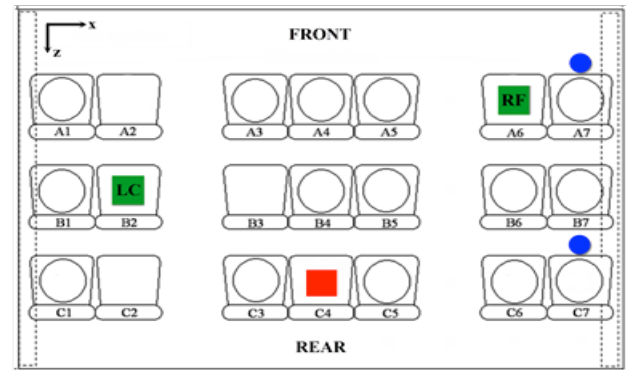
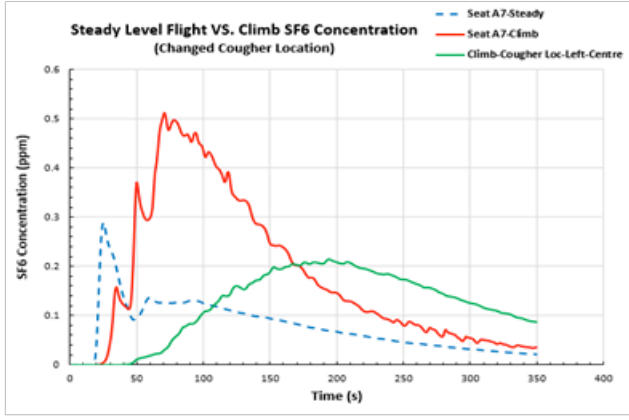
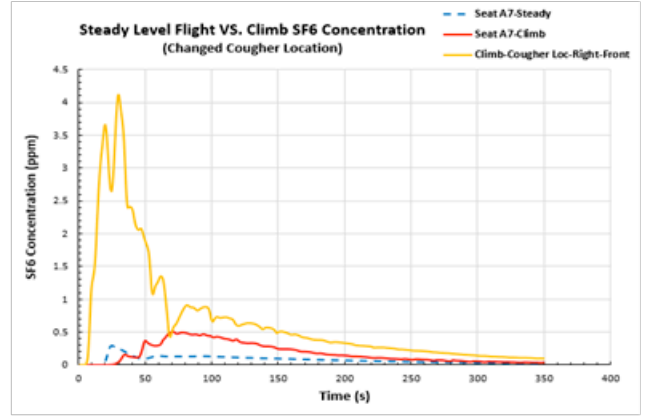


Figure 10. New cougher locations with respect to the original location (red square). First location is on the left side at centre row (LC), and the second is on the right side at the front row (RF)

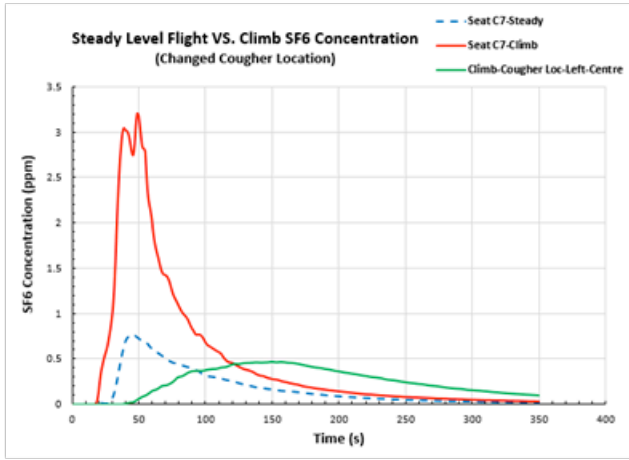
other hand, the contaminant released from the cougher in the RF position follows along the frontal cabin separator to the left before it disperses to the back rows by the help of airflow mix-



(a)

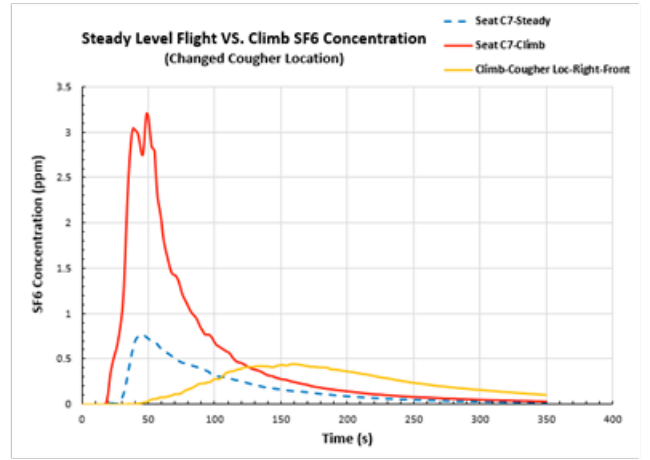


(a)



(b)

Figure 11. Comparison of the predicted SF_6 concentration time series between the steady level flight and climb legs with the cougher moved to the centre row on the left side of the cabin during climb; (a) at seat A7, and (b) at seat C7



(b)

Figure 12. Comparison of the predicted SF_6 concentration time series between the steady level flight and climb legs with the cougher moved to the front row on the right side of the cabin during climb; (a) at seat A7, and (b) at seat C7

ing. This leads to higher concentration regions in the central section of the cabin and more exposure of the occupants to the contaminant. According to the results, a coughing person aboard an aircraft may not only cause higher probability of exposure to the passengers close to him/her, but also to most occupants if this person is coughing in proximity to a wall or surface.

Finally, Table III lists the passenger exposure ratio between the various cases studied in the current paper and the baseline climb case. The lowest exposure ratios are presented in bold font.

The exposure values were calculated using equation 4 (previously presented) and applying the composite Simpson's and the composite trapezoidal rules in determining the area under the curve for each concentration time series.

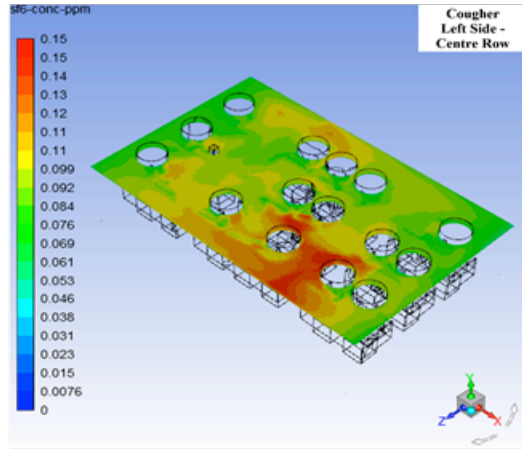
From Table III, the lowest average exposure ratio between the two seats is found for the left-side, centre-row relocation of the cougher case. The case that come after is the $\alpha = 30^\circ$ airflow direction case followed by the $\alpha = 30^\circ$ and $\beta = 30^\circ$ (to the back) airflow direction case. Such airflow design and/or s-

ource control strategies could be implemented to reduce the exposure of the aircraft occupants to expiratory contaminants released in this aircraft cabin, especially during the climb leg.

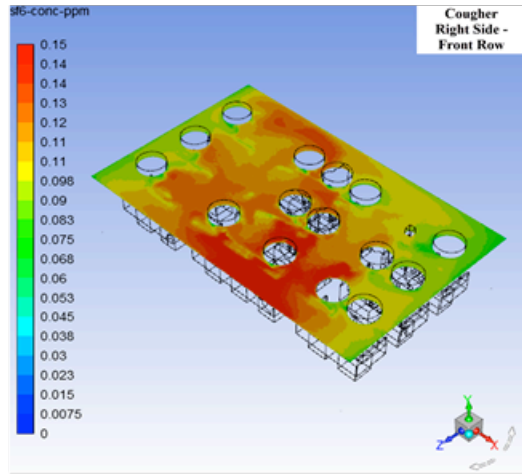
Nevertheless, the highest exposure ratio is noticed for the $\alpha = 60^\circ$ airflow direction scenario. This is followed by the occurrence in which the cougher is moved to the front row on the right side of the cabin, and later by setting $\alpha = 30^\circ$ and $\beta = 30^\circ$ to the front as airflow direction condition.

IV. CONCLUSIONS

In the current study, airflow design and source control strategies were investigated for their potential for reducing cough-released airborne contaminant exposure in the cabin of a passenger aircraft (Boeing 767-300) during the climb leg. Sulfur Hexafluoride (SF_6) was used to mimic the airborne contaminant in the cabin, representing cough-released particles in the size range from 1.6 to 3 μm in diameter. The SF_6 dispersion behavior was analyzed by calculating the concentration time series at two monitoring locations, seats A7 and C7 in the cabin, in addition to SF_6 concentration contour p-



(a)



(b)

Figure 13. SF_6 concentration contour plots at passenger breathing level during climb at 350 s; with (a) the cougher located at the centre row on the left side of the cabin, and (b) the cougher located at the front row on the right side of the cabin

lots at the breathing level of the occupants. The concentration time series were further used to infer the passenger exposure to the contaminant through determining the area under each curve.

In this paper, the airflow design strategies researched involved altering the supply airflow direction, while the source control strategies involved moving the cougher to different locations in the cabin.

Changing the airflow supply angle from the ceiling only, or α , from 20° to 60° had different effects on the SF_6 dispersion behavior in the cabin. The angle $\alpha = 30^\circ$ led to the lowest passenger exposure. However, the 20° and 60° angles, compared to passenger exposure of the baseline climb case with $\alpha = 0^\circ$, resulted in similar and higher exposures, respectively. Further, when a three-dimensional approach was introduced to the airflow redirection scenario through the angle β , directing the airflow to the back of the cabin proved to be better than directing it to the front.

TABLE III.
RATIO OF PASSENGER EXPOSURE BETWEEN DIFFERENT PARAMETRIC SENSITIVITY CASES AND THE BASELINE CLIMB CASE AT THE TWO MONITORING LOCATIONS

Case	Passenger Exposure Ratio to Baseline Climb Case	
	Seat A7	Seat C7
$\alpha = 20^\circ$ Airflow	1.1 : 1	0.9 : 1
$\alpha = 30^\circ$ Airflow	0.7 : 1	0.5 : 1
$\alpha = 60^\circ$ Airflow	3.5 : 1	2.1 : 1
$\alpha = 30^\circ$ and $\beta = 30^\circ$ to Front	2.4 : 1	0.7 : 1
$\alpha = 30^\circ$ and $\beta = 30^\circ$ to Back	0.9 : 1	0.4 : 1
Cougher at Left Side-Centre Row	0.7 : 1	0.5 : 1
Cougher at Right Side-Front Row	3.9 : 1	0.5 : 1

For the source control strategies, relocating the cougher to other locations in the cabin other than the original back-row-centre-seat position had a quantifiable effect on the dispersion behavior of the contaminant, and consequently, the exposure.

Moving the cougher to the left side of the cabin in the centre row (LC location) led to decreased passenger exposure both locally at the monitoring locations, and as an average in the whole cabin. On the other hand, moving the cougher to the right side of the cabin in the front row (RF location) caused a substantial increase in the exposure at seat A7, as it is very close to it, while it halved the exposure at seat C7. In addition to the proximity of the cougher to specific passengers, the airflow patterns in the cabin, the body forces on the aircraft during climb, and the existence of walls and/or surfaces near the cougher all have confounding effects on the resulted contaminant dispersion behavior from different cougher locations.

Generally, the cases that showed most promising reduction in passenger exposure as an average between the two monitoring locations at seats A7 and C7 with respect to the baseline climb case were: the left-side, centre-row relocation of the cougher, the $\alpha = 30^\circ$ airflow direction, and the $\alpha = 30^\circ$ and $\beta = 30^\circ$ (to the back) airflow direction cases. The exposure ratios are 0.7:1 at seat A7 and 0.5:1 at seat C7 for the first case, 0.7:1 at seat A7 and 0.5:1 at seat C7 for the second case, and 0.9:1 at seat A7 and 0.4:1 at seat C7 for the third case, respectively. On the other hand, the highest exposure in average between the two seats belongs to the $\alpha = 60^\circ$ airflow direction case with 3.5:1 at seat A7 and 2.1:1 at seat C7.

For future work, and to generalize the findings of the current study, similar parametric analysis needs to be implemented on other models of passenger aircraft with different cabin configurations. Moreover, different ventilation strategies, other than the conventional mixing ventilation used

in this study, such as underfloor and personalized ventilation systems can be implemented. Finally, further combinations and/or additions to the proposed airflow design and source control strategies in the current work can be investigated for possible enhancements in the in-cabin air quality. This work warrants the need for multiple detailed investigations related to the influence of aircraft acceleration-induced body forces on ventilation performance of aircraft, an issue that has been neglected in the literature for a long time.

ACKNOWLEDGMENT

The authors thank the Government of Ontario, Canada, for providing the funding for this work in the form of the Ontario Trillium Scholarship (OTS) for the lead author. The authors also thank Natural Sciences and Engineering Research Council (NSERC) of Canada, specifically the Discovery Grant program (Grant Number 401231) supporting the lead author knowledge dissemination in conferences and journals. They also thank Dr. Christopher Y. H. Chao (Department of Mechanical Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China), and Dr. Man Pun Wan (Department of Mechanical & Aerospace Engineering, Nanyang Technological University, Singapore) for providing their full experimental research data. Additionally, the authors recognize CMC Microsystems for providing the Academic Research license of ANSYS. Finally, they thank the Centre for Advanced Computing (CAC) at Queen's University (Kingston, ON, Canada) for allowing access to the Frontenac Compute Cluster, and SHARCNET for providing the Graham cluster to run the simulations.

REFERENCES

- [1] H. A. Elmaghraby, Y. W. Chiang, and A. A. Aliabadi, "Are Aircraft Acceleration-induced Body Forces Effective on Contaminant Dispersion in Passenger Aircraft Cabins?," *Sci. Technol. Built Environ.*, accepted, 2019.
- [2] J. W. Drake and D. E. Johnson, "Measurements of certain environmental tobacco smoke components on long-range flights," *Aviat. Space. Environ. Med.*, vol. 61, no. 6, pp. 531–542, 1990.
- [3] M. Dechow, H. Sohn, and J. Steinhilber, "Concentrations of Selected Contaminants in Cabin Air of Airbus Aircrafts," *Chemosphere*, vol. 35, pp. 21–31, 1997.
- [4] F. Li, J. Liu, J. Ren, X. Cao, and Y. Zhu, "Numerical investigation of airborne contaminant transport under different vortex structures in the aircraft cabin," *Int. J. Heat Mass Transf.*, vol. 96, pp. 287–295, 2016.
- [5] L. Yang, X. Li, Y. Yan, and J. Tu, "Effects of cough-jet on airflow and contaminant transport in an airliner cabin section," *J. Comput. Multiph. Flows*, vol. 10, no. 2, pp. 72–82, 2018.
- [6] F. Haghighat, F. Allard, A. C. Megri, P. Blondeau, and R. Shimotakahara, "Measurement of thermal comfort and indoor air quality aboard 43 flights on commercial airlines," *Indoor Built Environ.*, vol. 8, no. 1, pp. 58–66, 1999.
- [7] M. B. Hocking, "Passenger aircraft cabin air quality: Trends, effects, societal costs, proposals," *Chemosphere*, vol. 41, no. 4, pp. 603–615, 2000.
- [8] N. L. Nagda and M. Hodgson, "Low Relative Humidity and Aircraft Cabin Air Quality," *Indoor Air*, vol. 11, no. 3, pp. 200–214, 2001.
- [9] M. A. Waters, T. F. Bloom, B. Grajewski, and J. Deddens, "Measurements of Indoor Air Quality on Commercial Transport Aircraft," in *The 9th International Conference on Indoor Air Quality and Climate*, 2002, pp. 782–787.
- [10] C.-H.- Lin, K. H. Dunn, R. H. Horstman, J. L. Topmiller, M. F. Ahlers, J. S. Bennett, L. M. Sedgwick, and S. Wirogo, "Numerical Simulation of Airflow and Airborne Pathogen Transport in Aircraft Cabins--Part I: Numerical Simulation of the Flow Field," *ASHRAE Trans.*, vol. 111, no. 1, pp. 755–764, 2005.
- [11] Z. Zhang, X. Chen, S. Mazumdar, T. Zhang, and Q. Chen, "Experimental and numerical investigation of airflow and contaminant transport in an airliner cabin mockup," *Build. Environ.*, vol. 44, no. 1, pp. 85–94, 2009.
- [12] S. B. Poussou, S. Mazumdar, M. W. Plesniak, P. E. Sojka, and Q. Chen, "Flow and contaminant transport in an airliner cabin induced by a moving body: Model experiments and CFD predictions," *Atmos. Environ.*, vol. 44, no. 24, pp. 2830–2839, 2010.
- [13] S. S. Isukapalli, S. Mazumdar, P. George, B. Wei, B. Jones, and C. P. Weisel, "Computational fluid dynamics modeling of transport and deposition of pesticides in an aircraft cabin," *Atmos. Environ.*, vol. 68, pp. 198–207, 2013.
- [14] A. A. Aliabadi, S. N. Rogak, K. H. Bartlett, and S. I. Green, "Preventing airborne disease transmission: review of methods for ventilation design in health care facilities," *Adv. Prev. Med.*, vol. 2011, pp. 1–21, 2011.
- [15] ASHRAE, "Standard 161-2013: Air Quality within Commercial Aircraft," American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, USA, 2013.
- [16] H. A. Elmaghraby, Y. W. Chiang, and A. A. Aliabadi, "Ventilation strategies and air quality management in passenger aircraft cabins: A review of experimental approaches and numerical simulations," *Sci. Technol. Built Environ.*, vol. 24, no. 2, pp. 160–175, 2018.
- [17] D. G. Hull, *Fundamentals of Airplane Flight Mechanics*. New York: Springer Berlin Heidelberg, 2007.
- [18] G. N. Sze To, M. P. Wan, C. Y. H. Chao, L. Fang, and A. Melikov, "Experimental Study of Dispersion and Deposition of Expiratory Aerosols in Aircraft Cabins and Impact on Infectious Disease Transmission," *Aerosol Sci. Technol.*, vol. 43, no. 5, pp. 466–485, 2009.
- [19] M. P. Wan, G. N. Sze To, C. Y. H. Chao, L. Fang, and A. Melikov, "Modeling the Fate of Expiratory Aerosols and the Associated Infection Risk in an Aircraft Cabin Environment," *Aerosol Sci. Technol.*, vol. 43, no. 4, pp. 322–343, 2009.
- [20] P. Strøm-Tejsten, D. P. Wyon, L. Lagercrantz, and L. Fang, "Passenger evaluation of the optimum balance between fresh air supply and humidity from 7-h exposures in a simulated aircraft cabin," *Indoor Air*, vol. 17, no. 2, pp. 92–108, 2007.

NUMERICAL INVESTIGATIONS OF A COMB-LIKE FILM-COOLING SCHEME

Hao-Ming Li, Wahid Ghaly
 MIE, Concordia University
 Montreal, QC

Ibrahim Hassan
 MIE, Texas A&M University at Qatar
 Dohar, Qatar

Abstract—A new cooling scheme with both perfect performance and practical structural integrity was developed and investigated. The slot has perfect performance but without structural integrity. Based on the critical effect of the counter-rotating vortex pair (CRVP) on the film cooling effectiveness (η), a new cooling scheme named Comb scheme was developed. It is composed of a comb-like structure and a slot-like exit. 14 configurations of the Comb scheme have been numerically investigated. The steady Reynolds-Averaged Navier-Stokes simulation was performed, and the realizable $k-\varepsilon$ model was selected as the turbulence model. The simulation results were validated with the published data and the measurements. The new scheme displayed high performance carrying comparison with the slot. The geometrical parameters have been analyzed in terms of the η and the CRVP intensity, respectively. Not only were the effects of the geometrical parameters demonstrated, but the relations of the η and the CRVP intensity have also been revealed. Two typical Comb configurations were further analyzed. It was demonstrated that their CRVP occurred on the bottom of the blind slot, and was buried in the blind slot, led to the perfect performance. The success of the Comb scheme has pointed out a direction to develop the advanced film cooling schemes, and in turn proven the critical effect of the CRVP.

I. INTRODUCTION

Cooling technologies have been essential in the advanced gas turbine, in which film cooling is the mainstay. For film cooling geometries, discovering the superior performance of the shaped hole in several decades ago by Goldstein *et al.* [1] was a revolutionary progress. After that, even tremendous and sustained efforts have been made in the past decades, no comparable improvement with this breakthrough has been achieved [2]. The clarification on the complex fundamentals of the film cooling heat transfer has become a much greater urgency.

Bogard and Thole [3] summarized the factors affecting the film cooling effectiveness (η). Many flow conditions affect η , including blowing ratio (BR), density ratio (DR), turbulence, and so on. The most prevailing opinion is the momentum flux ratio (I) is the most crucial factor on η [1, 3, 4]. Sinha *et al.* [4] classified the film cooling flow into attached, detached and reattached flows. They were controlled by the I . Goldstein *et al.* [1] thought the improved performance of the shaped holes was due to the expanded exit reducing the jet velocity. However, this opinion is contradictory with some research results. For example, Gritsch *et al.* [5] demonstrated that the area ratio had no significant effect on η .

Models were built to predict η with pertaining parameters, in which the crucial factors have to be included. The correlation of Baldauf *et al.* [6] included the parameters of DR , hole pitch (P),

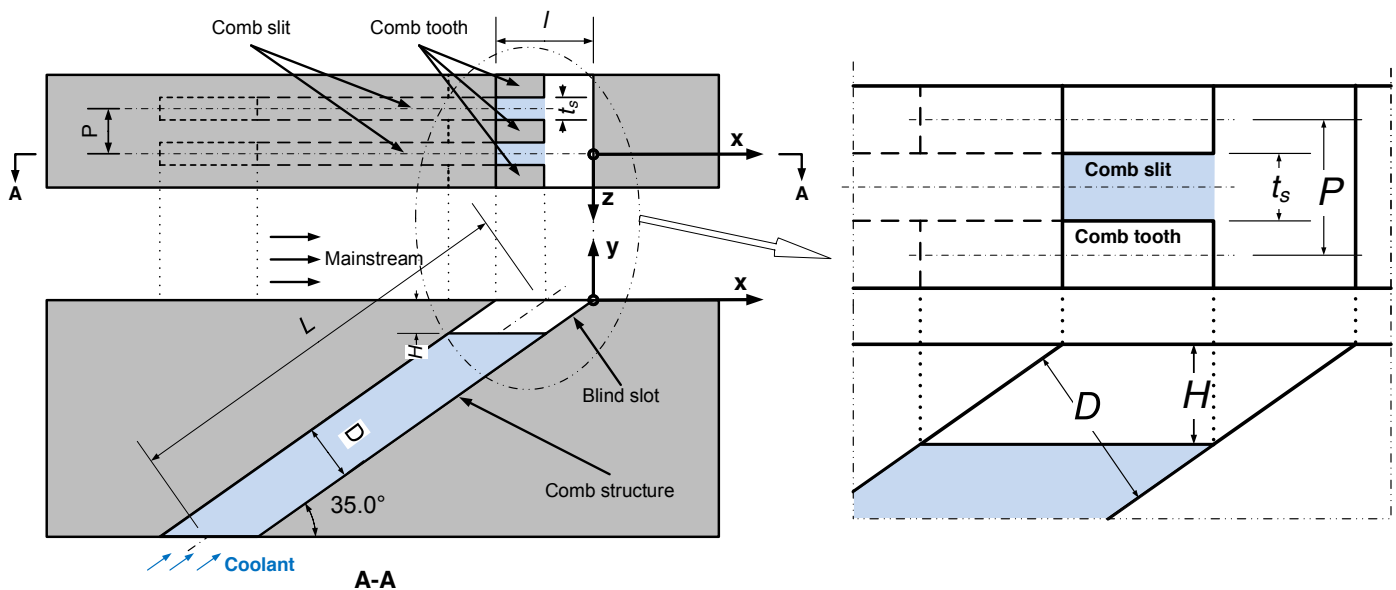


Figure 1 The Comb scheme

II. METHODOLOGY AND MODELS

A new film cooling scheme named Comb scheme is developed. It is investigated with steady RANS simulations.

A. The Structure of the New Scheme

Figure 1 shows the structure of the new scheme. It has two components. The exit is a blind slot, which is fed by a comb-like structure. In the comb-like structure, the coolant comes out through the comb slits; and the comb teeth maintain the structural integrity. Because of the comb-like structure, the new scheme is named Comb scheme.

The inclined angle of the Comb scheme is 35 degree, shown in the left side of Figure 1. The width of the blind slot is $D = 12.7$ mm, which is the characteristic length in the present investigations. The axial length of the configurations is $L = 7D$. To show the geometrical parameters, the exit is enlarged into the right side of the Figure 1. t_s is the width of the comb slit, and P is the lateral space between two comb slits. The vertical depth of the blind slot is H . The structural integrity value is determined by the equation of

$$MCSA = \left(1 - \frac{t_s}{P}\right) * \left(1 - \frac{H}{L * \sin 35^\circ}\right) \quad (1)$$

The definition of the Cartesian coordinates is also shown in the figure. The origin is on the trailing edge of the exit and the central plane of a comb slit. The x, y and z axes are aligned with the mainstream, vertical and lateral directions, respectively.

B. CFD Setup and Experimental Validation

The simulation domain is shown in Figure 2. The simulated tunnel length is $80D$, including the $60D$ of the downstream length. As a steady simulation, the domain includes a half element, its width is $0.5P$. Two lateral sides were set to symmetry planes. Therefore, the present simulation did not account of the end effect. The domain had two velocity inlets. One was the wind tunnel inlet, where the velocity was set at 20 m/s and temperature was 300 K. Another was the plenum inlet, where the temperature is 150 K, and the velocity varied on the blowing ratio. The simulated DR was 2. A pressure outlet was set at the wind tunnel exit. The wind tunnel top was set to free slip and adiabatic. Then all remaining walls were set as adiabatic and no slip.

Zhang and Hassan [21] compared several steady turbulence models. The realizable $k-\varepsilon$ model (RKE) model was suggested for this kind of simulations. So the RKE model was used in the present paper. The FLUENT packaged in ANSYS was used as the CFD solver.

The meshes were generated with the ICEM-CFD packaged in ANSYS. A Comb configuration was chosen to test the grid independence, in which H is $1D$, t_s is $0.1D$ and P is $2t_s$. Five grids, which includes approximately 0.70 to 1.48 million elements, were created for the configuration. The results of the film-cooling centerline effectiveness (η_c) indicated slight changes after grid size larger than 0.81 million. The grid of 1.27M cells was finally selected, to compromise the computing costs and the capturing of the vortical structures. The mesh generations for all the rest Comb configurations complied with the identical criterion.

The Comb scheme has also been investigated experimentally. It has been report in Li et al. [22], in which the η of two classic geometries and the Comb scheme have been measured. The results are used in the present paper to validate the simulations.

The simulating test matrix is shown in Table 1. Cases 1 to 8 were compared with the open published data and the

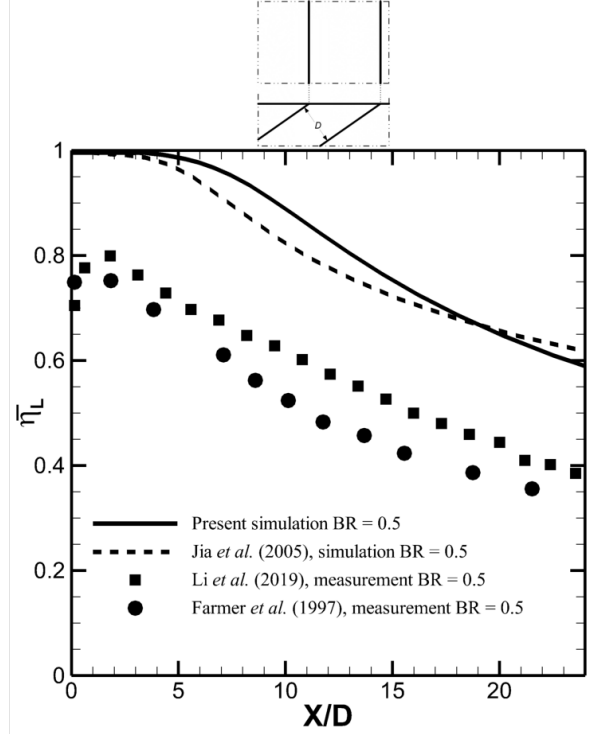


Figure 3 Comparisons of slot at $BR = 0.5$

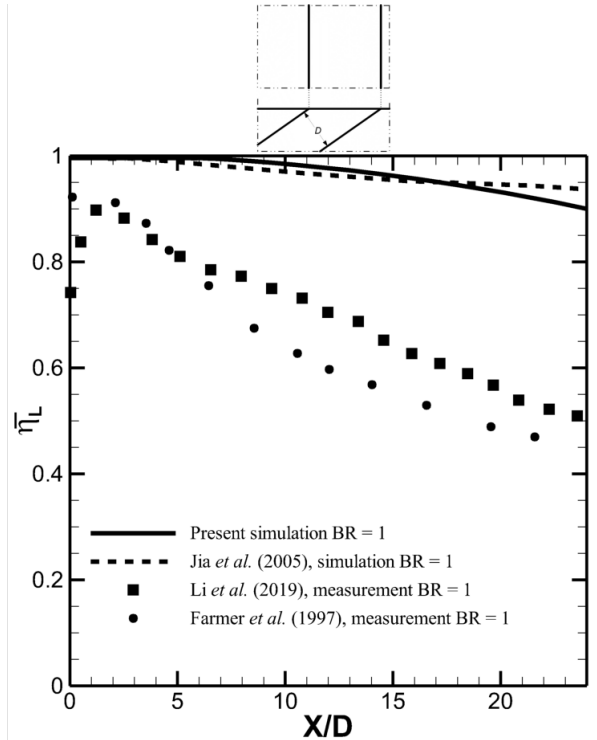


Figure 4 Comparisons of slot at $BR = 1$

measurements for validations. Cases 4 to 20 are the Comb configurations exhibited in the present investigations.

C. Parameter Normalization

The present investigation normalized two parameters. The normalized temperature, θ , of which the definition is similar to η . However, θ is valid in the whole domain, instead of the test plate only.

$$\theta = \frac{T - T_m}{T_j - T_m} \quad (2)$$

According to the analysis in [20], the CRVP intensity is mostly x direction. And ω_x is normalized as:

$$\omega_x = \frac{\omega_x^*}{U_m/D} \quad (3).$$

Here $U_m = 20 \text{ m/s}$, is the mainstream velocity, and $D = 12.7 \text{ mm}$.

III. RESULTS AND DISCUSSIONS

The simulation results are compared with published data and experimental results for validations. The performance of the new scheme is also presented. Then the main parameters of the new scheme are analyzed. Details of two typical Comb configurations are looked insight, to demonstrate the mechanism underlying the film-cooling performance.

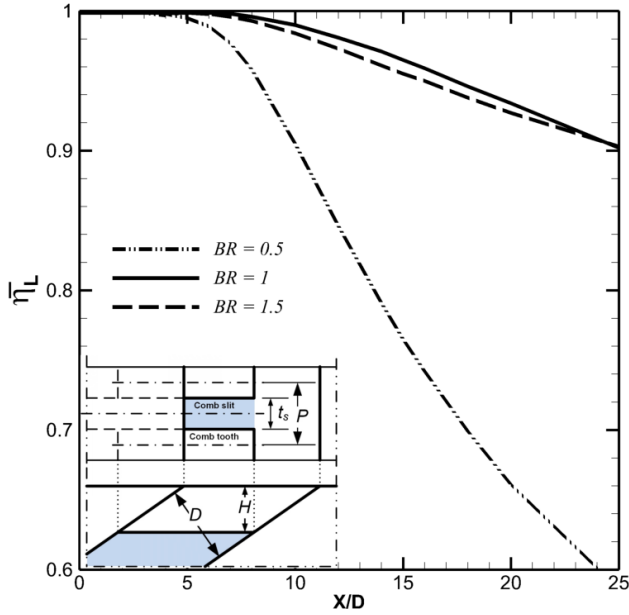


Figure 6 BR effect at $H = 1D$, $t_s = 0.1D$ and $P = 2t_s$

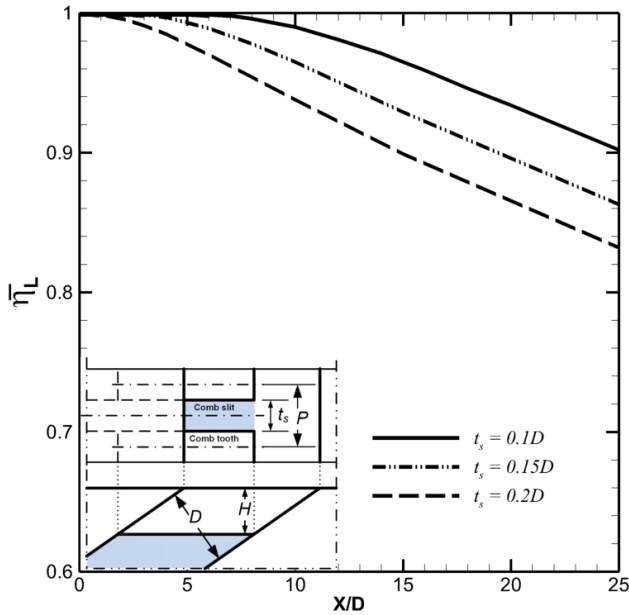
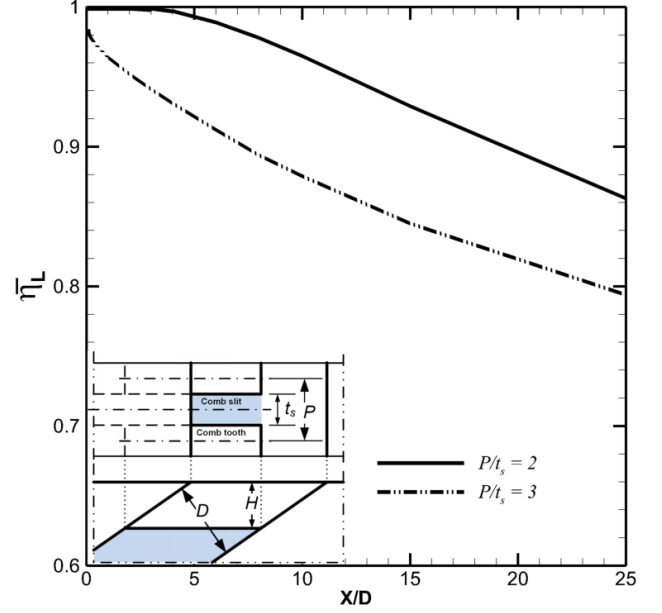
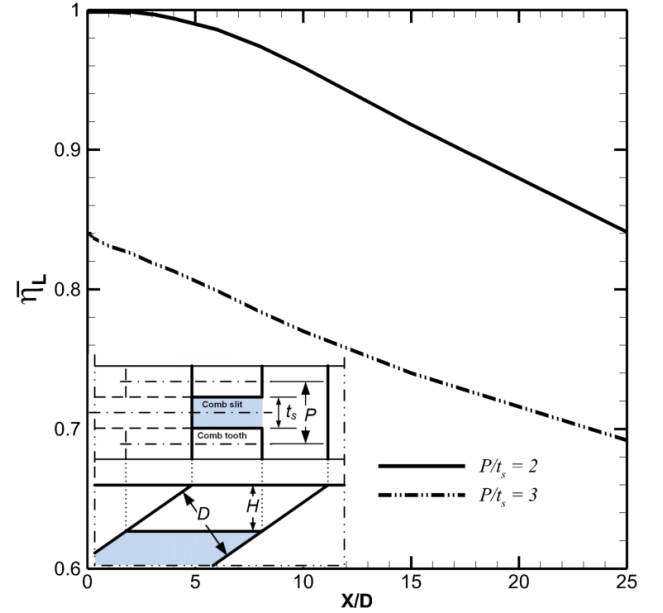


Figure 7 t_s effect at $H = 1D$ and $P = 2t_s$, for $BR = 1$



a) Configurations of $H = 1D$, $t_s = 0.15D$ and $P = 2t_s$ and $3t_s$



b) Configurations of $H = 0.5D$, $t_s = 0.15D$ and $P = 2t_s$ and $3t_s$

Figure 5 P/t_s effect at $BR = 1$

A. Validations of the Simulation

In addition to the Comb scheme, the classic geometries, the cylindrical hole and the simple slot, were simulated and compared with the published data. The data of the cylindrical hole were compared with the simulation in [21], and the measurements in [22, 23]. The two measurements agree well. The two simulations have acceptable agreement. Generally, the simulation is considered to qualitatively agree with the measurements.

The comparisons of the simple slot are shown in Figure 3 and Figure 4 for $BR = 0.5$ and 1, respectively. Two simulations and two measurements are shown in each figure. The present simulation is comparing with the simulation in Jia *et al.* [24], in which the inclined angle was 30 degree, a little bit smaller than the current 35 degree. The present measurement in [22] is comparing with the measurement in Farmer *et al.* [25]. For $BR = 0.5$ in Figure 3, the two measurements have decent agreement, and the two simulations too. However, even the simulations reflect the tendency of the measurements well, the simulations do not capture the liftoff-reattachment, and over predict the measurements. Rather than the DR difference in the cylindrical hole, it is mostly because of the end effect. The boundary conditions in the simulations assumed the infinite lateral dimension, did not account of the end effect in the experiments. Generally, the simulations qualitatively agree with the measurements.

Figure 4, which is for $BR = 1$, is similar to Figure 3. The agreements of the two simulations, as well as the two measurements, are decent. Mostly because of the end effect, the simulations miss the liftoff-reattachment, and over predict the measurements. The simulations qualitatively agree with measurements.

The comparisons of the classic geometries have demonstrated that the present simulations and the present measurements have decent agreements with their corresponding published data, and the simulations have qualitatively predicted the measurements. The present simulation is shown to be reliable.

B. BR Effect of the Comb Scheme

Three blowing ratios were simulated for the configuration of $H = 1D$, $t_s = 0.1D$. The results is shown in Figure 5, where they

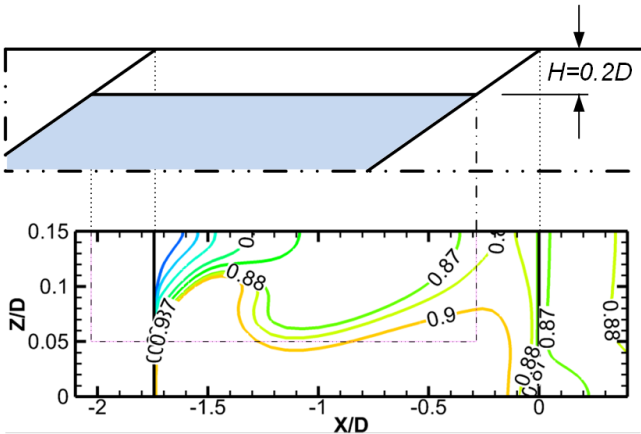


Figure 9 Temperatures around the exit for the configuration of $H = 0.2D$, $t_s = 0.1D$ and $P = 3t_s$, at $BR = 1$

are all very high, and start at 1. These performance is similar to the slot showing in Figure 3 and Figure 4. The so call ideal performance has been demonstrated for the new scheme. The $\bar{\eta}_L$ distributions for the $BR = 1.5$ and 1 are similar, and decrease slowly. But the $\bar{\eta}_L$ distributions of the $BR = 0.5$ decreases fast. Similar regularity occurs in other Comb configurations and the simple slot. As the main objective is about the ideal performance, the ensuing discussions are focused on the cases of $BR = 1$, except as otherwise noted.

C. Geometrical Parameter Analysis on the Performance

The Comb scheme has demonstrated both high performance and practical structural integrity. The effects of the geometrical parameters, t_s , P/t_s and H have been investigated. The t_s effect is showing in Figure 6 for the configurations of $H = 1D$, $t_s = 0.1-0.2D$, and $P = 2t_s$. It was investigated at three t_s values, $t_s = 0.1$,

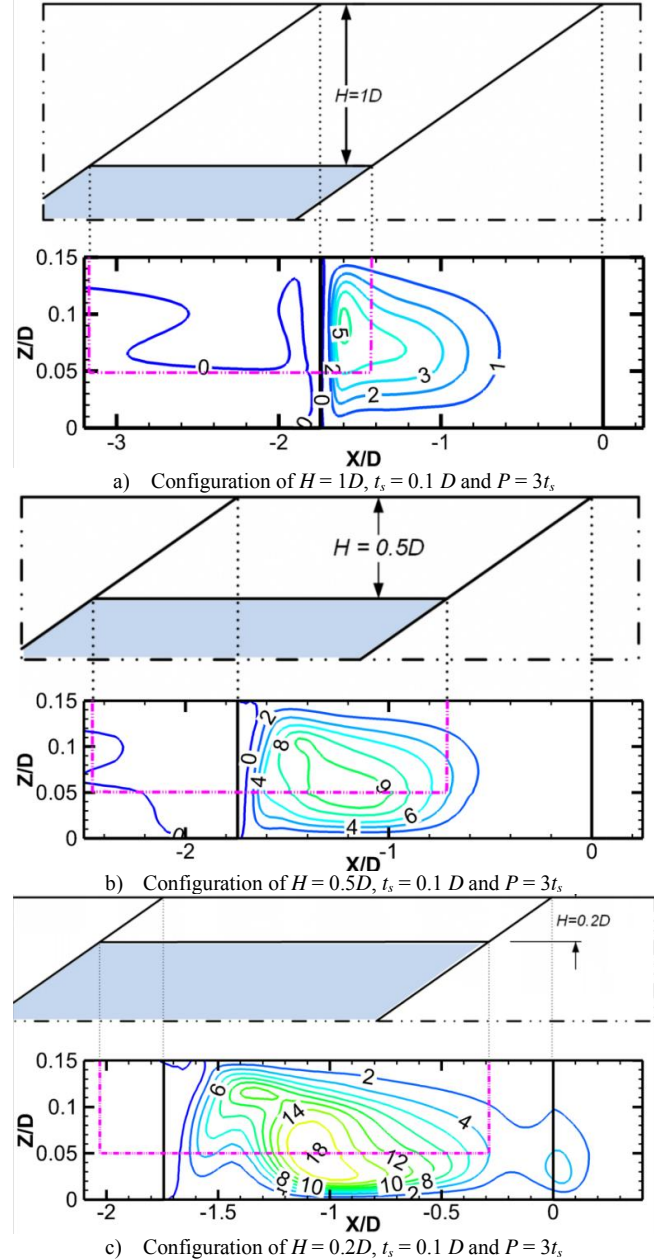


Figure 8 The H effect on ω_x for $H = 0.2-1D$, $t_s = 0.1D$, $P = 3t_s$ at $BR = 1$

0.15 and $0.2D$. The performance decreases when the t_s increases. All $\overline{\eta}_L$ start at 1, the $\overline{\eta}_L$ declines faster at bigger t_s , finally end at 0.9, 0.86 and 0.83, respectively.

The present investigations have investigated two P/t_s values, $P/t_s = 2$ and 3. The P/t_s effect is shown in Figure 7. Figure 7a is for the configurations of $H = 1D$, $t_s = 0.15D$, and $P/t_s = 2, 3$. At $P/t_s = 2$, the $\overline{\eta}_L$ starts at 1, and ends at 0.86. At $P/t_s = 3$, the $\overline{\eta}_L$ starts at less than 1, indicates the mainstream has penetrated the coolant film. Finally, the $\overline{\eta}_L$ ends at approximate 0.8. The similar regularity but stronger P/t_s effect is shown in Figure 7b, which is for the configurations of $H = 0.5D$, $t_s = 0.15D$, and P/t_s

$= 2, 3$. It is demonstrated that the cases of $P/t_s = 3$ has lower performance and stronger CRVP effect than the $P/t_s = 2$.

Three H values, $H = 0.2, 0.5$ and $1D$, were investigated, but the results were not presented here. Generally, the performance increases with the H increasing. This effect is weak in the range of $H = 0.5-1D$.

It has been demonstrated that, reducing t_s and P/t_s , increasing H are favor for the performance. However, according to the equation (1), which is for the MCSA, increasing H and reducing P/t_s decrease the MCSA. A compromise has to be reach. In terms of reducing t_s , it has to be feasible. A too small value will restrict the application.

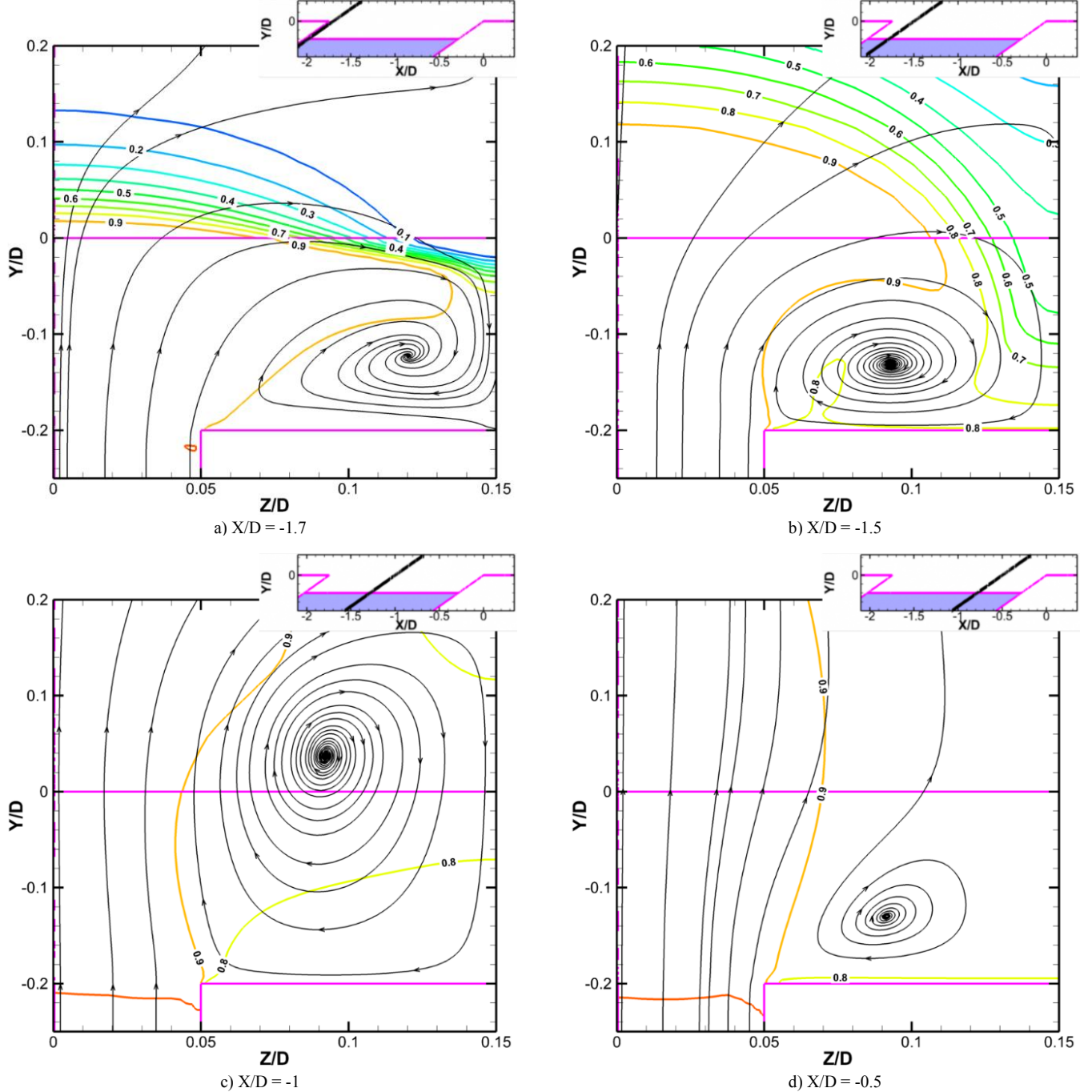


Figure 10 The CRVP for the configuration of $H = 0.2D$, $t_s = 0.1D$ and $P = 3t_s$ at $BR = 1$

D. Temperature Distributions Around Exit

The temperature distributions around the exit were visualized. Figure 8 is for the configuration of $H = 0.2D$, $t_s = 0.1D$ and $P = 3t_s$, which has typical low performance, and the $\overline{\eta}_L$ distributions have liftoff-reattachment. On the top of this figure is a sketch of the Comb scheme. In the contour chart, the leading edge and the trailing edge of the blind slot, and the edges of the comb tooth, are also shown on the background. This figure style is applied in all the figures showing contours around the exit. The contours on the exit show the mainstream is mixed into the coolant. The coolant is not continuous near the slot trailing edge. The sequence of the contour lines near the trailing edge is 0.9, 0.88, 0.87, slot trailing edge line, 0.87 and 0.88, denotes a small liftoff-reattachment.

In contrast, in the distributions of a high performance case, the θ contour lines of 0.9 and 0.1 were overlapping with the slot leading edge. The whole slot was full covered by the coolant, produces a very thick, continuous coolant film (Results did not present here). So the $\overline{\eta}_L$ distributions started at 1, did not have liftoff.

E. Geometrical Parameter Analysis on the CRVP Intensity

The design of the Comb scheme is intended to eliminate the CRVP effect on the exit. It has been pointed out that ω_x is the main component of the CRVP intensity in the cylindrical hole film cooling [14, 20], and it is also true in the present new scheme. The effect of the geometrical parameters on the $\overline{\eta}_L$ has been analyzed in the subsection III.C. Their effect on ω_x was also demonstrated by visualizing the ω_x around the exit.

Three ω_x contour charts are shown in Figure 9 for the configurations of $H = 0.2-1D$, $t_s = 0.1D$ and $P = 3t_s$. The figure style is the same as Figure 8. A sketch of the Comb scheme exit is shown on the top of the figure, and some edges are drawn on the background to indicate the positions. The most relevant place is on the mainstream/coolant interface. Its position is varied. The plane above the exit and at $Y/D = 0.05$ is considered to be very close to the interface, and was selected for visualizations. At Figure 9a, in which $H = 1D$, the local peak ω_x is 5.2, locates close to the exit leading edge, but not right above the comb-tooth edge. At Figure 9b, in which $H = 0.5D$, the local peak ω_x is 9.7, locates right above the comb-tooth edge. . At Figure 9c, in which $H = 0.2D$, the local peak ω_x is 18.4, locates just right above the comb-tooth edge. It is clear that the ω_x decreases as the H increases.

Table 2 CRVP intensity and the core positions

Nominal Position at Y=0	X/D=-1.7	X/D=-1.5	X/D=-1.0	X/D=-0.5
CRVP Core Position	0.08, -0.97	0.07, -0.95	0.07, -0.95	0.06, -0.99
Case 5				
ω_x	10	23	13	4
CRVP Core Position	0.12, -0.12	0.09, -0.13	0.09, 0.04	0.09, -0.13
Case 19				
ω_x	11	32	12	4

Similar visualizations demonstrated that increasing t_s or P/t_s increases ω_x . The results did not present here.

The geometrical parameter analysis demonstrated that, decreasing H , increasing t_s or P/t_s lead to the increasing ω_x . The comparisons of the geometrical parameter effects on ω_x and the effects on η , which has been presented in the subsection III.C, indicates the heat transfer mechanism underlying the high performance of the Comb scheme.

As the Comb scheme has a slot-like exit, its lateral coverage has already been 1. According to [14], η depends on the CRVP effect, or on other words, the relevant velocity gradients [20]. At the bottom of the blind slot, where $Y = -H$, large lateral velocity gradients occur when the coolant jets out from the comb-slit. As the coolant going up, it laterally expands, the velocity gradients gradually decrease. Usually, the mainstream/coolant interface is on a place higher than $Y = 0$. Therefore, larger H gives longer distance for the coolant to expand, or the velocity gradients to decrease, hence lower ω_x is on the interface, leads to higher uniformity and the high performance. On other words, increasing H is favor for the η . However, after the coolant reaches lateral uniformity, continuously increasing H has no significant effect, just like the $H = 0.5-1D$ in the aforementioned results.

In terms of reducing t_s or P/t_s , they decrease the lateral space for the CRVP formation, thus constricts the CRVP size and the intensity. So the CRVP effect decrease, leads to the high performance.

F. CRVP Intensity Around the Exit

As the link of the performance and the CRVP intensity has been demonstrated, CRVP details are worth being further investigated. The CRVPs of the cases with typical high and typical low performance have been visualized with streamlines on inclined cut planes, overlapping with the θ contours. The inclined angle is 35 degrees, parallel with the slot axial direction. The nominal position of the plane is their X value at $Y = 0$. For example, the cut plane of the nominal $X/D = -0.5$ is the plane in which $X = -0.5D$ when $Y = 0$.

The CRVP of the configuration of $H = 0.2D$, $t_s = 0.1D$ and $P = 3t_s$, which has the typical low performance, is visualized on the cut planes of $X/D = -1.7, -1.5, -1$ and -0.5 , and shown in Figure 10. As the X/D of the exit leading edge is -1.74 , the cut plane of $X/D = -1.7$ is very close to the leading edge. On the top right of the figures, a sketch indicates the cut plane position. On the bottom right of the figure, the edge of the comb tooth is shown on the background. Figure 10a and b show the two cut planes at $X/D = -1.7$ and -1.5 . They clearly demonstrate that the CRVP impacts the mainstream/coolant interface, and some mainstream is entrained into the coolant. The kidney shape is apparent, in particular the $\theta 0.9$ contour line in Figure 10b. Figure 10c and d show the blind slot is full of the mainstream/coolant mixture. The entrained mainstream mixing with the coolant causes the liftoff-reattachment shown in Figure 8.

In contrast, for a case with typical high performance, such as the configuration of $H = 1D$, $t_s = 0.1D$ and $P = 2t_s$, the CRVP cores are lower than $Y = -0.95D$. No CRVP impacts on the mainstream/coolant interface (Results did not present here). For

the comparison, the CRVP core positions of this two configurations, and in which the ω_x values, are listed in Table 2.

A characteristic of the new scheme is worth noting. Rather than the velocity gradients of the mainstream/coolant shear layer in the film cooling flow of the cylindrical hole, in the present geometry, the velocity gradients of the coolant are the source of the CRVP. It is driven by the relevant velocity differences between the coolant jetting out the comb slit and the coolant above the comb tooth. So the CRVP strongly links to the geometrical parameters like t_s and P/t_s . However, the present investigations choosing the comb structure is mostly aimed to simplify the simulation, because of the simple shape. Otherwise, any geometry provides structural integrity and high uniformity could be chosen.

The CRVP core positions of these two configurations were extracted. They are shown in Figure 11. The effect of the distance between the CRVP and the mainstream/coolant interface is demonstrated. The mainstream/coolant interface is usually higher than $Y = 0$. For the typical low performance configuration, its core positions of Y/D are ranged from -0.13 to 0.4, the CRVP is very likely to entrain the mainstream. For the typical high performance configuration, its core positions of Y/D are lower than -0.95, the CRVP is so far away from the mainstream/coolant interface that it is hard to impact the interface.

The comparison of the two configurations demonstrates that, instead of the mainstream/coolant interface, the CRVP occurs on the bottom of the blind slot, which was designed to be lower than $Y = 0$. Meanwhile, the mainstream/coolant interface is usually higher than $Y = 0$. Their distance eliminates or decreases the CRVP effect. On other words, the CRVP is buried by the blind slot in the Comb scheme.

With respect to the geometrical parameters, it has been revealed that, the distance between the CRVP and the mainstream/coolant interface is mostly linked to H . It is consistent with the aforementioned parameter analysis.

In film cooling heat transfer, advection is the main manner, and the diffusion can be ignored. The near field vortical structure like CRVP enhances the advection when the rotating velocity is perpendicular to the mainstream/coolant interface [14, 20]. In the previous researches on the CRVP effect, only the CRVP intensity was highlighted, because in which the CRVP is naturally close to the mainstream/coolant interface. However, in the Comb scheme, the CRVP effect is not inevitable any more. The special structure makes the CRVP away from the mainstream/coolant interface. Therefore, two factors have to be considered for the CRVP effect: intensity and position. First, the CRVP has to be strong enough to mix the mainstream and the coolant. And, the CRVP has to locate near the mainstream/coolant interface, then the rotating velocity can penetrate the interface.

IV. CONCLUSIONS

A new film cooling scheme named Comb scheme was developed based on the research results about the critical effect of the CRVP intensity. Its lateral coverage ratio was 1, and the CRVP intensity was negligible on the mainstream/coolant interface. The comb-like structure provided similar structural integrities to the current commercial using film cooling

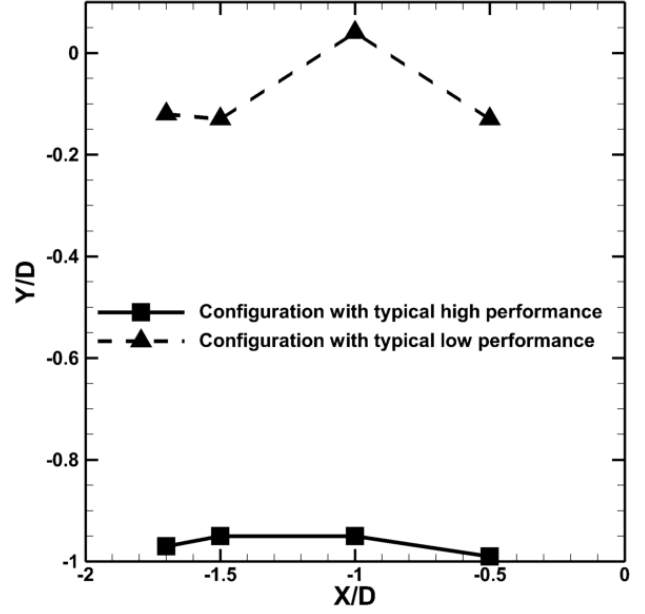


Figure 11 The CRVP core positions for the configuration of $H = 1D$, $t_s = 0.1D$ and $P = 2t_s$ (typical high performance) and the configuration of $H = 0.2D$, $t_s = 0.1D$ and $P = 3t_s$ (typical low performance) at $BR = 1$

geometries. The investigated results demonstrated the high performance of the new scheme, which is bearing comparison with the simple slot.

The investigations were performed with steady numerical simulations. The comparisons with the open published data and the experimental results have validated the simulations, indicated the reliability.

The effects of the main parameters of the Comb scheme, H , t_s , and P/t_s , have been analyzed on the η and the ω_x . The effects of higher H , lower t_s or P/t_s increasing the η are clear. The opposite effects have been shown on the ω_x . That implied the relation between the ω_x and the η . The CRVP forms on the bottom of the blind slot when the coolant jets out. The deeper blind slot increases the distance to the mainstream/coolant interface, hence reduces the CRVP effect, leads to higher performance.

Two configurations with typical high and typical low performance were selected to further analyzed the CRVP details. In addition to exhibiting their local θ and local ω_x around the exit, their CRVPs were visualized in the blind slot, and the CRVP core positions were extracted. The distance between the CRVP and the mainstream/coolant interface has been shown to play an important role. It has been clearly demonstrated that, the CRVP occurred on the end of the comb structure, and being buried by the blind slot when it is deep enough. So the CRVP is far away from the mainstream/coolant interface, leads to a continuous coolant film, and the ideal performance.

The CRVP in the Comb scheme has shown a special characteristic. Its source is mostly related to the coolant, is the coolant velocity gradients occurring in the blind slot, rather than

the source of the conventional CRVP, which is driven by the velocity gradients of the mainstream/coolant shear layer.

The new scheme was an application of the proposal that the film cooling effectiveness depended on the CRVP intensity. The CRVP is hard to avoid in the discrete film cooling geometry. The Comb scheme successfully achieves the perfect film cooling effectiveness by burying CRVP in the designed slot. It has pointed out a direction to develop the advanced film cooling schemes, and in turn proven the previous proposed mechanism of the film cooling heat transfer.

REFERENCES

- [1] Goldstein, R. J., and Eckert, E. R. G., 1974, "Effects of Hole Geometry and Density on Three-Dimensional Film Cooling," *International Journal of Heat and Mass Transfer*, **17**(5) pp. 595-607.
- [2] Bunker, R. S., 2007, "Gas Turbine Heat Transfer: Ten Remaining Hot Gas Path Challenges," *Transactions of the ASME, the Journal of Turbomachinery*, **129**(2) pp. 193-201.
- [3] Bogard, D., and Thole, K., 2006, "Gas Turbine Film Cooling," *Journal of Propulsion and Power*, **22**(2) pp. 249-270.
- [4] Sinha, A. K., Bogard, D. G., and Crawford, M. E., 1991, "Film-Cooling Effectiveness Downstream of a Single Row of Holes with Variable Density Ratio," *Journal of Turbomachinery*, **113**(3) pp. 442-449.
- [5] Gritsch, M., Colban, W., Schar, H., 2005, "Effect of Hole Geometry on the Thermal Performance of Fan-Shaped Film Cooling Holes," *Transactions of the ASME, the Journal of Turbomachinery*, **127**(4) pp. 718-25.
- [6] Baldauf, S., Scheurlen, M., Schulz, A., 2002, "Correlation of Film-Cooling Effectiveness from Thermographic Measurements at Enginelike Conditions," *Journal of Turbomachinery*, **124**(4) pp. 686-698.
- [7] Colban, W. F., Thole, K. A., and Bogard, D., 2011, "A Film-Cooling Correlation for Shaped Holes on a Flat-Plate Surface," *Journal of Turbomachinery*, **133**(1) pp. 011002 (11 pp.).
- [8] Fric, T. F., and Roshko, A., 1994, "Vortical Structure in the Wake of a Transverse Jet," *Journal of Fluid Mechanics*, **279**pp. 1-47.
- [9] Haven, B. A., Yamagata, D. K., Kurosaka, M., 1997, "Anti-kidney pair of vortices in shaped holes and their influence on film cooling effectiveness," *ASME 1997 International Gas Turbine and Aeroengine Congress and Exhibition, GT 1997*, June 2, 1997 - June 5, Anonymous American Society of Mechanical Engineers (ASME), Orlando, FL, United states, **3**, pp. International Gas Turbine Institute.
- [10] Hyams, D. G., and Leylek, J. H., 2000, "A Detailed Analysis of Film Cooling Physics: Part III- Streamwise Injection with Shaped Holes," *Journal of Turbomachinery*, **122**(1) pp. 122-132.
- [11] Kusterer, K., Bohn, D., Sugimoto, T., 2007, "Double-Jet Ejection of Cooling Air for Improved Film Cooling," *Journal of Turbomachinery*, **129**(4) pp. 809-815.
- [12] Heidmann, J. D., and Ekkad, S., 2007, "A novel anti-vortex turbine film cooling hole concept," *2007 ASME Turbo Expo*, May 14, 2007 - May 17, Anonymous American Society of Mechanical Engineers, Montreal, Que., Canada, **4 PART A**, pp. 487-496.
- [13] Dhungel, A., Lu, Y., Phillips, W., 2009, "Film Cooling from a Row of Holes Supplemented with Antivortex Holes," *Journal of Turbomachinery*, **131**(2) pp. 021007 (10 pp.).
- [14] Li, H. M., and Hassan, I., 2015, "The Effects of Counterrotating Vortex Pair Intensity on Film-Cooling Effectiveness," *Heat Transfer Engineering*, **36**(16) pp. 1360-70.
- [15] Sargison, J. E., Guo, S. M., Oldfield, M. L. G., 2002, "A Converging Slot-Hole Film-Cooling Geometry-Part 1: Low-Speed Flat-Plate Heat Transfer and Loss," *Journal of Turbomachinery*, **124**(3) pp. 453-460.
- [16] Bunker, R. S., 2002, "Film cooling effectiveness due to discrete holes within a transverse surface slot," *ASME TURBO EXPO 2002:Heat Transfer, Manufacturing Materials and Metallurgy*, June 3, 2002 - June 6, Anonymous American Society of Mechanical Engineers, Amsterdam, Netherlands, **3 A**, pp. 129-138.
- [17] Bunker, R. S., 2011, "A Study of Mesh-Fed Slot Film Cooling," *Journal of Turbomachinery*, **133**(1) .
- [18] Davidson, F. T., Bogard, D. G., Bruce-Black, J., 2008, "Adiabatic effectiveness on the suction side of a turbine vane and the effects of curvature at the point of film injection," *2008 ASME Turbo Expo*, June 9, 2008 - June 13, Anonymous American Society of Mechanical Engineers (ASME), Berlin, Germany, **4**, pp. 1137-1146.
- [19] Bruce-Black, J., Davidson, F. T., Bogard, D. G., 2011, "Practical Slot Configurations for Turbine Film Cooling Applications," *Journal of Turbomachinery*, **133**(3) pp. 031020 (8 pp.).
- [20] Li, H. M., Ghaly, W., and Hassan, I., 2016, "The Formation of Counter-Rotating Vortex Pair and the Nature of Liftoff-Reattachment in Film-Cooling Flow," *Fluids*, **1**(4) pp. 39.
- [21] Zhang, X. Z., and Hassan, I., 2006, "Film Cooling Effectiveness of an Advanced-Louver Cooling Scheme for Gas Turbines," *Journal of Thermophysics and Heat Transfer*, **20**(4) pp. 754-763.
- [22] Li, H., Ghaly, W., and Hassan, I., 2019, "EXPERIMENTAL INVESTIGATIONS OF A COMB-LIKE FILM-COOLING SCHEME," *ASME - JSME - KSME Joint Fluids Engineering Conference 2019*, July 28 - August 1, 2019, Anonymous ASME - JSME - KSME Joint Fluids Engineering Conference 2019, Hyatt Regency, San Francisco, CA, **AJKFLUIDS2019-4695**, .
- [23] Wright, L. M., McClain, S. T., and Clemenson, M. D., 2011, "Effect of Density Ratio on Flat Plate Film Cooling with Shaped Holes using PSP," *Journal of Turbomachinery*, **133**(4) pp. 041011 (11 pp.).
- [24] Jia, R., Sunden, B., Miron, P., 2005, "A Numerical and Experimental Investigation of the Slot Film-Cooling Jet with various Angles," *Transactions of the ASME, the Journal of Turbomachinery*, **127**(3) pp. 635-45.
- [25] Farmer, J. P., Seager, D. J., and Liburdy, J. A., 1997, "Effect of shaping inclined slots on film cooling effectiveness and heat transfer coefficient," *Proceedings of the 1997 International Gas Turbine & Aeroengine Congress & Exposition*, June 2, 1997 - June 5, Anonymous ASME, Orlando, FL, USA, pp. ASME.

Darcy's Law for Moving Packed Beds

Pedro A. Isaza, Markus Bussmann*

Department of Mechanical and Industrial Engineering
University of Toronto

Stephen Whitaker

Department of Chemical Engineering
University of California at Davis

Abstract—The application of a Galilean transformation and the method of volume averaging are employed as a means of identifying and solving the momentum equations associated with a moving packed bed. Under the well-established assumption that the solids move with constant velocity, the results in this work identify the Darcy's law type relation for the fluid motion subject to a corrective shift that accounts for the solids velocity and the fluid void fraction. The derivation proves that under the conditions explored, the results of the closure problem and the permeability correlations of fixed packed beds can be used for moving systems. This work provides a proof of the intuitive results that would be obtained following a Galilean transformation of the moving bed problem, while simultaneously identifying important and subtle intricacies. The results are also validated with experimental data, creating a firm platform for further research.

Keywords—*Moving packed beds; Volume averaging; Darcy's Law; Closure problem, Galilean transformation*

I. INTRODUCTION

Over the last few decades, interest has grown in the application of moving beds to many processes including chemical separations [1], drying of biomass and food products [2, 3], removal of flue gas by-products [4], and solar radiation capture [5]. Moving beds are characterized by low investment costs, low energy consumption, and low maintenance requirements [6], which give them an important advantage over competing technologies.

Unlike fixed beds, moving beds consist of granular solids and an interstitial fluid which are both in motion. Packed bed transport phenomena have been studied extensively with investigations on momentum transfer yielding theoretical derivations of Darcy's law [7] and the Forchheimer equation [8, 9]. Rigorously established thermal models have also been developed to ascertain when local thermal equilibrium is an appropriate assumption and a single temperature properly describes the solids and interstitial fluid temperatures [10, 11]. Constraints associated with local thermal equilibrium in packed beds have been presented [11, 12], which provide inequalities quantifying the need for separate solids and interstitial fluid energy equations [13,14].

For volume averaging of moving bed problems, the analysis remains in its infancy. Studies to date have postulated and solved

empirical transport equations which successfully describe various experimental conditions [1, 15, 16], and more recently, analytical solutions have been presented for the differential equations associated with energy transport in moving bed heat exchangers [17-20]. Much of this analysis, however, hinges on the assumptions that underlie these empirical equations. To date, only a few studies have made use of volume averaging in an effort to quantify the implications behind the empirical correlations. These initial studies have focused primarily on the heat transfer problem [6, 21]; however, when it comes to momentum, the analysis still relies partly on intuition and has yet to demonstrate the mathematical rigor of fixed bed developments.

An important first step is the application of volume averaging to the moving bed momentum problem. Such a mathematical development would identify the intricacies associated with the problem and create a firm platform for subsequent research. This work presents that development by means of volume averaging and a Galilean transformation, and the results are substantiated by comparing the theoretical predictions with experimental data.

II. THEORETICAL DEVELOPMENT

The moving bed system under consideration and two representative averaging volumes are illustrated in Fig. 1. In the schematic, the macroscopic length of the moving bed is identified as L . Note that the problem studied in this work is associated with the mechanics of the bulk of the bed (i.e. section (b) in Fig. 1). More complex volume averaging developments near boundary walls (i.e. section (c) in Fig. 1) remain an open area of moving bed research and are the subject of on-going investigations. Analogous developments of this kind have been presented in [22, 23] for boundaries between porous media and homogeneous fluids.

To make the analysis tractable, the following assumptions are made at the local (or pore) level:

1. Constant solid thermo-physical properties (i.e. thermal conductivity k , density ρ , and specific heat capacity C_p).

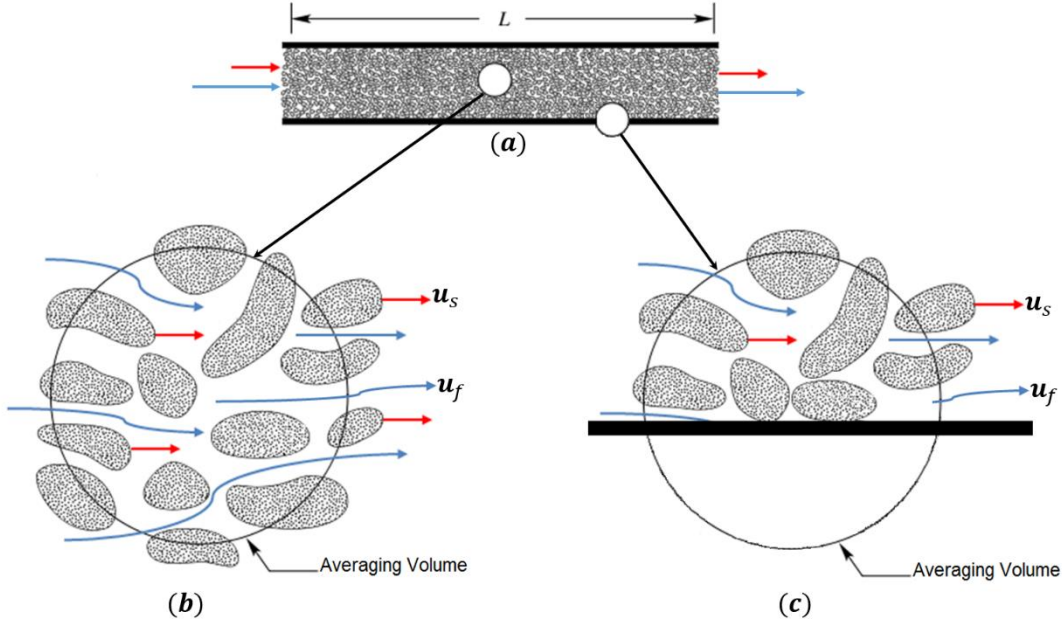


Figure 1. Moving bed schematic and averaging volumes

2. Constant fluid thermo-physical properties (i.e thermal conductivity k , density ρ , specific heat capacity C_p , and viscosity μ_f).
3. The solid particles move with constant velocity. This simplification creates a baseline for moving bed investigations, which has been experimentally validated in parallel-plate and cylindrical geometries [1, 16, 24], while also being applied in recent volume averaging developments [6, 21]. Subsequent research could relax the assumption and incorporate complex mechanisms like random particle motion [25].

The assumption of constant thermo-physical properties decouples the momentum and energy equations [13]. Based on Fig. 1, and the above assumptions, a pore-level moving bed momentum model can be formulated as follows:

Fluid Phase Governing Equations:

$$\nabla \cdot \mathbf{u}_f = 0 \quad (1)$$

$$\rho_f \frac{\partial \mathbf{u}_f}{\partial t} + \rho_f \mathbf{u}_f \cdot \nabla \mathbf{u}_f = -\nabla P_f + \rho_f \mathbf{g} + \mu_f \nabla^2 \mathbf{u}_f \quad (2)$$

Solids Phase Governing Equations:

$$\mathbf{u}_s = \text{constant} \quad (3)$$

where P_f defines the pressure in the fluid phase, while \mathbf{u} denotes the velocities in the laboratory frame of reference. Note that the fluid and solid phases are identified by the subscripts f and s respectively.

The boundary conditions that complement the model are as follows:

BCs 1 & 2: Impermeable and Moving Fluid-Solids Interfaces:

$$\rho_f (\mathbf{u}_f - \mathbf{w}) \cdot \mathbf{n}_{fs} = 0 \quad \text{at } \mathbf{A}_{fs} \quad (4)$$

$$\rho_s (\mathbf{u}_s - \mathbf{w}) \cdot \mathbf{n}_{sf} = 0 \quad \text{at } \mathbf{A}_{fs} \quad (5)$$

where \mathbf{w} is the displacement speed of the fluid-solids interfaces, \mathbf{n}_{fs} and \mathbf{n}_{sf} define the normal unit vectors pointing from the fluid to the solids phase and from the solids to the fluid phase respectively, and \mathbf{A}_{fs} denotes the interfacial area between the fluid and solids in the macroscopic region (see Fig. 1). Note that these jump conditions imply that no mass is exchanged at the interfaces. Analogous equations have been formulated for packed bed drying problems [26].

BC 3: Moving Bed Entrance and Exit Boundary Conditions:

$$\mathbf{u}_f = \mathbf{f}(\mathbf{r}, t) \quad \text{at } \mathbf{A}_{fe} \quad (6)$$

where \mathbf{r} is the position vector, $\mathbf{f}(\mathbf{r}, t)$ are the arbitrary entrance/exit velocity functions with respect to position and time, and \mathbf{A}_{fe} denotes the entrance and exit areas of the macroscopic region in Fig. 1. Note that if one were solving the transient form of the problem, a fluid velocity initial condition would be required (while not being needed for the solid particles given their assumed constant velocity). The solution of that problem will be explored in subsequent research.

To move into the realm of Stokes flow and Darcy's law, the assumptions of steady-state and negligible convective effects are applied [7]. Doing so simplifies the problem as follows:

$$\nabla \cdot \mathbf{u}_f = 0 \quad (7)$$

$$0 = -\nabla P_f + \rho_f \mathbf{g} + \mu_f \nabla^2 \mathbf{u}_f \quad (8)$$

$$\mathbf{u}_s = \text{constant} \quad (9)$$

Assuming no-slip at the interfaces collapses the impermeability BCs, and one obtains:

$$\mathbf{u}_f = \mathbf{u}_s \quad \text{at } \mathbf{A}_{fs} \quad (10)$$

where velocity equality in both the normal and tangential directions (i.e. \mathbf{n}_{fs} and $\boldsymbol{\lambda}_{fs}$ are the normal and tangential vectors at the fluid-solid interfaces shown in Fig. 2) is required. The final entrance condition also reduces accordingly to:

$$\mathbf{u}_f = \mathbf{f}(\mathbf{r}) \quad \text{at } \mathbf{A}_{fe} \quad (11)$$

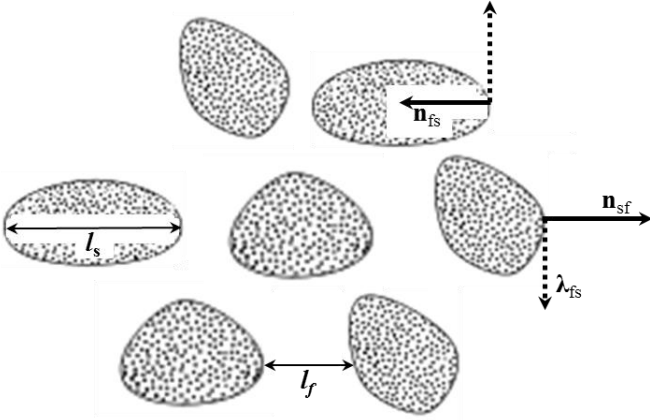


Figure 2. Normal and tangential vectors defined for the moving porous media

A. Galilean Transformation

Since we have assumed that the solids move with constant velocity, a Galilean transformation (see Chapter 14 in [27]) can be applied to shift the problem to the moving solids frame of reference. Intuitively, the transformation should result in governing equations analogous to those of packed beds (subject to a relative description of the fluid phase velocity). Provided that a reverse transformation is available, existing packed bed results can then be applied to moving beds.

To impose the transformation, the laboratory frame of reference velocities for both phases are expressed as:

$$\mathbf{u}_f = \mathbf{v}_f + \mathbf{w}_o ; \quad \mathbf{u}_s = \mathbf{v}_s + \mathbf{w}_o \quad (12)$$

where \mathbf{v}_f and \mathbf{v}_s represent the relative velocities, and \mathbf{w}_o the constant velocity transformation. Choosing the special case where $\mathbf{w}_o = \mathbf{u}_s$ yields the following relative descriptions:

$$\mathbf{v}_f = \mathbf{u}_f - \mathbf{u}_s ; \quad \mathbf{v}_s = 0 \quad (13)$$

Substituting Eq. (13) into the laboratory frame Eqs. (7) - (11), one obtains the following:

Shifted Governing Equations:

$$\nabla \cdot \mathbf{v}_f = 0 \quad (14)$$

$$0 = -\nabla P_f + \rho_f \mathbf{g} + \mu_f \nabla^2 \mathbf{v}_f \quad (15)$$

$$\mathbf{v}_s = 0 \quad (16)$$

Shifted Boundary Conditions:

$$\mathbf{v}_f \cdot \mathbf{n}_{fs} = 0 ; \quad \mathbf{v}_f \cdot \boldsymbol{\lambda}_{fs} = 0 \quad \text{at } \mathbf{A}_{fs} \quad (17)$$

$$\mathbf{v}_f = \mathbf{f}(\mathbf{r}) - \mathbf{u}_s = \mathbf{g}(\mathbf{r}) \quad \text{at } \mathbf{A}_{fe} \quad (18)$$

As required, the governing equations are invariant to the Galilean transformation, and yield a fluid velocity problem subject to a relative description. Additionally, a zero-velocity condition is obtained at the fluid-solids interfaces. Most importantly, the transformed equations (i.e. Eqs. (14) - (17)) are analogous to those of the packed bed problem (see section 4.1 in [10]). Note that although the macroscopic BC, given by Eq. (18), *does not* transform to the classical packed bed problem, the shift is only by a constant thereby remaining time invariant. We can now proceed to volume average the shifted problem with all of the packed bed methods and findings at our disposal.

B. Volume Averaging Fundamentals

Volume averaging begins by associating an averaging volume with every point in the global domain [10, 26]. A position vector \mathbf{x} locates the centroid of the volume, while a relative position vector \mathbf{y} identifies points within the volume. Subscripts on \mathbf{y} specify the phase. Figure 3 depicts a sample averaging volume along with its position vectors.

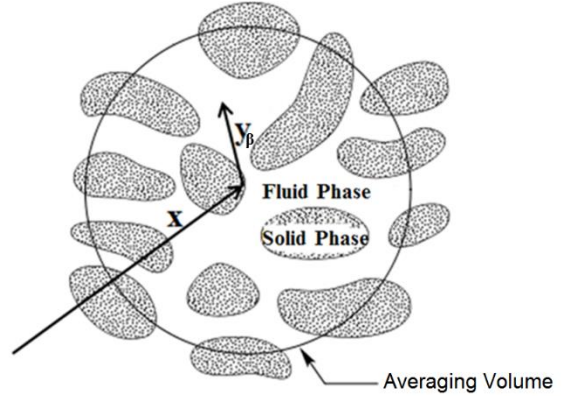


Figure 3. Averaging volume and position vectors.

Given the two-phase nature of the system, the averaging volume is defined as:

$$V = V_f(\mathbf{x}, t) + V_s(\mathbf{x}, t) \quad (19)$$

where V_f and V_s are the fluid and solid volumes which depend on both position and time.

As discussed in [14], two types of volume averaged properties can be defined: *superficial* and *intrinsic*. Volume averaging defines a *superficial average* as:

$$\langle \psi_\beta \rangle \big|_{\mathbf{x}, t} = \frac{1}{V} \int_{V_\beta(\mathbf{x}, t)} \psi_\beta(\mathbf{x} + \mathbf{y}_\beta, t) dV \quad (20)$$

where β denotes the phase of interest, \mathbf{y}_β locates β within the averaging volume, $V_\beta(\mathbf{x}, t)$ is the volume of β in the averaging volume, $\psi_\beta(\mathbf{x} + \mathbf{y}_\beta, t)$ is the pore-level property of the β -phase, and $\langle \psi_\beta \rangle \big|_{\mathbf{x}, t}$ is the *superficial averaged* property at the centroid. To simplify the notation, it is standard procedure to re-write the expression in the following way (see Sec. 1.2 in [10]):

$$\langle \psi_\beta \rangle = \frac{1}{V} \int_{V_\beta} \psi_\beta dV \quad (21)$$

As discussed in [10], *intrinsic* properties are preferred over *superficial* properties, as they better represent the variables associated with the phase of interest. The relation between these averages is the following:

$$\langle \psi_\beta \rangle = \varepsilon_\beta \langle \psi_\beta \rangle^\beta \quad (22)$$

where $\langle \psi_\beta \rangle^\beta$ is the *intrinsic* property, while ε_β is the volume fraction of the β -phase in the averaging volume defined as:

$$\varepsilon_\beta = \frac{V_\beta(\mathbf{x}, t)}{V} \quad (23)$$

Spatial Averaging Theorem:

During the process of volume averaging the governing equations, the order of integration and differentiation needs to be interchanged. This is achieved by means of the spatial averaging theorem, which is a three-dimensional representation of the Leibniz rule [28-31]. For a scalar property ψ associated with the β -phase, in contact with a second phase α in the averaging volume, the theorem states:

$$\langle \nabla \psi_\beta \rangle = \nabla \langle \psi_\beta \rangle + \frac{1}{V} \int_{A_{\beta\alpha}} \mathbf{n}_{\beta\alpha} \psi_\beta dA \quad (24)$$

Here $\mathbf{n}_{\beta\alpha}$ is the normal unit vector pointing from the β to the α phase, and $A_{\beta\alpha}$ is the interfacial area between the β and α phases in the averaging volume. In vector form, the theorem can be written as:

$$\langle \nabla \cdot \mathbf{d}_\beta \rangle = \nabla \cdot \langle \mathbf{d}_\beta \rangle + \frac{1}{V} \int_{A_{\beta\alpha}} \mathbf{n}_{\beta\alpha} \cdot \mathbf{d}_\beta dA \quad (25)$$

where \mathbf{d}_β is the vector property \mathbf{d} of the β -phase.

C. Volume Averaging the Continuity Equation

The basics of volume averaging can now be applied to the shifted problem. Before doing so, it is important to reinforce once more that Eqs. (14) - (17) are identical in form to those describing packed beds. Volume averaging of that problem has been presented in detail in Chapter 4 of [10] and [7]. For the sake of completeness, the process of volume averaging the continuity equation is presented in detail; however, only a few equations for the momentum problem are presented as they follow the work in [7] and [10].

The fluid continuity equation (Eq. (14)) can be volume averaged as:

$$\frac{1}{V} \int_{V_f} (\nabla \cdot \mathbf{v}_f) dV = \langle \nabla \cdot \mathbf{v}_f \rangle = 0 \quad (26)$$

Applying the spatial averaging theorem (see Eq. (25)), the expression can be written as:

$$\langle \nabla \cdot \mathbf{v}_f \rangle = \nabla \cdot \langle \mathbf{v}_f \rangle + \frac{1}{V} \int_{A_{fs}} \mathbf{n}_{fs} \cdot \mathbf{v}_f dA \quad (27)$$

Enforcing zero velocity at the solids-fluid boundaries required by Eq. (17), yields the superficially averaged continuity equation:

$$\nabla \cdot \langle \mathbf{v}_f \rangle = 0 \quad (28)$$

Equation (22) can now be applied to obtain:

$$\nabla \cdot \langle \mathbf{v}_f \rangle = \nabla \cdot (\varepsilon_f \langle \mathbf{v}_f \rangle^f) = \varepsilon_f \nabla \cdot \langle \mathbf{v}_f \rangle^f + \nabla \varepsilon_f \cdot \langle \mathbf{v}_f \rangle^f = 0 \quad (29)$$

Rearranging yields the intrinsic average continuity equation:

$$\nabla \cdot \langle \mathbf{v}_f \rangle^f = -\varepsilon_f^{-1} \nabla \varepsilon_f \cdot \langle \mathbf{v}_f \rangle^f \quad (30)$$

Applying the supplementary assumption of “homogeneous” flow (i.e. constant volume fractions of all phases) the expression simplifies to:

$$\nabla \cdot \langle \mathbf{v}_f \rangle^f = 0 \quad (31)$$

This assumption will typically be valid in the free-flowing bulk of a moving bed, but not near walls and obstructions. As discussed in [8], regions of significant porosity variation are normally found at porous media boundaries with homogeneous fluids or solids (i.e. walls like that shown in Fig. 1(c)). These locations require the development of jump conditions like those of [32], which are outside the scope of this work.

D. Volume Averaging the Momentum Equation

Volume averaging the momentum equation (Eq. (15)) yields:

$$0 = -\langle \nabla P_f \rangle + \langle \rho_f \mathbf{g} \rangle + \langle \mu_f \nabla^2 \mathbf{v}_f \rangle \quad (32)$$

Spatial averaging the pressure gradient term (see Eq. (24)) results in:

$$-\langle \nabla P_f \rangle = -\nabla \langle P_f \rangle - \frac{1}{V} \int_{A_{fs}} \mathbf{n}_{fs} P_f dA \quad (33)$$

where as proposed in [33] the local pressure can be decomposed as:

$$P_f = \langle P_f \rangle^f + \tilde{P}_f \quad (34)$$

and \tilde{P}_f represents the fluid phase pressure deviation between the microscopic variable and the intrinsic average.

Substituting Eqs. (34) and (22) into Eq. (33), one obtains:

$$-\langle \nabla P_f \rangle = \varepsilon_f \nabla \langle P_f \rangle^f - \frac{1}{V} \int_{A_{fs}} \mathbf{n}_{fs} \tilde{P}_f dA \quad (35)$$

subject to the length scale constraint:

$$\frac{r_o^2}{L_{\varepsilon_f} L_{p1}} \ll 1 \quad (36)$$

where r_o is the radius of the averaging volume, and L_{ε_f} and L_{p1} are the characteristic lengths of the porosity variation and of the pressure deviation respectively (see Chapter 4 in [10] for details).

A system assumed to be comprised of homogeneous porous media (i.e. constant porosity) will inherently have an infinite characteristic length for the porosity variation. This is equivalent to an invariance of the volume fraction, which automatically satisfies Eq. (36).

Expanding the average of the gravitational term, one obtains:

$$\langle \rho_f \mathbf{g} \rangle = \rho_f \mathbf{g} \langle 1 \rangle = \varepsilon_f \rho_f \mathbf{g} \quad (37)$$

since the density and the gravitational field are assumed constant.

Finally, spatial averaging the viscous term, introducing the velocity decomposition (analogous to Eq. (34)), and applying the fluid-solids boundary condition (Eq. (17)), one can write (see [10] for more details):

$$\langle \nabla^2 \mathbf{v}_f \rangle = \nabla^2 \langle \mathbf{v}_f \rangle - \nabla \varepsilon_f \cdot \nabla \langle \mathbf{v}_f \rangle^f + \frac{1}{V} \int_{A_{fs}} \mathbf{n}_{fs} \cdot \nabla \tilde{\mathbf{v}}_f dA \quad (38)$$

subject to the constraint:

$$\frac{r_o^2}{L_{\varepsilon_f} L_{v2}} \ll 1 \quad (39)$$

where $\tilde{\mathbf{v}}_f$ is the local spatial deviation of the fluid velocity (i.e. $\mathbf{v}_f - \langle \mathbf{v}_f \rangle^f$), L_{v2} is the characteristic length of the spatial derivative of velocity (see Eq. (4.1-25) in [10]), and the inequality is satisfied when porosity is assumed constant.

Substituting Eq. (22) into Eq. (38), expanding, and applying the constant porosity assumption one finds:

$$\langle \nabla^2 \mathbf{v}_f \rangle = \varepsilon_f \nabla^2 \langle \mathbf{v}_f \rangle^f + \frac{1}{V} \int_{A_{fs}} \mathbf{n}_{fs} \cdot \nabla \tilde{\mathbf{v}}_f dA \quad (40)$$

Substituting Eqs. (35), (37) and (40) into Eq. (32), and dividing by ε_f , yields the volume averaged momentum equation for the shifted moving bed problem:

$$0 = -\nabla \langle P_f \rangle^f + \rho_f \mathbf{g} + \mu_f \nabla^2 \langle \mathbf{v}_f \rangle^f + \frac{1}{V_f} \int_{A_{fs}} \mathbf{n}_{fs} \cdot (-\mathbf{I} \tilde{P}_f + \mu_f \nabla \tilde{\mathbf{v}}_f) dA \quad (41)$$

where \mathbf{I} is the unit tensor.

E. Closure Problem

To obtain a closed form of the volume averaged problem, representations for $\tilde{\mathbf{v}}_f$ and \tilde{P}_f are needed. This is commonly done by deriving governing equations and boundary conditions for the deviation variables, as presented in much of the volume averaging work to date (e.g. [7, 8, 10, 11]). The resulting equations, referred to as the closure problem, are then solved through the method of superposition in terms of *closure variables*. This provides a means for determining the Darcy's law permeability tensor, as will be seen subsequently. The derivation of the problem and its solution is discussed in detail in Sec. 4.2 of [10]. Below we present the closure problem resulting from the subtraction of the pore-level equations from the volume averaged equations:

$$\nabla \cdot \tilde{\mathbf{v}}_f = 0 \quad (42)$$

$$0 = -\nabla \tilde{P}_f + \mu_f \nabla^2 \tilde{\mathbf{v}}_f - \frac{1}{V_f} \int_{A_{fs}} \mathbf{n}_{fs} \cdot (-\mathbf{I} \tilde{P}_f + \mu_f \nabla \tilde{\mathbf{v}}_f) dA \quad (43)$$

$$\text{BC\#1} \quad \tilde{\mathbf{v}}_f = -\underbrace{\langle \mathbf{v}_f \rangle^f}_{\text{source}} \quad \text{at } A_{fs} \quad (44)$$

$$\text{BC\#2} \quad \tilde{\mathbf{v}}_f = \underbrace{\mathbf{k}(\mathbf{r})}_{\text{source}} \quad \text{at } A_{fe} \quad (45)$$

$$\text{Average} \quad \langle \tilde{\mathbf{v}}_f \rangle^f = 0 \quad (46)$$

where $\mathbf{k}(\mathbf{r})$ are the arbitrary entrance and exit boundary condition functions for the deviation variable, and the average of

the velocity deviation equals zero on the basis of the length scale constraints presented in Sec. 4.2.4 of [10]. To be precise, we note that Eq. (43) is obtained by subtracting Eq. (41) from Eq. (15).

Clearly one does not wish to solve the above problem over the entire macroscopic region; however, the global boundary condition (Eq. (45)) can be discarded if a suitable replacement can be found. It is of some interest for the replacement boundary to be associated with the averaging volume. This naturally leads to a description of the averaging volume as a representative region of the porous media, or a unit cell in a spatially periodic system [10]. Under such conditions, the macroscopic boundary can be replaced with the periodicity conditions:

$$\begin{aligned} \text{Periodicity} \quad \tilde{P}_f(\mathbf{r} + l_i) &= \tilde{P}_f(\mathbf{r}) \\ \tilde{\mathbf{v}}_f(\mathbf{r} + l_i) &= \tilde{\mathbf{v}}_f(\mathbf{r}) \quad \text{for } i = 1, 2, 3 \end{aligned} \quad (47)$$

where l_i represents the three non-unique lattice vectors required to describe a spatially periodic porous medium. The use of periodicity as a replacement has been supported and is extensively studied in [34-38].

The solution for the spatial deviation problem can now be proposed as a function of the single source term $\langle \mathbf{v}_f \rangle^f$ by the principle of superposition (see Sec. 4.2.5 in [10]) as follows:

$$\tilde{\mathbf{v}}_f = \mathbf{B} \cdot \langle \mathbf{v}_f \rangle^f \quad (48)$$

$$\tilde{P}_f = \mu_f \mathbf{b} \cdot \langle \mathbf{v}_f \rangle^f \quad (49)$$

where \mathbf{B} and \mathbf{b} are the tensor and vector closure variables respectively. Substituting this description into Eqs. (42) - (47) (where Eq. (45) is replaced by (47) as discussed), one arrives at a closure problem analogous to that derived for packed beds (see Eqs. (3.39) - (3.43) in [7]):

$$\nabla \cdot \mathbf{B} = 0 \quad (50)$$

$$0 = -\nabla \mathbf{b} + \nabla^2 \mathbf{B} - \frac{1}{V_f} \int_{A_{fs}} \mathbf{n}_{fs} \cdot (-\mathbf{I} \mathbf{b} + \nabla \mathbf{B}) dA \quad (51)$$

$$\text{BC\#1} \quad \mathbf{B} = -\mathbf{I} \quad \text{at } A_{fs} \quad (52)$$

$$\begin{aligned} \text{Periodicity} \quad \mathbf{b}(\mathbf{r} + l_i) &= \mathbf{b}(\mathbf{r}) \\ \mathbf{B}(\mathbf{r} + l_i) &= \mathbf{B}(\mathbf{r}) \quad \text{for } i = 1, 2, 3 \end{aligned} \quad (53)$$

$$\text{Average} \quad \langle \mathbf{B} \rangle^f = 0 \quad (54)$$

At this point, we make use of the last term in Eq. (41) along with Eqs. (48) and (49) to obtain:

$$\begin{aligned} &\frac{1}{V_f} \int_{A_{fs}} \mathbf{n}_{fs} \cdot (-\mathbf{I} \tilde{P}_f + \mu_f \nabla \tilde{\mathbf{v}}_f) dA \\ &= \left\{ \frac{1}{V_f} \int_{A_{fs}} \mathbf{n}_{fs} \cdot [-\mathbf{I} \mathbf{b} + \nabla \mathbf{B}] dA \right\} \cdot \mu_f \langle \mathbf{v}_f \rangle^f \end{aligned} \quad (55)$$

Since Darcy's law is traditionally represented in terms of the superficial velocity, one is led to the following definition of the permeability tensor (\mathbf{K}):

$$\frac{1}{V_f} \int_{A_{fs}} \mathbf{n}_{fs} \cdot [-\mathbf{I} \mathbf{b} + \nabla \mathbf{B}] dA = -\varepsilon_f \mathbf{K}^{-1} \quad (56)$$

Using this result along with the two additional definitions:

$$\mathbf{b} = \mathbf{d} \cdot (\varepsilon_f \mathbf{K}^{-1}) \quad (57)$$

$$\mathbf{B} = -\mathbf{I} + \mathbf{D} \cdot (\varepsilon_f \mathbf{K}^{-1}) \quad (58)$$

one can express Eqs. (50) through (53) as:

$$\nabla \cdot \mathbf{D} = 0 \quad (59)$$

$$0 = -\nabla \mathbf{d} + \nabla^2 \mathbf{D} + \mathbf{I} \quad (60)$$

$$\text{BC\#1} \quad \mathbf{D} = 0 \quad \text{at } A_{fs} \quad (61)$$

$$\begin{aligned} \text{Periodicity} \quad \mathbf{d}(\mathbf{r} + l_i) &= \mathbf{d}(\mathbf{r}) \\ \mathbf{D}(\mathbf{r} + l_i) &= \mathbf{D}(\mathbf{r}) \quad \text{for } i = 1, 2, 3 \end{aligned} \quad (62)$$

where \mathbf{D} and \mathbf{d} are alternative tensor and vector fields simplifying the closure problem.

From Eqs. (54) and (58) one can conclude that the permeability tensor is given by:

$$\mathbf{K} = \varepsilon_f \langle \mathbf{D} \rangle^f \quad (63)$$

and the closure problem represented by Eqs. (59) to (62) has been solved (either directly or indirectly) by numerous researchers [39-43].

F. Closed Form of the Volume Averaged Equations

Making use of Eqs. (48) and (49), Eq. (41) can be written as:

$$\begin{aligned} 0 &= -\nabla \langle P_f \rangle^f + \rho_f \mathbf{g} + \mu_f \nabla^2 \langle \mathbf{v}_f \rangle^f \\ &+ \mu_f \left\{ \frac{1}{V_f} \int_{A_{fs}} \mathbf{n}_{fs} \cdot (-\mathbf{I} \mathbf{b} + \nabla \mathbf{B}) dA \right\} \cdot \langle \mathbf{v}_f \rangle^f \end{aligned} \quad (64)$$

where the volume averaged property can be removed from the area integral due to the length scale restrictions of the problem [10].

On the basis of Eq. (56), we obtain Darcy's law with the *Brinkman correction*:

$$0 = -\nabla \langle P_f \rangle^f + \rho_f \mathbf{g} + \mu_f \nabla^2 \langle \mathbf{v}_f \rangle^f - \mu_f \varepsilon_f \mathbf{K}^{-1} \cdot \langle \mathbf{v}_f \rangle^f \quad (65)$$

which is identical to that of packed beds [7], with the exception that the fluid velocity is relative (i.e. from the solids frame of reference).

Rearranging, and expressing the equation in terms of the superficial velocity (typical of Darcy's law) yields:

$$\langle \mathbf{v}_f \rangle = -\frac{\mathbf{K}}{\mu_f} \cdot (\nabla \langle P_f \rangle^f - \rho_f \mathbf{g}) + \mathbf{K} \cdot \nabla^2 \langle \mathbf{v}_f \rangle^f \quad (66)$$

On the basis of the length-scale constraints imposed on the problem (i.e. the assumption of homogeneous porous media), the *Brinkman correction* can be ignored [10] because:

$$\langle \mathbf{v}_f \rangle \gg \mathbf{K} \cdot \nabla^2 \langle \mathbf{v}_f \rangle^f \quad (67)$$

which simplifies Eq. (66) to:

$$\langle \mathbf{v}_f \rangle = -\frac{\mathbf{K}}{\mu_f} \cdot (\nabla \langle P_f \rangle^f - \rho_f \mathbf{g}) \quad (68)$$

It is important to recognize that the permeability tensor can thus be estimated from packed bed numerical and experimental investigations, subject to a relative velocity description for the fluid. As such, under the assumption of solids moving with

constant velocity, the permeability data for packed beds can be used to describe fluid flow in moving beds.

Finally, we develop the procedure for shifting the solution back to the laboratory frame of reference. To do so, recall the Galilean transformation imposed on the fluid problem, given by Eq. (13):

$$\mathbf{u}_f = \mathbf{v}_f + \mathbf{u}_s \quad (69)$$

Volume averaging the expression yields:

$$\langle \mathbf{u}_f \rangle = \langle \mathbf{v}_f \rangle + \langle \mathbf{u}_s \rangle \quad (70)$$

Extracting \mathbf{u}_s from the average (since it is a constant), and recognizing that $\langle 1 \rangle = \varepsilon_f$, one finds that:

$$\langle \mathbf{u}_f \rangle = \langle \mathbf{v}_f \rangle + \varepsilon_f \mathbf{u}_s \quad (71)$$

Substituting Eq. (68) into Eq. (71), one finds Darcy's law for the moving bed in the laboratory frame of reference:

$$\langle \mathbf{u}_f \rangle = -\frac{\mathbf{K}}{\mu_f} \cdot (\nabla \langle P_f \rangle^f - \rho_f \mathbf{g}) + \varepsilon_f \mathbf{u}_s \quad (72)$$

For spherical particles, the Blake-Kozeny equation presented in [9, 44, 45] yields an estimate for the permeability in homogeneous isotropic media:

$$\mathbf{K} = \frac{d_p^2 \varepsilon_f^3}{150(1-\varepsilon_f)^2} \quad (73)$$

Substituting this expression into Eq. (72), yields Darcy's law for moving beds composed of homogeneous isotropic media:

$$\langle \mathbf{u}_f \rangle = -\frac{d_p^2 \varepsilon_f^3}{150(1-\varepsilon_f)^2 \mu_f} \cdot (\nabla \langle P_f \rangle^f - \rho_f \mathbf{g}) + \varepsilon_f \mathbf{u}_s \quad (74)$$

This demonstrates that under the assumption of constant solids motion, the permeabilities of packed beds can be extended to moving beds. Future research into the closure problem of packed beds can now be translated to moving beds through the above relationships. Finally, recall that Eq. (74) applies only in the bulk region of the moving bed; regions near the wall require a special analysis of the kind presented in [32]. This will be the subject of future studies.

III. EXPERIMENTAL VALIDATION

As discussed in previous work [17-20], experimental data of pressure drop and thermal transport in moving beds is scarce. An important exception exists in the experimental data presented by Sissom [46]. Although their investigation was primarily oriented to examine the thermal transport between a moving bed of aluminum granules and compressed air, valuable pressure drop data was collected as a function of the relative velocity. In their system, aluminum granules moved downwards by gravity while the interstitial fluid (i.e. dry compressed air) flowed counter-currently. Equation (74) can thereby be written in the following way for the one-dimensional case explored in that work:

$$\frac{d \langle P_f \rangle^f}{dx} = -\frac{150(1-\varepsilon_f)^2 \mu_f}{d_p^2 \varepsilon_f^3} \cdot [\langle \mathbf{u}_f \rangle - \varepsilon_f \mathbf{u}_s] - \rho_f \mathbf{g} \quad (75)$$

Note that the influence of the gravitational field is reversed due to the counter-flow nature of the experiment. Also, the Cartesian x -direction has its origin at the bottom of the bed and increases in the upwards direction.

Since temperature variations in the region where the pressure drop was monitored were small, density and viscosity variations can be assumed to be negligible. Equation (75) can therefore be integrated to generate the following pressure drop relationship as a function of the experimental length:

$$\Delta\langle P_f \rangle^f(x) = \langle P_f \rangle^f(0) - \langle P_f \rangle^f(x) = \left[\frac{150(1-\varepsilon_f)^2 \mu_f}{d_p^2 \varepsilon_f^3} \cdot [\langle \mathbf{u}_f \rangle - \varepsilon_f \mathbf{u}_s] + \rho_f \mathbf{g} \right] \cdot \mathbf{x} \quad (76)$$

Extracting the plenum-related losses from the raw data reveals the moving bed pressure drops. Figure 4 presents the profiles observed experimentally, along with those predicted by Eq. (76), at the low, middle, and high relative velocities explored in that work (i.e. $\langle \mathbf{v}_f \rangle$ equal to 1.8, 2.7 and 3.4 ft/s).

From Fig. 4, it is clear that the predictions made by the volume averaged expressions are in good agreement with the experimental data, over the range of superficial velocities explored. Also as expected, higher relative velocities result in larger pressure drops. This behaviour is exhibited by both the experimental data and the theoretical predictions. Finally, the expected linear relationship between pressure drop and superficial relative velocity (per Eq. (75)) is exhibited by the experimental data.

The complete data set presented by Sissom [46] was examined in an analogous manner. Figure 5 below summarizes the rates of pressure drop observed experimentally vs. those predicted theoretically.

In Fig. 5, the data presented in red outlines the 15 cases which yield good agreement between the experimental and theoretical predictions. The data in blue outlines the eight cases where the theoretical prediction overestimates the experimental observations. From the figure, it is clear that a good degree of correlation is observed between the majority of the observations and the theoretical predictions. For the other data, variations in the particle diameter can readily account for the deviation. As outlined by Sissom [46] in Table D5, manufacturing variability yielded aluminum particles that could vary in diameter by 25 %, and Sissom [46] did not characterize the mean particle size for each of the 23 runs. Future experimental work could repeat the analysis, subject to a more stringent particle size characterization (i.e. minimize the source of error introduced by this parameter). Finally, although not explicitly presented, all 23 cases exhibited the linear behaviour predicted theoretically between pressure drop and superficial velocity, per the trends in Fig. 4.

IV. CONCLUSIONS

The derivation of a Darcy's law relationship for a moving packed bed is presented by means of a Galilean transformation followed by the application of volume averaging. This analysis demonstrates that under the well-established assumption that the solids move with constant velocity, the fluid equation (i.e. Darcy's law) takes a form analogous to that of fixed packed beds

subject to a shift that accounts for the solids velocity and the fluid void fraction. The resulting closure problem for these moving systems (following the Galilean transformation) is analogous to that of packed beds. As a result, the moving bed solution presented here can be coupled to the packed bed permeability tensors/correlations found in the literature. Finally, the rigorous derivation presented here stands as a proof of the intuitive results expected following the transformation, while simultaneously identifying important but subtle intricacies associated with the equations. The validity of the approach is further substantiated by demonstrating good agreement with experimental data in the literature. This paper represents a firm platform for more complex volume averaging research to follow.

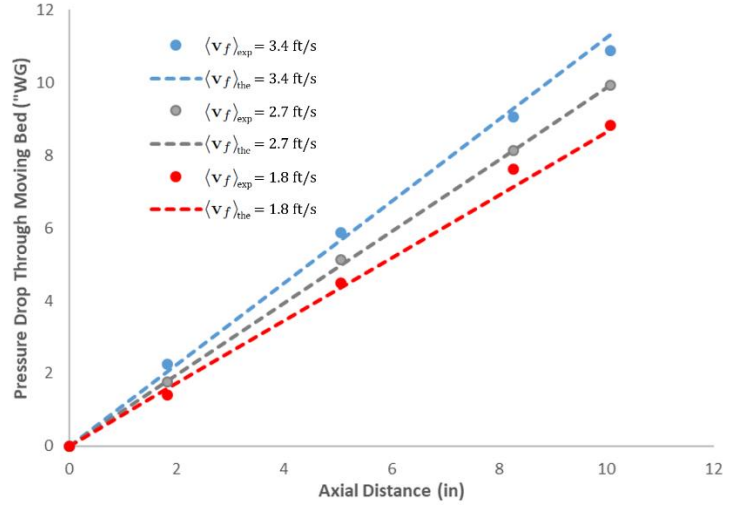


Figure 4. Experimental data vs. theoretical pressure drop predictions for a moving bed.

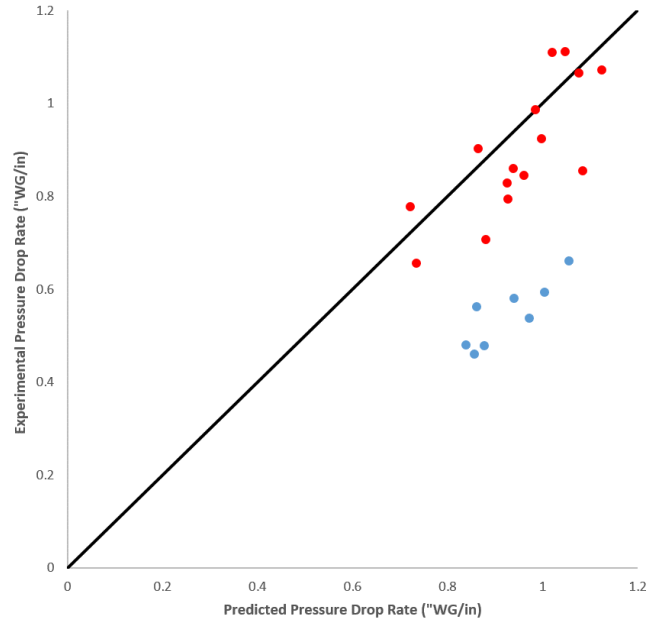


Figure 5. Experimental data vs. theoretical pressure drop rate predictions for a moving bed.

REFERENCES

- [1] M. Colakyan and O. Levenspiel, "Heat transfer between moving bed of solids and immersed cylinders," *AICHE Symposium Series*, vol. 80, p. 241, 1984.
- [2] J. J. Saastamoinen, "Heat exchange between two coupled moving beds by fluid flow," *International Journal of Heat and Mass Transfer*, vol. 47, no. 6, pp. 1535-1547, 2004.
- [3] J. Yrjola and J. J. Saastamoinen, "Modelling and practical operation results of a dryer for wood chips," *Drying Technology*, vol. 20, no. 6, pp. 1077-1099, 2002.
- [4] H. W. Pennline and J. S. Hoffmann, "Flue gas cleanup using the moving bed copper oxide process," *Fuel Processing Technology*, vol. 114, pp. 109-117, 2013.
- [5] T. Baumann and S. Sunft, "Theoretical and experimental investigation of a moving bed heat exchanger for solar central receiver power plants," *Journal of Physics: Conference Series*, vol. 395, pp. 1-9, 2012.
- [6] A. C. Pivem and M. J. S. de Lemos, "Laminar heat transfer in a moving porous bed reactor simulated with a macroscopic two-energy equation model," *International Journal of Heat and Mass Transfer*, vol. 55, no. 7, pp. 1922-1930, 2012.
- [7] S. Whitaker, "Flow in porous media I: A theoretical derivation of Darcy's law," *Transport in Porous Media*, vol. 1, no. 1, pp. 3-25, 1986.
- [8] S. Whitaker, "The Forchheimer equation: A theoretical development," *Transport in Porous Media*, vol. 25, no. 1, pp. 27-61, 1996.
- [9] I. Macdonald, M. El-Sayed, K. Mow, and F. Dullien, "Flow through porous media – the Ergun equation revisited," *Industrial and Engineering Chemistry Fundamentals*, vol. 18, no. 3, pp. 199-208, 1979.
- [10] S. Whitaker, *The Method of Volume Averaging*, vol. 13. Springer, 1999.
- [11] M. Quintard and S. Whitaker, "Local thermal equilibrium for transient heat conduction: Theory and comparison with numerical experiments," *International Journal of Heat and Mass Transfer*, vol. 38, no. 15, pp. 2779-2796, 1995.
- [12] S. Whitaker, "Improved constraints for the principle of local thermal equilibrium," *Industrial and Engineering Chemistry Research*, vol. 30, no. 5, pp. 983-997, 1991.
- [13] M. Quintard, M. Kaviani, and S. Whitaker, "Two-medium treatment of heat transfer in porous media: Numerical results for effective properties," *Advances in Water Resources*, vol. 20, no. 2, pp. 77-94, 1997.
- [14] M. Quintard and S. Whitaker, "Theoretical analysis of transport in porous media," in *Handbook of heat transfer in porous media* (K. Vafai, ed.), pp. 1-52, New York, USA: Marcel Dekker Inc, 2000.
- [15] J. K. Spelt, C. E. Brennen, and R. H. Sabersky, "Heat transfer to flowing granular material," *International Journal of Heat and Mass Transfer*, vol. 25, no. 6, pp. 791-796, 1982.
- [16] W. N. Sullivan and R. H. Sabersky, "Heat transfer to flowing granular media," *International Journal of Heat and Mass Transfer*, vol. 18, no. 1, pp. 97-107, 1975.
- [17] P. A. Isaza, W. D. Warnica, and M. Bussmann, "Co-current parallel-plate moving bed heat exchanger: An analytical solution," *International Journal of Heat and Mass Transfer*, vol. 87, pp. 616-624, 2015.
- [18] P. A. Isaza, W. D. Warnica, and M. Bussmann, "Counter-current parallel-plate moving bed heat exchanger: An analytical solution," *International Journal of Heat and Mass Transfer*, vol. 87, pp. 625-635, 2015.
- [19] P. A. Isaza, Y. Cai, W. D. Warnica, and M. Bussmann, "Co-current and counter-current vertical pipe moving bed heat exchangers: Analytical solutions," *International Journal of Heat and Mass Transfer*, vol. 95, pp. 1115-1128, 2016.
- [20] P. A. Isaza, A. O'Brien, W. D. Warnica, and M. Bussmann, "Assessing axial heat conduction in moving bed heat exchangers," *International Journal of Thermal Sciences*, vol. 120, pp. 303-313, 2017.
- [21] M. J. de Lemos and M. B. Saito, "Computation of turbulent heat transfer in a moving porous bed using a macroscopic two-energy equation model," *International Communications in Heat and Mass Transfer*, vol. 35, no. 10, pp. 1262-1266, 2008.
- [22] J. A. Ochoa-Tapia and S. Whitaker, "Momentum transfer at the boundary between a porous medium and a homogeneous fluid I. Theoretical development," *International Journal of Heat and Mass Transfer*, vol. 38, no. 14, pp. 2635-2646, 1995.
- [23] J. A. Ochoa-Tapia and S. Whitaker, "Momentum transfer at the boundary between a porous medium and a homogeneous fluid II. Comparison with experiment," *International Journal of Heat and Mass Transfer*, vol. 38, no. 14, pp. 2647-2655, 1995.
- [24] S. I. Park, "Performance analysis of a moving-bed heat exchanger in vertical pipes," *Energy*, vol. 21, no. 10, pp. 911-918, 1996.
- [25] R. Jackson, *The Dynamics of Fluidized Particles*. Cambridge University Press, New York, USA, 2000.
- [26] S. Whitaker, "Coupled transport in multiphase systems: A theory of drying," *Advances in Heat Transfer*, vol. 31, pp. 1-104, 1998.
- [27] K. Greider, *Invitation to physics*. Harcourt Brace Jovanovich Inc., 1973.
- [28] T. Anderson and R. Jackson, "Fluid mechanical description of fluidized beds," *Industrial & Engineering Chemistry Fundamentals*, vol. 6, no. 4, pp. 527-539, 1967.
- [29] C. Marle, "Ecoulements monophasiques en milieu poreux," *Rev. Inst. Francais du Petrole*, vol. 22, no. 10, pp. 1471-1509, 1967.
- [30] J. C. Slattery, "Flow of viscoelastic fluids through porous media," *AICHE Journal*, vol. 13, no. 6, pp. 1066-1071, 1967.
- [31] S. Whitaker, "Diffusion and dispersion in porous media," *AICHE Journal*, vol. 13, no. 3, pp. 420-427, 1967.
- [32] J. A. Ochoa-Tapia and S. Whitaker, "Heat transfer at the boundary between a porous medium and a homogeneous fluid," *International Journal of Heat and Mass Transfer*, vol. 40, no. 11, pp. 2691-2707, 1997.
- [33] W. Gray, "A derivation of the equations for multi-phase transport," *Chemical Engineering Science*, vol. 30, no. 2, pp. 229-233, 1975.
- [34] M. Quintard and S. Whitaker, "Transport in ordered and disordered porous media I: The cellular average and the use of weighting functions," *Transport in Porous Media*, vol. 14, no. 2, pp. 163-177, 1994.
- [35] M. Quintard and S. Whitaker, "Transport in ordered and disordered porous media II: Generalized volume averaging," *Transport in Porous Media*, vol. 14, no. 2, pp. 179-206, 1994.
- [36] M. Quintard and S. Whitaker, "Transport in ordered and disordered porous media III: Closure and comparison between theory and experiment," *Transport in Porous Media*, vol. 15, no. 1, pp. 31-49, 1994.
- [37] M. Quintard and S. Whitaker, "Transport in ordered and disordered porous media IV: Computer generated porous media for three-dimensional systems," *Transport in Porous Media*, vol. 15, no. 1, pp. 51-70, 1994.
- [38] M. Quintard and S. Whitaker, "Transport in ordered and disordered porous media V: Geometrical results for two-dimensional systems," *Transport in Porous Media*, vol. 15, no. 2, pp. 183-196, 1994.
- [39] L. Snyder and W. Stewart, "Velocity and pressure profiles for Newtonian creeping flow in regular packed beds of spheres," *AICHE Journal*, vol. 12, no. 1, pp. 167-173, 1966.
- [40] J. Sørensen and W. Stewart, "Computation of forced convection in slow flow through ducts and packed beds II. Velocity profile in a simple cubic array of spheres," *Chemical Engineering Science*, vol. 29, no. 3, pp. 819-825, 1974.
- [41] A. Eidsath, "Flow and dispersion in spatially periodic porous media: A finite element study," Master's thesis, University of California at Davis, 1981.
- [42] A. Zick and G. Homsy, "Stokes flow through periodic arrays of spheres," *Journal of Fluid Mechanics*, vol. 115, pp. 13-26, 1982.
- [43] A. Sangani and A. Acrivos, "Slow flow past periodic arrays of cylinders with application to heat transfer," *International Journal of Multiphase Flow*, vol. 8, no. 3, pp. 193-206, 1982.
- [44] M. J. MacDonald, C.-F. Chu, P. P. Guilloit, and K. M. Ng, "A generalized Blake-Kozeny equation for multisized spherical particles," *AICHE Journal*, vol. 37, no. 10, pp. 1583-1588, 1991.
- [45] H. E. Pacella, H. J. Eash, B. J. Frankowski, and W. J. Federspiel, "Darcy permeability of hollow fiber bundles used in blood oxygenation devices," *Journal of Membrane Science*, vol. 382, no. 1-2, pp. 238-242, 2011.
- [46] L. E. Sissom, "Fluid-to-particle heat transfer in a vertical moving bed of solids with interstitial fluid flow," PhD thesis, Georgia Institute of Technology, 1965.

ON THE AERODYNAMIC PERFORMANCE OF TAIL DEVICES FOR DRAG REDUCTION OF ROAD HEAVY VEHICLES

Mohammad Saeedi^{1,2}, Ali Tarokh¹, Alexander DERNY³, Kefu Liu¹

¹ Dept. of Mechanical Engineering, Lakehead University, Thunder Bay, ON, Canada

² Dept. of Aerospace Engineering, Amirkabir University of Technology, Tehran, Iran

³ Transtex-llc, Montreal, Quebec, Canada

Emails: mohammad.saeedi@aut.ac.ir, atarokh@lakeheadu.ca,

ABSTRACT— In the current research, a series of numerical simulations of the flow field around a typical tractor-trailer combination with and without a drag reduction device have been performed. The results of the simulations have been validated through comparison against available wind-tunnel data conducted at the National Aerodynamic LAB in Ottawa. It has been shown that the trailer wake region in proximity of the trailer rear surface is significantly influenced by the so-called trailer boat tail. This resulted in about 8.6% decrease in the vehicle aerodynamic drag and 4.6% fuel saving based on the traveling speed of 65 mph.

Keywords—*Drag Reduction Devices; Aerodynamic of Heavy Truck; Fuel Saving; Boat Tail; CFD*

I. INTRODUCTION

With the incredibly fast growth rate of energy demand in the world, increasing population of urban areas with dense human habitation and the influence of human activities on the global climate change, it is critically important to optimize and reduce the fuel consumption in different sectors as much as possible. One of the important sectors with a considerable share in fossil fuel consumption and greenhouse gas emission is heavy duty road transportation. In view of this, optimizing fuel consumption for this sector can be meaningfully influential and has been the subject of several studies.

One of the key elements to reduce the vehicle fuel consumption is to reduce the aerodynamic resistance (Drag) of the vehicle which can be accomplished through applying the geometric modification to the vehicle. Several fundamental researches have been

conducted to analyze and investigate the fundamental flow physics around simplified models. For instance, Serre et al. [1] conducted a brief review of which compared the results of several numerical simulations of turbulent flow around a standardized simplified car model (the so-called Ahmed body) based on large-eddy simulation (LES) and direct numerical simulation (DNS). Gurlek et al. [2] investigated the flow field and turbulent structures behind a bus model using PIV measurements. Atlaf et al. [3] numerically investigated the flow field around a simplified truck model and proposed an elliptical flap to be mounted on the back of their model. It was shown in their study that up to 11.1% of drag reduction can be achieved due to use of the optimal flap. McArthur et al. [4] also analyzed the flow field in the wake region of a generic truck and identified dominant frequencies in the wake region corresponding to a wake pumping mechanism and vortex shedding. Krajnovic and Fernandes [5] studied the flow field over a simplified vehicle model with active flow control. They observed a strong influence of side flaps installed on the model to the near-wake flow structures and the resulting drag.

It has been tried in different researches to aerodynamically optimize the vehicle and reduce its drag coefficient using different add-ons such as deflectors, side skirts, boat tails, and different trailer fairings. Hyams et al. [6] numerically investigated the turbulent flow field around a Generic Conventional tractor-trailer model based on the Reynolds-averaged Navier-Stokes (RANS) approach. They studied the effect of front spoilers, mud flaps and base flaps and validated their results through comparison of available wind-tunnel measurement data. Mohamed-Kassim and Filippone [7] numerically analyzed the potential drag reduction and consequent fuel saving of a heavy

road truck due to using different aerodynamic devices. They showed that the performance of the aerodynamic devices for vehicles in long-haul routes was twice as good as that for those in urban areas. They also showed that 1% to 9% of fuel cost saving could be obtained for different configurations of trucks and aerodynamic devices. Hwang et al. [8] conducted wind-tunnel experiments and numerical simulations and reported approximately 5% reduction in drag coefficient for different types of side skirts to be mounted on a typical tractor-trailer combination. Schmidt et al. [9] conducted wind-tunnel experiments and investigated the influence of using base flap to reduce the drag of a rectangular bluff body. They showed that the flap had to be sufficiently long to shift the wake structures far enough downstream of the bluff body to reduce the drag. It was shown in their study that the drag force could be reduced up to 13% due to using optimum values for the length and angles of base flaps. Lee and Lee [10] conducted wind-tunnel experiments and investigated the flow field around a specific 15-ton Hyundai truck and proposed a boat tail with lower inclined air deflector for the mentioned vehicle. They showed a drag reduction of approximately 9% could be reached when the boat tail was mounted on the truck.

In the current research, we aim to investigate the effectiveness of a specific trailer boat tail in reducing the aerodynamic drag and consequent fuel saving of a generic tractor-trailer model. The flow field and turbulent coherent structures will be also analyzed. In the remaining of the paper, the test case and details of the numerical simulation will be presented followed by qualitative and quantitative result analysis. Finally, major findings and conclusions are reviewed.

II. TEST CASE, NUMERICAL SIMULATION, AND BOUNDARY CONDITIONS

The test case for the current research is based on a recent wind-tunnel experiment conducted at the National Research Council (NRC) Canada wind tunnel using a 30% scaled generic tractor-trailer combination with and without tail devices which will be referred to as boat tail henceforth [11]. Boat tail can be identified by three or four flat panels attached to the rear edges of the trailer to form a closed cavity. The trailer has 6.58 m length, 1.3 m height and 0.966 m width and the Reynolds number based on the free-stream velocity $u_\infty = 29.06$ m/s and the square root of its frontal area $d = 0.98$ m is 1.89×10^6 . According to standards of the Society of Automotive Engineers (SAE) [12], the Reynolds number of the test case must be above one million which has been considered in the

current research. Also, the size of the computational domain was chosen according to the same SAE standards [12]. Figure 1 schematically shows the computational domain, boundaries and both models tested. At the inlet of the flow domain and also for side boundaries (span wise direction), Dirichlet boundary condition has been utilized. At the outlet, Neumann boundary condition is used. For the upper boundary, zero gradient normal to the boundary plane is applied. No slip boundary condition is applied to all solid surfaces including the ground plane and also the tractor-trailer surfaces.

The numerical simulations are carried out by accurate time-dependent solution of the set of coupled Reynolds-averaged Navier-Stokes (RANS) equations presented below:

$$\frac{\partial \bar{u}_j}{\partial x_j} = 0, \quad (1)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\nu \frac{\partial \bar{u}_i}{\partial x_j} - \overline{u'_i u'_j} \right), \quad (2)$$

where \bar{u}_i and \bar{p} are the mean (or, ensemble-averaged) velocity and pressure, ρ and ν are the density and kinematic viscosity of the fluid and $\overline{u'_i u'_j}$ is the so-called Reynolds stress tensor. In order to close the set of RANS equations, the K-omega SST model has been used to model the Reynolds stress tensor and close the system of governing equations.

Numerical simulations have been performed using the open-source package of Field Operation and Manipulation (open-Foam). A second-order finite volume scheme based on a non-uniform collocated grid system was applied to discretize the governing equations and the SIMPLE algorithm has been used for coupling the velocity and pressure fields. The second-order Backward-Euler scheme was used to advance the velocity field over a single time step. A line Gauss-Seidel (LGS) solver was used for solving the discretized momentum equations and the Poisson equation was solved using an algebraic multigrid method to obtain the new pressure field.

The domain has been discretized using approximately 5.2 and 5.7 million tetrahedral cells for the baseline and tail configurations, respectively. The grid independency has been also checked using the overall drag coefficient not to be more than 3% different for 5.5 and 6 million cells for the baseline and tail configurations, respectively. Figure 2 presents an isometric view of the discretized domain around the trailer and the lower ground. In order to perform

parallel computing, the computational domain is divided into 12 sub-domains using the Scotch domain decomposition scheme. Correspondingly, 280 CPU-Hours were spent to perform each simulation approximately. All the computations were performed using one of the cluster systems of Westgrid Canada (namely Orcinus) based at the University of British Columbia's campus.

III. RESULTS DISCUSSION

In order to provide a general understanding of the flow field, its evolution around the tractor-trailer and dominant features captured, the lateral and top view of the stream-wise velocity contours inside planes $y = 0$ and $z = 0.9$ passing through the tractor-trailer are shown in Figs. 3(a) and 3(b), respectively. It is evident in both figures that a stagnation region is formed over the windshield of the trailer as the free-stream flow strikes the trailer. The flow then accelerates over the curved surface of the trailer and reaches its maximum value. It is clearly grasped in the figure that a thin boundary layer grows over the trailer surface. The boundary layer over the trailer surface has a significant contribution to the skin friction drag of the trailer. The upper view of the rear wake flow and the low-velocity region inside the wake has been clearly shown in the Fig. 3(b). This rear wake behind the trailer is one of the main sources of the pressure drag acting on the tractor trailer. The objective of the design of the trailer boat-tail is mainly to reduce the vehicle aerodynamic drag and consequently its fuel consumption. Following the classical approach, the drag coefficient is defined as follows:

$$C_D = \frac{\text{Drag force}}{0.5 \rho u_\infty^2 A_{\text{ref}}}, \quad (3)$$

where C_D is the drag coefficient, ρ is the air density, u_∞ is the free-stream velocity and A_{ref} is the frontal area of the model. The drag coefficient of the baseline and tail configurations are calculated in two flow angles (0 and 4.5 degrees yaw angles, respectively) and then converted to the so-called Wind-Averaged Coefficient of Drag (WACD) using the following equation [13]:

$$WAC_D(U_g) = \frac{1}{2\pi} \int_0^{2\pi} C_{D,w}(\eta) \left[1 + \left(\frac{u_{\text{avg}}}{u_g} \right)^2 + 2 \left(\frac{u_{\text{avg}}}{u_g} \right) \cos \theta \right] d\theta, \quad (4)$$

in which $\eta = \tan^{-1} \left[\frac{(u_{\text{avg}}/u_g) \sin \theta}{1 + (u_{\text{avg}}/u_g) \cos \theta} \right]$, is the vehicle-referenced wind angle, U_{avg} is the mean terrestrial

wind speed which is approximately 7 mph (3.13 m/s) for North America, U_g is the vehicle ground speed and θ is the terrestrial wind angle. The flow angles are chosen according to the Greenhouse Gas Emissions and Fuel Efficiency Standards (GHG2) of the Environmental Protection Agency (EPA). The reduction in wind-averaged drag coefficient is a key parameter in order to have an estimation of potential fuel saving caused by the drag reduction and to indicate the efficiency of the boat tail model. The overall fuel saving due to the reduced drag force in different operating conditions is stated through the following equation.

$$\Delta f[\%] = \frac{1}{CF} \frac{\Delta WAC_D}{WAC_D} \times 100, \quad (5)$$

where CF is a conversion factor associated with the equivalent ground speed from which the wind-averaged-drag coefficients takes the value 1.85 for the free-stream velocity of 29.06 m/s tested in this research [10]. Table 1 summarizes the values of drag coefficient C_D for the baseline and tail configurations in two different flow angles, the wind-averaged coefficient of drag and also the drag reduction percentage obtained from the simulation and wind-tunnel experiments. Evidently, there is a very good agreement between the results of the numerical simulation and those of the experimental measurement. It is also observed that the boat tail model can reduce the aerodynamic drag coefficient of the tractor-trailer by approximately 8.6%. This would result in 4.6% fuel saving which is quite significant for road vehicles. This level of drag reduction is mainly due to the enhanced pressure recovery in the rear region of the trailer which occurs when the boat tail is mounted. In other words, the boat tail reorganizes the wake pattern behind the trailer and consequently decreases the vehicle form drag. In order to investigate this change of the vehicle wake flow, Fig. 4 presents the top view of time-averaged streamlines superimposed with the streamwise velocity contours behind the trailer and inside the plane $z = 0.9$. As it is evident in the figure, two large counter-rotating vortices are present within the recirculation bubble behind the trailer. Presence of two counter-rotating vortices inside the wake region behind the baseline configuration of the trailer is a characteristic feature of bluff body flows and has been clearly shown that it in turn creates a large suction region and enhances the form drag on the trailer. Also, these large vortices cause high-velocity gradient and shear stress in the boundary of the wake region and free-stream flow which will consequently consume the kinetic energy of the main flow. It is seen in Fig. 4b that the boat tail mounted on the trailer reduces the size of two counter-rotating vortices and also the recirculation bubble by

approximately 10%. Figure 5 presents the lateral view of the flow streamlines superimposed with the streamwise velocity contours at $y = 0.2$. Two large vortices behind the baseline configurations in the wake region are apparent in Fig. 4(a) in the wake region. However, inclusion of the boat tail dislocates the upper vortex to the top surface of the top boat tail panel with a considerably reduced size shown in Fig. 5(b). Also, the lower large vortex observed in Fig. 5(a) is broken up into several smaller vortices. This consequently reduces the size of the wake region behind the trailer (consistent with the previous discussion) and decreases the form drag on the trailer.

Figures 6(a) and 6(b) present the cross-stream profiles of the static gauge pressure behind the trailer at the selected elevation $z = 0.8$ and at different streamwise locations. As it is evident in the figure, close to the trailer rear surface at $x = 7.2$ m, the maximum pressure suction behind the baseline configuration has been decreased by approximately 20% when the boat tail is mounted behind the trailer (from -100 Pa to -80 Pa) which would result in a considerable reduction in the form drag. However, for further downstream after this location, no significant difference in the cross-stream profile for two cases is observed. This reflects the critical importance of the wake region very close to the trailer rear surface when the drag reduction is aimed.

The structure of the three-dimensional wakes can also be visualized using the so-called Q-criterion. The Q criterion defined in equation (6) is a useful tool for studying the coherent structures in turbulent flows especially from a qualitative point of view. This criterion identifies vortices as flow regions with positive second invariant of velocity gradient. It is actually a local measure of the excess rotation rate relative to the strain rate.

$$Q = 0.5(\mathbf{r}_{ij}\mathbf{r}_{ij} - \mathbf{s}_{ij}\mathbf{s}_{ij}) \quad (6)$$

where $\mathbf{r}_{ij} = 0.5(\frac{\partial u_j}{\partial x_i} - \frac{\partial u_i}{\partial x_j})$ and $\mathbf{s}_{ij} = 0.5(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j})$ are rotation and strain rate tensors respectively, and \mathbf{u}_i represents instantaneous velocity field. Figure 7 shows the vortical structures using the Q-criterion around the entire tractor trailer model in both configurations. It is also evident in the figure that when the boat tail is mounted behind the trailer, the large vortical structures behind the trailer is broken to several small vortices. As discussed before, this would result in a smaller recirculation region and consequently is expected to reduce the induced pressure drag on the trailer.

IV. CONCLUSIONS

The flow field around a typical tractor-trailer combination with and without a boat tail mounted on the rear surface of the trailer was numerically investigated based on the RANS approach. The simulations were performed for the ground speed of 65 mph and two different yaw angles of 0 and 4.5 degrees. It was shown that based on the obtained results inclusion of the trailer boat tail can lead to approximately 8.6% reduction in the vehicle drag which was equivalent to 4.6% decrease in the vehicle fuel consumption. This level of drag reduction was mainly due to the decreased wake region in proximity of the trailer rear surface and re-organized wake structures. Also, it was observed that pressure suction behind the trailer was reduced by approximately 20% which leads to significant decrease in the vehicle form drag.

ACKNOWLEDGMENT

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

- [1] E. Serre, M. Mingue, R. Pasquetti, E. Guilmineau, G. B. Deng, M. Kornhaas, M. Schäfer, J. Fröhlich, C. Hinterberger, W. Rodi. On simulating the turbulent flow around the Ahmed body: A French-German collaborative evaluation of LES and DES. *Computer and Fluids*, Vol 78, pp. 10-23, 2013.
- [2] C. Gurlek, B. Sahin and G. M. Ozkan. PIV studies around a bus model. *Experimental Thermal and Fluid Science*, Vol. 38, pp. 115-126, 2012.
- [3] A. Altaf, A. A. Omar and W. Asrar. Passive drag reduction of square back road vehicles. *Journal of Wind Engineering and Industrial Aerodynamic*. Vol. 134, pp. 30-43, 2014
- [4] D. McArthur, D. Burton, M. Thompson and J. Sheridan. On the near wake of a simplified heavy vehicle. *Journal of Fluids and structures*, Vol. 66, pp. 293-3146, 2016.
- [5] S. Krajnovic' and J. Fernandes. Numerical simulation of the flow around a simplified vehicle model with active flow control. *International Journal of Heat and Fluid Flow*, Vol 32, pp.192-200, 2011.
- [6] D. G. Hyams, K. Sreenivas, R. Pankajakshan, D. S. Nichols, W. R. Briley and D. L. Whitfield. Computational simulation of model and full scale Class 8 trucks with drag reduction devices. *Computer and Fluids*, Vol. 41, pp. 27-40, 2011
- [7] Z. Mohamed-Kassim and A. Filippone. Fuel savings on a heavy vehicle via aerodynamic drag reduction. *Transportation Research Part D*, Vol. 15, pp. 275-284, 2010.

- [8] B. Geun Hwang, S. Lee, E. Lee, J. J. Kim, M. Kim, D. You and S. Lee. Reduction of drag in heavy vehicles with two different types of advanced side skirts. *Journal of Wind Engineering and Industrial Aerodynamics*, Vol.155, pp. 36-46, 2016.
- [9] H.-J. Schmidt, R. Woszidlo, C. N. Nayeri and C. O. Paschereit. Drag reduction on a rectangular bluff body with base flaps and fluidic oscillators. *Experiments in Fluids*, Vol. 56, pp.151-167, 2015
- [10] E. J. Lee and S. J. Lee. Drag reduction of a heavy vehicle using a modified boat tail with lower inclined air deflector. *Journal of Visualization*, Vol. 20, pp. 743-752, 2017.
- [11] B. R. McAuliffe, A. S. Wall and A. J. Kirchhefer. SmartWay Verification Test Results for Transtex Trailer Skirts and Tail Devices using the NRC 9 mWind Tunnel. Technical report LTR-AL-2018-0006, 2018.
- [12] SAE Technical report J2966, Guidelines for Aerodynamic Assessment of Medium and Heavy

Commercial Ground Vehicles Using Computational Fluid Dynamics, 2013.

- [13] SAE Technical report J1252, SAE Wind Tunnel Test Procedure for Trucks and Busses, 2012.

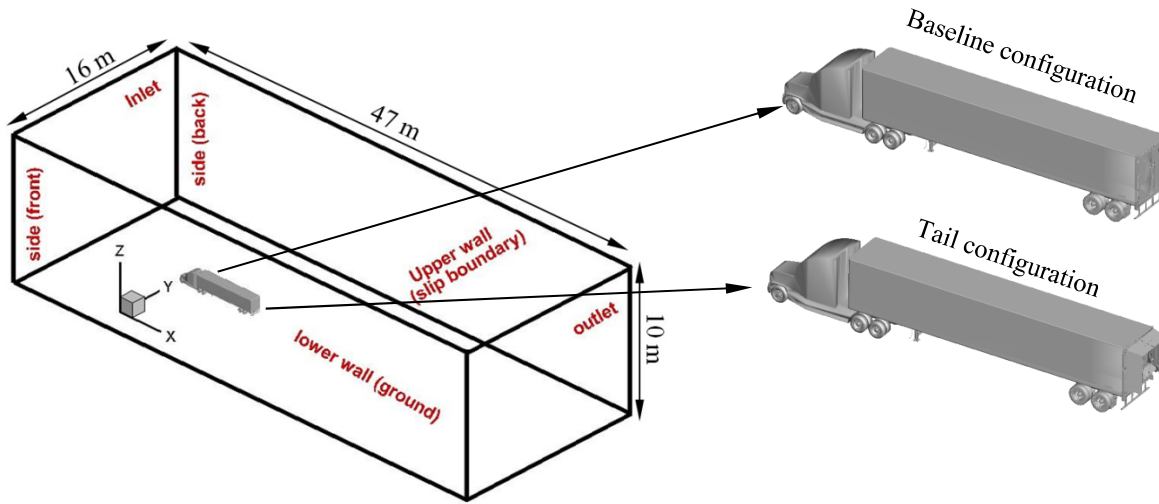


Fig. 1: Schematic of the computational domain, boundaries, coordinate system and models

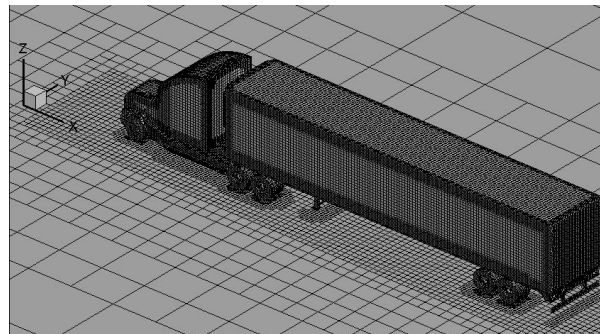


Fig. 2: An isometric view of the domain discretization for the baseline tractor-trailer and the lower ground

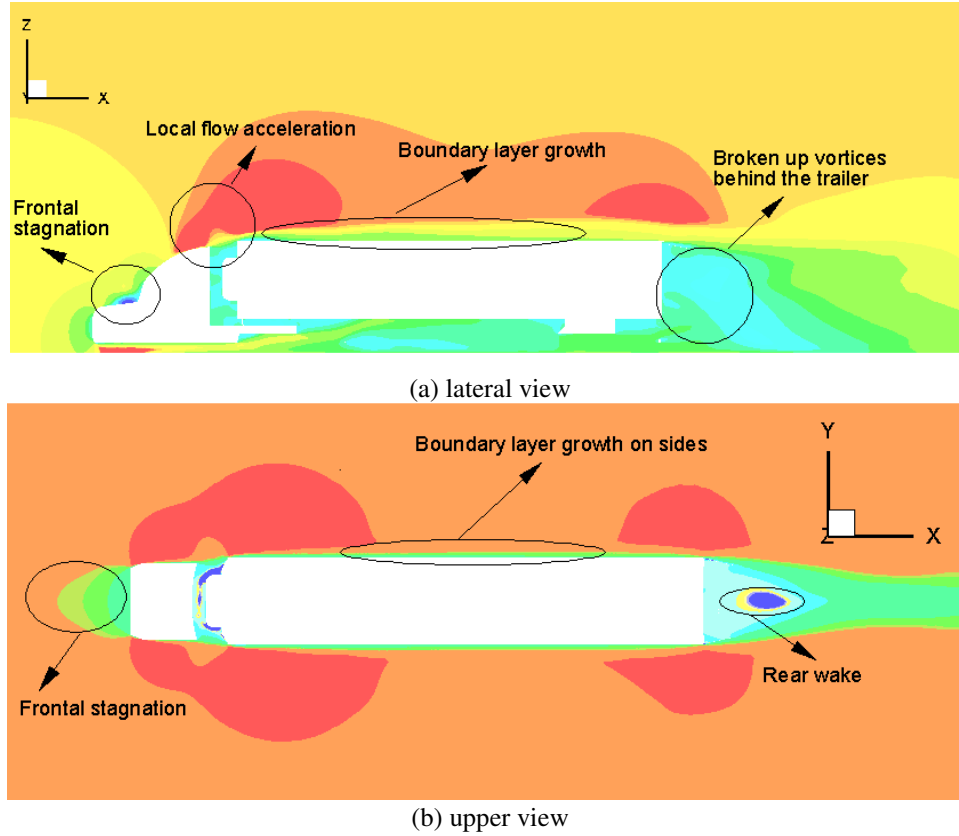


Fig. 3 Streamwise velocity contours around the baseline configuration and corresponding stagnation and wake regions

Table 1. Values of drag coefficients for the baseline and tail configurations obtained from the simulation and the experiment

Flow angle (°)	C_D (baseline, simulation)	WAC_D (baseline, simulation)	WAC_D (baseline, experiment)	C_D (boat tail, simulation)	WAC_D (boat tail, simulation)	WAC_D (boat tail, experiment)	ΔWAC_D [%] (simulation)	ΔWAC_D [%] (experiment)
0	0.564	0.567	0.565	0.513	0.517	0.520	8.6 %	8.1%
+4.5	0.566			0.520				

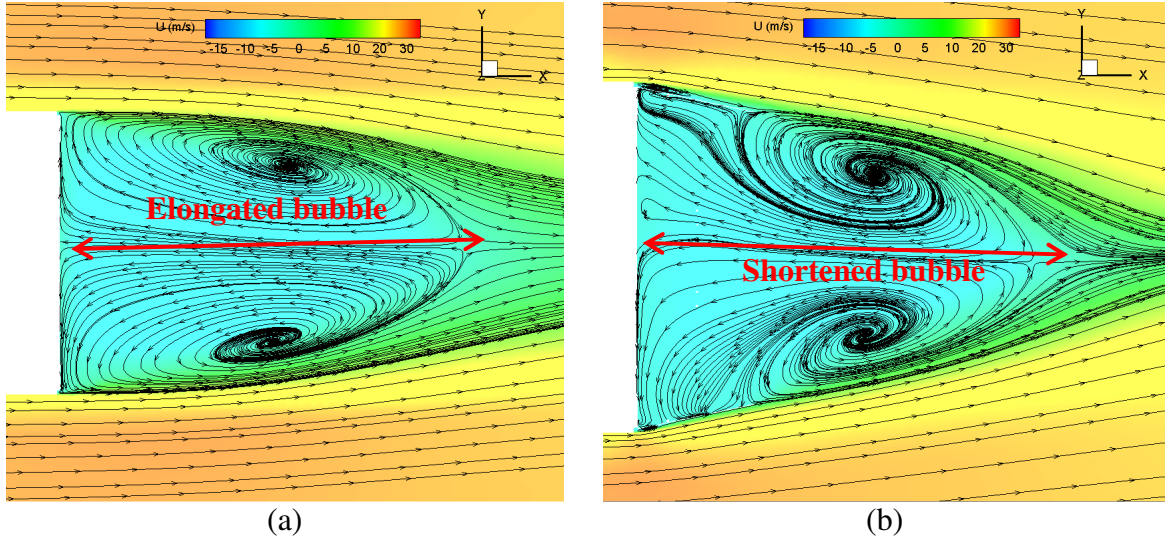


Fig. 4 Streamlines and velocity contours on the z -plane at behind the trailer (a) baseline, (b) boat tail

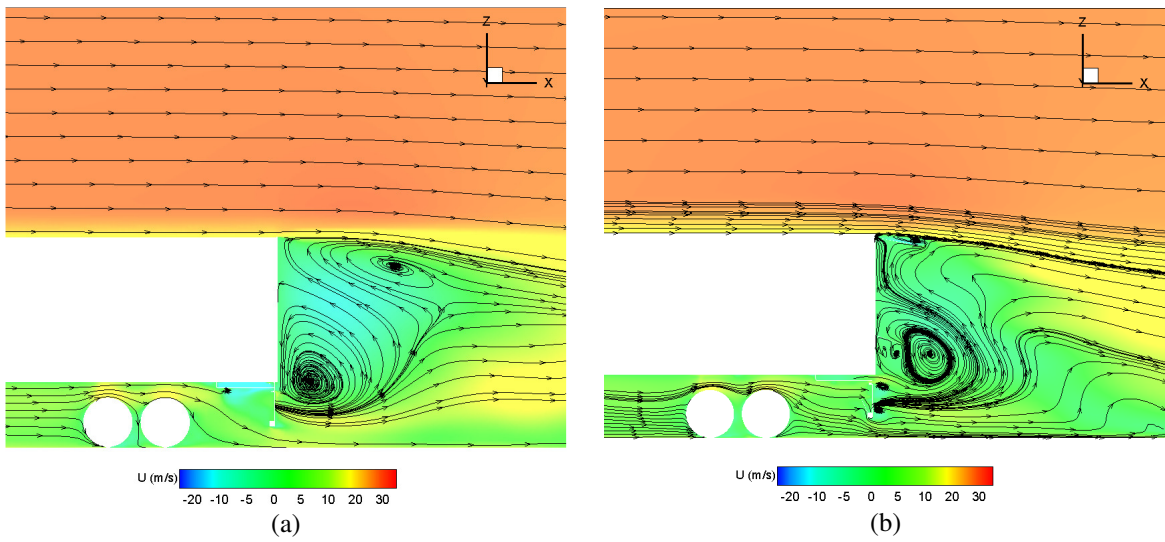
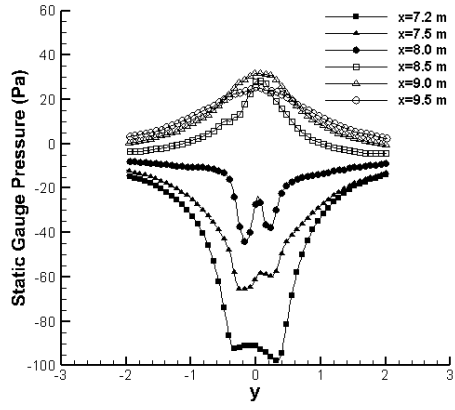
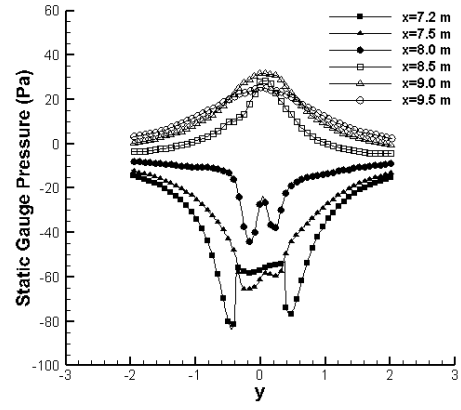


Fig. 5 Streamlines and velocity contours on the y -plane (a) Baseline, (b) Boat tail

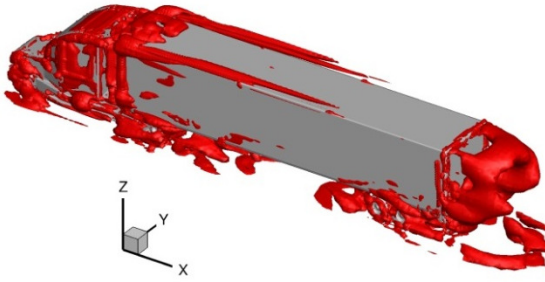


(a)

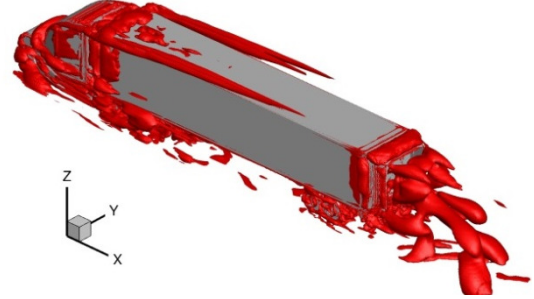


(b)

Fig. 6: Coherent flow structures visualized using iso-surfaces of Q-criterion (a) baseline, (b) boat tail



(a)



(b)

Fig. 7: Coherent flow structures visualized using iso-surfaces of Q-criterion (a) baseline, (b) boat tail

Experimental Investigation of an Archimedes Screw Pump

Lyons, Murray; Simmons, Scott; Lubitz, William¹

School of Engineering
 University of Guelph
 Guelph, Ontario, Canada
¹wlubitz@uoguelph.ca

Abstract—Tests were run on a laboratory-scale Archimedes screw pump and data collected regarding the pumping efficiency of the screw pump at various lower end (supply basin) water levels and screw rotation rates. Comparisons are made to the empirically and analytically-derived guidelines found in the literature regarding the optimal lower end water levels and rotation rate at which to run an Archimedes screw pump. The guidelines regarding optimal lower end water levels were found to hold for a subset of cases roughly matching those found in the literature. The guidelines regarding rotation rates were found not to hold for a screw on the scale of the one tested, which is smaller than normally used in operational screw pumps.

Keywords—*Archimedes screw; hydrodynamic screw; Archimedian screw; screw pump; efficiency*

I. INTRODUCTION

The Archimedes screw has been used since antiquity as a pump, notably by Archimedes himself in Ptolemaic Egypt in the mid-third century BCE [1], and some evidence suggests that it was also used much earlier than this in the Assyrian Empire under the reign of King Sennacherib (704-681 BCE) in the 7th century BCE [1].

Today, Archimedes screws are used extensively, for example as pumps for low-lying land drainage systems [2], wastewater treatment plant pumps, pumps for irrigation systems, and fish conveyors [3]. Recently, (within the last 20 years), the Archimedes screw has also seen use as a hydropower generator [4][5][6].

An Archimedes screw is constructed by winding one or more helical surfaces (“flights”) around a central shaft (Fig. 1). Typically 3 or more flights are used. The screw sits in a cylindrical trough, which may be open at the top or a fully closed tube. In some cases the trough is connected to the outer edge of the helical flights such that the trough rotates with the screw, however most applications use a fixed trough with a small gap between it and the outer edges of the flights, due to construction and sealing concerns related to the former setup. When used as a pump, the lower end of the inclined screw is partially immersed in water, and the screw is rotated so that water is trapped between two consecutive sets of flights. The volume between two adjacent flights is termed a “bucket”. As the screw continues to

rotate, this “bucket” of water is translated along the screw and therefore lifted upward, until it reaches the top of the screw and is emptied into the upper reservoir.

A screw may be defined by the dimensions shown in Fig. 1, and defined in Table 1.

TABLE I. ARCHIMEDES SCREW DIMENSIONS

Symbol	Dimension
D	Outer Diameter, m
d	Inner Diameter, m
L	Length, m
S	Pitch, m
N	Number of flights, -
β	Inclination angle from horizontal, °

Despite their extensive use, there is relatively little literature examining the performance of screw pumps. Among the most notable literature known to the authors is [7], which gives detailed instructions on how to design a screw pump, based on heuristics and field experience, though no field data is presented in it. The work cited is an English translation of a document originally written in German. Another notable work is [8], which presents a mathematical optimization of an Archimedean screw pump design, assuming a few idealizations of the screw geometry (for example, infinitely thin flights). An earlier work [9] gives guidance on recommended operating conditions for Archimedes screw pumps, and is in turn referenced extensively by many later authors, such as [7].

A key assumption throughout the literature is that the Archimedean screw pump should be operated with a lower basin water level which does not vary significantly, and which is sufficient to submerge approximately 75 percent of the lower, or intake, end. This is thought to be required to allow the flights to efficiently entrap large volumes of water as the screw turns, in order to lift it [7].

To the authors’ knowledge, only [9] presents any data regarding the effects of lower basin level on screw efficiency, and this is a small dataset looking at a single screw at a fixed rotation rate and a relatively small range of lower end fills. In this paper, data showing mechanical efficiency of a laboratory-

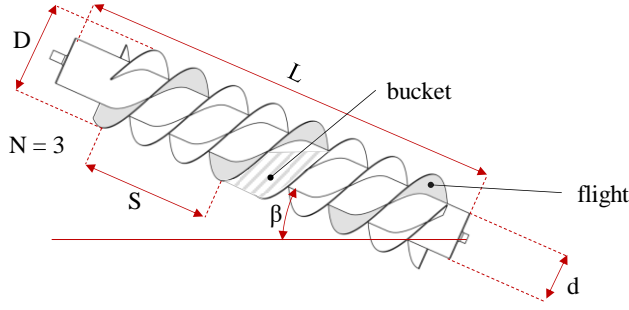


Figure 1 - Archimedes Screw Dimensions

scale Archimedes screw pump is presented as a function of both lower basin level and rotation rate. Comparisons are made between the collected data and empirically and analytically derived guidelines found in the literature regarding the optimal lower end conditions and rotation rate at which to run an Archimedes screw pump.

The depth that the lower end of the screw is submerged may be represented using a ratio with the height of the screw intake opening, so that the lower water fill ratio, ψ_L (Fig. 2), may be computed as

$$\psi_L = \frac{h_L}{D \cos(\beta)} \quad (1)$$

where h_L is the lower water level in meters, from a base aligned with the lowest point on the screw inlet. The term $D \cos(\beta)$ gives the vertical height of the intake end of the screw. Similarly, the depth that the upper end of the screw is submerged in the receiving basin may be represented as

$$\psi_U = \frac{h_U}{D \cos(\beta)} \quad (2)$$

where h_U is the upper water level in meters, and ψ_U is the upper water fill ratio.

II. EXPERIMENTAL SETUP

Experiments were conducted at the University of Guelph Archimedes screw testing facility. This facility is set up to allow screws to be tested both as pumps and as generators. Figs. 3 and 4 show the testing system. In pumping mode, the screw (A) is placed in the experimental setup such that the lower end is submerged in the lower basin (B), and the upper end extends to the upper basin (C). The screw is rotated by a SEW gearmotor (model: FA27B DRS71S4) controlled by a SEW frequency inverter (SEW MCLTEB00042B1140) capable of maintaining a wide range of selected rotation speeds regardless of load. This causes the screw to pump water from the lower basin to the upper basin. An adjustable weir (D) at the back of the upper basin allows for fine control over the depth of water in the upper basin. Water flows over this weir into a well (E), and then through an inline flowmeter (F) (Omega FTB740) before reentering the lower basin.

The gear motor is supported by a metal frame with a load cell (Omega LC703-25) placed such that the torque of the motor turning the screw is transmitted through the load cell. This load

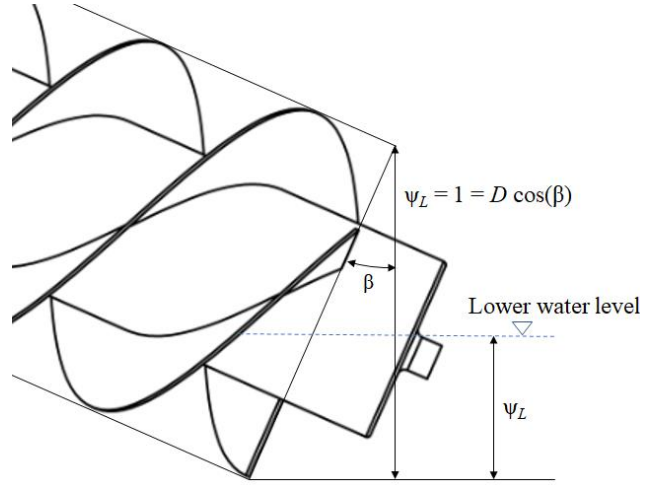


Figure 2 – Definition of non-dimensional intake water depth ψ_L

cell is positioned at a known distance from the axis of rotation and is used to measure the mechanical torque required to turn the screw (Fig. 5).

The rotation rate of the screw was measured by placing a magnet on the shaft of the screw such that every rotation the magnet would trip a solid-state magnetic switch mounted on the fixed frame of the screw. Readings were taken for 30 seconds per point, and pulse signals from the magnetic switch were recorded. The spacing between leading edges in this signal allowed the rotation rate to be calculated. The flowmeter created a similar pulse-width signal, which allowed the flow to be calculated via an empirical calibration curve created through tests prior to collecting experimental data. The upper and lower basin levels were recorded using Keller Acculevel depth gauges, as well as manually checked via graduated measures mounted in each basin. All electrical signals were recorded using a data acquisition device (DAQ; National Instruments USB-6009) connected to a laptop computer. For non-pulsed signals (the depth transducers and the load cell), signals were averaged over the 30 second recording time, to remove possible bias due to the screw not being perfectly radially symmetric.

III. EXPERIMENTAL METHOD

The screw was installed in the experimental setup inclined at 24.45° from the horizontal. The upper water level was maintained such that $\psi_U = 0.21 \pm 0.014$ for all cases. This upper water level was chosen such that the water exiting the screw with $\psi_L = 0.6$ and a moderate rotation rate (taken as the “midpoint” of testing conditions) would smoothly exit into the upper basin at the same depth as the upper basin. This also generally agrees with [10], which states that the upper fill should be between

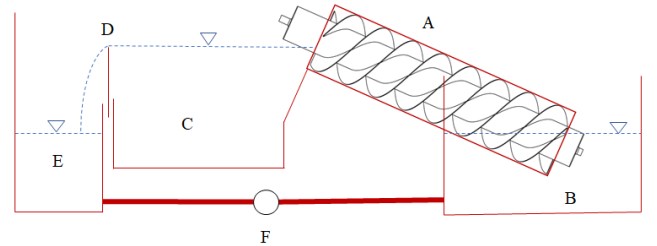


Figure 3 – Pumping mode diagram



Figure 4 – Image of screw in testing facility

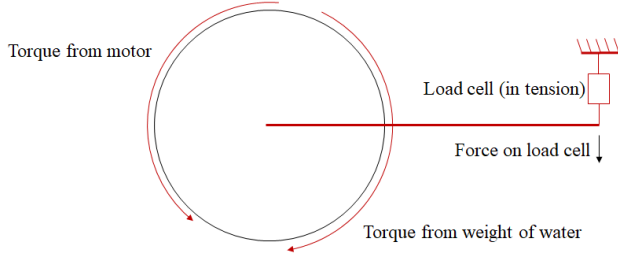


Figure 5 – Diagram of load cell measuring torque

$0.15 D$ to $0.33 D$ above the point where $\psi_U = 0$; this equals $0.165 \leq \psi_U \leq 0.363$ for the current setup. For each set of measurements in the series, ψ_L was set to one of 20, 40, 60, 80, or 100 percent, and the rotation rate, n , varied through a range from 5 rev/min to 100 rev/min, in increments of 5 rev/min. At each rotation rate, the water levels were manually verified, and a reading was taken for 30 seconds using the DAQ. The range of ψ_L tested were chosen to give good resolution into the nature of the effects of ψ_L while keeping the total time required to perform the lab tests reasonable.

For low lower basin depths the very low rotation rates were not able to be achieved while also maintaining the upper water level, due to minor leakage around the adjustable weir. For higher lower basin levels, the very high rotation rates were not able to be achieved while also maintaining the upper water level, due to the flow exceeding the maximum flow rate which could be realized through the return pipe.

The mean uncertainty for the upper and lower basin depth measures was 0.8 % (based on manufacturer specs). The mean uncertainty in the torque measurement was 1.46 %. Due to the high sampling frequency (1000 Hz) used to measure the rotation rate, the uncertainty on the rotation rate measure was considered negligible. The uncertainty in the mechanical power calculation was 1.46 %. The uncertainty in the flow rate is estimated at 5 %. The uncertainty on the efficiency calculation was calculated as

5.4 % of the total efficiency value, based on the percent uncertainties for the depths, torque, rotation rate, and flow measurements.

IV. RESULTS

The power required to lift the water pumped by the Archimedes screw may be calculated using the formula

$$P = \rho g h Q \quad (3)$$

where ρ is the density of the water, g is the gravitational constant, h is the difference in height in meters between the free surface of the upper basin and the free surface of the lower basin, and Q is the volumetric flow rate in cubic meters per second through the screw as it is lifting the water. The efficiency, η , of the screw at lifting the water may be calculated using

$$\eta = \frac{P}{P_{in}} \quad (4)$$

where P_{in} is the mechanical power applied to the screw by the motor. This may be calculated using

$$P_{in} = \omega \tau \quad (5)$$

where ω is the rotation rate of the screw in radians per second, and τ is the torque in Newton-meters transmitted to the screw from the motor.

Fig. 6 shows the rotation rate vs flow (Q) through the screw for all 5 lower end conditions tested. Fig. 7 shows n vs η , and Fig. 8 shows Q vs η . These are the three main views of the data collected.

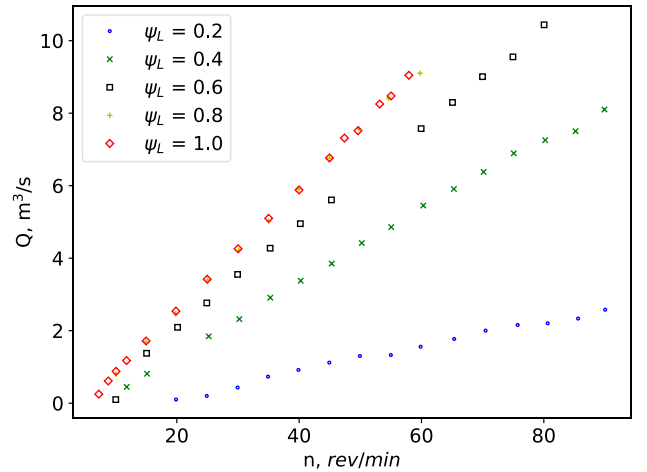


Figure 6 – Rotation Rate vs Flow Rate

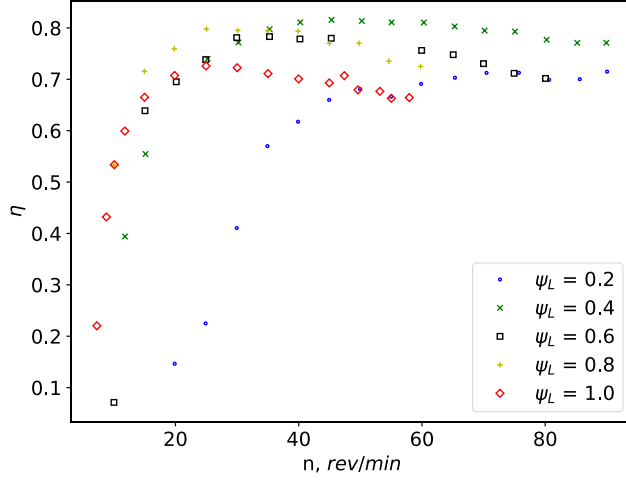


Figure 7 – Rotation Rate vs Efficiency

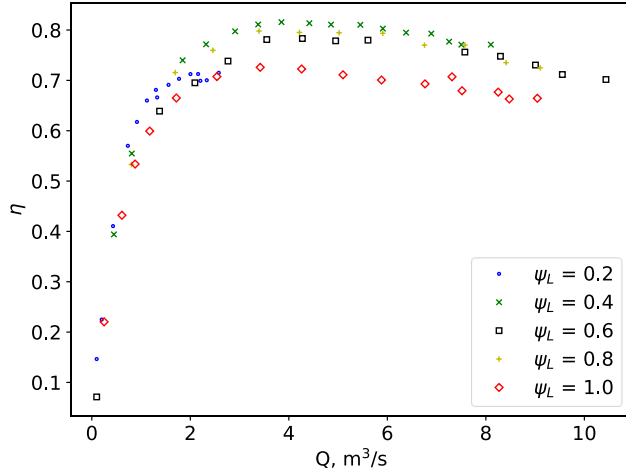


Figure 8 – Flow Rate vs Efficiency

V. DISCUSSION

Examining Fig. 5, for $\psi_L < 0.8$, it appears that greater ψ_L means that a lower rotation rate is required to achieve a given flow through the screw. At $\psi_L \geq 0.8$, the flow rate becomes independent of ψ_L and $Q = f(n)$, suggesting that a lower basin level above $\psi_L \geq 0.8$ does not give any practical benefit. As the head at $\psi_L = 1.0$ is less than at $\psi_L = 0.8$, but the same flow rate is achieved at a given rotation rate, it is expected that the $\psi_L = 1.0$ condition will be less efficient than the $\psi_L = 0.8$ condition. Fig. 6 shows that this is the case. Additionally, for each ψ_L there is a clear drop in efficiency as the rotation rate decreases. Interestingly, for $\psi_L = 0.2$, this drop seems to occur at a higher rotation rate (approx. 40 rev/min) than the others (approx. 25-30 rev/min). This is may be due to the $\psi_L = 0.2$ case having very low flow rates even at moderate rotation rates. This occurs because the end of the screw is unable to capture water effectively at this very low supply basin depth, resulting in less efficient pumping.

Reference [7] states that that the optimum value for ψ_L occurs such that the water level in the lower end just touches the top edge of the inner cylinder. This point may be found using

$$\psi_{LO} = \frac{D+d}{2D} \quad (6)$$

For the screw used in these experiments, this computes to $\psi_{LO} = 0.766$.

Based on (6), it is expected that the highest peak efficiency of the cases tested should occur at $\psi_L = 0.8$. The peak efficiency at $\psi_L = 0.8$ ($\eta = 0.798$) is higher than the efficiencies at $\psi_L = 0.2$ ($\eta = 0.715$), $\psi_L = 0.6$ ($\eta = 0.783$), and $\psi_L = 1.0$ ($\eta = 0.721$), however the efficiency at $\psi_L = 0.4$ (0.816) slightly exceeds the efficiency at $\psi_L = 0.8$. It should be noted that [7] used a fixed rotation rate to come up with (6). Plotting ψ_L vs η for a selection of n (Fig. 9), it is clear that at high and low ψ_L , there is a decrease in efficiency, particularly at lower rotation rates. This matches the observations given in [7], in particular that the peak efficiency occurs near $\psi_L = 0.8$ and drops sharply with decreasing ψ_L . As the rotation rate increases, the change in efficiency with changing ψ_L is less drastic. At higher rotation rates, the value of ψ_L where peak efficiency occurs shifts towards $\psi_L = 0.4$. This suggests that the optimal ψ_L is dependent on the screw rotation rate, though the commonly reported convention (by [7] and [10], for example) that ψ_L be kept below 1.0 for optimum performance does still hold.

When efficiency is plotted against flow (Fig. 7), for $\psi_L < 1$ all the curves overlap to some degree. This suggests that η is not directly dependent on Q , but rather that Q and η are both dependent on the rotation rate and ψ_L . It also reaffirms that fills above $\psi_L = 0.8$ are less efficient than those ≤ 0.8 .

Reference [7] recommends using a rotation rate presented by [9] as the maximum rotation rate for a screw pump as a function of the outer diameter:

$$n_{max} = \frac{50}{\sqrt[3]{D^2}} \quad (7)$$

Note that this is a non-homogenous equation. The units of D must be in meters, and the units of n_{max} will be in rev/min.

Computing this recommended maximum rotation speed for the screw tested here gives $n_{max} = 107.7$ rev/min. This is significantly higher than the rotation rates at which peak efficiency was found in the lab; this implies that the maximum rotation rate for the screw tested should be much lower than that given by (7). The lab screw tested is much smaller than any of the screws examined by [9]. This is likely why the empirically-derived n_{max} to keep a screw running near peak efficiency does not match well with the data collected. For the screw that was tested, a maximum rotation rate closer to 35 rev/min, would be more effective in keeping near peak efficiency across all ψ_L . According to (7), this rotation rate would be suitable for a screw approximately five times as large as the one tested. This suggests that (7) is not a good model for screws the size of the one tested.

VI. CONCLUSIONS

The experiment presented here provides data on the effects of lower water level and rotation rate on the efficiency of Archimedes screw pumps. This expands on data available in the literature by examining multiple rotation rates and low fill levels.

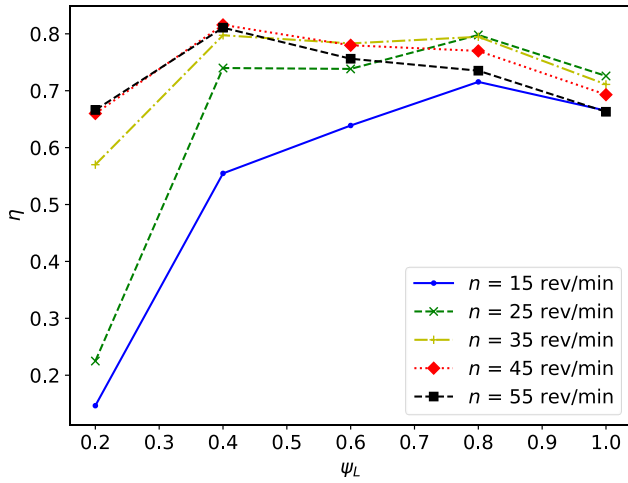


Figure 9 – ψ_L vs η for a selection of n

The experimental results were consistent with the commonly accepted guideline that screw pumps should be operated with $\psi_L \leq 0.8$ for optimal efficiency. However, the recommendations in the literature regarding optimal rotation rate and optimal lower end fill are shown to be valid for only a subset of conditions. In particular, the recommended n_{max} is shown to not scale well to small screw pumps, and the recommended ψ_{LO} is shown not to hold for high rotation rates.

While the trends observed in the data are well behaved, there is not enough data present to model these trends as a function of screw geometry, which would be required in order to build a robust predictive model for picking an optimal combination of rotation rate and lower end fill for any given screw geometry. In particular, more data of a similar nature would need to be collected for various geometries and scales of screw pump. Future work will focus on increasing the resolution of the data

collected here, as well as collecting it for other geometries of screws.

ACKNOWLEDGEMENT

Aspects of this work were financially supported by the Natural Sciences and Engineering Research Council (NSERC) Collaborative Research and Development (CRD) program (grant # CRDPJ 433740-12) and Greenbug Energy Inc. (Delhi, ON, Canada). The assistance of Tony Bouk and Brian Weber of Greenbug Energy Inc. is gratefully acknowledged.

REFERENCES

- [1] S. Dalley and J. P. Oleson, "Sennacherib, Archimedes, and the Water Screw: The Context of Invention in the Ancient The Context of Invention in the Ancient World," *Technol. Cult.*, vol. 44, no. 1, pp. 1–26, 2003.
- [2] K. Vaughan, "Windmills of Holland," *PSA J.*, no. April, pp. 30–33, 2006.
- [3] S. R. Waters and G. A. Aggidis, "Over 2000 years in review: Revival of the Archimedes Screw from Pump to Turbine," *Renew. Sustain. Energy Rev.*, vol. 51, pp. 497–505, 2015.
- [4] K.-A. Radlik, "Hydrodynamic screw for energy conversion -uses changes in water supply to regulate energy output," DE4139134A1, 1997.
- [5] A. Lashofer, W. Hawle, and B. Pelikan, "State of technology and design guidelines for the Archimedes screw turbine," *Univ. Nat. Resour. Life Sci. Vienna*, no. October, pp. 1–8, 2012.
- [6] S. J. Williamson, B. H. Stark, and J. D. Booker, "Low head pico hydro turbine selection using a multi-criteria analysis," *Renew. Energy*, vol. 61, no. 0, pp. 43–50, 2014.
- [7] G. Nagel, "Archimedean Screw Pump Handbook." 1968.
- [8] C. Rorres, "The Turn of the Screw: Optimal Design of and Archimedes Screw," *J. Hydraul. Eng.*, vol. 126, no. January, pp. 72–80, 2000.
- [9] J. Muysken, "Calculation of the effectiveness of the auger." 1932. *De Ingenieur*, 21, 77–91 (in Dutch).
- [10] W. Boiten, A. Dommerholt, and M. Soet, "Handbook on flow measurements in open watercourses." January 1995. Dept. of Water Management PA Wageningen (in Dutch), ISBN 90.74476.13.9

PERFECTLY SUBCRITICAL GRAVITY CURRENTS: LABORATORY AND NUMERICAL EXPERIMENTS AND TWO-LAYER SHALLOW WATER SOLUTIONS

Mitchel S. Baker and Morris R. Flynn
 Department of Mechanical Engineering
 University of Alberta
 Edmonton, Canada
 mrflynn@ualberta.ca

Marius Ungarish
 Department of Computer Science
 Technion – Israel Institute of Technology
 Haifa, Israel

Abstract—Gravity current flows in two-layer stratified ambient layers are well understood when the gravity current front propagates at supercritical speeds i.e. when the gravity current front advances more quickly than the disturbance along the ambient interface. Less well understood are the subcritical gravity currents, which must, at some point, be overtaken by the interfacial disturbance. As a starting point, we herein explore the limit of a perfectly subcritical flow i.e. where the gravity current density matches the lower layer density. Gaining insight into the dynamics of this limiting case is an important first step in exploring the subcritical regime more generally. Our study considers two-layer lock-release Boussinesq gravity current experiments, which were performed in the laboratory and using direct numerical simulations (DNS). In both cases, the speed of the advancing gravity current, as well as that of the leading interfacial wave were well-predicted by a two-layer shallow water (SW) theory, while the speed of the trailing interfacial wave, formed

from the reflection of the left-propagating disturbance from the end of the lock, showed somewhat less consistent agreement with SW theory. More precisely, measured data showed good qualitative agreement with the analogue SW prediction, but the theory under-predicted the speed by about 20%. Conversely, leading and trailing wave amplitudes corresponded well with SW theory when measured near the beginning of their evolution. Following the initial measurements, both waves exhibited a time-dependent decrease in amplitude.

Keywords—gravity current; Boussinesq; shallow water; tailwater; subcritical; buoyancy-driven flow

I. INTRODUCTION

Gravity currents, which are horizontal flows driven by density differences, are ubiquitous in nature and in industry. Common examples are advancing frontal weather systems, saline wedges in estuaries, and the effluent discharged from a pipe in a marine environment.

Recognizing that gravity currents often propagate through ambients that are density-stratified (rather than uniform), much attention has been given to *supercritical* flows. In the present lock exchange context, in which the ambient consists of two distinct layers, a supercritical gravity current is defined as one whose front speed exceeds the speed of the ambient interfacial disturbance that is excited by the removal of the lock gate and the subsequent slumping of the gravity current fluid. By contrast, when the gravity current is sub-critical, its front speed is comparatively slow and the interfacial disturbance can propagate ahead of the front in the form of an interfacial wave or an interfacial bore [6]. Owing to this added complexity, subcritical gravity currents have been less thoroughly studied. (An exception is [10], where the authors present hydraulic theory and Navier-Stokes simulations of subcritical gravity currents.)

The perfectly subcritical regime (where the density of the advancing gravity current matches that of the lower layer, i.e. tailwater, in which it is propagating) has been explored from the point of view of predicting the propagation of downstream

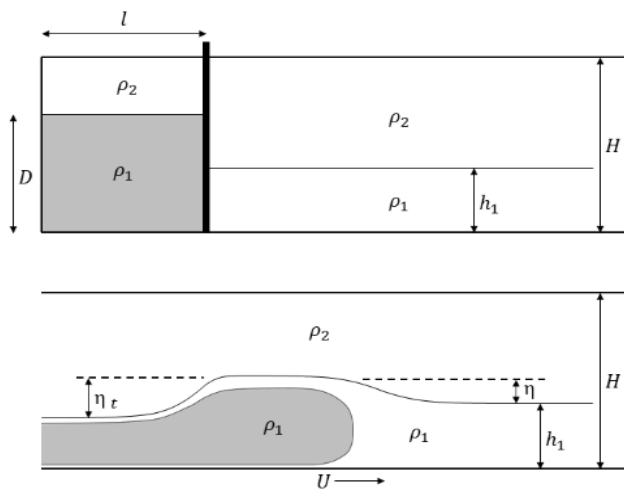


Figure 1. Definition sketch of gravity current experiments showing the lock-end of the tank (top) before the initiation of flow, and a downstream portion of the tank (bottom) with propagating gravity current.



Figure 2. Laboratory equipment. Video recording camera (bottom left) and camera's position relative to tank (top left). Layering fresh water on top of salt water using the sponge float (top and bottom right.)

disturbances (what we shall term the “leading wave” below) [8] but not from the point of view of predicting the evolution of the dense lock fluid (what we shall term the “gravity current”).

Note that the gravity currents studied in this perfectly subcritical investigation are unique in that the lower ambient layer fluid immediately ahead of the gravity current front travels at the same speed as the front itself. (Such behaviour is not anticipated when the gravity current density exceeds that of the lower ambient layer.) So, to reiterate, what defines the gravity current front in figures such as Fig. 3 below is colour (passive tracer) rather than salinity (active tracer).

In this investigation a two-layer perfectly subcritical gravity current model was studied using both laboratory and DNS lock-release experiments and results were compared to predictions made using a two-layer SW theory. Both experiments and theory applied the Boussinesq approximation according to which density differences were small ($\sim 4\%$ or less) and so impacted only the body force term in the equations of motion.

The Reynolds number was defined as

$$Re = \frac{\rho_2 U H}{\mu}$$

where the variables in the numerator are defined in Fig. 1 and the dynamic viscosity was taken to be 0.01002 g/cm/s . Although Re varied from 400 to 23 000, viscous effects were in any case small.

Fig. 1 summarizes the experimental geometry and parameters. Densities are represented by ρ and the horizontal speed of the gravity current front by U , while the remainder of the parameters have dimensions of length.

II. THEORY

Use of a SW model is here justified because the flow in question is, as noted above, approximately inviscid and vertical

accelerations are generally small, this owing to the fact that the fluid depth is less than 10% of its length. The standard SW continuity equation reads

$$\frac{\partial h}{\partial t} + u \frac{\partial h}{\partial x} + h \frac{\partial u}{\partial x} = 0 \quad (1)$$

where h is the depth of the lower layer/dense lock fluid and u is the corresponding horizontal velocity. Because $u \gg w$, the two-dimensional Euler equations can be combined into a single momentum equation. Using the Boussinesq approximation, and following the analysis of Rottman and Simpson [4] leads to

$$\left(\frac{H}{H-h}\right) \frac{\partial u}{\partial t} + \left[\frac{H(H-3h)}{(H-h)^2}\right] u \frac{\partial u}{\partial x} + \left[g' - \left(\frac{H}{H-h}\right)^3 \frac{u^2}{H}\right] \frac{\partial h}{\partial x} = 0 \quad (2)$$

where H is the channel height, x is the horizontal distance from the end of the lock, and t is time. The SW equations are made non-dimensional by scaling heights with the initial height, $D (\leq H)$, of dense fluid inside the lock, horizontal lengths with lock-length l , horizontal velocities with the long wave speed

$$U = \sqrt{g'D} \quad \text{where} \quad g' = g \frac{\rho_1 - \rho_2}{\rho_1} \quad \text{is the reduced gravity,}$$

and time with $T = l/U$ (see Fig. 1). The resulting non-dimensional coupled system of hyperbolic partial differential equations can then be solved using the method of characteristics (see [9]), which results in the following ODE:

$$\frac{dh}{du} = \frac{1}{D} \left[Au \mp \sqrt{(Au)^2 + Dh} \right] \quad (3)$$

on characteristic trajectories defined by

$$\frac{dx}{dt} = c_{\pm} = u(1-A) \pm \sqrt{(Au)^2 + Dh} \quad (4)$$

where $A = h/(H-h)$ and $D = 1 - h/H - H/(H-h)^2 u^2$.

Solving (3) along the characteristics results in a spectrum of horizontal velocities that varies with height. However, determining the shape and speed of the gravity current still requires the specification of the front (nose) and initial conditions. The type of current nose depends, for instance, on whether the disturbance propagating into the lock is a bore or a rarefaction wave. Following the seminal analysis of Benjamin [1], the following expression for the nose velocity was derived for the Boussinesq case by Klemp et al. [3]:

$$V_N = Fr h_N^{1/2} \quad (5)$$

in which the Froude number, Fr , is defined as

$$Fr = \sqrt{\frac{a(2-a) - b(2-b)}{(a-b) \left(\left(1 + \frac{b}{a} - 2b\right) + 2a \frac{1-b}{1-a} \right)}} \quad (6)$$

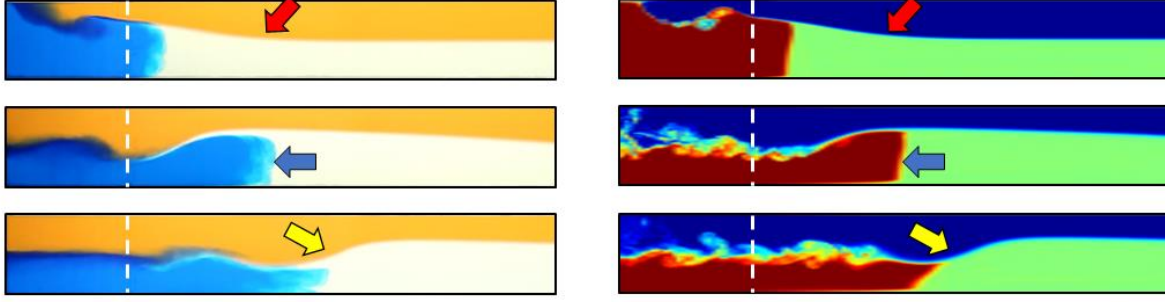


Figure 3. Timeseries of a gravity current propagating in a laboratory experiment (left) and the analogous DNS experiment (right). The dense lock fluid depth was $D/H = 1$, and dense ambient fluid depth $h_1/H = 0.5$. The domain shown in both cases is about 20×160 cm. Non-dimensional times ($t\sqrt{g'H}/H$) after beginning of current motion are 2 (top), 9 (middle) and 15 (bottom). Also shown are the positions of the gravity current front (blue arrow), leading wave (red arrow), trailing wave (yellow arrow), and lock gate position (white dashed line).

and where h_N is nose height, $a = h_N/H$, and $b = h_1/H$. Note that (6) reduces to Benjamin's classical formula in the limit of vanishingly thin tailwaters, i.e. $b \rightarrow 0$.

When the leading wave is a bore, the speed of the gravity current can be predicted based on the speed of the leading wave. Specifically,

$$U = \frac{V_N}{1 + \frac{h_1}{\eta}} \quad (7)$$

where V_N is the speed of the leading wave and the other variables are defined in Fig. 1. Note that this formula is not applicable when the leading wave takes the form of a rarefaction wave. In these cases the (7) predicts a gravity current speed of zero (as $\eta \rightarrow 0$) even though experiments produce a gravity current that advances with non-zero speed.

Using the preceding equations together with mass continuity arguments allows for the calculation of the gravity current and wave speeds and heights (even when (7) does not make correct predictions - see [8,9] for details).

III. EXPERIMENTS (LABORATORY AND NUMERICAL)

A. Experimental Procedure

Laboratory experiments were prepared by filling to a depth $> h_1$ a 227 cm long by 25 cm wide by 34 cm high glass tank with salt water with the lock gate up. The lock gate was then lowered into the salt water to within approximately 1 cm of the tank bottom. Fresh water dyed red with food coloring was then slowly added to the main side of the tank, pushing up the salt water in the lock.

The lock gate was then closed and a small amount of blue food coloring was added to the dense lock fluid. This method allowed for a nearly exact match between the density (~ 1.04 g/cm³) of the dense lock fluid and that of the lower ambient layer. Fig. 2 shows images from the laboratory corresponding to this preparatory stage of the experimental setup.

Laboratory experiments were initiated by slowly removing the lock gate, which was situated 36 cm from the left end wall of the tank. Experimental videos were collected using a tripod-mounted DSLR camera, also shown in Fig. 2. The camera, which was aligned with the center of the front face of the tank, was located a sufficient distance away (i.e. 375 cm) to

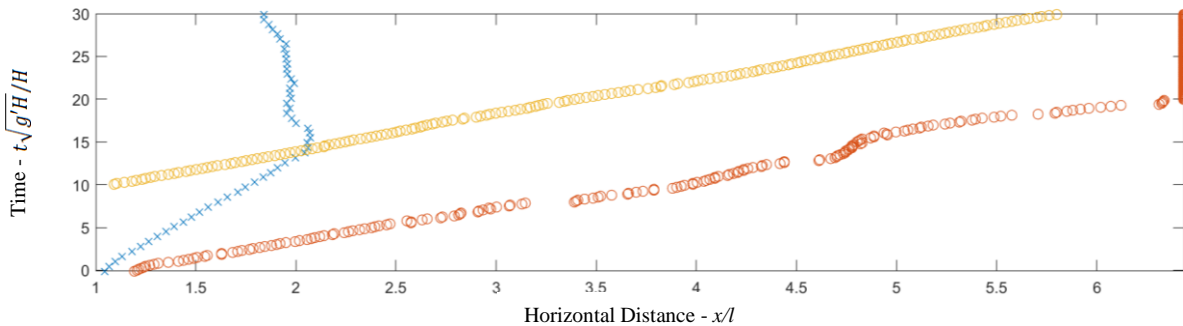


Figure 4. Horizontal timeseries of the laboratory experiment shown in Fig. 3. The measured horizontal position, in lock lengths, of the gravity current front (blue xs), leading wave front (red circles) and trailing wave front (yellow circles) are as indicated.

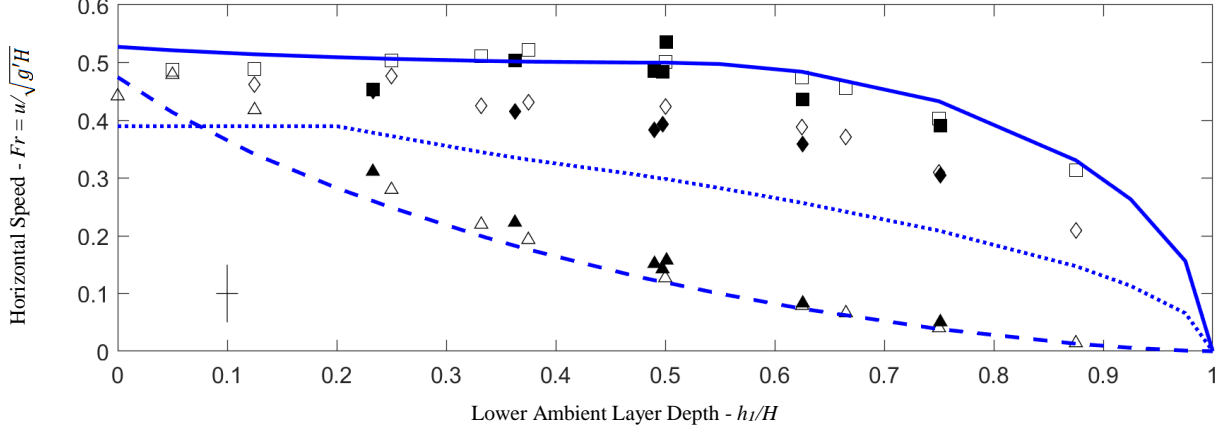


Figure 5. Leading wave speed (squares), trailing wave speed (diamonds), and gravity current speed (triangles) versus initial lower ambient layer depth. Open symbols represent DNS data while filled symbols represent laboratory data. Blue lines are SW predictions for the leading wave (solid), trailing wave (dotted), and gravity current speeds (dashed). The crossed vertical and horizontal black lines represent an estimate of uncertainty in the experimental measurements.

minimize parallax errors.

Analogous DNS experiments were performed using Diabolo [7], an open-source mixed-spectral finite difference algorithm that employs a staggered grid with spatial resolutions of 0.2 cm and 0.08 cm, respectively, in the horizontal and vertical directions. Although the former grid spacing is relatively large for DNS studies, we justify this choice by pointing out that our density contrasts and Reynolds numbers are relatively modest, i.e. just large enough to ensure a turbulent gravity current flow. Moreover, and as we shall discuss in Section IV, our principal interest concerns the speeds of the gravity current and the leading and trailing waves. These speeds are largely unaffected by the fine-scale details of diapycnal mixing or turbulent dissipation, much of which occurs well behind even the trailing wave -- see e.g. Fig. 3 above.

For most of the numerical experiments the domain measured 20 cm tall by 400 cm long, the latter distance including a 40 cm lock. The light fluid density was 1.00 g/cm³ and the heavy fluid density was 1.02 g/cm³; though lower (for reasons of computational efficiency) than in the analogue laboratory experiments, the density contrast was still enough to guarantee the appearance of turbulent flow features such as Kelvin-Helmholtz billows.

A total of 10 DNS experiments and seven laboratory experiments were performed with full-depth dense fluid in the lock ($D = H$), with varying depths, h_1 , of the dense fluid tongue outside of the lock. Consistent with e.g. Fig. 2 of Flynn et al. [2], the qualitative agreement between laboratory experiment and numerical simulation is typically robust as confirmed by Fig. 3 above, which shows a representative comparison for a full-depth lock release experiment having $h_1/H = 0.5$. Comparisons such as those associated with Fig. 3 confirm that the even relatively coarse, 2D numerical simulations are sufficient for capturing the essential physics of the flow, particularly as these relate to the motion of the gravity current and the leading and trailing waves.

B. Post-processing

Laboratory video was first cropped and resampled to 12 frames per second to reduce file size, typically resulting in approximately 500 frames/video. Video frames were then converted to arrays of 8-bit RGB values. Typical array size corresponding to one frame was 225 x 1600 x 3. Pixel resolution was determined by selecting points a known distance apart on the starting frame.

Typical DNS data consisted of 51 frames of double-precision density value arrays of size 256 x 2048. The relatively low number of frames output by the DNS code kept output file size manageable, with a time increment between frames of 1.25 s. Conversely, the time step used in the numerical simulations was 0.005 s.

The position of the gravity current and wave fronts in laboratory and DNS data were detected using MATLAB-based image-processing tools. On this basis, we were able to generate horizontal timeseries images like that of Fig. 4. It shows the relative positions of the gravity current (blue), leading wave¹ (red) and trailing wave front (yellow) for the experiment shown in Fig. 3 and confirms a well-defined interval where constant speeds were realized. Note that the leading wave position was only measured until its arrival at the far end wall while measurement of the trailing wave position was limited by interference with the reflected leading wave at large values of x/l .

Linear fits to the position data were then used to calculate the gravity current front speed and the speeds of the leading and trailing waves.

Measured front position data was then overlaid on laboratory and DNS images to verify its accuracy.

¹ In this manuscript, the term “leading wave” or “trailing wave” is used to indicate a disturbance that might alternatively be considered a bore in the SW theory.

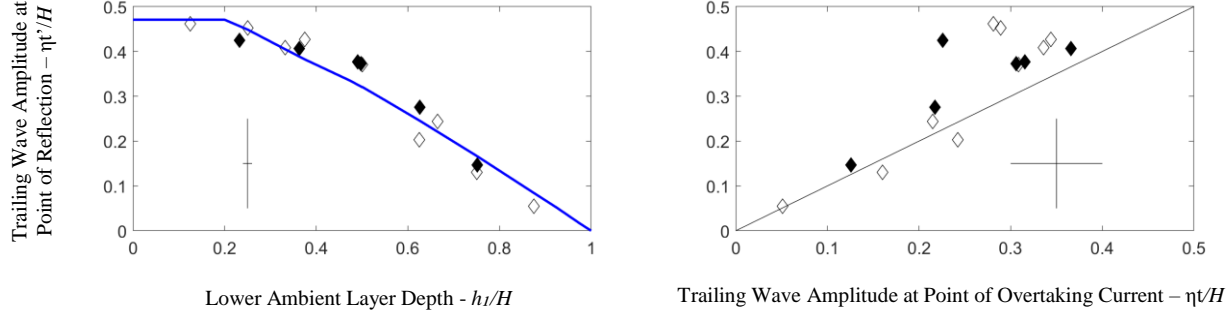


Figure 6. Trailing wave amplitude measured immediately following reflection from the lock wall vs. lower ambient layer depth (left, symbols) and vs trailing wave amplitude measured at point of overtaking current (right, symbols). Open symbols represent DNS data and filled symbols represent laboratory data. Lines are the predictions made by SW theory for trailing wave amplitude (thick blue line) and a reference line of 45° slope (thin black line). The crossed vertical and horizontal black lines represent an estimate of uncertainty in the experimental measurements.

Finally, wave amplitudes were measured by determining the boundary between light and dark fluid layers, again using MATLAB. Leading wave amplitudes were measured at a series of fixed positions. Trailing wave amplitudes were measured at the point where the trailing wave overtook the gravity current front, unless this was limited by end wall effects. In this case, the trailing wave amplitude was measured just prior to any interactions of the leading wave with the far end wall. For comparison's sake, a second set of trailing wave amplitude measurements were collected immediately following the reflection of the trailing wave from the end of the lock.

IV. RESULTS AND DISCUSSION

The horizontal time series in Fig. 4 highlights the typical gravity current behaviour observed in laboratory and DNS experiments. The gravity current fluid (blue) exhibits a period of constant speed advance that begins with lock release and the initiation of the leading wave (red) and is terminated by the overtaking of the trailing wave (yellow). After the overtaking, the gravity current front speed drops significantly and the eventual motion of the gravity current fluid becomes dominated by the reflection of the leading wave from the right end wall of the tank.

This research focused on the period between lock release

and the arrest of the gravity current fluid following the passage of the trailing wave, i.e. the period of steady advance corresponding to $2 < t\sqrt{g'H}/H < 15$ in Fig. 4. The speeds measured during this interval, including that of the gravity current, leading wave and trailing wave, were plotted vs. the depth of the lower ambient layer and compared to predictions made by SW theory. These results are shown in Fig. 5.

In the case of the leading wave (open and filled squares, solid curve), and in good agreement with qualitative observations from our experiments, SW theory predicts the appearance of a bore for lower layer depths below half the channel height ($h_1/H < 0.5$) and a rarefaction/long wave for lower layer depths above this critical value [8]. (The sequence of panels in Fig. 3 illustrates a leading wave at the crossover value of $h_1/H = 0.5$). The SW prediction for the leading wave speed of Fig. 5 is nonetheless continuous as confirmed by the solid curve, which agrees well with both laboratory (filled squares) and DNS data (open squares) for $h_1/H < 0.5$ and $h_1/H > 0.5$.

Measured gravity current speeds are shown in Fig. 5 as triangles while the analogue SW prediction is given by the dashed curve. Both approach Benjamin's prediction of $Fr = 0.5$ as the ambient fluid height approaches zero [1]; both also approach $Fr = 0$ as $h_1/H \rightarrow 1$. In between these extremes, the correspondence of SW theory and measurement is both robust

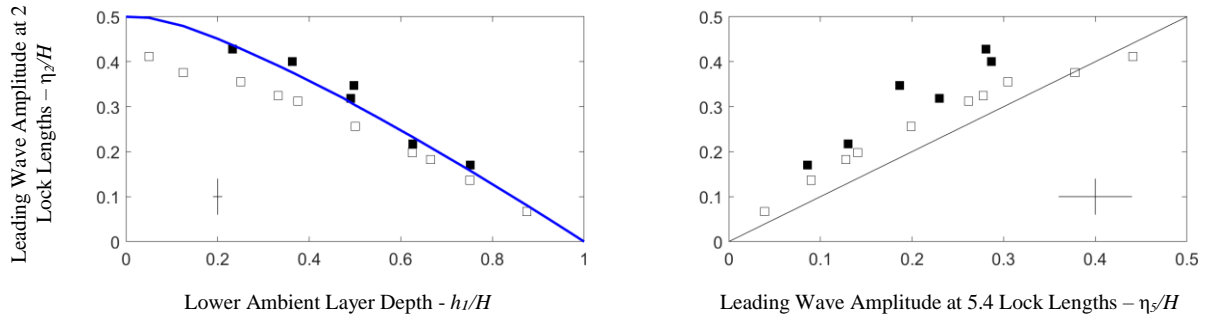


Figure 7. Leading wave height measured at two lock lengths vs. lower ambient layer depth (left, symbols) and vs leading wave height measured at 5.4 lock lengths (right, symbols). Open symbols represent DNS data and filled symbols represent laboratory data. Lines are the predictions made by SW theory for leading wave amplitude (thick blue line) and a reference line of 45° slope (thin black line). The crossed vertical and horizontal black lines represent an estimate of uncertainty in the experimental measurements.

and roughly comparable to the agreement observed Tan et al. [6], who focused primarily on supercritical flows and derived a corresponding theory that was not SW in nature.

Also included in Fig. 5 are trailing wave speed data (diamonds) and SW theory (dotted line). The theory predicts a bore rather than a rarefaction wave in all cases shown. The agreement between experiment and theory is not as good as in the case of the leading wave and gravity current fluid. In part, this may be because trailing wave speeds must be measured some distance away from the back of the lock owing to the enhanced turbulent mixing that accompanies reflection. Between this point of reflection and measurement, the trailing wave speed may modify in such a way as to deviate from the analogue SW prediction.

There is evidence for such an evolution in the right-hand side panel of Fig. 6, which shows measurements of the trailing wave amplitudes made close to the point of reflection from the lock end wall vs. at the point where the trailing wave overtakes the gravity current (open and filled diamonds). The solid (1:1) line, which represents the case of equal amplitudes, roughly bisects the experimental data except with $\eta_l/H > 0.25$ where trailing wave amplitudes measured at the point of overtaking are consistently smaller than those immediately following reflection. The parametric range in question corresponds to experiments with relatively small h_l/H , where the point of overtaking occurred later due to the increased speed of the gravity current. To the extent that a time-dependent decrease in wave amplitude seems to be at play in these experiments, the decrease in amplitude would have the longest duration over which to manifest, which could explain why a more substantial deviation from the 1:1 line is seen for the largest values of η_l/H in the right-hand panel of Fig. 6.

Meanwhile, the left panel of Fig. 6 shows the variation of trailing wave amplitude (measured immediately after reflection from the lock wall) with the initial depth of the lower ambient layer (open and filled diamonds). The thick blue line represents the amplitudes predicted by shallow water theory. Whereas the agreement between experiments and theory was weaker for trailing wave speeds, in the case of trailing wave amplitudes the agreement between experiments and SW theory is excellent.

Fig. 7 displays information analogous to Fig. 6 for the leading wave. The left-hand side panel shows the relationship between the leading wave amplitudes and the depth of the lower ambient layer. Again, both measured and SW data are considered. SW theory correctly captures the qualitative trend of the measured data, and there is good quantitative agreement between the theory and the laboratory experiments (filled squares). However, there is a small offset between the theory and the numerical experiments (open squares.) This is the opposite of the agreement of the speed data (Fig. 5) and suggests that the leading wave may likewise be modified in some way as it travels. Indeed, each leading wave's amplitude was measured at several successive locations and small, progressive decreases in amplitude were measured.

Evidence for this is seen in the right-hand panel of Fig. 7, which plots leading wave amplitudes measured at two lock

lengths of horizontal travel vs. leading wave amplitudes measured at 5.4 lock lengths (the end of the laboratory tank.) For most of the data, a decrease in amplitude is measured over the distance of 3.4 lock lengths. As with the trailing wave, this is evidenced by the data (filled and solid squares) falling above the 1:1 solid line.

V. CONCLUSION

The present laboratory experiments substantially validate the newly-presented SW theory and DNS experiments for perfectly subcritical gravity currents (i.e. gravity currents with tailwaters) [8,9]. The new experiments and theory provide a starting point from which trends into the (non-perfectly) subcritical regime can be inferred, and linked to previously obtained data and corresponding SW and non-SW theories [4,6,8].

Future work will include cases where the dense lock fluid will span less than the full lock depth ($D < H$), as well as cases where the ambient fluid density is continuously stratified instead of prepared in two layers.

Examples of applications of the perfectly subcritical regime studied here are dam-break calculations with tailwaters, or in the analysis of nocturnal thunderstorm downdrafts where the downdraft expands horizontally into a layer of equally cold air already adjacent to the ground. The precise application of our results to these and other environmental scenarios will be explored in forthcoming studies.

REFERENCES

- [1] Benjamin, T. B. (1968). Gravity currents and related phenomena. *Journal of Fluid Mechanics*, 31(2), 209-248.
- [2] Flynn, M. R., Ungarish, M., & Tan, A. W. (2012). Gravity currents in a two-layer stratified ambient: The theory for the steady-state (front condition) and lock-released flows, and experimental confirmations. *Physics of Fluids*, 24(2), 26.
- [3] Klemp, J. B., Rotunno, R., & Skamarock, W. C. (1997). On the propagation of internal bores. *Journal of Fluid Mechanics*, 331, 81-106.
- [4] Longo, S., Ungarish, M., Di Federico, V., Chiapponi, L., & Maranzoni, A. (2015). The propagation of gravity currents in a circular cross-section channel: Experiments and theory. *Journal of Fluid Mechanics*, 764, 513-537.
- [5] Rottman, J. W., & Simpson, J. E. (1983). Gravity currents produced by instantaneous releases of a heavy fluid in a rectangular channel. *Journal of Fluid Mechanics*, 135, 95-110.
- [6] Tan, A. W., Nobes, D. S., Fleck, B. A., & Flynn, M. R. (2011). Gravity currents in two-layer stratified media. *Environmental Fluid Mechanics*, 11(2), 203-223.
- [7] Taylor, J. R. (2008). Numerical simulations of the stratified oceanic bottom boundary layer (Doctoral dissertation, UC San Diego).
- [8] Ungarish, M., Borden, Z., & Meiburg, E. (2014). Gravity currents with tailwaters in boussinesq and non-boussinesq systems: Two-layer shallow-water dam-break solutions and Navier-Stokes simulations. *Environmental Fluid Mechanics*, 14(2), 451-470.
- [9] Ungarish, M. (2009). *An Introduction to Gravity Currents and Intrusions* Chapman and Hall/CRC.
- [10] White, B. L., & Helfrich, K. R. (2012). A general description of a gravity current front propagating in a two-layer stratified fluid. *Journal of Fluid Mechanics*, 711, 545-575.

Comparison of Commonly Used Atomization Models

Siyu Chen

Department of Mechanical and Industrial Engineering
 University of Toronto
 Toronto, ON, Canada
 siyuchen@mie.utoronto.ca

Nasser Ashgriz

Department of Mechanical and Industrial Engineering
 University of Toronto
 Toronto, ON, Canada
 ashgriz@mie.utoronto.ca

Abstract—Theoretical models for the atomization process are mainly based on a two-step process: (i) Formation and growth of interface waves resulting in the primary breakup of the liquid that issues from a nozzle, and (ii) the secondary breakup of the ligaments and droplets that form from the primary breakup. This paper reviews several atomization models and compares their predicted droplet sizes to identify their model deficiencies. The models are tested for a range of air and liquid velocities, as well as several fluid properties.

Keywords; *Atomization and sprays, Liquid instability; Primary breakup; Secondary breakup;*

I. INTRODUCTION

Sprays have a wide variety of applications, such as painting, tablet coating, fuel injection in engine, agricultural pesticide application. As a liquid issues from a spray nozzle, it experiences aerodynamic instability and disintegrate into droplets with different sizes. The size distribution of the droplets and the mean droplet diameter (e.g. Sauter mean diameter, SMD) determine the quality of the spray, which, in turn, influence the quality of the process or product in the application. As a result, predicting a mean droplet size for certain operating conditions is crucial for the design. Multiple factors have influences on the droplet size in a spray, such as the nozzle type (flat fan nozzle, swirl nozzle, twin-fluid nozzle), injection pressure, orifice diameter, fluid properties (viscosity, surface tension, density). There has been a significant theoretical and experimental effort to determine the relationship between operating conditions and the mean droplet size of a spray (for a detailed discussion of atomization and sprays see [1]). The present work reviews seven more commonly used atomization models and compares the results as applied on certain flow conditions and fluid properties.

II. INSTABILITY MODEL

Small disturbances at liquid-gas interfaces may grow resulting in interface instability and breakup of the liquid. Several different mechanisms can cause such instabilities. The

mechanisms that are commonly used in the atomization process are capillary, aerodynamically, and/or acceleration induced instabilities. Linear instability analysis for such cases provides a dispersion equation for the growth rate of different wave numbers. The wave with the wavenumber that had the highest growth rate among the spectrum of waves is the dominant surface wave and is responsible for the breakup of liquid. The instability models that are commonly used to describe the atomization process are provided next.

A. Rayleigh-Plateau (RP) Instability

The Rayleigh-Plateau instability is a capillary induced instability that describes the temporal instability of a wave on a liquid jet. The dispersion equation for an inviscid jet is [1]

$$\omega^2 = \frac{\sigma k}{\rho_l a^2} (1 - k^2 a^2) \frac{I_1(ka)}{I_0(ka)} \quad (1)$$

where I_0, I_1 are Bessel functions.

The wave with the largest growth rate ω_{max} is given by

$$\omega_{max} = 0.34 \left(\frac{\sigma}{\rho_l a^3} \right)^{\frac{1}{2}} \quad (2)$$

The corresponding wave number is

$$k_{max} = \frac{0.69}{a} \quad (3)$$

which is only a function of jet radius.

For a viscous liquid jet,

$$\omega^2 + \frac{2\nu k^2}{I_0(ka)} \left(I_1'(ka) - \frac{2kl}{k^2 + l^2} \frac{I_1(ka)}{I_1(la)} I_1'(la) \right) \omega = \frac{\sigma k}{\rho_l a^2} (1 - k^2 a^2) \frac{I_1(ka)}{I_0(ka)} \frac{l^2 - k^2}{k^2 + l^2} \quad (4)$$

where $l^2 \equiv k^2 + \frac{\omega}{\nu}$.

For long waves, $ka \ll 1$ and the dispersion equation can be simplified to

$$\omega^2 + \frac{3\mu k^2}{\rho_l} \omega - \frac{\sigma}{2\rho_l a^3} (1 - k^2 a^2) k^2 a^2 = 0 \quad (5)$$

The jet is unstable when $ka < 1$ and the maximum growth rate is given by

$$\omega_{max} = \left[\left(\frac{8\rho_l a^3}{\sigma} \right)^{\frac{1}{2}} + \frac{6\mu a}{\sigma} \right]^{-1} \quad (6)$$

The corresponding wave number on the jet is

$$k_{max}^2 = \frac{\sqrt{2\rho_l a \sigma}}{2a^2(\sqrt{2\rho_l a \sigma} + 3\mu)} \quad (7)$$

B. Kelvin-Helmholtz (KH) Instability

The Kelvin-Helmholtz instability is an aerodynamic instability based on the velocity difference between the liquid and gas. If the densities and velocities for the two inviscid fluids are ρ_1, U_1, ρ_2, U_2 respectively, and the surface tension and body forces are neglected, the dispersion equation is given by

$$\omega = \frac{ik(\rho_1 U_1 + \rho_2 U_2)}{\rho_1 + \rho_2} \pm \frac{\sqrt{\rho_1 \rho_2}}{\rho_1 + \rho_2} |U_1 - U_2| \quad (8)$$

A liquid sheet issuing from a nozzle can be modelled with a Kelvin-Helmholtz instability applied on two sides of the liquid sheet. For an inviscid liquid moving in the gaseous medium, the dispersion equation is given as

$$\frac{\omega}{kU_o} = \left[\frac{\rho_r K}{(\rho_r + K)^2} - \frac{ka}{We} \frac{1}{\rho_r + K} \right]^{\frac{1}{2}} \quad (9)$$

where $\rho_r = \frac{\rho_g}{\rho_l}$, $U_o = |U_l - U_g|$, and $K = \tanh(ka)$ for sinuous wave on the interfaces and $K = \coth(ka)$ for varicose wave. At low density ratio ρ_r , which is common for liquid sheet in gaseous medium, the growth rate of sinuous waves always outweighs the growth rate of varicose wave at the same wave number k . As a result, the dominant wave on the interface is the sinuous waves with the maximum growth rate.

For long waves, $K = \tanh(ka) \approx ka$. If $\rho_r \ll K$, the dispersion equation simplifies to

$$\frac{\omega}{kU_o} = \left(\frac{\rho_r}{ka} - \frac{1}{We} \right)^{\frac{1}{2}} \quad (10)$$

For short waves, $K = \tanh(ka) \approx 1$. If $\rho_r \ll K$, the dispersion equation simplifies to

$$\frac{\omega}{kU_o} = \left(\rho_r - \frac{ka}{We} \right)^{\frac{1}{2}} \quad (11)$$

Senecal et al. [2] provided the following dispersion equation for a viscous liquid sheet:

$$\omega = -\frac{2\nu k^2 K}{\rho_r + K} + \frac{\left[4\nu^2 k^4 K^2 - \rho_r^2 U_o^2 K^2 - (\rho_r + K) \left(\rho_r U_o^2 K^2 + \frac{\sigma k^3}{\rho_1} \right) \right]^{\frac{1}{2}}}{\rho_r + K} \quad (12)$$

For long waves, the dispersion equation simplifies to [2]

$$\omega = -2\nu k^2 + \left(4\nu^2 k^4 - \frac{\rho_r U_o^2 k}{a} + \frac{\sigma k^2}{a\rho_1} \right)^{\frac{1}{2}} \quad (13)$$

For short waves, the dispersion equation simplifies to

$$\omega = -2\nu k^2 + \left(4\nu^2 k^4 - \rho_r U_o^2 k^2 + \frac{\sigma k^3}{\rho_1} \right)^{\frac{1}{2}} \quad (14)$$

There is no exact solution for the maximum growth rate and corresponding wavenumber. Numerical method or graphs need to be applied for viscous sheets.

C. Rayleigh-Taylor (RT) Instability

The Rayleigh-Taylor instability is an acceleration induced instability that describes the instability of waves on an interface between a heavier and a lighter fluid with an acceleration in the direction toward the heavier fluid. For a liquid-gas interface, if there is an acceleration a toward the liquid phase and the effect of surface tension of the liquid is considered, the dispersion equation can be found as

$$\omega^2 = \frac{(\rho_l - \rho_g)ak - \sigma k^3}{\rho_l + \rho_g} \quad (15)$$

If $\rho_l \gg \rho_g$, the equation can be simplified to

$$\omega^2 = ak - \frac{\sigma k^3}{\rho_l} \quad (16)$$

The maximum growth rate can be found as

$$\omega_{max}^2 = \frac{2a}{3} \sqrt{\frac{\rho_l a}{3\sigma}} \quad (17)$$

The corresponding wavelength is

$$k_{max} = \sqrt{\frac{\rho_l a}{3\sigma}} \quad (18)$$

III. PRIMARY BREAKUP MODEL

When there is a relative movement between a liquid sheet or a liquid jet and a surrounding gas, small disturbances may grow at the interface due to the instability and eventually break the sheet or jet into ligaments and then into droplets. By examining the dominant wavelength on the interface, the size of droplets that generates from the breakup can be predicted by the mass conservation. When the liquid moves at very high velocity, the spray becomes turbulent. Turbulence theories are combined with instability to predict the droplet size after primary breakup. Table I provides different atomization models and their predicted drop diameters that are considered in this review.

A. Jet Breakup Model

The jet breakup model applies RP instability to predict the diameter of droplets generated from the breakup of a cylindrical liquid jet. The wave growth pinches the jet at the troughs of surface waves. The volume of the liquid within each portion of jet forms a spherical droplet. The length between troughs is given by the dominant wavelength from RP instability. Then, for a jet with diameter the diameter d_j , the generated droplet d_D can be calculated from the mass conservation equation as:

$$d_D = (6a^2 \lambda_{RP})^{\frac{1}{3}} = \left(\frac{3\pi d_j^2}{2k_{RP}} \right)^{\frac{1}{3}} \quad (19)$$

B. Wave Breakup Model

Reitz and coworkers [3, 4, 5] developed an instability model for a jet including the effect of surrounding gas. Since there was no explicit solution for the wavenumbers with the maximum growth rates, they curve fit their numerically calculated growth rates for a wide range of parameters, and provided the following relationship for the maximum growth rate and the corresponding wavelength:

$$\frac{\omega_{max}}{a} = 9.02 \frac{(1 + 0.45Oh^{0.5})(1 + 0.40h^{0.7}We_g^{0.35})}{(1 + 0.87We_g^{1.67})^{0.6}} \quad (20)$$

$$\lambda_{max} = \left(\frac{\sigma}{\rho_l a^3}\right)^{0.5} \frac{0.34 + 0.38We_g^{1.5}}{(1 + Oh)(1 + 1.40h^{0.6}We_g^{0.3})} \quad (21)$$

The breakup behavior of the jet largely depends on the We_g , which determines the dominant wavelength. As We_g increases, the dominant wavelength decreases, generating finer sprays. In order to consider such effects, the dominant wavelength is compared with the radius of the jet. If $B_0\lambda_{max} \leq a$ (where B_0 is a constant and equals 0.61), the jet generates small droplets due to small dominant wavelength, whose radius is given by

$$r = B_0\lambda_{max} \quad (22)$$

If $B_0\lambda_{max} > a$, the jet will generate large droplets. Two conditions are considered: One droplet is generated per wavelength or the droplet size is found by balancing the frequency of droplet generation and the mass flow rate of jet. The predicted droplet radius is the minimal value of the two, which is

$$r = \min\left(\left(\frac{3a^2\lambda_{max}}{4}\right)^{\frac{1}{3}}, \left(\frac{3\pi a^2 U}{2\omega_{max}}\right)^{\frac{1}{3}}\right) \quad (23)$$

C. Sheet Breakup Model

Dombrowski and coworkers [6, 7] developed a model to predict the droplet diameter from the breakup of a liquid sheet in air. The sheet breakup model applies both KH instability and RP instability to predict the diameter of droplets generated by a liquid sheet. The droplets are formed by two steps: the breakup of liquid sheet into cylindrical ligaments and the breakup of the ligaments into droplets. In this model, it is assumed that two ligaments are generated per KH wavelength. The diameter of these ligaments can be found in terms of dominant wavenumber k_{max} from the mass conservation as

$$d_L = \left(\frac{8a}{k_{max}}\right)^{\frac{1}{2}} \quad (24)$$

According to the RP instability, the droplet diameter generated by the ligament is

$$d_D = \left(\frac{3\pi d_L^2}{2k_{RP}}\right)^{\frac{1}{3}} = 1.882 \left(\frac{8a}{k_{max}}\right)^{\frac{1}{2}} (1 + 3Oh)^{\frac{1}{6}} \quad (25)$$

D. Flat Fan Atomizer Breakup Model

Qin et al. [8] developed a model to predict the droplet diameter from the breakup of a liquid sheet generated by a flat fan nozzle. The model applies KH, RT and RP instability to provide an expression for a mean droplet diameter. They assumed that their mean droplet diameter represents the Sauter Mean Diameter (SMD) of a spray. When the liquid sheet issues from a nozzle, KH instability generates the surface waves. The drag force exerted by the air on the wave crests creates an acceleration and RT instability is then responsible for the breakup of the wave crest into ligaments. Finally, RP instability breaks the cylindrical ligaments into droplets. The final droplet diameter created from these three instabilities is given by

$$d_D = C(1 + m_r)X \quad (26)$$

$$X = \frac{2v_l \ln \frac{k\sigma}{\rho_g U^2} + \sqrt{\frac{\rho U^2}{k^2} - \frac{\sigma}{\rho_l k}} - \frac{\sigma}{U\sqrt{\rho_l \rho_g}} \ln \left(\sqrt{\frac{\rho_g U^2}{\sigma k}} + \sqrt{\frac{\rho_g U^2}{\sigma k} - 1} \right)}{4v_l \left(\sqrt{\frac{k}{2\pi}} - \sqrt{\frac{\rho_g U^2}{2\pi\sigma}} \right) + \sqrt{\frac{\sigma}{2\pi\rho_l}} \sqrt{\frac{\rho_g U^2}{\sigma k} - 1} - \sqrt{\frac{8\sigma\pi^3}{\rho_l}} \arctan \left(\sqrt{\frac{\rho_g U^2}{\sigma k} - 1} \right)}$$

where $X = \frac{1}{(1-\sqrt{\rho})u_g - u_l}$, $\rho = \rho_g/\rho_l$, $k = \pi/a$ and C is a constant which equals 0.33 given from experiments.

E. Coaxial jet and Twin Fluid Breakup Model

Aliseda et al. [9] develops a breakup model for the coaxial twin fluid nozzle. The model predicts the diameter of droplets from the breakup of a viscous jet generated by the nozzle. The model applies KH and RT instability to provide an explicit expression for the droplet size. The KH instability causes the surface waves in the longitudinal direction. At the crests of the wave, due to the drag force exerted by the air, the RT instability is responsible for the breakup of the crests into droplets. The droplet diameter is given by

$$\frac{d_D}{d_{jl}} = C_1(1 + m_r) \left(\frac{b_g}{d_{jl}}\right)^{\frac{1}{2}} \left(\frac{\rho_l/\rho_g}{Re_{bg}}\right)^{\frac{1}{4}} \frac{1}{\sqrt{We_{jl}}} \left\{ 1 + C_2 \left(\frac{d_{jl}}{b_g}\right)^{\frac{1}{6}} \left(\frac{\rho_l/\rho_g}{Re_{bg}}\right)^{-\frac{1}{12}} We_{jl}^{\frac{1}{6}} Oh^{\frac{2}{3}} \right\} \quad (27)$$

where C_1 and C_2 are constants with magnitude of 1.

F. Eddy breakup Model

Faeth and coworkers [10,11] developed a breakup model describing the jet breakup in turbulent conditions. When the kinetic energy of an eddy with size λ is large enough to provide the surface energy of a droplet at that size, a droplet will form. The droplet diameter is relevant with the turbulent formed at the nozzle, which is given by

$$\frac{d_D}{\Lambda} = C_{st} \left(\frac{\bar{u}_o}{\bar{v}_o'}\right)^{\frac{6}{5}} We_{\Lambda}^{-\frac{3}{5}} \quad (28)$$

where \bar{u}_o is the mean velocity, \bar{v}_o' is the velocity fluctuation and Λ is the radial length scale.

For a pipe flow, $\frac{\bar{u}_o}{\bar{v}_0} \approx 17$ and C_{si} is a constant which equals 2.5 calculated experimentally.

G. Turbulence Breakup Model

Huh et al. [12] develop an atomization model based on the turbulent scales. There are two basic assumption made on this model: the atomization length scale L_A is proportional to turbulence length scale L_t and wavelength of surface perturbation L_w ; the atomization time scale τ_A is a linear sum of turbulence τ_t and wave growth time τ_w scale.

$$L_A = C_1 L_t = C_2 L_w, \quad \tau_A = C_3 \tau_t + C_4 \tau_w \quad (29)$$

where the constants are given experimentally as $C_1 = 2$, $C_2 = 0.5$, $C_3 = 1.2$, $C_4 = 0.5$. If the turbulence is homogeneous and

can be modelled by $k - \varepsilon$ model, the time evolution of L_t and τ_t can be solved as

$$L_t(t) = L_t^0 \left(1 + \frac{0.0828t}{\tau_t^0} \right)^{0.457}, \quad \tau_t(t) = \tau_t^0 + 0.0828t \quad (30)$$

τ_w is related to the KH wave growth as

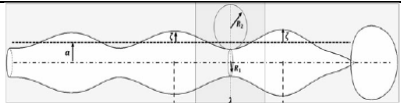
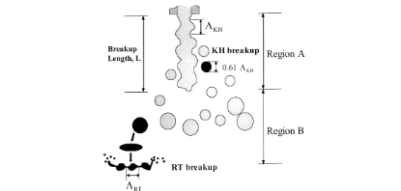
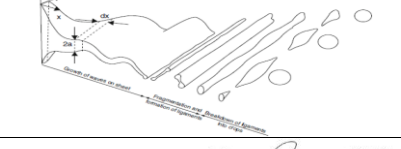
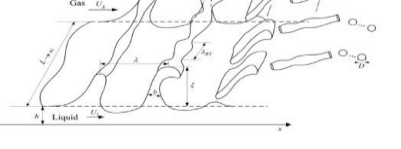
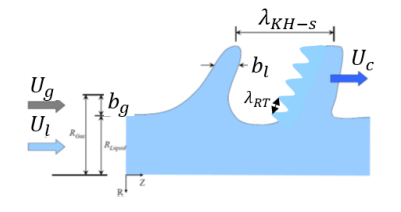
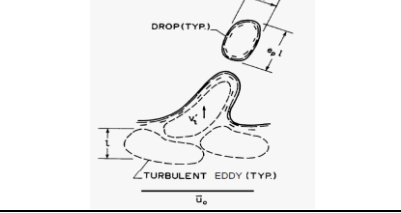
$$\tau_w = \frac{L_w}{U} \sqrt{\frac{\rho_l}{\rho_g}} \quad (31)$$

Finally, the droplet diameter can be calculated as a function of time t , which is the time since ejected from nozzle exit as

$$\frac{dd_D}{dt} = C_5 \frac{L_A}{t_A} \quad (32)$$

where C_5 is determined to be -0.25 experimentally.

TABLE I. ATOMIZATION MODEL

No	Model	Equation	Mechanism
1	Jet breakup	$d_D = \left(\frac{3\pi d_f^2}{2k_{RP}} \right)^{\frac{1}{3}}$	
2	Wave breakup [3]	$r = \begin{cases} B_0 \lambda_{max}, & \text{if } B_0 \lambda_{max} \leq a \\ \min \left(\left(\frac{3a^2 \lambda_{max}}{4} \right)^{\frac{1}{3}}, \left(\frac{3\pi a^2 U}{2\omega_{max}} \right)^{\frac{1}{3}} \right), & \text{if } B_0 \lambda_{max} > a \end{cases}$	
3	Sheet breakup [6]	$d_D = 1.882 \left(\frac{8a}{k_{max}} \right)^{\frac{1}{2}} (1 + 30h)^{\frac{1}{6}}$	
4	Flat fan nozzle breakup [8]	$d_D = c(1 + m_r)X$ $2v_l \ln \frac{k\sigma}{\rho_g U^2} + \sqrt{\frac{\rho U^2}{k^2} - \frac{\sigma}{\rho_l k}} - \frac{\sigma}{U \sqrt{\rho_l \rho_g}} \ln \left(\sqrt{\frac{\rho_g U^2}{\sigma k}} + \sqrt{\frac{\rho_g U^2}{\sigma k} - 1} \right)$ $4v_l \left(\sqrt{\frac{k}{2\pi}} - \sqrt{\frac{\rho_g U^2}{2\pi\sigma}} \right) + \sqrt{\frac{\sigma}{2\pi\rho_l}} \sqrt{\frac{\rho_g U^2}{\sigma k} - 1} - \sqrt{\frac{8\sigma\pi^3}{\rho_l}} \arctan \left(\sqrt{\frac{\rho_g U^2}{\sigma k} - 1} \right)$	
5	Coaxial twin fluid jet breakup [9]	$\frac{d_D}{d_{jl}} = C_1(1 + m_r) \left(\frac{b_g}{d_{jl}} \right)^{\frac{1}{2}} \left(\frac{\rho_l}{\rho_g} \right)^{\frac{1}{4}} \frac{1}{\sqrt{We_{jl}}}$ $\left\{ 1 + C_2 \left(\frac{d_{jl}}{b_g} \right)^{\frac{1}{6}} \left(\frac{\rho_l}{\rho_g} \right)^{-\frac{1}{12}} We_{jl}^{\frac{1}{5}} Oh^{\frac{2}{3}} \right\}$	
6	Eddy breakup model [10]	$\frac{d_D}{\Lambda} = C_{si} \left(\frac{\bar{u}_o}{\bar{v}_0'} \right)^{\frac{6}{5}} We_{\Lambda}^{-\frac{3}{5}}$	
7	Turbulence breakup [12]	$\frac{dd_D}{dt} = -0.25 \frac{L_A}{t_A}$	

IV. SECONDARY BREAKUP MODEL

Some of the droplets generated from the primary breakup have very large diameter, which may become unstable at high velocities and break up into smaller droplets. Different models are developed to predict the droplet sizes after the secondary breakup.

A. Weber number based breakup

The Weber number secondary breakup criterion is based on either a pressure balance or a kinetic energy balance on a droplet moving in a gas. When the droplet with diameter R moving with a velocity u in a gas with density ρ_g , the droplet will experience a secondary breakup if the pressure of the gas on the droplet surface exceeds the surface tension of the droplet

$$\frac{1}{2}\rho_g u^2 > \frac{2\sigma}{R} \quad (33)$$

or when the kinetic energy exceeds the surface tension energy

$$\frac{1}{2}mu^2 = \frac{\rho_g \pi d^3}{12} u^2 > \pi d_o^2 \sigma \quad (34)$$

A critical Weber number $We_c = \frac{\rho_g d u^2}{\sigma}$ can be calculated, exceeding which the droplet will deform and breakup. The critical value is 8 for the pressure balance and 12 for the energy balance.

If we consider the effect of viscosity, the critical Weber number will be

$$We_c = We_{c_{Oh \rightarrow 0}}(1 + 1.0770h^{1.6}) \quad (35)$$

From the critical Weber number, we can calculate the maximum possible diameter under certain velocity and gaseous medium.

B. Taylor Analogy Breakup (TAB Model)

The Taylor analogy breakup [13] considers a distorting and oscillating droplet as a spring-damper-mass system, where the surface tension forces is considered as the restoring force of the spring, the viscosity forces is considered as the damping force and the drag force is the external force on the mass. The forces are given as

$$\frac{k}{m} = C_k \frac{\sigma}{\rho_l r^3}, \quad \frac{d}{m} = C_d \frac{\mu_l}{\rho_l r^2}, \quad \frac{F}{m} = C_F \frac{\rho_g u^2}{\rho_l r} \quad (36)$$

The constants are determined experimentally as $C_k = 8$, $C_d = 5$, $C_F = \frac{1}{3}$

According to the Newton's Second Law,

$$F - kx - d \frac{dx}{dt} = m \frac{d^2 x}{dt^2} \quad (37)$$

When the displacement is greater than a half of the droplet radius, the droplet will experience a secondary breakup. If we non-dimensionalize the displacement x to $y = x/C_b r$, where $C_b = 0.5$, the droplet will breakup when $y > 1$. The displacement as a function of time can be found as:

$$y(t) = We_c + e^{-\left(\frac{t}{t_d}\right)} \left[(y_o - We_c) \cos(\omega t) + \frac{1}{\omega} \left(\frac{dy_o}{dt} + \frac{y_o - We_c}{t_d} \right) \sin(\omega t) \right] \quad (38)$$

where $\frac{1}{t_d} = \frac{C_d}{2} \frac{\mu_l}{\rho_l r^2}$, $\omega^2 = C_k \frac{\sigma}{\rho_l r^3} - \frac{1}{t_d^2}$.

The size of child droplets is given by an energy conservation on the parent droplet and child droplets as

$$r_{32} = \frac{r}{1 + \frac{8Ky^2}{20} + \frac{\rho_l r^3 (dy/dt)^2}{\sigma} \left(\frac{6K-5}{120} \right)} \quad (39)$$

C. Boundary Layer Breakup

When a droplet is moving at a high relative velocity in air, the droplet deforms into a thin disk and child droplets are formed downstream on the edge of the disk. The size of the child droplets is proportional to the boundary layer thickness when the child droplets are about to form [14]. For a parent droplet with diameter d_o moving at u_o in the air, the relationship is given as

$$\frac{d_p}{d_o} = C_s \left(\frac{\rho_l}{\rho_g} \right)^{\frac{1}{4}} \left(\frac{\mu_l}{\rho_l d_o u_o} \right)^{\frac{1}{2}} \quad (40)$$

where C_s is a constant and equals to 6.2 determined experimentally.

V. COMPARISON OF DIFFERENT MODELS

Three different fluid with different surface tensions, viscosities and densities are used to compare the prediction given by different models as provided in Table II.

TABLE II. PROPERTIES OF FLUIDS USED IN THE MODELS

Fluid	σ (N/m)	μ (mPa.s)	ρ (kg/m ³)
Water	0.072	0.89	1000
Acetone	0.025	0.316	784
CA-PEG 10% solids	0.022	152	800

The gaseous phase is assumed to be air under room temperature with density ρ_g of 1.225kg/m³. The thickness of the sheet or the diameter of jet ejected from the nozzle are set as 1 mm in the calculation for the ease of comparison. The result for each model with the jet velocity up to 100m/s are discussed below.

Figure 1 compares jet breakup model (equation 1 in table 1) and wave breakup model (equation 2 in table 1) by Reitz [3]. The jet breakup model is independent of liquid jet velocity and ambient conditions. Therefore, for a fixed liquid and a jet diameter, the resulting droplet diameter is fixed. The results on Fig. 1 show that for a jet diameter of 1mm, the droplet sizes for water, acetone, and CA-PEG are 1.88mm, 1.88mm and 2.41mm, respectively. In addition, the results show the stabilizing effect of viscosity (increase in size with increase in viscosity) and destabilizing effect of surface tension (decrease in size with increase in surface tension).

The wave model includes the effect of jet velocity. Figure 1 shows that the wave model results in the same droplet size as the jet model at jet velocities close to zero. For the wave model, as the relative velocity between liquid and air increases, the resulting droplet size decreases. The viscosity of the fluid help to stabilize the jet, resulting in a larger droplet diameter. CA-PEG, the fluid with a much larger viscosity than water and acetone, has a droplet size of 1.88mm at $U=10\text{m/s}$, which is greater than that of water (1.73mm) and acetone (1.39mm). As the jet velocity increases, the droplet size of CA-PEG is reduced to 0.59mm at $U=100\text{m/s}$, which is greater than that of water (0.45mm) and acetone (0.32mm). The surface tension in the wave model is stabilizing, therefore, water droplet size is always larger than that of acetone. In addition, although the wave model considers the velocity and property of both gas and liquid phases, the absolute value of gas velocity has no effect on the droplet size.

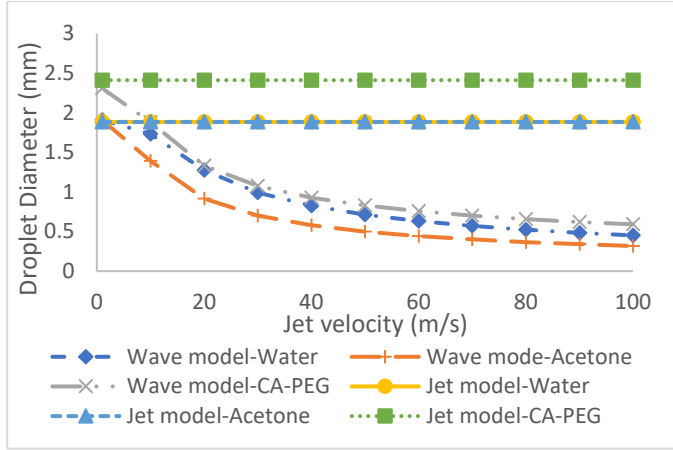


Figure 1. Comparison of jet and wave breakup models (equations 1 and 2 in table 1, respectively).

The droplet size produced by wave breakup model is large. As a result, the wave model is usually applied with the TAB model for secondary breakup. Figure 2 shows the resulting droplet size by TAB model. At low jet velocity (below 20 m/s), the droplets given by wave model is stable and will not experience a secondary breakup. At high jet velocity, the secondary breakup of the droplets creates much smaller droplets. For example, at $U=100\text{m/s}$, the water and acetone droplet diameters are 0.06mm and 0.023mm. The surface tension has a strong effect in stabilizing the droplet: the water droplet diameter after secondary breakup is 13% of that after primary breakup while the acetone droplet diameter is only 7% of that after primary breakup. The TAB model fails to give valid value for droplet diameter of CA-PEG after secondary breakup because for small parent droplets, the high viscosity will result in a negative value for ω^2 and the droplet system becomes overdamped.

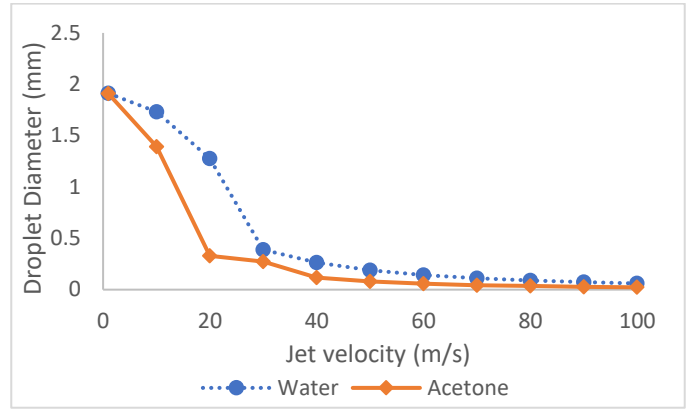


Figure 2. Droplet size for different liquid jet velocities based on wave breakup model (Eq. 2 of Table 1) with TAB secondary breakup.

Figure 3 shows the droplet sizes that are obtained for different sheet velocities using the sheet breakup model of Dombrowski [6], Eq. 3 in table 1. For the sheet breakup model, the droplet size decreases as the jet velocity increase. Figure 3 shows the stabilizing effect of surface tension in this model, as the water droplet size is larger than those of acetone and CA-PEG.

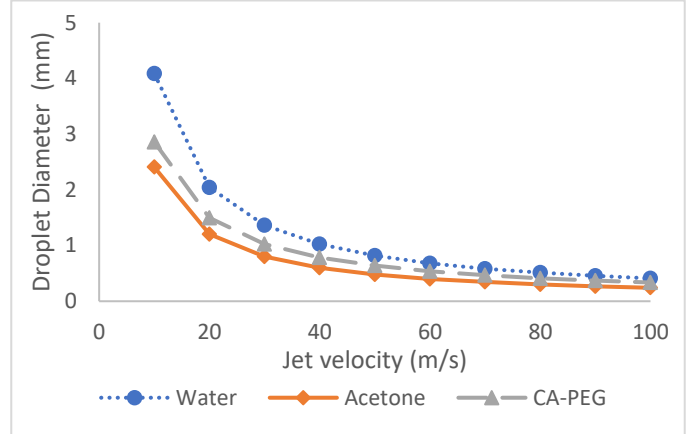


Figure 3. Droplet size for different liquid jet velocities based on sheet breakup model (Eq. 3 of Table 1).

The breakup models of Qin et al. [8] and Aliseda et al. [9] consider the effects of properties and velocities of both liquid and air phase. In these two models, not only the relative velocity between liquid and air phase but also the absolute value of air velocity generated by different air pressure determine the droplet size. As a result, the effects of relative velocity and the air velocity are investigated in the calculation. In addition, the ratio between the cross-sectional area of air exit and liquid exit changes the ratio of mass flow rate between two phases and, consequently, changes the droplet size. Nevertheless, this parameter depends on the nozzle design and is set to 1 in the calculation for the ease of comparison.

Figure 4 and 5 show the droplet diameter versus liquid jet velocity and air velocity, respectively, using Eq. 4 in Table 1. In Fig. 4, the air velocity is set as 120m/s and the jet velocity is

changed from 0m/s to 100 m/s to investigate the effect of relative velocity. As the jet velocity increases, the relative velocity decreases and the droplet size increases, as expected. However, when the jet velocity approaches the air velocity, there is a sudden drop in droplet diameter predicted by the model.

In Fig. 5, the jet velocity is set to 10m/s and air velocity is changed between 50m/s and 250m/s to investigate the effect of air velocity. The droplet size decreases as the air velocity increases. However, the relationship between the size and air velocity is nonlinear, namely, the rate of droplet size change decreases with the magnitude of air velocity. The droplet diameter of acetone is reduced from 0.86mm to 0.46mm by 0.4mm when jet velocity increases from 50m/s to 100m/s, while it is reduced from 0.22mm to 0.17mm by 0.05mm when jet velocity increases from 200m/s to 250m/s. In addition, there are only two curves on each diagram because the model fails to give a valid droplet size with the data for highly viscous CA-PEG fluid, which returns a negative value. In fact, Qin et al. [8] only compare their theory with the experimental data of water at low jet velocity (below 4m/s) and high air velocity (60 m/s to 180 m/s), where the relative velocity is always large. Therefore, this model cannot be applied to high viscous fluids and at high jet velocity close to the air velocity.

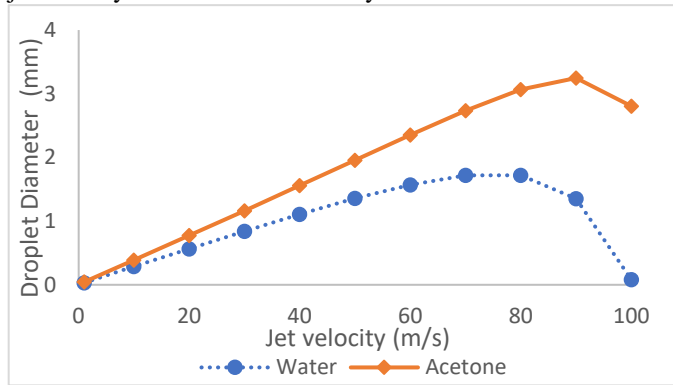


Figure 4. Flat fan model of Qin et al. (Eq. 4 in Table 1): Effect of liquid jet velocity when air velocity equals 120m/s.

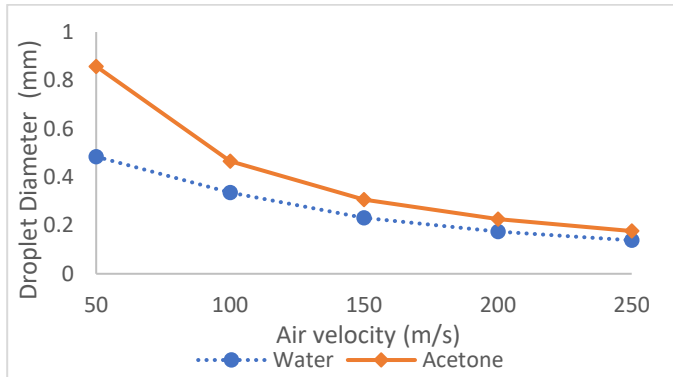


Figure 5. Flat fan model of Qin et al. (Eq. 4 in Table 1): Effect of air velocity when jet velocity equals 10m/s.

Figures 6 and 7 show the droplet diameter versus liquid jet velocity and air velocity, respectively, using Eq. 5 in Table 1. In Fig. 6, the air velocity is set to 120m/s. As the liquid jet velocity increases, the relatively velocity between the liquid and gas jets decreases. This model shows that droplet size increases with increasing liquid jet velocity, however, this increase is unrealistic. The droplet size for water jet increases from 22mm to 108mm as the jet velocity increases from 60m/s to 100m/s. In fact, Aliseda et al. [9] only compare their results at very low jet velocity (0.4m/s) and high air velocity (220m/s) with the experimental data.

In Fig. 7, the liquid jet velocity is set at 10m/s while changing the gas velocity from 50m/s to 250m/s. Increasing the air velocity from 50m/s to 100m/s result in a rapid drop in the droplet diameter from 26mm to 6mm for CA-PEG jet, whereas increasing the air velocity from 200m/s to 250m/s results in a decrease of droplet diameter from 1.65mm to 1.08mm. In this model, the stabilizing effect of viscosity is greater than that of surface tension, since under same condition, CA-PEG spray have a larger droplet size compared to water.

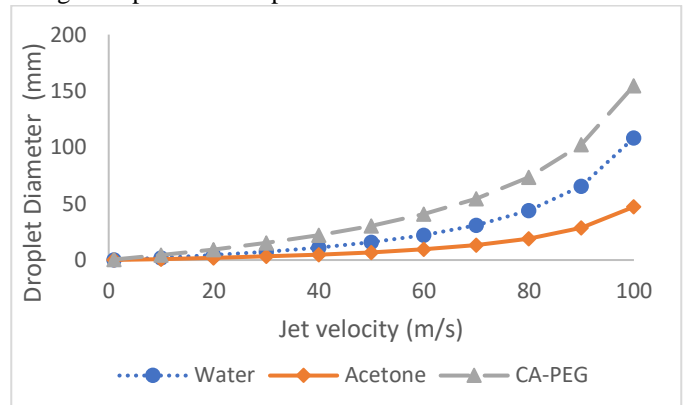


Figure 6. Coaxial jet model of Aliseda et al. (Equation 5 in Table 1): Effect of liquid jet velocity when air velocity equals 120m/s.

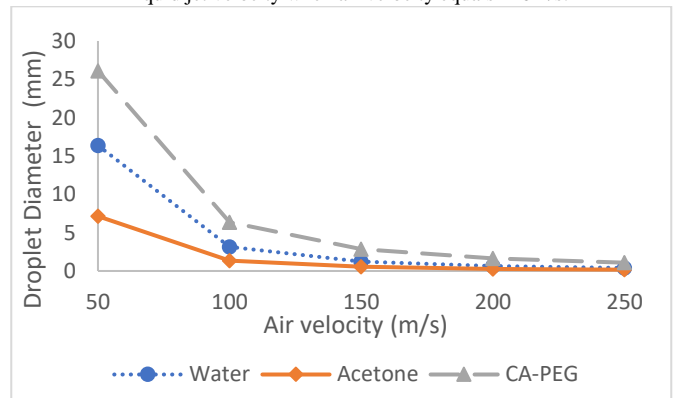


Figure 7. Coaxial jet model of Aliseda et al. (Equation 5 in Table 1): Effect of air velocity when jet velocity equals 10m/s.

When comparing Figure 4 and 6 for the same condition, the droplet sizes are one to two order of magnitude different at larger jet velocities. In Figure 4, the droplet sizes reduce when the relative velocity decreases over certain value due to the arctan

term in the denominator in Eq. 4. In Figure 6, the rapid increase in droplet size at small relative velocity is caused by the small We_{jl} in the denominator.

Figure 8 shows the result of the eddy breakup model by Faeth and coworkers [10, 11] as provided in Eq. 6 in Table 1. As the jet velocity goes up the droplet size decreases. Since the model considers the effect of turbulence, the predicted value is below 0.5mm at a jet velocity of 10m/s, which is almost ten times smaller than the other models. Since the model does not consider the effect of viscosity, it predicts the same droplet sizes for CA-PEG as for acetone.

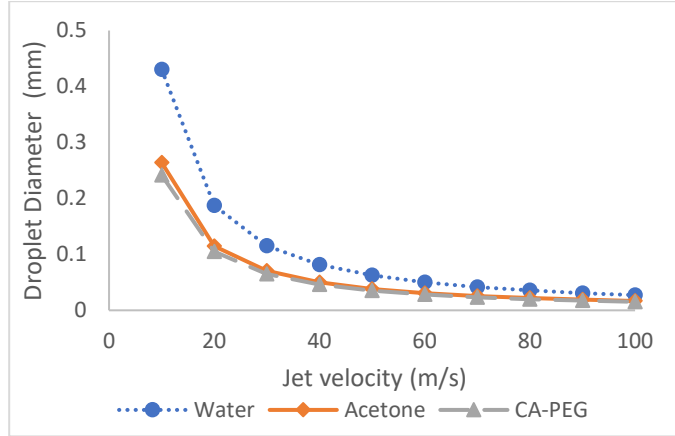


Figure 8. Droplet size for different liquid jet velocities based on eddy breakup model (Equation 6 of Table 1).

VI. CONCLUSION

There are significant amount of effort made to develop physics-based models and correlations to predict a spray droplet sizes at certain operating conditions. Different models have different hypothesis on the breakup process of the spray. For example, the jet breakup model and sheet breakup model assume the breakup across the entire jet or sheet. Wave model, Qin's model, and Aliseda's model assume the generation of droplets from the surface of the liquid sheet or jet. In fact, the physics of the atomization process can change at different operating conditions. Therefore, a model that developed for high relative velocity between liquid and gas may behave poorly at low relative velocities. Qin's model shows a sudden drop in droplet diameter as the relative velocity decreases. Aliseda's model predicts unrealistic growth in droplet diameter as relative velocity decreases. It is critical to examine if these models reflect the real physics of the atomization process properly. Such effort is necessary as it will help understand the spray and develop a model that predicts the spray quality under a wide variety conditions in the industry.

LIST OF NOTATIONS

a	liquid jet radius or half sheet thickness
A_g	cross-sectional area of air
A_l	cross-sectional area of liquid jet or sheet
b_g	the thickness of air in a coaxial jet

k	wavenumber
k_{RP}	Rayleigh-Plateau dominant wavenumber
m_r	ratio of mass flow rate between air and liquid $\frac{\rho_l U_l A_l}{\rho_g U_g A_g}$
Oh	Ohnesorge number $\frac{\mu}{\sqrt{\rho_l \sigma a}}$
Re_{bg}	Reynolds number with respect to bg $\frac{\rho_g U_g b_g}{\mu_{gas}}$
U	relative velocity between air and liquid
U_g	velocity of air
U_l	velocity of liquid jet or sheet
We	Weber number $\frac{\rho_l U^2 a}{\sigma}$
We_g	Weber number with gas $\frac{\rho_g U^2 a}{\sigma}$
σ	surface tension coefficient
μ	dynamic viscosity
ν	kinematic viscosity
ρ_l	density of liquid jet
ρ_g	density of gas
ω	growth rate of wave
λ_{RP}	Rayleigh-Plateau dominant wavelength

REFERENCES

- [1] Ashgriz, N. (2011). Handbook of Atomization and Sprays. Boston, MA: Springer US.
- [2] Senecal PK, Schmidt DP, Nouar I, Ruthland CJ, Reitz RD, Corradini ML, "Modeling high speed viscous liquid sheet atomization", Int. J. Multiphase Flows, 25,1073–1097, 1999.
- [3] Reitz, R.D., "Modeling atomization processes in high pressure vaporising sprays," Atomization and spray technology, 3, pp. 309-337, 1987
- [4] Lin, S. and Reitz, R. (1998). "Drop and spray formation from a liquid jet". Annual Review of Fluid Mechanics, 30(1), pp.85-105.
- [5] Kong, S., Senecal, P. and Reitz, R. (1999). "Developments in Spray Modeling in Diesel and Direct-Injection Gasoline Engines." Oil & Gas Science and Technology, 54(2), pp.197-204.
- [6] Dombrowski N, Johns WR, "The aerodynamic instability and disintegration of viscous liquid sheets", Chem, Eng. Sci. 18, 203–214, 1963.
- [7] Clark CJ, Dombrowski N, "Aerodynamic instability and disintegration of inviscid liquid sheets", Proc. Roy. Soc. A 329, 467–478, 1972.
- [8] Qin, L., Ran, Y., Yang, L., "Theoretical breakup model in the planar liquid sheets exposed to high-speed gas and droplet size prediction," International Journal of Multiphase Flow 98 (2018) pp.158–167.
- [9] Aliseda, A., E.J. Hopfinger, J.C. Lasheras, D.M. Kremer, A. Berchielli, E.K. Connolly, "Atomization of viscous and non-Newtonian liquids by a coaxial, high-speed gas jet. Experiments and droplet size modeling", International Journal of Multiphase Flow, 34, 2, pp.161-175, 2008.
- [10] P-K. Wu, L-K. Tsene, and G.M. Faeth. "Primary breakup in Gas/Liquid Mixing Layers for Turbulent Liquids," Atomization and Spray, Vol. 2, 1992, pp. 295-317.
- [11] Wu, P. and Faeth, G. (1993). "Aerodynamic effects on primary breakup of turbulent liquids". Atomization and Sprays, 3(3), pp.265-289.
- [12] Huh, K., Gosman, A.D., A phenomenological model of diesel spray atomization, in: Proceedings of The International Conference on Multiphase Flows, Tsukuba, Japan, 1991.
- [13] O.J. O'Rourke, A.A. Amsden, "The TAB Method for Numerical Calculation of Spray Droplet Breakup", SAE Technical Paper 872089, 1987.
- [14] Wu, P. and Faeth, G. (1995). Onset and end of drop formation along the surface of turbulent liquid jets in still gases. Physics of Fluids, 7(11), pp.2915-2917.

EFFECTS OF VARYING INCLINATION ANGLE ON ARCHIMEDES SCREW GENERATOR POWER PRODUCTION WITH CONSTANT HEAD

Scott Simmons¹, Murray Lyons, and William Lubitz

School of Engineering
 University of Guelph
 Guelph, Canada

¹ssimmons@uoguelph.ca

Guilhem Dellinger

ICube Laboratory
 ENGEES, INSA Strasbourg
 Strasbourg, France

guilhem.dellinger@engees.unistra.fr

Abstract—Archimedes screw generators are small-scale hydropower devices that are usually installed as diversion systems at sites with low head and moderate flow rates. Screw generators are usually designed empirically – they have been in use as pumps for millennia, and the same practical design principles are commonly implemented on generator schemes. For example, most Archimedes screws are installed at inclination angles at or about 25° - however, there is a lack of evidence to prove that this is best practice. This experiment tested three different screws, identical in all parameters except overall length so that the inclination angle could be varied while maintaining a constant head difference across the screw. Each of the screws were set in a test rig to measure power production. Since the head across the screws was the same, the screws were installed at varying inclination angles to meet this constraint. It was found that the longest screw (with the smallest inclination angle of 15°) performed the best, followed closely by the screw set at the common inclination angle of 25°. The shortest screw (at the steepest inclination angle of 33.8°) performed the poorest. Some suggestions are made with regards to the performance of the screws, and how to improve future power prediction models for Archimedes screw generators.

Keywords—Archimedes screw generator; computational fluid dynamics; microhydro generation; inclination angle; number of blades; efficiency; bucket fill height;

I. INTRODUCTION

An Archimedes screw generator (ASG) is a small-scale hydropower device that is usually installed as a diversion (or run-of-river) system. ASGs have been implemented for hydroelectric energy conversion since the 1990s [1]. They have historically been used as pumping devices with evidence of use dating back as far as the 7th century BCE Assyrian Empire under the reign of King Sennacherib [2]. Due to its simple, robust design, the Archimedes screw has been used to pump a variety of mediums including granular solids in agriculture and mixtures of liquids and solids in wastewater treatment.

The Archimedes screw is now widely accepted as an “eco-friendly” hydropower option – since it allows for sediment, debris, and even aquatic wildlife to pass through the screw unharmed when operating as a generator [3]–[5]. Additionally, as a run-of-river device, up- and downstream conditions in the waterway will be negligibly affected by the installation when compared to the impacts of conventional, impounded hydropower systems [6]. Altogether, this suggests that the device has limited impacts on riverine ecology.

The simple, robust design of the screw also lends itself to low costs of manufacturing and installation [7], [8], and maintenance [7]–[10] – especially when compared to other turbine options with similar operating ranges. Specifically, ASGs operate best under low head and moderate flow rates [11], and can reach efficiencies upwards of 80% in these ranges [12].

The construction of Archimedes screws is considered simple because it is made up of a set of helical blades that are welded around a central cylindrical tube. Screws are usually classified based on their geometries or design flows. The geometry of the screw is shown in Fig. 1.

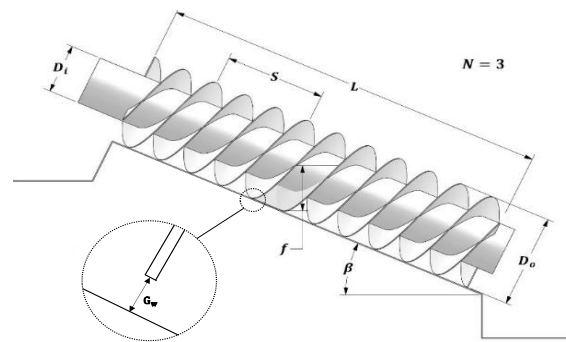


Figure 1. Archimedes screw geometric parameters

The screw rotates within a trough with a small gap between the trough and the blade tips – termed the gap width G_w . The screw has an inner and outer diameter (D_i and D_o , respectively), pitch S , number of blades N , and overall flighted length L . When operating as a generator, water fills in the space between a set of adjacent blades (termed a “bucket” [13]), and a torque

due mostly to the static water pressure on the screw blades causes the screw to rotate. The screw can either have a fixed or variable rotational speed (ω) which, along with the available flow Q and head at the inlet and outlet of the screw, determines the fill height ratio (f) of the buckets.

The fill height ratio is a dimensionless parameter that defines the amount at which a bucket is “full”. The full case, or $f = 1$, can be seen in Fig. 1 – the water level reaches the top of the inner cylinder without spilling over into the next bucket along that flight. Experience suggests that a screw can generate more energy with a small amount of overflow (i.e. if $f > 1$). The fill height ratio is discussed in more detail in the literature [14].

The literature also contains a few models of ASG performance and power losses [15]–[19], as well as some data collected experimentally on laboratory-scale devices [20], [21] or through numerical simulation [22], [23]. However, there is a lack of evidence to definitively assure that the models are valid, particularly since the issue of scaling has not been rigorously addressed (i.e. most of the experimental evidence for ASG power production has been carried out at the laboratory-scale level, not on full-sized installations).

The University of Guelph’s Archimedes Screw Laboratory has an inventory of 16 unique laboratory-scale screws for experimentation. The screws vary in their diameter ratios, pitch, number of blades, and length. The laboratory setup allows for the flow rate, rotational speed, inlet and outlet water levels, and inclination angle to be adjusted.

Experience suggests that most installations have an inclination angle set around $\beta = 25^\circ$ [12], but there is a lack of empirical evidence to back this claim. A recent study showed the effect of variable inclination angles on ASG performance [24]. In the study, a screw was installed in the University of Guelph’s laboratory and the inclination angle of the screw was varied, and the power production recorded. The experimental dataset was then extended with a computational fluid dynamic (CFD) simulation of the screw, which allowed the dataset to be extended to screw geometries beyond what was available in the laboratory. The head of the screw changed as the inclination angle of the screw changed in these experiments since the screw’s geometry remained constant (cf. Fig 2). This meant that the power production could not be directly compared between the different orientations (since the available power in the water changes with the head), but the efficiencies of the screw orientations could be compared.

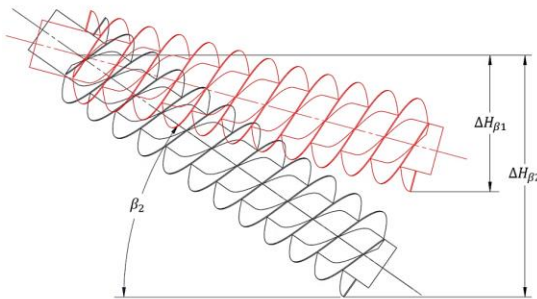


Figure 2. Head changes with inclination angle in these experiments, as the screw length and geometry are kept constant [24].

The objective of the study described in this paper was to compare the effect of inclination angle to power production and performance of an ASG, while varying inclination angle and keeping all other geometric parameters constant.

A couple hypotheses were made towards the goal of this paper. Firstly, as the inclination angle increases it is expected that the overflow and gap leakage losses will increase; however, as the rotational speed of the system increases they are expected to decrease at any inclination angle. For example, it is expected that a screw at the steepest inclination and slowest rotational speed will have the most leakage and overflow losses.

Secondly, as the inclination angle decreases, it is expected that the minor losses in the system will become more evident. The lower inclination angle lends itself to a longer screw – which has more surface area to exhibit frictional losses.

An experiment was carried out to gather performance data for laboratory-scale ASGs under a range of inclination angles and rotational speeds. The data was then analyzed and used to draw conclusions related to the hypotheses.

II. EXPERIMENTAL METHODS

There is a set of three screws in the University of Guelph’s Archimedes Screw laboratory that have identical parameters except their flighted length – allowing them to be set at varying inclination angles that correspond to the same flow and head conditions at the screw’s inlet and outlet. For the purposes of this paper, the three screws will be called the “short screw”, “medium screw”, and “long screw”. They are shown drawn to scale in Fig. 3.

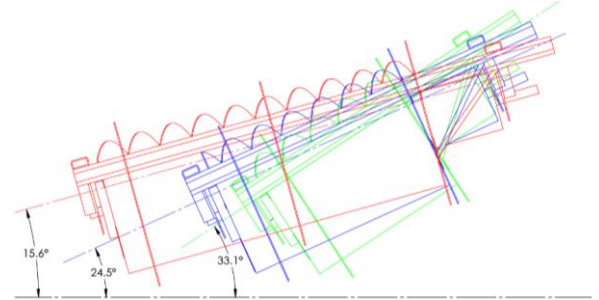


Figure 3. Angles of three screws with constant head. The short screw (green) has the steepest inclination angle, the long screw (red) has the shallowest inclination angle, and the medium screw (blue) has been set at an inclination angle of $\beta = 24.5^\circ$ to match with previously collected data and the most common inclination angles of real-world ASG installations.

The dimensions of the three lab-screws are shown in Table 1. The inclination angle is found with respect to the flighted length of the screw and the constant head difference between the set of screws as demonstrated in Fig. 4.

The tests were started with the medium screw, since it was the closest representation to a typical “real-world” installation’s inclination angle. The screw was installed in the Archimedes screw test rig at the University of Guelph shown in Fig. 5. The system has settings to adjust rotational speed and the flow of water through the screw. The rotational speed of the screw is controlled with a variable frequency drive (VFD). The VFD controller allows the motor to turn at a specified frequency and

will either apply a brake or accelerate the screw to maintain the set speed. A torque arm is attached to the VFD, and force is measured between the torque arm and the fixed frame of the setup. The water flows through a recirculating loop, and flow rate is controlled with a variable water pump in the lower basin. The system is also designed to utilize weirs to control the water level in the lower basin – allowing for the outlet water level and head difference in the system to be controlled.

Table 1. Experimental screw dimensions with corresponding inclination angles.

	Symbol	Short Screw	Medium Screw	Long Screw
Number of Blades	N	4	4	4
Inner Diameter (mm)	D_i	168	168	168
Outer Diameter (mm)	D_o	381	381	381
Pitch (mm)	S	381	381	381
Flighted Length (mm)	L	478	617	952
Inclination Angle (deg)	β	33.8	24.5	15.6

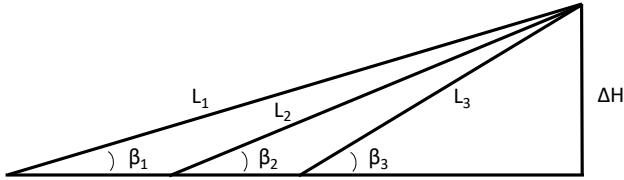


Figure 4. As flighted length of the screw changes, the inclination required to maintain a constant head ΔH also changes.

The laboratory setup has sensors that measure the flow rate, torque, rotational speed, bucket fill height, and upper and lower basin water levels. The flow rate is measured with a propeller-type flow meter (Omega FTB-740) and verified with a Cipoletti weir on the inlet end of the system. Both the upper and lower basin have depth sensors (Keller Series 26 Y) in stilling wells that measure water level – these are verified with manual readings of the water level. The rotational speed of the screw is measured with an optical tachometer and verified with a separate handheld unit. Finally, the torque was sampled by mounting a load cell (Omega LC703-25) to a torque arm attached on the system’s VFD motor/generator.



Figure 5. Laboratory test setup.

The process for each experimental run was as follows. The screw and recirculation pump were set to a specified rotational speed and flow rate, respectively. Next, the system was given a few minutes to reach an equilibrium condition at the new settings. It was noted that the screw usually required less than a minute to reach equilibrium conditions after a change of settings. After equilibrium conditions were met, a data acquisition program was run to sample the sensors on the screw setup for one minute at a frequency of 1000 Hz – a frequency much higher than the Nyquist rate for the apparatus. The one-minute sample was then time-averaged to more accurately predict the system’s response to the given flow, rotational speed, and head difference. Next, the data from the sample run was reviewed to see if the flow rate, rotational speed, and outlet water level were within 1% of their target values – if it met the criteria, the data was accepted and recorded. The flow rate and rotational speed of the next run were then set, the screw was allowed to reach equilibrium at the new conditions, and the process was repeated.

It is important to note that all of the tests in these experiments were carried out at outlet water levels of 60%, since previous testing has shown this to be the most optimal outlet condition for these screws [14], [21].

The medium length screw was run through a range of five flow rates (9, 9.5, 10, 10.5, and 11 L/s), and, at each of these flow rates, ten different rotational speeds (30 to 39 RPM) to find the setting that corresponded to the highest power production and efficiency. It was found that the most optimal flow rate for this screw was 10 L/s. Afterwards, the short and long screws were run through the same range of rotational speeds at the 10 L/s flow rate. The experiment was carried out this way to maintain the available power in the water. The flow rate corresponding to the most optimal power production in the medium screw was selected as the base case as a point of comparison (i.e. all the screws must be tested at the same flow rate and head difference, so the most optimal setting was chosen arbitrarily for the medium screw since it was most reflective of a real-world installation).

III. RESULTS

As mentioned above, tests were first carried out on the medium screw to determine the optimal flow rate for the experiment. The plot in Fig. 6 shows the efficiency of the screw with respect to rotational speed for each of the five flow rates tested. An uncertainty analysis was carried that found an average uncertainty for efficiency values based on measured data of about $\delta\eta_{\text{avg}} \approx 0.054$ or 5.4%. The magnitude of the uncertainty is due to propagation of error in subsidiary measurements – as will be shown shortly.

In Fig. 6, the highest efficiencies occur along the curve for 10 L/s; the 10.5 L/s curve seems to have similar maximum efficiencies but tends to drop off faster at the lower rotational speeds. It is interesting to note that the maximum efficiency points seem to shift to different rotational speeds as the flow rate changes. As flow rate increases, the maximum efficiency occurs at higher rotational speeds. Since the data is so close together for all curves (noting that the Fig. 6 ordinate (“y-axis”) spans a range of less than 10% efficiency – the error bars also

help to reinforce the closeness of the datapoints), the points near the highest efficiencies were run multiple times: multiple sets of one-minute time-averaged data were collected and averaged to verify that the points were not outliers and were representative of the performance of the screw under those conditions. It was determined that the medium screw performed best at a rotational speed of 36 RPM when experiencing a flow rate of 10 L/s.

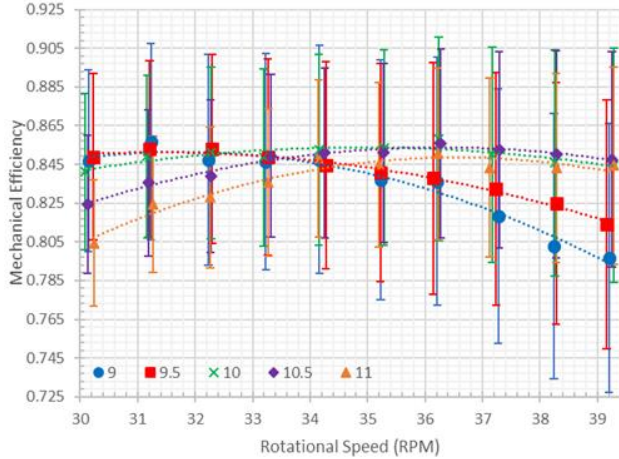


Figure 6. Mechanical efficiency of Screw 15, varying rotational speed from 30 to 39 RPM for 5 different flow rates (9, 9.5, 10, 10.5, 11 L/s). The plot has approximate trend lines superimposed to make it easier to visualize the data trends.

The short and long screws were then run at 10 L/s for the same range of rotational speeds, the results of these tests are shown in Fig. 7. As the figure shows, the long screw (with the smallest inclination angle) proved to be the best performing screw – it generated the most power and was thusly the most efficient of the three inclination angles. The medium screw had similarly high efficiencies and the two trends seemed to converge as the rotational speed increased. The short screw (i.e. the steepest inclination angle) was the least efficient option but performed better at higher rotational speeds as it has less drop-off near the end of the curve.

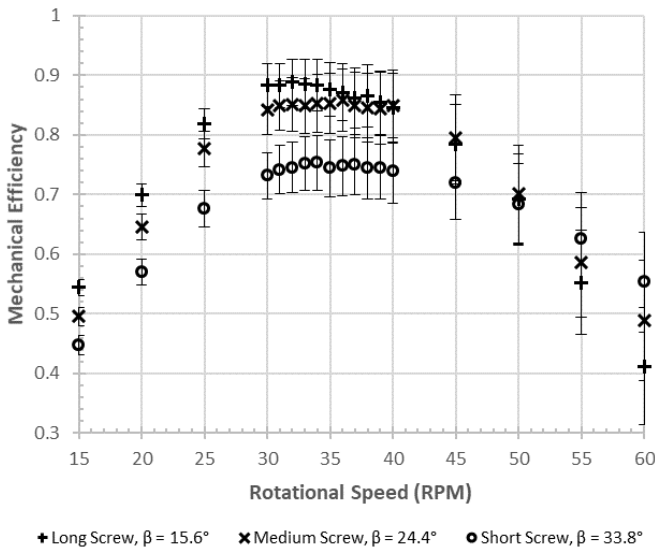


Figure 7. Mechanical efficiency of screw 14, 15, and 16 for varying rotational speeds.

Fig. 7 includes error bars calculated for each test run. The uncertainty in the mechanical efficiency was found by carrying out an uncertainty analysis given the accuracies of the equipment used in the experiment to calculate the power and efficiency of the screw. The power produced by the screw is a function of the torque and the rotational speed. Since the data was sampled at a frequency of 1000 Hz, and the rotational speed of the screw was less than 60 RPM, the uncertainty in the rotational speed measurements was neglected. The torque measurements were taken with a strain gauge that was shown to have an average error of $\delta T = 0.22 \text{ N}\cdot\text{m}$ [14].

Efficiency is calculated by dividing the mechanical power produced by the screw by the power available in the flow (i.e. $\eta = P / P_w$, where $P_w = \rho g h Q$). Two Keller Series 26 Y depth sensors were used to measure the head (h) each having an accuracy of 0.25 %FS. An Omega FTB-740 Inline Turbine Flow Meter was used to measure the flow rate, having an accuracy of 1 %FS. After carrying out uncertainty analysis it was shown that the average error in the efficiency measurements was about $\delta \eta_{\text{avg}} \approx 0.054$ or 5.4%, as mentioned above.

The better performance of the long screw relative to the other two is likely due to the filling scenarios and the scale of the buckets. Past experience suggests that short screws will exhibit more gap leakage, more overflow leakage, will not contain as many buckets (since they are shorter), and that each bucket will have less volume due to the cylindrical and helical shape of the bucket geometry [25]. It was observed during testing that the short screw only has a fully formed bucket of water for a brief time before the water exits the screw. Since it is so short, there is only a very brief period where the height of the water in a bucket is fully independent of the upper or lower basin.

Interestingly, the long screw was predicted to underperform the medium screw – this was shown false in the experiments. While the power output of a short screw (at steep inclination) tends to be limited by geometric constraints, longer screws (at shallow inclinations) will experience more minor losses and bearing losses [21], [25]. In this case, since the screws are all laboratory-scale, it may be that the minor losses and bearing losses will make up a much larger proportion of power loss than in a real-world installation. In any case, the long screw was able to fully form more buckets during power production and seemed to have less overflow than the medium screw. Power production is dominantly based on the static pressure distributions within the buckets. The fill heights within each of the three screws during their most efficient tests are shown in Fig. 8 to lend more insight into the performance curves found in Fig. 7. Since there are four blades for these screws (i.e. $N = 4$) the curves of Fig. 8 have four separate peaks and valleys for each full rotation, and therefore a full rotation of all the screw blades is shown in Fig. 8.

Fig. 8 shows that the buckets of the medium and long screws have fully formed, with bucket fill height ratios above $f = 1$, indicating that some overflow leakage is occurring in these runs. Given that the error of the Omega PX309 pressure sensor was 2%, and, taking into account error in measuring

offset, the error in the fill height measurements was found to be an average of $\delta f_{\text{avg}} = 0.017$.

The short screw seems to have fill heights much lower than a full bucket, which is further evidence that it does not have fully formed buckets along its length. It is interesting to note that the slope of the sawtooth (cf. Fig. 8) is proportional to the inclination angle of the installation. This means that the bucket will be filled closer to level as the length of the screw increases (i.e. the shallower incline corresponds to a larger bucket volume). It is also evident from Fig. 7, that the medium length screw experienced the most overflow leakage – partly leading to its lower-than-expected efficiency performance.

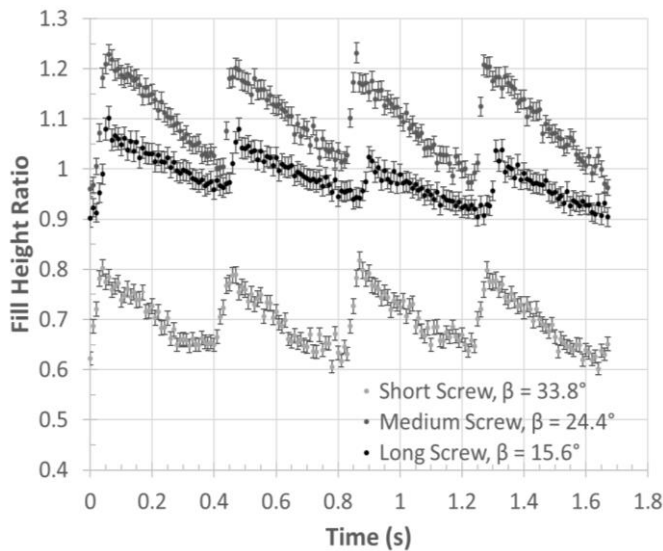


Figure 8. Fill height ratios for screws 14, 15, and 16 over one full screw rotation.

These experiments provided insight into the effects of inclination angle on power production in ASGs – they will serve to help further develop power loss modelling in screw generators.

IV. CONCLUSIONS

The experiment presented in this paper provides valuable insight into the effect of inclination angle on the performance of Archimedes screw generators. Experience has suggested that ASGs operate best at inclination angles around 25°, but the laboratory-scale data shows that this may not be true for all sizes of screws. Indeed, at this system scale the best performing screw had an inclination angle of $\beta=15.6^\circ$.

The experimental observations provide support for some of the hypothesis made in this paper. For example, as the inclination angle increases the overflow losses and leakage losses are expected to increase. The experiments demonstrated that this is likely true, since there is much less power produced in this screw. As well, if volume flow rate is held constant, gap leakage and overflow leakage tend to decrease as rotational speed increases [21], and since the short screw performs best in higher rotational speeds, its likely because these effects are lessened in this range.

Another early hypothesis was that as the inclination angle decreased, it was expected that there would be an increase in minor losses in the system. This effect was not wholly evident in the data; however, it is very likely to be negligible in these laboratory-scale devices. In larger installations, a change in inclination angle can mean a screw that is meters longer, not centimeters (as seen in these experiments). Since the screws are all among a similar size and length at this laboratory-scale, it is suggested to continue to explore the relationship between inclination angle and power output in more detail on larger-scale installations or by means of CFD.

ACKNOWLEDGMENT

Aspects of this work were financially supported by the Natural Sciences and Engineering Research Council (NSERC) Collaborative Research and Development (CRD) program (grant # CRDPJ 433740-12) and Greenbug Energy Inc. (Delhi, ON, Canada). The assistance of Tony Bouk and Brian Weber of Greenbug Energy Inc. is gratefully acknowledged.

The authors also gratefully acknowledge Pierre-André Garambois, Nicolas Dellinger, Abdelali Terfous, and Abdellah Ghenaim for their assistance with this collaborative research.

REFERENCES

- [1] K.-A. Radlik, "Hydrodynamic screw for energy conversion - uses changes in water supply to regulate energy output," DE4139134A1, 1997.
- [2] S. Dalley and J. P. Oleson, "Sennacherib, Archimedes, and the Water Screw: The Context of Invention in the Ancient The Context of Invention in the Ancient World," *Technol. Cult.*, vol. 44, no. 1, pp. 1–26, 2003.
- [3] P. Kibel, "Archimedes Screw Turbine Fisheries Assessment. Phase II: Eels and Kelts," *Report. Fishtek Consult. Ltd.*, 2008.
- [4] C. D. McNabb, C. R. Liston, and S. M. Borthwick, "Passage of Juvenile Chinook Salmon and other Fish Species through Archimedes Lifts and a Hidrostal Pump at Red Bluff, California," *Trans. Am. Fish. Soc.*, vol. 132, no. 1985, pp. 326–334, 2003.
- [5] T. B. Havn *et al.*, "Downstream migration of Atlantic salmon smolts past a low head hydropower station equipped with Archimedes screw and Francis turbines," *Ecol. Eng.*, vol. 105, pp. 262–275, 2017.
- [6] P. Fearnside, "Greenhouse Gas Emissions from a Hydroelectric Reservoir (Brazil's Tucurui Dam) and the Energy Policy Impactions," *Water Air Soil Pollut.*, 2002.
- [7] O. M. Dada, I. A. Daniyan, and O. . Adaranola, "Optimal Design of Micro Hydro Turbine (Archimedes Screw Turbine) in Arinta Waterfall in Ekiti State, Nigeria," vol. 4, no. 2, pp. 34–38, 2014.
- [8] Spaans Babcock Hydro Power, "Archimedean Screw Turbine." Spaans Babcock Hydro Power, Balk, 2012.
- [9] D. Bennion, "Maintaining Archimedes Screw Pumps," *ECS Engineering Services*, 2013. [Online]. Available: <http://www.ecsengineering-services.com/maintaining-archimedes-screw-pumps/>.
- [10] ECS Engineering Services, "Archimedes Screw Pumps." ECS Engineering Services, Sutton-in-Ashfield, 2016.

- [11] S. J. Williamson, B. H. Stark, and J. D. Booker, "Low head pico hydro turbine selection using a multi-criteria analysis," *Renew. Energy*, vol. 61, no. 0, pp. 43–50, 2014.
- [12] A. Lashofer, W. Hawle, and B. Pelikan, "State of technology and design guidelines for the Archimedes screw turbine," *Univ. Nat. Resour. Life Sci. Vienna*, pp. 1–8, 2012.
- [13] C. Rorres, "The Turn of the Screw: Optimal Design of and Archimedes Screw," *J. Hydraul. Eng.*, vol. 126, pp. 72–80, 2000.
- [14] K. Songin, "Experimental Analysis of Archimedes Screw Turbines," Masters Thesis, University of Guelph, 2017.
- [15] A. Kozyn, "Power Loss Model for Archimedes Screw Turbines," Masters Thesis, University of Guelph, 2016.
- [16] W. D. Lubitz, M. Lyons, and S. Simmons, "Performance Model of Archimedes Screw Hydro Turbines with Variable Fill Level," *J. Hydraul. Eng.*, vol. 140, no. 10, pp. 1–11, 2014.
- [17] G. Müller and J. Senior, "Simplified theory of Archimedean screws Théorie simplifiée de la vis d'Archimède," *J. Hydraul. Res.*, vol. 47, no. 5, pp. 666–669, 2009.
- [18] D. M. Nuernbergk and C. Rorres, "An Analytical Model for the Water Inflow of an Archimedes Screw Used in Hydropower Generation," *J. Hydraul. Eng.*, vol. 139, p. 120723125453009, 2012.
- [19] D. Nuernbergk, *Wasserkraftschnecken - Berechnung und optimaler Entwurf von archimedischen Schnecken als Wasserkraftmaschine (Hydropower screws - Calculation and Design of Archimedes Screws used in Hydropower)*, 1st ed. Detmold: Verlag Moritz Schäfer, 2012.
- [20] G. Dellinger, A. Terfous, P.-A. Garambois, and A. Ghenaim, "Experimental investigation and performance analysis of Archimedes screw generator," *J. Hydraul. Res.*, vol. 54 (2), pp. 197–209, 2016.
- [21] S. Simmons, K. Songin, and W. Lubitz, "Experimental investigation of the factors affecting Archimedes screw generator power output," in *Proc. of HYDRO 2017*, Seville, Spain, Oct. 9–11, 2017.
- [22] G. Dellinger *et al.*, "Computational fluid dynamics modeling for the design of Archimedes Screw Generator," *Renew. Energy*, vol. 118, pp. 847–857, 2018.
- [23] S. Simmons, "A Computational Fluid Dynamic Analysis of Archimedes Screw Generators by," Masters Thesis, University of Guelph, 2018.
- [24] G. Dellinger, S. Simmons, W. D. Lubitz, P.-A. Garambois, and N. Dellinger, "Effect of slope and number of blades on Archimedes screw generator power output," *Renew. Energy*, vol. 136, pp. 896–908, 2019.
- [25] K. J. Songin and W. D. Lubitz, "Measurement of fill level and effects of overflow in power-generating Archimedes screws," *J. Hydraul. Res.*, 2019.

A NUMERICAL INVESTIGATION ON THE FUNDAMENTALLY DIFFERENT MECHANISMS OF CIRCULAR HYDRAULIC JUMP FOR HIGH AND LOW-VISCOSITY LIQUIDS

Yunpeng Wang, Roger E. Khayat
 Mechanical & Materials Engineering
 Western University
 London, Canada
 ywan988@uwo.ca

Abstract—Circular hydraulic jump affects numerous industrial processes such as, but not limited to, jet cooling, jet quenching and surface cleaning etc. The prediction of the jump location is of vital importance in such applications. In the present work, the formation of circular hydraulic jump in liquid jet impingement flow is investigated with a numerical approach. The numerical model is first validated against existing experiment and then adopted to study the effects of different parameters involved in the flow configuration. The influences of viscosity, gravity and surface tension are comprehensively explored. Some important new findings are reported, and their implications are discussed

Keywords—*impinging jet; hydraulic jump; multiphase flow; surface tension; free-surface flow*

I. INTRODUCTION

When a liquid jet of radius impacts a solid surface, it spreads out radially as a thin film until the height of the liquid layer rises abruptly at a radius, where a circular hydraulic jump is formed as illustrated in Fig. 1. The region before jump is characterized by having high speed and commonly referred to as the super-critical region. In flow field after the jump, which is also known as the sub-critical region, the velocity of the liquid film significantly drops due to the sudden increase of the fluid depth. Circular hydraulic jump can be daily observed when the tap water hits the bottom of a kitchen sink.

Impinging jet flow is important in numerous industrial applications such as jet cooling, jet quenching and surface cleaning etc. The fast motion of liquid inside the jump provides high rate of heat and mass transfer whereas the low velocity caused by hydraulic jump dramatically harms the performance [1]. Consequently, the prediction of location (radius) of the hydraulic jump is crucial in the design of relevant processes. In the last few decades, extensive theoretical, experimental and numerical studies are devoted to find the cause of hydraulic jump and consequently predict its location but until now success is still limited.

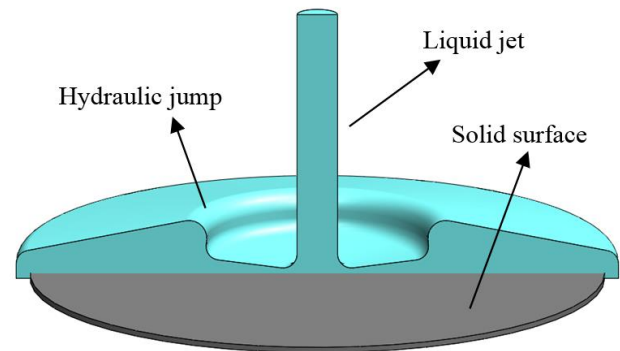


Figure 1. Illustration of circular hydraulic jump.

To the authors' knowledge, two branches of studies exist in the literature on the circular hydraulic jump. On one hand, it is naturally expected that the height after the jump should have a tangible effect on the location of the jump. In other words, a larger sub-critical depth would result in a smaller jump radius due to a larger reverse hydrostatic pressure gradient at the jump. The first major contribution based on this idea is from Watson [2]. Neglecting gravity and surface tension before the jump, Watson developed a similarity solution using boundary layer approach. Assuming the liquid height downstream of the jump is known, Watson obtained the location of the jump by considering the force and momentum balance across the jump. Watson's theoretical method yielded reasonably good agreements with his experimental result and became the basis for numerous later studies. Bush and Aristoff [3] found that Watson's theory is less accurate in the limit of relative weak jump for which surface tension effect should be important. In this respect, they incorporated surface tension in the relation of momentum and force balance at the jump and achieved better agreements with experiments. It is important to note that, in the Watson-Bush approach, the sub-critical depth is assumed to be known. In their experiments this height is controlled by

mounting a barrier downstream of the jump. However, we note that hydraulic can occur without any barrier as we can see every day in a kitchen sink. Consequently, to use this approach, one has to measure the height of the jump from experiment and use it as an input. To resolve this issue, Wang and Khayat [4] proposed a theory that the height after jump can be approximated by the thickness of large liquid puddle in static condition as the fluid velocity after the jump is indeed small under most circumstances. Therefore, the height of the jump can be determined with the knowledge of the static contact angle between the liquid and solid surface using an energy minimization theory [5]. While this method works well for low viscosity liquids, it is less satisfying for liquids of high viscosity due to the dominant viscous effect [6].

The other branch of studies originates from the work of Tani [7] and it was comprehensively explored by Bohr, Dimon and Putkaradze [8]. In these works, hydrostatic pressure is included in the boundary layer equations and a divergence point was found from the super-critical region. Based on this divergence location, a scaling relation, which is free from the knowledge of sub-critical depth, for the jump location is deduced in terms of the viscosity, flow rate and gravity. While some good agreements are achieved, discrepancies were found especially when compared with the data of water which has low viscosity but high surface tension. Recently, Bhagat, Jha, Linden and Wilson [9] highlighted the importance of the role of surface tension in determining the location of the jump using a surface energy approach in an approximate manner. Based on their analysis and experiments, Bhagat et al. concluded that, for circular hydraulic jump, surface tension is the dominant effect on the formation of the jump and gravity plays little role. In another recent study, however, Wang and Khayat [6] found that for high-viscosity liquids, the location of the jump can be well predicted without the inclusion of surface tension and the cause for the jump can be mostly attributed to gravity for such liquids. They concluded that surface tension is more important for low viscosity liquids, or for liquids with high surface tension, on triggering the hydraulic jump.

It is worth noting that though these works seem to be contradictory, their predictions are all validated by experiments. We here infer that these different findings in the literature should be the result of the boundary layer approach which assumes that the radial flow varies much slower than the vertical flow and the local curvature of the liquid film is small [10]. This assumption however does not strictly hold near the hydraulic jump region where flow varies significantly in both the radial and vertical directions. In this regard, numerical method has the advantage as it solves the full Navier Stokes equation throughout the domain with no truncation. However, due to the complexity of multiphase flow, limited number of numerical works exist in the literature for impinging liquid jet and less are on the formation of circular hydraulic jump. Ellegaard, Hansen, Haaning and Bohr [11] studied the flow separation phenomenon under the hydraulic jump. To circumvent the difficulties caused by the free surface, they replaced the liquid-air interface by a fixed, but stress-free surface at prescribed locations in their simulations. In other words, they fixed the free surface and only solved for the flow. Passandideh-Fard, Teymourtash and Khavari [12] proposed a

numerical method to determine the hydraulic jump location. The free surface was tracked by a volume-of-fluid [13] approach. In their calculation domain, however, the downstream depth was imposed at the disk edge.

To the best of our knowledge, most of existing works focused on the combined effect of gravity, surface tension on hydraulic jump. Few of them have singled out these physical properties and explored their individual contribution. Also, very few numerical works, if there are any, exist on the circular hydraulic jump on a flat plate, i.e. for the case where there is no barrier after the jump. More importantly, from the fore-mentioned contradictory findings, we infer that, for high and low-viscosity liquids, the mechanism for hydraulic jump might be different. In this context, the current work is focused on investigating the effects of gravity and surface tension on the hydraulic jump for both high and low-viscosity liquids. The combined effects of these physical properties are also explored.

II. GOVERNING EQUATIONS AND NUMERICAL METHOD

For laminar flow, the governing equations for mass and momentum in the axis-symmetric domain are given by [14]

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial r}(\rho v_r) + \frac{\rho v_r}{r} = 0, \quad (1)$$

$$\begin{aligned} & \frac{\partial}{\partial t}(\rho v_x) + \frac{1}{r} \frac{\partial}{\partial x}(r \rho v_x v_x) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v_r v_x) \\ &= -\frac{\partial p}{\partial x} + \frac{1}{r} \frac{\partial}{\partial x} \left[r \mu \left(2 \frac{\partial v_x}{\partial x} - \frac{2}{3} (\nabla \cdot \vec{v}) \right) \right] \\ &+ \frac{1}{r} \frac{\partial}{\partial r} \left[r \mu \left(\frac{\partial v_x}{\partial r} + \frac{\partial v_r}{\partial x} \right) \right] + F_x \end{aligned} \quad (2)$$

and

$$\begin{aligned} & \frac{\partial}{\partial t}(\rho v_r) + \frac{1}{r} \frac{\partial}{\partial x}(r \rho v_x v_r) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v_r v_r) \\ &= -\frac{\partial p}{\partial r} + \frac{1}{r} \frac{\partial}{\partial x} \left[r \mu \left(\frac{\partial v_r}{\partial x} + \frac{\partial v_x}{\partial r} \right) \right] \\ &+ \frac{1}{r} \frac{\partial}{\partial r} \left[r \mu \left(2 \frac{\partial v_r}{\partial r} - \frac{2}{3} (\nabla \cdot \vec{v}) \right) \right] \\ &- 2\mu \frac{v_r}{r^2} + \frac{2}{3} \frac{\mu}{r} (\nabla \cdot \vec{v}) + \rho \frac{v_z^2}{r} + F_r \end{aligned} \quad (3)$$

where

$$\nabla \cdot \vec{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_r}{\partial r} + \frac{v_r}{r}. \quad (4)$$

Here F_x and F_r are the components of body force in x and r directions. The liquid-air interface is tracked by the volume of fluid (VOF) method [15] and governed by

$$\frac{\partial}{\partial t} \alpha_q + \nabla \cdot (\alpha_q \vec{v}_q) = 0, \quad (5)$$

where α_q is the volume fraction of the liquid phase. ANSYS Fluent 15.0 is used to solve the 2-D axis-symmetric governing equations. The SIMPLEC scheme is selected for pressure correction for better convergence, and QUICK method is adopted for the spatial discretization. We emphasize here that the flow after the jump should be elliptic, so the QUICK scheme is more appropriate than an upwind scheme. The volume Fraction equation is discretized by a comprehensive scheme [14]. The convergence criteria for mass, momentum and volume fraction are 10^{-4} , 10^{-6} and 10^{-6} respectively. As we do not have distinct boundaries for each individual phase, the transient solver is preferred over the steady solver for better convergence. In fact, the flow always diverges under steady solver based on our test. In this case, a second-order bounded time scheme is adopted.

III. MESH DOMAIN AND BOUNDARY CONDITIONS

The 2-D axis-symmetric domain and adopted mesh of the impinging jet are illustrated in Fig. 2. A uniform inlet velocity is imposed based on the average velocity of the incoming jet. The pressures are set to be atmospheric at the outlets. To capture the edge effect, a side wall near the edge of the domain is constructed. The upper boundary is set to be a no-slip wall instead of pressure outlet as we found that over-imposing the pressure might drive the flow to unrealistic state. Therefore, it is better to use a wall condition and allow the pressure to freely adjust itself based on momentum transport. In fact, using a wall instead of outlet should not affect the flow in practical system as the viscosity of gas phase is much smaller than that of a liquid. Surface tension effect is modeled using the continuum surface force model proposed by Brackbill, Kothe and Zemach [16]. The mesh is generated in a way such that flow near the disk surface can be well captured. In this case, the grid is highly refined near the bottom wall as shown in the zoom-in view in Fig. 2. The grid independence test was based on the criteria that the variation of jump location is less than 0.5%.

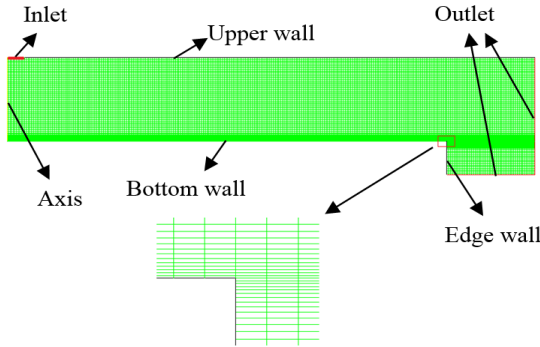


Figure 2. Physical domain and grid configuration.

IV. RESULTS AND DISCUSSIONS

Before investigating the effects of gravity and surface tension, we first validate the adopted numerical model by comparing the result with existing experiment and investigate the transient behavior of hydraulic jump. After the validation of the model, the influences of the aforementioned physical properties are explored, and the corresponding implications are discussed for high and low-viscosity liquids.

A. Validation of Model and The Transient Behavior

To validate the current model, we compare our numerical result with the experiments from Duchesne, Libon and Limat [17]. The liquid in use is silicon oil with density of 950 kg/m^3 . The viscosity and surface tension are $19 \text{ mPa}\cdot\text{s}$ and 20 dyn/cm respectively. The liquid is injected into the domain with a volume flow rate of 17 mL/s through a nozzle of radius 1.6 mm . The radius of the disk is 150 mm and the disk-to-nozzle distance is 16 mm . As illustrated in Fig. 3, both the location of the jump and the overall liquid thickness agree reasonably well with the experimental measurements. Here the radius r and the film thickness h are both scaled by the nozzle radius a .

The transient behavior of the hydraulic jump is demonstrated in Fig. 4. Interestingly, the location of the hydraulic jump seems not to vary much once it is formed. In other words, the location of the jump is not much affected by the flow after the jump if the flow is allowed to drain freely at the edge of the disk. This automatically confirms the findings of Wang and Khayat [6].

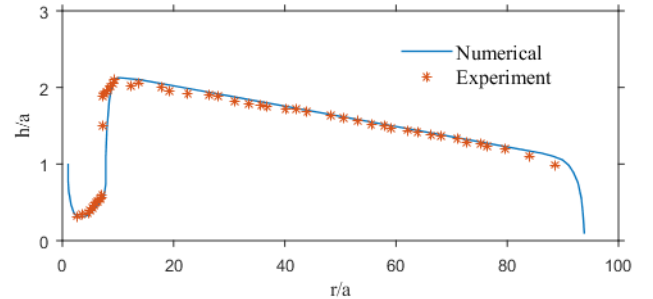


Figure 3. Validation against experiment.

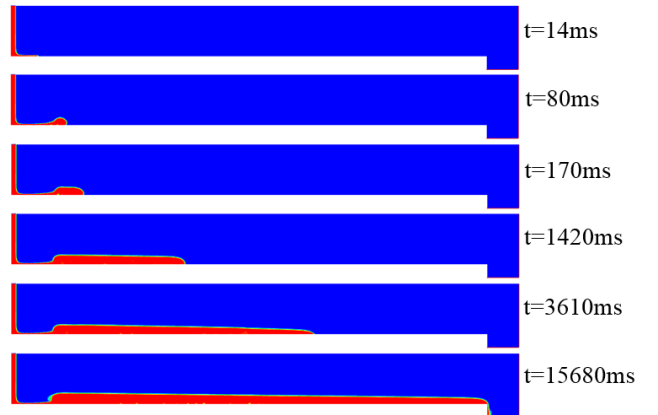


Figure 4. The transient behavior of circular hydraulic jump.

From the observation in Fig. 4, we conclude that, the location of the jump can, as it should, be determined without the knowledge of the flow in the sub-critical region unless a barrier is artificially imposed somewhere downstream of the jump.

B. The Influence of Gravity and Surface Tension

The effects of gravity and surface tension on circular hydraulic jump for a high-viscosity liquid are illustrated in Fig. 5. The liquid in use is still the same silicon oil and the flow is under the same inlet condition and nozzle radius as that in our validation. To lower the computational cost, we have shortened the radius of the domain to 25mm and reduced the nozzle-to-nozzle distance to 8mm. As can be seen, when the flow is free from both surface tension and gravity, no hydraulic jump forms. The thickness of the liquid film gradually increases due to the accumulating viscous effect until it leaves the edge of the disk where a bending is observed due to the removal of wall shear stress. Interestingly, in the single presence of either gravity or surface tension, the circular hydraulic jump will form which confirms that both gravity and surface tension can cause hydraulic jump. Here we note that when surface tension is present along (i.e. without gravity), the flow will not be steady. The thickness after the hydraulic jump will continue to increase. Therefore, for this particular case, we only presented a transient solution after the jump is formed. It should be emphasized that the location of the jump may change after it forms under this condition (i.e. with only the effect of surface tension). However, this is not the focus of the current work.

A more important observation is the fact that when both gravity and surface tension are incorporated, the location the jump does not seem to vary much, compared with the case where only gravity is included. This indicates that, for high-viscosity liquid, there is no simple argument regarding whether gravity or surface tension is the dominant effect in triggering the jump. However, gravity is indeed dominant on the overall film thickness in both the super and sub-critical region as illustrated in Fig. 5. In other words, the surface tension does not have much effect on the hydraulic jump for a high-viscosity liquid as long as gravity is present. However, surface tension does dominate the flow when gravity is neglected as already discussed. The thickness after the jump in this case can be infinite just like a liquid drop in the outer space where gravity is absent.

Now we investigate the effect of gravity and surface tension on the hydraulic jump of a low-viscosity liquid, under the same flow configuration. Fig. 6 depicts the influence of gravity and surface tension on the circular hydraulic jump of water. Similarly, the hydraulic jump does not occur in the absence of both gravity and surface tension. However, as we can see, the hydraulic jump is still not identifiable even when gravity presents. The behavior is significantly different than that of high-viscosity liquid for which gravity can trigger the jump on its own. In our case, only a bending is observed after the liquid leaves the disk edge in the presence of gravity. In contrast, a significant jump appears once surface tension is included and adding gravity to the flow field does not affect the location of the jump much even though it does significantly depress the height after the jump.

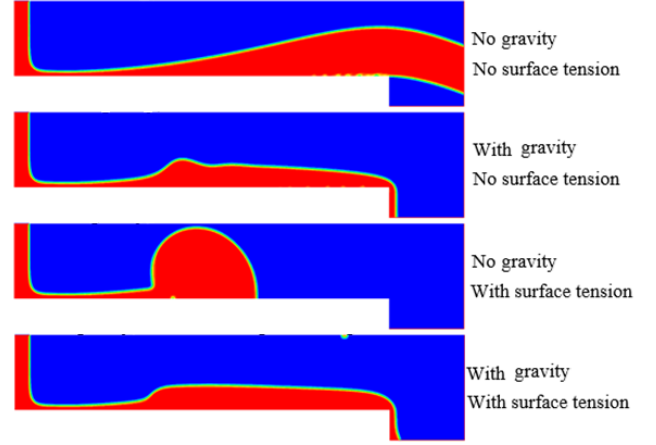


Figure 5. The effects of gravity and surface tension on the hydraulic jump of high-viscosity liquid (silicon oil).

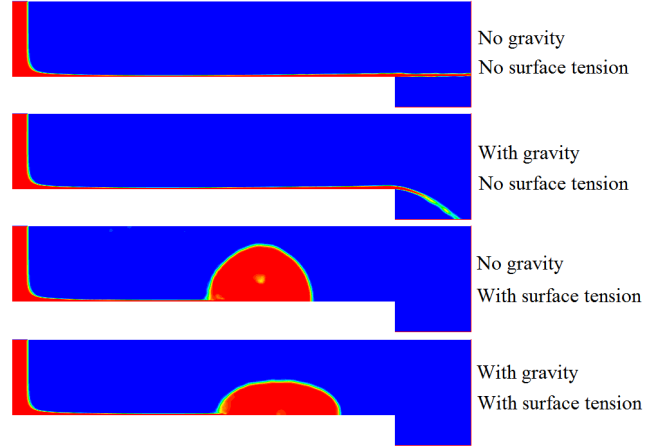


Figure 6. The effects of gravity and surface tension on the hydraulic jump of low-viscosity liquid (water).

V. CONCLUSION AND FUTURE WORKS

In the current study, the influences of gravity and surface tension on the formation of circular hydraulic jump are comprehensively explored using a numerical method. From the computational result, we find that the mechanism that triggers the jump is significantly different for low and high-viscosity liquid. The hydraulic jump for high-viscosity liquid is predominantly caused by gravity due to strong reverse hydrostatic pressure gradient caused by the accumulating viscous effect. On the other hand, the hydraulic jump for low-viscosity liquid is mainly caused by surface tension (i.e. reverse Laplace pressure gradient), which agrees with the recent findings [9].

Furthermore, the work of Maynes, Johnson and Webb [18] found that the contact angle has a tangible effect on the location of the jump for low-viscosity liquids whereas the recent work [9] suggested that the influence of contact angle is negligible. In this case, we will investigate the influence of the contact angle between the solid and liquid in the near future.

REFERENCES

- [1] B. Mohajer, and R. Li, "Circular hydraulic jump on finite surfaces with capillary limit," *Phys. Fluids*, vol. 27, pp. 117102, 2015.
- [2] E. Watson, "The spread of a liquid jet over a horizontal plane," *J. Fluid Mech*, vol. 20, pp. 481-499, 1964.
- [3] J. W. M. Bush and J. M. Aristoff, "The influence of surface tension on the circular hydraulic jump," *J. Fluid Mech*, vol. 489, pp. 229-238, 2003.
- [4] Y. Wang and R.E. Khayat, "Impinging jet flow and hydraulic jump on a rotating disk," *J. Fluid Mech*, vol. 839, pp. 525-560, 2018.
- [5] V. Lubarda and K. A. Talke, "Analysis of the equilibrium droplet shape based on an ellipsoidal droplet Model," *Langmuir*, vol. 27, pp. 10705–10713, 2011.
- [6] Y. Wang and R.E. Khayat, "The role of gravity in the prediction of the circular hydraulic jump radius for high-viscosity liquids," *J. Fluid Mech*, vol. 862, pp. 128-161, 2019.
- [7] I. Tani, "Water Jump in the Boundary Layer," *J. Phys. Soc. Japan*, vol. 4, pp. 212-215, 1949.
- [8] T. Bohr, P. Dimon, and V. Putkaradze, "Shallow-water approach to the circular hydraulic jump," *J. Fluid Mech*, vol. 254, pp. 635-648, 1993.
- [9] R. K. Bhagat, N. K. Jha, P. F. Linden, and D. I. Wilson, "On the origin of the circular hydraulic jump in a thin liquid film," *J. Fluid Mech*, vol. 851, R5, pp. 1-11, 2018.
- [10] H. Schlichting, *Boundary-layer theory*. New York: Springer, 2000.
- [11] C. Ellegaard, A. E. Hansen, A. Haaning, and T. Bohr, "Experimental results on flow separation and transitions in the circular hydraulic jump," *Phys. Scr*, vol. T67, pp. 105–110, 1996.
- [12] M. Passandideh-Fard, A. R. Teymourtash, and M. Khavari, "Numerical study of circular hydraulic jump using volume-of-fluid method," *J. Fluids Eng*, vol. 133, pp. 011401, 2011.
- [13] M. Passandideh-Fard and E. Roohi, "Transient simulations of cavitating flows using a modified volume-of-fluid (VOF) Technique," *Int. J. Comput. Fluid Dyn*, vol. 22(1&2), pp. 97–114, 2008.
- [14] Ansys Fluent. *Fluent 14.0 Theory Guide*, 2011.
- [15] C.W. Hirt and B. D. Nichols, "Volume of Fluid (VOF) Method for the Dynamics of Free Boundaries," *J. Comput. Phys*, vol. 39, pp. 201–225, 1981.
- [16] J. U. Brackbill, D. B. Kothe, and C. Zemach, "A Continuum Method for Modeling Surface Tension," *J. Comput. Phys*, vol. 100, pp. 335–354, 1992.
- [17] A. Duchesne, L. Lebon, and L. Limat, "Constant Froude number in a circular hydraulic jump and its implication on the jump radius selection," *Europhys. Lett*, vol. 107, pp. 54002, 2014.
- [18] D. Maynes, M. Johnson, and B. W. Webb "Free-surface liquid jet impingement on rib patterned superhydrophobic surfaces," *Phys. Fluids*, vol. 23, pp. 052104, 2011.

DESIGN OF PASSIVE WIND-TUNNEL GRIDS USING A PREDICTION MODEL FOR TURBULENCE INTENSITY

Dwaipayan Sarkar

Mechanical & Materials Engineering
The University of Western Ontario
London, Canada
dsarkar3@uwo.ca

Eric Savory

Mechanical & Materials Engineering
The University of Western Ontario
London, Canada
esavory@uwo.ca

Abstract— A design methodology for constructing square mesh passive grids is proposed in the current study. The methodology is based on a prediction model of turbulence intensity formulated using the geometrical parameters b and M of a grid, where b is the bar-width and M is the centre spacing between the grid elements. The derived model is validated using relevant experimental data and can be used as an efficient design tool to estimate the magnitudes of the grid parameters, b and M in order to achieve desired turbulence intensity downstream of a typical wind-tunnel section.

Keywords- design; passive grid; mesh width, bar-width, wind-tunnel.

I. INTRODUCTION AND BACKGROUND

Turbulence in a wind-tunnel section can be generated by various methods (e.g. spires, perforated plates, trip-wires, vortex-generators, turbulent jets and grids). Among these turbulence generated using grids in a wind-tunnel facility (refer to fig. 1 for grid-generated turbulence) remains one of the simplest and primary methods of reproducing and analyzing different classes of turbulence [1-3] for e.g. atmospheric turbulence [4], and freestream turbulence [5]. Turbulence generated downstream of a regular and simple grid is often nearly isotropic and reasonably homogeneous in nature and this property of turbulence is beneficial in characterizing the boundary layer dynamics around simple bluff-bodies [6,7] such as cylinders [8] or spheres [9]. In the absence of continuous turbulence production, the turbulent kinetic energy (referred to as TKE in the text hereafter) within the flow decays continuously, mainly due to inertial interaction at larger scales and viscous effects at smaller scales. Therefore, to analyze the effects of decaying turbulence on flow around bluff-bodies, it becomes necessary to quantify those turbulence levels based on grid geometrical parameters. The magnitude of turbulence scales downstream for a flow through a grid is not known *a priori* and, therefore, attempts have been made to address this issue in order to be able to predict turbulence levels based on the upstream grid geometrical parameters. The present work highlights the importance of designing grids in order to

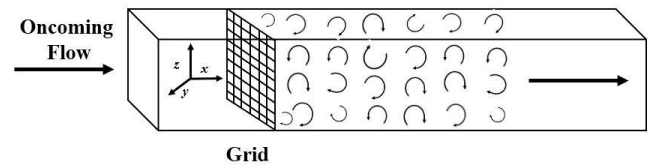


Figure 1 Schematic of the grid-generated turbulence

achieve desired turbulence levels downstream.

Grids can be of different types ranging from conventional regular or fractal, simple or multi-scale, passive or active, to grid elements having round or square cross-section. Experimental studies have been carried out in each of these different classes of grids to investigate the flow pattern of eddies downstream and compare it with the classical theories of turbulence decay by Batchelor [1], Taylor [10], Kolmogorov [11], Batchelor [12], and Saffman [13]. However, attention has only been given to regular, passive, conventional grids with rectangular cross-sectional bars. The reason for choosing such conventional grids over other complicated grid structures (fractal and multiscale) is driven by its constructional simplicity and its generation of near-isotropic and homogenous turbulence flow fields with less complexity [5]. Regular passive grids are also the most commonly used in bluff-body wind tunnel studies [4,14].

However, there are certain classifications of the passive regular grids that must be mentioned before any other further discussion is made. The classifications are mostly based on the geometrical configurations of a grid that generates different wake-structure patterns behind it. Different genres of passive turbulence generators are mentioned in Roach [15] and are not referred to again here. However, a general schematic of three different types of passive grids discussed in the current text is shown in figure 2 for better understanding by the reader.

The focus of this work is primarily given to the design of conventional “square-mesh grids” built with rectangular cross-sectional bars. There are several reasons for making this choice most of which are associated with the cross-section and layout of the grid-bars on the entire mesh unit.

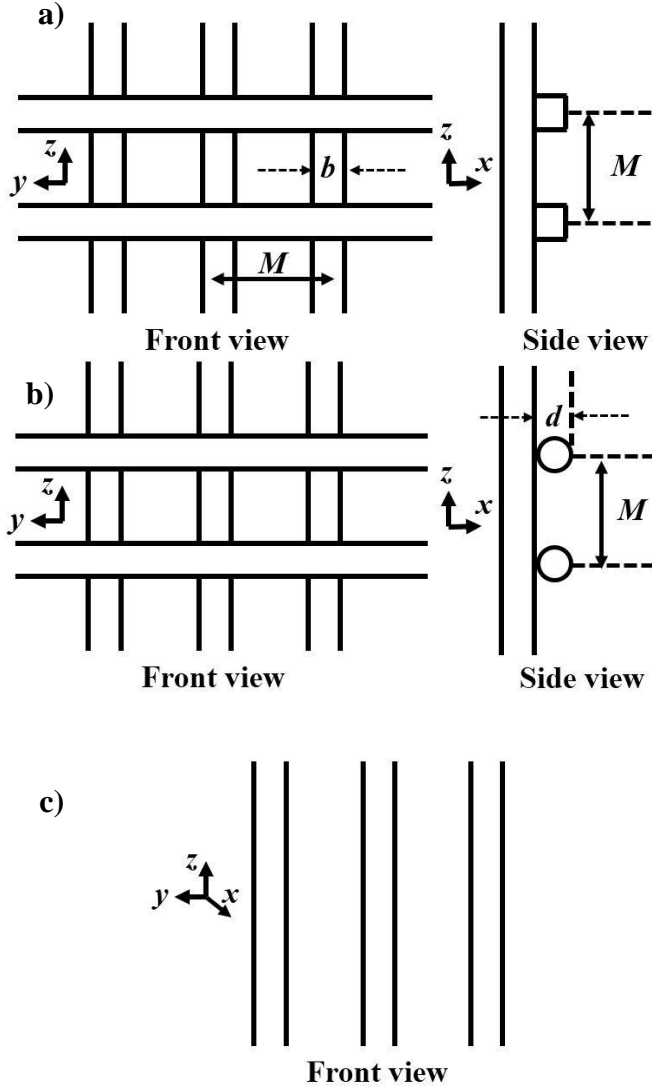


Figure 2 Schematic of different grid geometries a) Square mesh with rectangular cross-sectional bars b) Square mesh with cylindrical cross-sectional bars c) Parallel arrays of bars

Firstly, a single set of parallel bars (either cylindrical or rectangular cross-section) have been shown to exhibit non-uniformity in the mean velocity profiles for a longer period of time after the start of flow and the turbulence statistics are often contaminated by the presence of noise from the residual turbulence levels present in the test section [6]. Secondly, wake patterns developed from the separated flow around cylindrical rods have shown a strong dependency on the mesh-grid Reynolds number, $Re_d = Ud/\nu$, where U and ν are the mean flow velocity and kinematic viscosity of the fluid, respectively and d is the cylindrical rod-diameter of the grid. The dependency of the turbulent statistics with Re_d is probably due to the flow being separated at different angles around such structures relative to the mean flow direction [16]. The measured angle of separation also depends on the relative roughness of the cylindrical surfaces and, therefore, careful choices among the available data sets must be made before comparing and analyzing turbulence for cylindrical rods. Considerable

inhomogeneity was also observed in the measured normal and spanwise velocity fluctuations for such grid topology [17], although Uberoi and Wallis [17] have linked those variations to the inaccuracies in the construction of the circular grid bars.

Square meshes, on the other hand, built with square or rectangular cross-sectional bars have a fixed flow separation location at the trailing edge of the geometric corners and, therefore, have a weak Reynolds number dependency [15]. Nevertheless, the range of the grid Reynolds number Re_d mostly covered in the literature related to square-bar grid turbulence is between 10^2 and 10^4 and all the further analysis are made within this range only. Rectangular cross-sectional bars for generating homogeneous isotropic turbulence have been used in many previous studies (Nakamura et al. [7]; Vickery [18]; Bearman [19]; Kiya et al. [20]; Nakamura and Ohya [21]; Krogstad & Davidson [22]) with no reports of significant anomalies in the flow behind the grid. However, Hancock & Bradshaw [23] pointed out that bi-planar grids with a rectangular 2:1 cross-section (with larger dimension being normal to the mean flow direction) may impart substantial non-uniformity in the mean flow and, therefore, are not recommended for generating freestream turbulence. Similar observations were also made by Cherry [24] on three different kinds of cross-sectional bars (rectangular, square and mono-planar) and, based on those arguments, it may be concluded that bars of square-cross section are more suitable for generating uniform turbulence downstream.

Mono-planar grids (orthogonal bars in the same plane), on the other hand, are difficult to construct (Tanjil et al. [25] private communication) and it also imparts considerable non-uniformity and unsteadiness in the mean flow [23]. The reason is likely due to the larger size of the separated wakes behind each intersection which take longer to coalesce and, thereby, the non-uniformity in the flow decays slowly. Henceforth, mono-planar grids are also not recommended for generating high freestream turbulence for bluff-body studies [23]. However, based on physical reasoning it can be hypothesized that turbulence after getting fully developed should decay in a similar manner irrespective of the upstream geometrical anatomy of the grid [15].

A single set of parallel square-bars, either in a horizontal or vertical orientation could also act as a suitable grid for the generation of freestream turbulence but the spectra of the u-component velocity at low-wavenumbers has shown some disparity and, therefore, its usage is not recommended [23]. All of the above-mentioned arguments indicate that bi-planar conventional square-mesh grids are the most suitable configuration for generating well-developed homogeneous turbulence.

Although much grid-turbulence data exists in the literature, only a few authors discuss complete guidelines for designing such equipment. Even after 60 years of research, there is not a single document that gives adequate and simple design guidelines for constructing grids that can generate desired turbulence intensity (referred to as TI in the text hereafter where $TI = u'/U$ and u' is the characteristic velocity fluctuation of the flow) and integral length scales (L_{ii}) independently of one another, downstream of a grid. The pioneering works by Batchelor & Townsend [26] and Batchelor [1], used a

characteristics mesh scale M in their decay formulation to normalize the streamwise distance x , thereby establishing the relationship between the TI and the centre-spacing M of the grid-elements as

$$TI = A \left(\frac{x-x_0}{M} \right)^{-n_1} \quad (1)$$

where x_0 is the virtual origin, A is the decay-coefficient and n_1 is the decay-exponent. The magnitudes of these empirical constants depend on several factors including Reynolds number of the grid flow $Re_M = UM/\nu$, and geometrical variations of the grids such as surface roughness and solidity [3]. Mohamed and Larue [3] also showed that previous studies on turbulence decay (Comte-Bellot and Corrsin [5]; Uberoi and Wallis [17]; Batchelor and Townsend [26]; Wyatt [27]; Sirivat and Warhaft [28]) demonstrated non-uniformity in the estimated magnitudes of n_1 , most of which originates from the improper estimates of the virtual origin x_0 and inaccurate identification of the isotropic and homogeneous regime of the turbulence. Nevertheless, a decay exponent of $n_1 \sim (1.2-1.4)$ and a growth exponent $m_1 \sim (0.28-0.4)$ has generally been accepted as the classical predictions of the decay and the growth exponents based on the benchmarked studies of Batchelor [1] and Saffman [13]. A higher bound of the decay exponent with $n_1 \sim 2$ [29] and a lower bound with $n_1 \sim 1$ ([30,31]) has also been reported which suggest that even after years of investigation, a well-defined magnitude of decay exponent is not available.

Frenkiel [32], on the other hand, showed that the streamwise component of turbulence intensity downstream of a grid when scaled to the characteristic bar-width of a grid, b (for high Reynolds number flows) gives better agreement to all the available results, rather than the centre-spacing M . The relationship between given TI and b is

$$TI = B \left(\frac{x}{b} \right)^{-n_2} \quad (2)$$

where the constant B represents the decay-exponent in equation (2) and is a function of both the grid geometry and the Reynolds number of the flow and n_2 represents the decay exponent. Frenkiel argued that the grid-element width b is the appropriate dimension to scale the distance downstream of a grid since there is no need to estimate the pressure losses encountered in a grid flow in order to compute the downstream turbulence energy as opposed to the method in [1] where pressure losses needs to be estimated first from the grid solidity correlation [15] to compute the turbulence energy. The turbulence energy is proportional to the pressure drop across the grid [16]. However, this function also suffers from the inconsistent estimates of the decay exponents since Baines & Peterson [33]; Vickery [34] found n_2 to be around 0.71 while, Laneville [35] suggested $n_2 \sim 0.89$.

It is quite evident from the above-mentioned discussion that studies relating to grid turbulence can be categorized into two groups based on the use of the geometrical scaling parameter of the grid (refer to equation 1 & 2). Whilst the benchmark studies of Batchelor and Townsend [26], Comte-Bellot and Corrsin [5], Hancock [36], Mohamed and Larue [3] and Hurst and Vasillicos [29] have used the mesh-width M as a scaling parameter, studies by Frenkiel [32], Vickery [18], Roach [15], Mara [37] and Vita [4] have adopted the bar-width b as a scaling parameter. The reason for favoring one scaling parameter over the other is not

discussed in any of the previous studies which makes it difficult to draw conclusion at this point in the context of grid design. However, the objective of this present discussion is not to debate over the two theories of turbulence decay and its related exponents but, rather, to highlight the fact that there exists two distinct, yet similar power law expressions in the literature for the expressions of TI . A search in the previous literature did not reveal a single study where the same set of experimental data has been compared against these two existing decay functions. This research gap initiates further evaluation and analyses of the available experimental data sets. It is hypothesized that TI could be a function of both b and M and, in order to establish that, attempts have been made in the present study to reformulate a generic correlation of TI involving both b and M (i.e. $TI = f(b, M)$).

Therefore, the aim of this study is two-fold. Firstly, analyses have been carried out on the previous experimental data sets on turbulence decay to examine whether TI is a function of both b and M . The idea is to confirm the previously observed data trends reported in the literature and develop new correlations relating b and M to TI with global estimates of decay, growth coefficients and exponents. Secondly, the objective is to use those correlations equations as design tool to estimate the magnitudes of the two most important geometrical parameters: the width/diameter b of the grid bars and the mesh size M based on the turbulence intensity TI downstream at any position x . The novelty of this work lies in the fact that an update to the existing design correlations has been made with the inclusion of new relevant data sets since the work of Roach [15] keeping grid design criteria in mind. It is important to point out here that the studies with inconsistent reports of TI downstream with high uncertainties have been excluded in the present analysis. All the data sets included in the present analysis lie within the homogeneous and isotropic regime of a turbulent flow based on the evidence provided by the authors. Reference to all the studies has been made in the text (in the next section) for complete analysis of the data sets used in the present study.

II. METHODOLOGY AND ANALYSIS

In this section, the key metrics on the which the choice of the grid geometrical parameters (b and M) depend have been introduced and explained. In addition, an outline of the present approach for synthesizing the previous data has been provided.

The advantage of having a grid as a turbulence generator becomes evident from the fact that the turbulent scales downstream of a grid depend only on three primary variables i.e. the mesh size M , the bar-width b and the streamwise distance x along which the turbulence scales vary [18,24,35]. Laneville [35] noted that the choice of the mesh-size M and the bar-width b of a particular grid is governed by the available test section length and resultant drag of the grid. Vickery [18] suggested that $M = L/8$, where L is the available length of the wind-tunnel test section. Similarly, it was shown by Hinze [2] that the choice of the bar-size, b is regulated by the blockage ratio (b/M) of the grid which relates to the overall drag of the grid given by equation 3

$$C_D = \frac{b/M(2-b/M)}{(1-b/M)^4} \quad (3)$$

Laneville [35] found that for $b/M > 0.25$, C_D increases at a drastic rate (refer to fig. I.1 in Laneville [35]). This is not favored since higher C_D usually means the drag force is higher which, in turn, causes the mean velocity gradients to be higher in the immediate wake behind the grid. This results in higher initial turbulence production and, therefore, higher dissipation of turbulence downstream of a grid. However, the ratio of b/M suggested by Laneville (1973) is consistent with the earlier observations by Baines & Peterson [33] who suggested that $b/M < 0.29$, while Vickery [18] suggested that $b/M \approx 0.25$. Both these recommendations yield a drag-coefficient $C_D \leq 2$ (using equation 3) which is an accepted value found in the research on bluff-body aerodynamics [6, 14, 20, 24]. The grid-drag relates to the porosity (β) and solidity (σ) of the grid given by

$$\beta = (1 - b/M)^2 \quad (4)$$

$$\sigma = 1 - \beta \quad (5)$$

Roach [15] provided evidence that for $\sigma > 0.5$, significant flow instability may occur and therefore the use of grids with solidity greater than 50% is not recommended. However, it is seen that for $\sigma < 0.5$, $b/M < 0.29$ (using equation 4) which is consistent with the previous observations made by Baines & Peterson [33] and Vickery [18]. Summarizing the above observations, it is concluded that the choice of the bar-width b and the mesh-size M should be such that the ratio (b/M) should always be smaller than 0.29 (0.25 is ideal) satisfying the drag and the solidity conditions. There is one other aspect of grid turbulence which must be made before proceeding towards investigations of the previous literature. The flow past a grid has been seen to become nearly homogeneous and isotropic at a distance of $10M$ (effective virtual origin) from the physical location of the grid (Roach [15]), although this is highly debatable within the published literature. It has been observed that a downstream distance of $5M$ [38], $9M$, [39], $20M$ [3, 40-41], $40M$ [2] and $50M$ [42]) have been used as an effective virtual origin distance where the flow conditions display isotropy and effective homogeneity. Although the present study does not investigate further on predicting the correct limiting distance for such conditions, care has been taken to include the data sets within the isotropic and homogeneous turbulence regime irrespective of the different magnitudes of virtual origin reported by the previous authors. It is postulated that the different sets of data lying in the uniform turbulence regime should not exhibit much difference in the decay rate of TI .

Based on the above-mentioned conditions, attempts have been made to bring all the data sets related to square mesh grids together for further analysis. These data sets include those that are referred to in Roach [15] as well as those produced since then.

References in Roach [15] include studies by Nakamura and Ohya [21]; Sirivat and Warhaft [28]; Baines and Peterson [33];

Hancock [36]; Sato [43]; Sirivat and Warhaft [44] and Castro [45]. Additionally, data-sets from Vita et al [4]; Tanjil [25] unpublished data; Lee [38]; Manini et al. [39]; Torrano et al. [40] and Krogstad & Davidson [42]; have been included in the present analysis. The studies are represented in a tabular form (refer to Table 1) along with the magnitudes of the virtual origin (x_0), mesh-width M , bar-width b , solidity and TI at x_0 to give an idea of the range of the TI limits within which the proposed design correlations will hold. The references are represented by acronyms in the table and in the following figures due to space limitations.

Since the focus of the paper is mainly on the design methodology of the grids, no discussion regarding the evolution of the turbulent statistics (i.e. isotropy, Gaussianity, spectra) downstream are made since those are available in relevant literatures cited in this work.

The next section analyses the results on turbulence decay from the previous experiments with respect to the downstream distance x and the grid-flow parameters b and M in order to facilitate design of regular passive grids.

Table 1 Different key parameters of grid turbulence data from the previous experimental studies

References		x_0	M (m)	b (m)	σ	TI (%) at x_0
[4]	V (2018) (U = 5m/s)	10M	0.15	0.036	0.42	10.08
[4]	V (2018) (U = 10m/s)	10M	0.15	0.036	0.42	9.06
[4]	V (2018) (U = 15m/s)	10M	0.15	0.036	0.42	9.06
[4]	V (2018) (U = 20m/s)	10M	0.15	0.036	0.42	8.41
[21]	N & O (1983) (1 st grid)	10M	0.13	0.025	0.35	7.77
[21]	N & O (1983) (2 nd grid)	10M	0.6	0.15	0.44	10
[25]	T (2018)	10M	0.508	0.1016	0.36	8.17
[28]	S & W (1982)	24M	0.0125	0.0023	0.33	3.52
[33]	B & P (1951)	20M	0.0338	0.008	0.44	4.71
[36]	H (1980)	15M	0.152	0.0381	0.44	4.62
[38]	L (1975)	12M	0.076	0.013	0.31	4.50
[39]	M (2017)	9M	0.1	0.025	0.44	8.60
[40]	T (2015)	20M	0.05	0.01	0.36	4.42
[42]	K & D (2011)	50M	0.04	0.01	0.44	2
[43]	S (1951)	17M	0.025	0.005	0.36	3.60
[44]	S & W (1983) (U = 3.4m/s)	35M	0.025	0.0048	0.34	2.54
[44]	S & W (1983) (U = 6.3m/s)	39M	0.025	0.0048	0.34	2.42
[45]	C (1984) (1 st grid)	18M	0.05	0.0125	0.44	5.90
[45]	C (1984) (2 nd grid)	13M	0.1	0.025	0.44	7.13

III. RESULTS & DISCUSSIONS

In this section, new functional relationships between TI , b and M have been proposed based on the analysis of the previous experimental data sets. But before that, investigations have been carried out to see whether TI has a dependency on both b and M , when the other variable is kept constant. There are only a limited number of data sets available in the literature that report the downstream variation of TI for square-mesh grids with a fixed M and varying b and vice-versa (evident from the b and M values presented in table 1). Even if there are studies which matches such criteria, the virtual origin varies substantially which then limits the comparison for such data sets. A further critical search of the literature revealed that Baines & Peterson [33] and Sato [43] are the only two studies where grids of same b and different M and same M with different b have been used. Analysis of the data from those two studies are made in the forthcoming discussion.

Figure 3 shows the plot of TI (in percentage) versus (x/M) for different bar width b using the data sets from Sato [43]. The magnitude of turbulence intensity is seen to be dependent on the magnitude of b , where the grid with a larger b generates higher turbulence intensity over the course of the measurement section. This is expected since a higher b would usually give higher drag (refer to equation 5) that would eventually result in generating higher initial turbulence as observed for the first measurement point i.e. (x_0) . A similar kind of dependency of TI with M is also seen in figure 4 where TI is plotted against (x/b) for different M values, using the well-established results of Baines & Peterson [33]. A grid with higher M displays lower turbulence intensity at the first location point (refer to figure 4) in the isotropic homogeneous regime of turbulence

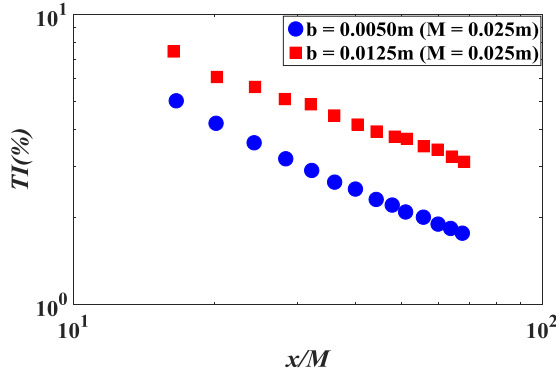


Figure 3 $TI(\%)$ vs x/M for varying b [43].

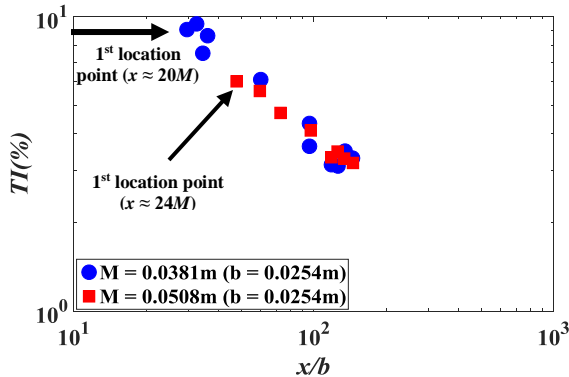


Figure 4 $TI(\%)$ vs x/b for varying M [33].

Both these figures establish the dependency of turbulence intensity on both b and M and, therefore, it is postulated that any function correlating the downstream distance x with TI should also account for the variability in both b and M . Which parameter among these has a stronger scaling influence on the magnitude of TI downstream is a matter of interest since both have been claimed by different researchers. Whilst the earlier studies on turbulence decay have used the relative distance (x/M) to be a scaling variable, others (Frenkiel (1948); Roach (1987)) have suggested that turbulence data is scaled better with the use of the variable (x/b) . It is difficult to comment on this aspect based on the previous experimental evidence since no single study comprising a large number of data sets (with a wide range of values) has tested these two scaling variables. Therefore, in the current study, all the relevant data sets from the previous experiments on turbulence decay have been plotted against both the non-dimensional parameters (x/b) and (x/M) in order to quantify its influence on the scaling. Figure 5 shows the plot of TI versus (x/M) for all the experiments listed in table 1, plotted in sequence with b increasing in magnitude. A regression analysis with a power law fit reveals $R^2 = 0.942$, $A = 0.55$ and $n_1 = 0.82$. Similar analyses have been performed over the same data sets, with TI plotted against (x/b) for different M , increasing in magnitude (fig. 6). The power law best-fit to the entire set of

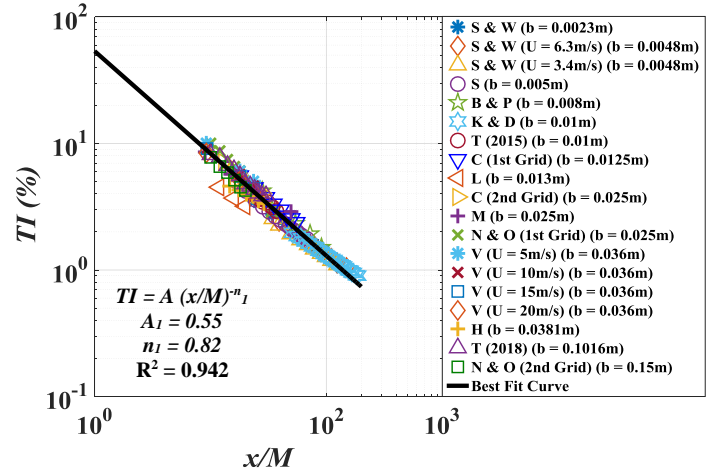


Figure 5 $TI(\%)$ vs x/M for all experimental data (refer to Table 1). The bold black line gives the best-fit curve.

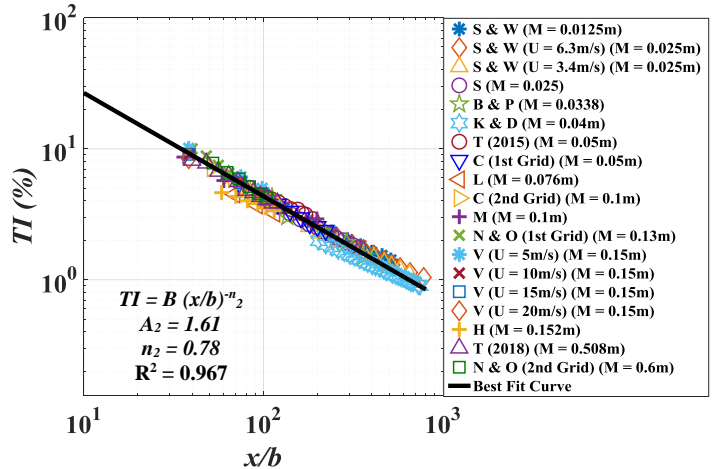


Figure 6 $TI(\%)$ vs x/b for all experimental data (refer to Table 1). The bold black line gives the best-fit curve.

data gives $R^2 = 0.967$ with the estimated coefficients of $B = 1.61$ and $n_2 = 0.78$. For both these plots the fitted coefficients lie within the 95% confidence bounds of the least-square fitting algorithm. The decay exponent lies within what is observed in [15] and [35]. Comparing the respective R^2 coefficients obtained from the best-fit straight line to the entire data set, plotted using the equations 1 and 2, it can be inferred that, x/b as a scaling variable behaves slightly better than x/M and collapses all the data points with an improved fit. These observations are at odds with [4] which claimed (x/M) to be a better scaling parameter than (x/b) . However, based on logical arguments it can be postulated that the wakes formed behind the flow past a grid would scale with the grid-width b rather than the grid-centre spacing M , as separation is caused at the upstream sharp-corners of the bar itself and, therefore, its characteristic dimension is of more relevance.

Now, since it has been established that TI is a function of both x/b and x/M , based on the dimensional analysis, and that TI decays as a power expression in x , attempts have been made to carry out multi-regression analysis on the entire set of data to come up with a generic correlation capable of estimating b and M from the known TI at a certain downstream distance x . But, first a regression analysis of the form of

$$TI = C \left(\frac{x^2}{bM} \right)^{-n_3} \quad (6)$$

has been attempted which assumes the same decay exponent n_3 for both x/b and x/M . This correlation imposes an equal influence of both b and M on TI and, therefore, plots of this function have been shown to evaluate the goodness of the fitted data in order to compare it with the fit using the multi-regression function discussed later in the section. The magnitude of C and n_3 estimated from the best-fit curve to the entire set of data are 0.98 and 0.41, respectively. The R^2 coefficient obtained was 0.962 with root-mean squared error estimate (RMSE) of 0.43. In order to show how well the predicted values, match up with the actual values, figure 7 is shown where TI (on the y-axis) is plotted against $(TI = (x^2/bM))$ (on the x-axis, using the best fit values) along with the percentage bounds that shows to within what accuracy the model function can predict an experimental value. Values which are in close proximity to the 45° line demonstrate that the predicted magnitudes are nearer to the original values. A second multi-regression analysis of the form

$$TI = D \left(\frac{x}{b} \right)^{-n_4} \left(\frac{x}{M} \right)^{-n_5} \quad (7)$$

was also carried out for the entire set of data where the correlation equation allows each variable to have a different exponent and, therefore, individual influences on the magnitude of TI downstream. Since R^2 cannot be computed for a non-linear multi-regression analysis due to statistical reasons, an estimate of RMSE is presented for this function which is 0.38. To see how good the model predicted values are in comparison to the experimental values, figure 8 is plotted comparing the two sets of values (actual and predicted) obtained by the model equation 7. The predicted values demonstrate very good agreement as seen from figure 8. The estimated magnitudes of the decay exponents, n_4 and n_5 are 0.72 and 0.07 with $D = 1.49$. Based on these estimates, a plot of TI (on the y-axis) versus $TI = D \left(\frac{x}{b} \right)^{-n_4} \left(\frac{x}{M} \right)^{-n_5}$ (on the x-axis, using the best-fit values) along

with the best-fit line and the percentage bounds are shown in figure 9. Comparing the RMSE values obtained for both the

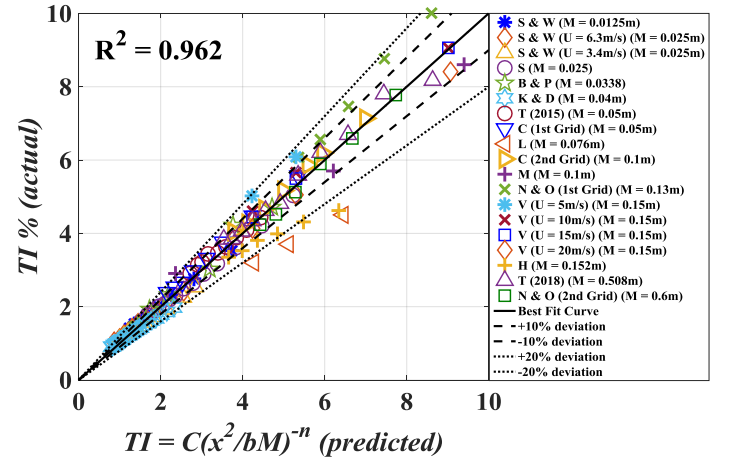


Figure 7 TI (actual) vs TI (predicted) for all the experimental data using equation 6. The solid black line represents the percentage bounds from the best-fit.

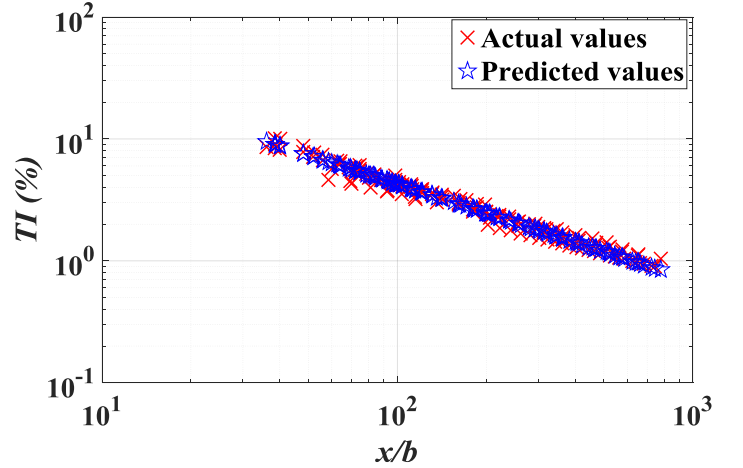


Figure 8a TI (%) vs x/b for the actual experimental values and the predicted values using the no-linear regression model in equation 7.

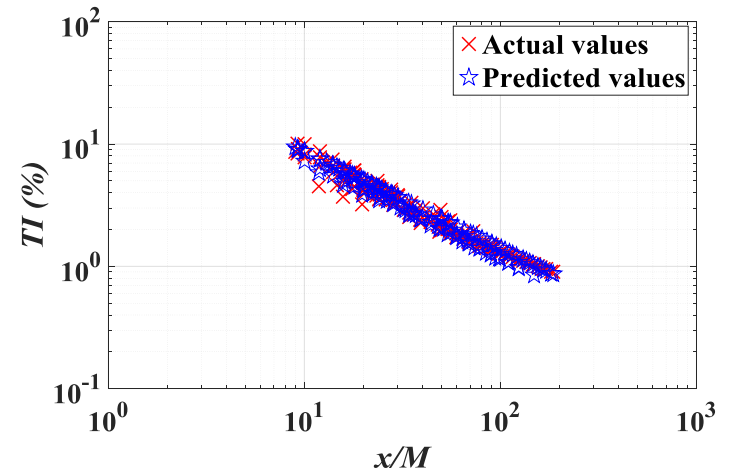


Figure 8b TI (%) vs x/M for the actual experimental values and the predicted values using the no-linear regression model in equation 7.

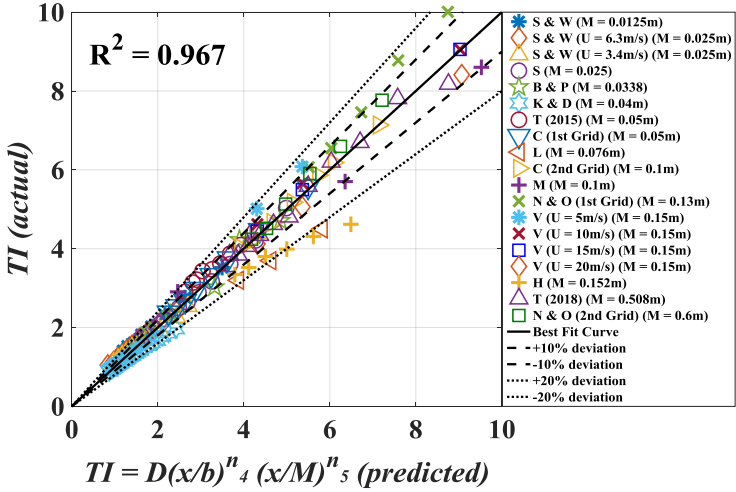


Figure 9 TI (actual) vs TI (predicted) for all the experimental data using equation 7. The solid black line represents the percentage bounds from the best-fit.

functions (eq. 8 and 9) reveals that, indeed, the model function with two exponents demonstrate a slightly better fit to all the different experimental data rather than the function with a single exponent. This is expected since each of these two variables effects the magnitude of TI differently, although (x/b) has a greater influence than (x/M) on TI (since $n_4 > n_5$). This is consistent with the previous observations made in figure 5 where the scaling variable (x/b) collapses the experimental data much better than (x/M) .

The following discussion is concluded by explaining the usage of the predictive functions to estimate the grid design parameters b and M . Taking the logarithm on both sides of the functions (eq. 8 and 9) gives a linear function involving b and M which is much simpler to work out mathematically. The functions are expressed as

$$\log(TI) = \log C + (-n_3 \log(x^2/bM)) \quad (8)$$

$$\log(TI) = \log D + (-n_4 \log(x/b)) + (-n_5 \log(x/M)) \quad (9)$$

the sum of the linear logarithmic functions. Provided that the decay coefficients and the exponents are already known, the desired TI at a desired downstream location ($x > 9M$) can then be substituted into these equations to estimate the magnitudes of b and M . Since b has a higher effect than M on TI , it is expected that the turbulence can be varied considerably by varying b at a fixed M . Whether the user wants to vary both b and M or just b to achieve desired TI downstream is a personal choice, although it is believed that both methods should estimate TI to within $\pm 10\%$ of the actual experimental value.

IV. CONCLUSIONS

A parametric study on the geometrical characteristics of a grid has been presented in terms of turbulence decay. The findings of this work deliver an important message to wind-tunnel experimentalists from the perspective of grid design. The primary findings from the present work are highlighted below:

- The magnitude of turbulence intensity downstream of a grid depends on both of the geometrical parameters b and M rather than each as an individual variable. The dependence on both of these variables is reflected through the drag coefficient of the grid bars.
- x/b is a better scaling variable than x/M for turbulence decay downstream, regardless of any set of experimental data.
- New prediction correlations that are developed, associating TI with b and M can help one to estimate the magnitudes of b and M in order to achieve desired turbulence levels at preferred downstream locations. The b and M values estimated can be re-iterated to confirm the limiting criteria for b/M . The magnitudes of b and M are helpful for constructing a passive regular square-mesh grid for carrying out specific wind-tunnel studies.

ACKNOWLEDGMENTS

The authors should like to thank Md. Mahbub Hossain and Greg Kopp for providing their grid turbulence data from the wind tunnel experiments performed at Boundary Layer Wind Tunnel Laboratory (BLWTL) at the University of Western Ontario. The support from Natural Sciences and Engineering Research Council is greatly acknowledged.

REFERENCES

- [1] Batchelor, G.K., 1953. The theory of homogeneous turbulence. Cambridge university press. pp. 1-197.
- [2] Hinze, J.O., 1975. Turbulence, 2nd ed., McGraw-Hill, USA, pp.1-790.
- [3] Mohamed, M.S. and Larue, J.C., 1990. The decay power law in grid-generated turbulence. Journal of Fluid Mechanics, 219, pp.195-214.
- [4] Vita, G., Hemida, H., Andrianne, T. and Baniotopoulos, C.C., 2018. Generating atmospheric turbulence using passive grids in an expansion test section of a wind tunnel. Journal of Wind Engineering and Industrial Aerodynamics, 178, pp.91-104.
- [5] Comte-Bellot, G. and Corrsin, S. (1966) "The use of a contraction to improve the isotropy of grid-generated turbulence," Journal of Fluid Mechanics. Cambridge University Press, 25(4), pp. 657-682
- [6] Bearman, P.W. and Morel, T., 1983. Effect of free stream turbulence on the flow around bluff bodies. Progress in Aerospace Sciences, 20(2-3), pp.97-123.
- [7] Nakamura, Y., Ohya, Y. and Ozono, S., 1988. The effects of turbulence on bluff-body mean flow. Journal of Wind Engineering and Industrial Aerodynamics, 28(1-3), pp.251-259.
- [8] Kestin, J. and Maeder, P.F., 1957. Influence of turbulence on transfer of heat from cylinders, NACA Tech. Note 4018. pp. 1-79.
- [9] Moradian, N., Ting, D.S.K. and Cheng, S., 2009. The effects of freestream turbulence on the drag coefficient of a sphere. Experimental Thermal and Fluid Science, 33(3), pp.460-471.
- [10] Taylor, G.I., 1935. Statistical theory of turbulence. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 151(873), pp.421-444.
- [11] Kolmogorov, A.N., 1941. Dissipation of energy in the locally isotropic turbulence. Proceedings: Mathematical and Physical Sciences, 434(1890), pp.15-17.
- [12] Batchelor, G.K., 1948. Energy decay and self-preserving correlation functions in isotropic turbulence. Quarterly of Applied Mathematics, 6(2), pp.97-116.

- [13] Saffman, P.G., 1967. The large-scale structure of homogeneous turbulence. *Journal of Fluid Mechanics*, 27(3), pp.581-593.
- [14] Vickery, B.J., 1965. On the flow behind a coarse grid and its use as a model of atmospheric turbulence in studies related to wind loads on buildings. NPL Aero Report 1143.
- [15] Roach, P.E., 1987. The generation of nearly isotropic turbulence by means of grids. *International Journal of Heat and Fluid Flow*, 8(2), pp.82-92.
- [16] Kistler, A.L. and Vrebalovich, T., 1966. Grid turbulence at large Reynolds numbers. *Journal of Fluid Mechanics*, 26(1), pp.37-47.
- [17] Uberoi, M.S. and Wallis, S., 1967. Effect of grid geometry on turbulence decay. *The Physics of Fluids*, 10(6), pp.1216-1224.
- [18] Vickery, B.J. 1968. Wind Effects on Structures. Ph.D. Thesis, University of Sydney, Sydney, Australia.
- [19] Bearman, P.W., 1972. Some measurements of the distortion of turbulence approaching a two-dimensional bluff body. *Journal of Fluid Mechanics*, 53(3), pp.451-467.
- [20] Kiya, M., Suzuki, Y., Arie, M. and Hagino, M., 1982. A contribution to the free-stream turbulence effect on the flow past a circular cylinder. *Journal of Fluid Mechanics*, 115, pp.151-164.
- [21] Nakamura, Y. and Ohya, Y., 1983. The effects of turbulence on the mean flow past square rods. *Journal of Fluid Mechanics*, 137, pp.331-345.
- [22] Krogstad, P.Å. and Davidson, P.A., 2010. Is grid turbulence Saffman turbulence?. *Journal of Fluid Mechanics*, 642, pp.373-394.
- [23] Hancock, P.E. and Bradshaw, P., 1983. The effect of free-stream turbulence on turbulent boundary layers. *ASME Journal of Fluids Engineering*, 105(3), pp.284-289.
- [24] Cherry, N.J., 1982. The effects of stream turbulence on a separated flow with reattachment. Ph.D. Thesis, Imperial College, London
- [25] Hossain, M.M., Corbin, M., Kopp, G., 2018 Generating atmospheric turbulence by a passive method in a boundary layer wind-tunnel. Unpublished Data. BLWTL, Western University.
- [26] Batchelor, G.K. and Townsend, A.A., 1948. Decay of isotropic turbulence in the initial period. *Proc. R. Soc. Lond. A*, 193(1035), pp.539-558.
- [27] Wyatt, L.A., 1955. Energy and spectra in decaying homogeneous turbulence (Doctoral dissertation, The University of Manchester).
- [28] Sirivat, A. and Warhaft, Z., 1982. The mixing of passive helium and temperature fluctuations in grid turbulence. *Journal of Fluid Mechanics*, 120, pp.475-504.
- [29] Hurst, D. and Vassilicos, J.C., 2007. Scalings and decay of fractal-generated turbulence. *Physics of Fluids*, 19(3), pp. 1-32.
- [30] Steward, R.W. and Townsend, A.A., 1951. Similarity and self-preservation in isotropic turbulence. *Philos. Trans. R. Soc. A*, 243(827), pp.359-386.
- [31] Portfors, E.A. and Keffer, J.F., 1969. Isotropy in initial period grid turbulence. *The Physics of Fluids*, 12(7), pp.1519-1521.
- [32] Frenkiel, F.N., 1948. The decay of isotropic turbulence. *Journal of Applied Mechanics-Transactions of the ASME*, 15(4), pp.311-321.
- [33] Baines, W.D. and Peterson, E.G., 1951. An investigation of flow through screens. *Transactions of the ASME*, 73, pp. 467-480.
- [34] Vickery, B.J., 1966. Fluctuating lift and drag on a long cylinder of square cross-section in a smooth and in a turbulent stream. *Journal of Fluid Mechanics*, 25(3), pp.481-494.
- [35] Laneville, A., 1973. Effects of turbulence on wind induced vibrations of bluff cylinders (Doctoral dissertation, University of British Columbia).
- [36] Hancock, P.E. 1980. The effect of freestream turbulence on turbulent boundary layers. Ph.D. Thesis, Imperial College, London, UK.
- [37] Mara, T.G., Galsworthy, J.K. and Savory, E., 2010. Assessment of vertical wind loads on lattice framework with application to thunderstorm winds. *Wind and Structures*, 13(5), pp.413-431.
- [38] Lee, B.E., 1975. Some effects of turbulence scale on the mean forces on a bluff body. *Journal of Wind Engineering and Industrial Aerodynamics*, 1, pp.361-370.
- [39] Mannini, C., Marra, A.M., Pigolotti, L. and Bartoli, G., 2017. The effects of free-stream turbulence and angle of attack on the aerodynamics of a cylinder with rectangular 5: 1 cross section. *Journal of Wind Engineering and Industrial Aerodynamics*, 161, pp.42-58.
- [40] Kang, H.S., Chester, S. and Meneveau, C., 2003. Decaying turbulence in an active-grid-generated flow and comparisons with large-eddy simulation. *Journal of Fluid Mechanics*, 480, pp.129-160.
- [41] Torrano, I., Tutar, M., Martinez-Agirre, M., Rouquier, A., Mordant, N. and Bourgoïn, M., 2015. Comparison of experimental and RANS-based numerical studies of the decay of grid-generated turbulence. *ASME Journal of Fluids Engineering*, 137(6), pp. 1-12.
- [42] Krogstad, P.Å. and Davidson, P.A., 2011. Freely decaying, homogeneous turbulence generated by multi-scale grids. *Journal of Fluid Mechanics*, 680, pp.417-434.
- [43] Sato, H., 1951. Experimental study of the spectrum of isotropic turbulence, I. *Journal of the Physical Society of Japan*, 6(5), pp.387-392.
- [44] Sirivat, A. and Warhaft, Z., 1983. The effect of a passive cross-stream temperature gradient on the evolution of temperature variance and heat flux in grid turbulence. *Journal of Fluid Mechanics*, 128, pp.323-346.
- [45] Castro, I.P., 1984. Effects of free stream turbulence on low Reynolds number boundary layers. *ASME Journal of fluids engineering*, 106(3), pp.298-306.

A NEW MULTIVARIATE CORRELATION FOR LIQUID / LIQUID SEPARATION

Weiwei E

Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, NL, Canada

Kevin Pope, Xili Duan

Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, NL, Canada

Abstract—The oil-water multiphase separator is a vital equipment in the oil industry. While various designs exist, a horizontal separator based on gravity is dominates the design. This research aims to develop a correlation for separator geometry design. A parametric study of numerical predictions is employed to investigate several critical variables that affect the separator design by using ANSYS FLUENT. The correlation results shows that oil droplet diameter, the density difference between water and oil, and the operation pressure has significant effect on separator design; however, the viscosity ratio of oil and water has negligible effect on separator geometry design.

Keywords—separator design; FLUENT; correlation; multiphase

I. INTRODUCTION

Oil-water two-phase separators are important in the oil and gas extraction industry. The separation efficiency directly affects the performance of downstream equipment. Therefore, various studies focus on this critical area by using both experimental and numerical methods. Most of the studies show that horizontal separators are more economical and provide higher separation efficiency.

Separation of a multiphase fluid system by using gravity is a physical process that is common in the chemical and petroleum industries [1-4]. Oil-water separation has been studied both experimentally and analytically by numerous investigators in the past few decades. Gravity-based separation relies on the difference in specific gravities of oil and water. The simplest form of a multiphase separator is a vertical or horizontal tank which provides a relative residence time for the light phase to coalesce and separate from the heavier phase.

A typical structure of a horizontal gravity separator is illustrated in Fig. 1. In this type of separator, the multiphase fluid stays in the settling section to be segmented. In the segment section, oil droplets coalesce. With sufficient residence time, oil and water will separate into two distinct phases. In the separation process, an oil droplets' rising velocity is perpendicular to the income flow

One of the most important criteria for the design of a horizontal separator is the volume of the liquid collection segment. To obtain a high separation efficiency, the segmented zone must provide enough residence time for oil droplets to rise

from the bottom of the separator to the top layer of the water phase. The separation efficiency of a gravity-based separator is related to the density difference of the phases. Keller [5] developed a method to study multiphase separation with oil density between 11 to 70 API. In this study, multiphase flow was injected into the separator, passing through a filter media system to collect oil droplets. The filter media has difference mesh sizes which are a function of gravity. The design of this filter formed oil droplets size distribution pattern increases uniformly and progressively from a small median size to large median size.

The structure of a gravity separator directly affects the performance; therefore, many experimental studies focus on the separator design [6, 7]. Rowley and Davies [8] proposed a sedimentation-oriented model to better represent oil droplets between parallel plates which enabled the smallest oil droplets to coalesce faster in a gravity separator. Lars Schlieper et al. [9] investigated the separation behavior of a horizontal gravity separator with three different inner components. Their study mainly focused on the effect of inflow to the plate, plate material, and distance between plates' influence on separation length of the separator. Based on the results, they showed that a separator with inner components, such as plates, can be an appropriate aid for separator design. In 2009, Fitnawan et al. [10] investigated inclined gravity downhole oil-water separation. Based on their study, the separator type has several advantages, including long-life potential due to its simplicity of components and a robust structure. In the paper, the performance effect of separator setting depth, setting inclination, tubing size, and tubing configuration were studied. Both experimental and simulation methods were used in the study. The simulation results of the study show that for a water volume fraction of approximately 81% to 87%, the inclined separator was able to increase separation efficiency up to 82%.

Although these studies have some scientific achievements, they still have some shortcomings. For instance, extensive studies have been focused on the separator internals structure design's effect on separation efficiency. However, as the results show that the internals structures play a minor role in improving the separation efficiency. The overall separator geometry design has not been fully analyzed. Therefore, the primary objective of this paper is to develop new correlations of gravity-based oil-water separator based on new numerical predictions and the Buckingham Pi method.

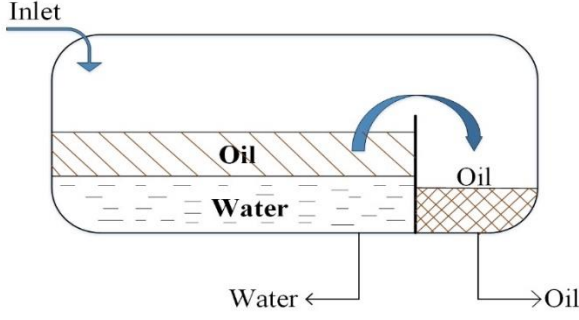


Figure 1. Schematic of the gravity horizontal separator

II. METHODOLOGY

A. CFD simulation method

A 2-D two-phase separator model is drawn in this study. Mesh dependence study and time dependence study has been performed to obtain relative high calculation accuracy. In order to validate simulation results, an experimental study is conducted in the same operating conditions as the base case in table I. All the simulation cases are list in table I.

B. Dimensional ananlysis method

According API separator design criteria, the following several parameters control the size of separator, the separator horizontal length (x) depends on the separator height (h), oil droplet diameter (d_{oil}), oil density (ρ_{oil}), water density (ρ_{water}), oil viscosity (μ_{oil}), water viscosity (μ_{water}), and operating pressure (P). Assume x is function of other parameters,

$$x = f(h, d_{oil}, \rho_{oil}, \rho_{water}, \mu_{oil}, \mu_{water}, P) \quad (1)$$

Each of the identified parameters is presented by a set of fundamental dimensions of mass (M), length (L), and time (T) in Table II. Based on the Buckingham Π Theorem, 5 dimensionless Π terms are determined, as presented in Table III.

The resulting relationship based on this analysis is as follows:

$$\Pi_1 = \phi(\Pi_2, \Pi_3, \Pi_4, \Pi_5) \quad (2)$$

or

$$\frac{x}{h} = C \cdot \left(\frac{d_{oil}}{h} \right)^{n_1} \left(\frac{\rho_{water}}{\rho_{oil}} \right)^{n_2} \left(\frac{\mu_{water}}{\mu_{oil}} \right)^{n_3} \left(\frac{P \cdot d_{oil}^2 \cdot \rho_{water}}{\mu_{water}^2} \right)^{n_4} \quad (3)$$

The values of C , n_1 , n_2 , n_3 , and n_4 will be evaluated by several numerical predictions of gravity-based oil-water separation.

III. RESULTS AND DISCUSSION

In this section, a new correlation is developed to predict the efficiency of oil/water separation. The data is from new numerical predictions using ANSYS FLUENT simulation results from table I. The correlation is presented and analyzed using

nonlinear regression on the x-y Cartesian coordinate system. To determine the coefficients in the general correlation relationship, manual iterations are performed. The simulation data was analyzed and plotted to calculate the values in equation 3.

TABLE I. THE SIMULATION STUDY CASES

Case	Inlet velocity (m/s)	Oil volume fraction	Temperature (°C)	Pressure (kPa)
Base case	0.0137	0.2	20	101.325
# 1	0.0274	0.2	20	101.325
# 2	0.0374	0.2	20	101.325
# 3	0.0137	0.5	20	101.325
# 4	0.0137	0.8	20	101.325
# 5	0.0137	0.2	32	101.325
# 6	0.0137	0.2	40	101.325
# 7	0.0137	0.2	50	101.325
# 8	0.0137	0.2	20	344.738
# 9	0.0137	0.2	20	698.476
# 10	0.0137	0.2	20	1,034.214

TABLE II. DESCRIPTION OF THE VARIOUS PARAMETERS IN FUNDAMENTAL DIMENSIONS

Variable	Description	SI unites	Dimensions
x	Separator horizontal length	m	L
h	Separator vertical length	m	L
d_{oil}	Oil droplets diameter	m	L
ρ_{oil}	Density of oil	kg/m^3	ML^{-3}
ρ_{water}	Density of water	kg/m^3	ML^{-3}
μ_{oil}	Dynamic viscosity of oil	$Pa \cdot s$	$ML^{-1} T^{-1}$
μ_{water}	Dynamic viscosity of water	$Pa \cdot s$	$ML^{-1} T^{-1}$
P	Operating pressure	Pa	$ML^{-1} T^{-2}$

TABLE III. DIMENSIONLESS Π TERMS

Π Terms	Dimensionless group
Π_1	$\frac{x}{h}$
Π_2	$\frac{d_{oil}}{h}$
Π_3	$\frac{\rho_{water}}{\rho_{oil}}$
Π_4	$\frac{\mu_{water}}{\mu_{oil}}$
Π_5	$\frac{P \cdot d_{oil}^2 \cdot \rho_{water}}{\mu_{water}^2}$

The first step of the manual iterations is performed for the first two Π terms.

$$\frac{x}{h} \text{ vs. } \frac{d_{oil}}{h} \quad (4)$$

Fig. 2 shows the resulting plot for the first two Π terms. The exponent for Π_2 term in the correlation is -0.192. The statistical R-squared value was 0.9188, which is a good indication that the data correlation well. This result shows that the ratio between oil droplet diameter and the height of the separator has a negative effect on the separator geometry design. That means with large oil droplets size, it requires small separator to separate oil from the water phase.

The second step is to determine the exponent for the ratio of densities of the two fluid at different operating temperatures. The effect of d_{oil}/h from the first step is considered for in this step.

$$\frac{\frac{x}{h}}{\left(\frac{d_{oil}}{h}\right)^{-0.192}} \text{ vs. } \frac{\rho_{water}}{\rho_{oil}} \quad (5)$$

Fig. 3 shows the resulting plot for the second step and provides the exponent and R-squared values. The result shows that Π_3 term has a significant effect on separator design. Increase the density difference between water and oil will significantly decrease the requirement of separator size. The driving force of oil/water separation is due to the density difference. Increasing the density difference will increase the driving force for two phases to separate.

Next step is to determine the correlation between separator geometry and the ratio of the viscosities of oil and water. As illustrated in Fig. 4, the exponent and R-squared values are small and close to zero, which means the fluid viscosities has a negligible effect on separator geometry design. Therefore, the impact of the ratio of viscosities is not considered in further iterations.

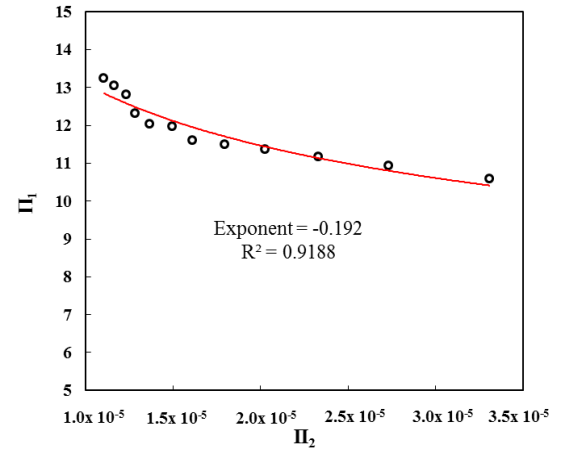


Figure 2. Plot for $\frac{x}{h}$ versus $\frac{d_{oil}}{h}$

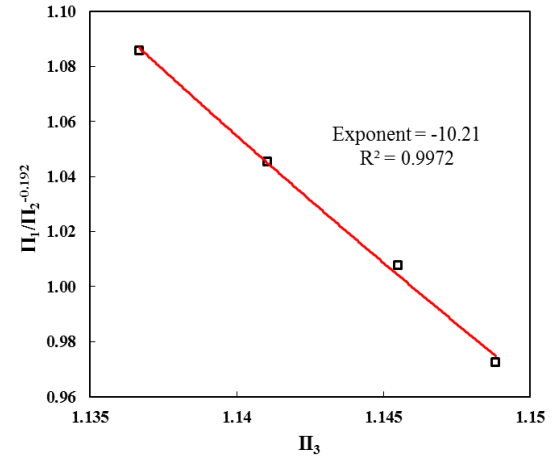


Figure 3. Determination of the exponent of the ratio of densities

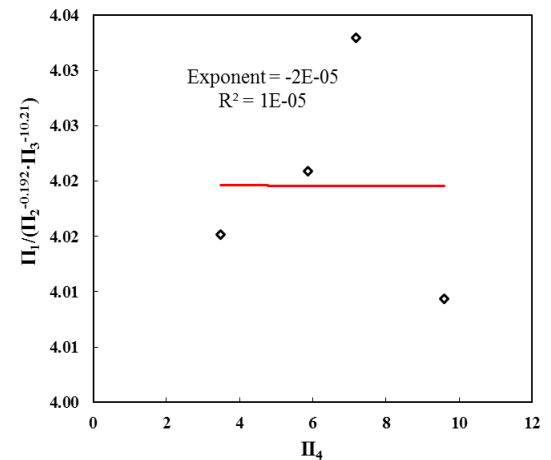


Figure 4. Determination of the exponent of the ratio of viscosities

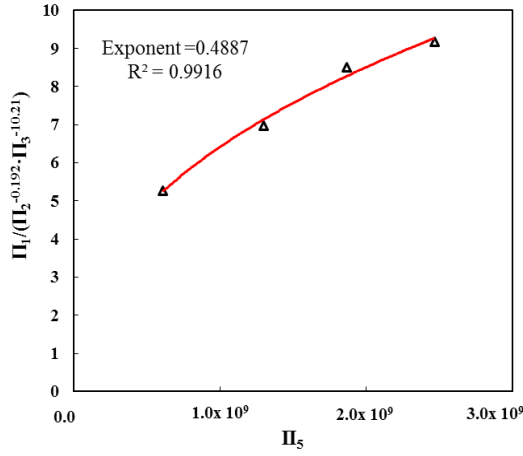


Figure 5. Determination of the exponent of the ratio of Π_5

$$\frac{\frac{x}{h}}{\left(\frac{d_{oil}}{h}\right)^{-0.192} \left(\frac{\rho_{water}}{\rho_{oil}}\right)^{-10.21}} \text{ vs. } \frac{\mu_{water}}{\mu_{oil}} \quad (6)$$

As with previous correlations, the exponent for the last dimensionless term Π_5 is studied. The operational pressure has a positive effect on separator geometry, increasing the operating pressure requires larger separator size. The result is presented in Fig. 5. Thus, the following correlation attained

$$\frac{x}{h} = C \cdot \left(\frac{d_{oil}}{h}\right)^{-0.192} \left(\frac{\rho_{water}}{\rho_{oil}}\right)^{-10.21} \left(\frac{P \cdot d_{oil}^2 \cdot \rho_{water}}{\mu_{water}^2}\right)^{0.4887} \quad (7)$$

A. Refine manual correlation

After the appropriate exponents have been established, further steps are conducted to increase the values of R-square of each exponent, thus increasing the accuracy of the correlation. The iteration equations, equation (8-10), new coefficients, and plots are shown in Fig.6 - 8.

Second iteration step 1:

$$\frac{\Pi_1}{\Pi_3^{-10.21} \cdot \Pi_5^{0.4887}} \text{ vs. } \Pi_2 \quad (8)$$

Second iteration step 2:

$$\frac{\Pi_1}{\Pi_2^{-0.180} \cdot \Pi_5^{0.4887}} \text{ vs. } \Pi_3 \quad (9)$$

Second iteration step 3:

$$\frac{\Pi_1}{\Pi_2^{-0.180} \cdot \Pi_3^{-17.87}} \text{ vs. } \Pi_5 \quad (10)$$

From the results of numerical simulation data analysis, the following final correlation for the geometry of an oil-water separator is determined,

$$\frac{x}{h} = 0.0255 \cdot \left(\frac{d_{oil}}{h}\right)^{-0.195} \left(\frac{\rho_{water}}{\rho_{oil}}\right)^{-17.23} \left(\frac{P \cdot d_{oil}^2 \cdot \rho_{water}}{\mu_{water}^2}\right)^{0.5388} \quad (11)$$

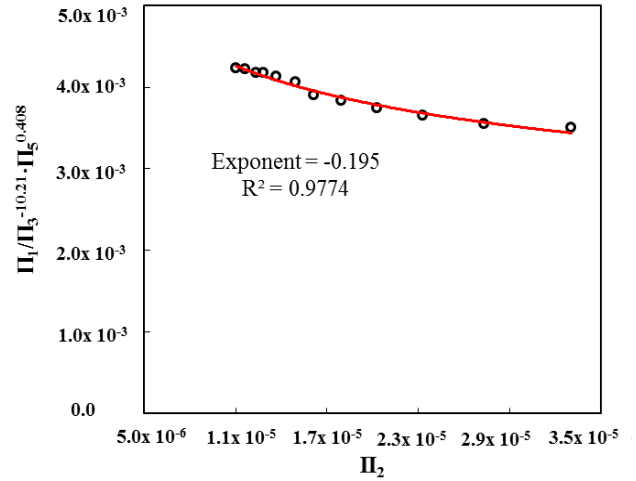


Figure 6. Determination of improved Π_2 exponent

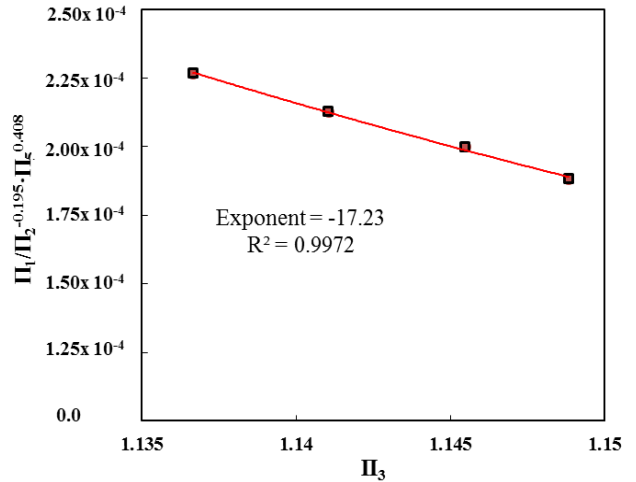


Figure 7. Determination of improved Π_3 exponent

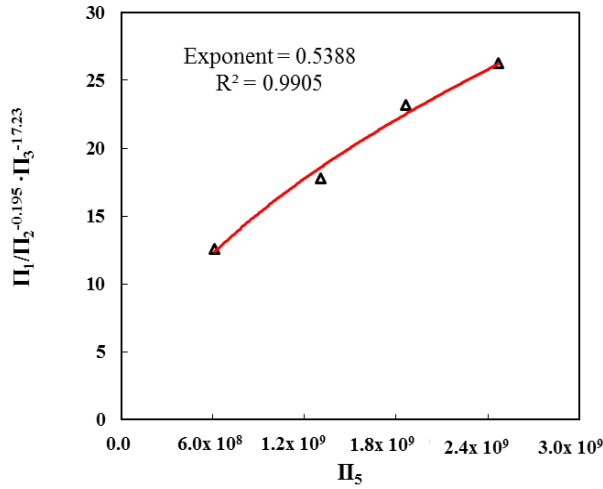


Figure 8. Determination of improved Π_5 exponent

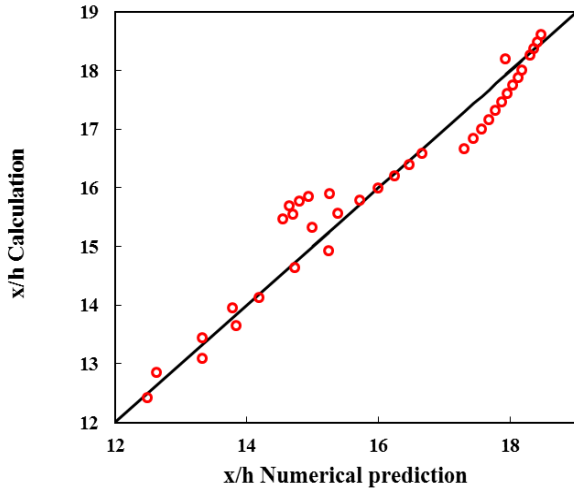


Figure 9. Calculated x/h vs. numerical prediction of x/h

B. Correlation validation

To verify the correlation, a comparison of numerically predicted data with calculated data from the new correlation is shown in Fig. 8. The figure is for a pressure of 101.325 kPa, water and oil properties at 20 °C, and varying droplet diameters from 20–100 μm . Fig. 9 provides scatter plots of the numerically obtained x/h versus the calculated x/h for the correlation. As can be seen from the figure, the new correlation shows good agreement with the simulated data.

IV. CONCLUSION

A new correlation has been developed based on main separator geometry design parameters, including separator length, separator height, oil droplet diameter, oil and water

density, operating pressure and oil / water viscosity ratio to analyze oil-water separator design. The correlation between five dimensionless groups has been developed by using the Buckingham Π Theorem. Results show that the viscosity ratio has a negligible effect on separator geometry design and increasing oil droplet diameter requires smaller separator geometry. Also, an increased density difference between oil and water can significantly reduce separator size, however, increased operating pressure increases the needed separator size. Under high operating pressure, a relatively large separator is needed to obtain a high separation efficiency.

ACKNOWLEDGMENT

The authors thank the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the NL Innovation Council.

REFERENCES

- [1] D.J. McClements, "Food emulsions: principles, practices, and techniques," CRC press, 2015. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [2] J. Sjöblom, N. Aske, I.H. Auflem, Ø. Brandal, T.E. Havre, Ø. Sæther, A. Westvik, E.E. Johnsen and, H. Kallevik, "Our current understanding of water-in-crude oil emulsions.: Recent characterization techniques and high pressure performance," Adv. Colloid Interface Sci., vol. 100, pp. 399–473, 2003.
- [3] J.O. Hinze, "Fundamentals of the hydrodynamic mechanism of splitting in dispersion processes," AIChE J., vol. 1(3), pp. 289–295, 1955.
- [4] L.R. Castilho, and R.A. Medronho, "A simple procedure for design and performance prediction of bradley and rietema hydrocyclones," Miner. Eng., vol. 3(2), pp.183–191, 2000.
- [5] J. H. F. Keller, U.S. Patent No. 3,867,285. Washington, DC: U.S. Patent and Trademark Office, 1975.
- [6] T. Husveg, O. Rambeau, T. Drengstig, and T. Bilstad, "Performance of a deoiling hydrocyclone during variable flow rates," Miner. Eng., vol. 20(4), pp.368–379, 2007.
- [7] Z.S. Bai, H.L. Wang, and S.T. Tu, "Oil–water separation using hydrocyclones enhanced by air bubbles," Chem. Eng. Res. Des., vol. 89(1), pp. 55–59, 2011.
- [8] M.E. Rowley, and G. Davies, "Design of plate separators for the separation of oil-water dispersions," Chem. Eng. Res. Des., vol. 66(4), pp.313–322, 1988.
- [9] L. Schlieper, M. Chatterjee, M. Henschke, and A. Pfennig, "Liquid–liquid phase separation in gravity settler with inclined plates," AIChE J., vol. 50(4), pp. 802–811, 2004.
- [10] E.A.Y. Fitnawan, R.M. Rivera, and M. Golan, "Inclined gravity downhole oil-water separator: using laboratory experimental results for predicting the impact of its application in high rate production wells," SPE, January 2009.
- [11] P.J. Pritchard, J.W. Mitchell, and J.C. Leylegian, "Fox and McDonald's Introduction to Fluid Mechanics," Binder Ready Version. John Wiley & Sons, 2016.

MEASUREMENT OF ICE ADHESION ON STAINLESS STEEL HYDROPHOBIC SURFACES

Sadman Sarar Rhythm, Kewei Shi, Xili Duan
Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, Newfoundland, A1B 3X5 Canada

Abstract—This paper presents the results of ice adhesion measurements on 17-4 pH stainless steel with different surface properties. The samples have a dimension of 30mm x 30mm x 1mm and are treated with sandblasting and electrochemical deposition of Zinc and stearic acid, which increases the hydrophobicity of the surface. An experimental setup is developed to measure ice adhesion forces on these surfaces. The objective is to investigate if there is a relationship between the surface wettability and the ice adhesion force. The measured results demonstrate a strong correlation of ice adhesion stress, advancing angle and contact angle hysteresis of water on the surfaces. Improved hydrophobicity seems to help reduce ice adhesion stress on these surfaces.

Keywords - *hydrophobic surface, stainless steel, wettability, ice adhesion.*

I. INTRODUCTION

Water droplets from the air can deposit on a surface and freeze in cold climate, forming ice. This icing can cause problems to power lines, runways, and turbines [1]. Icing can also affect operational performance on aircraft by reducing lift and thrust and increase air resistance and weight. Airframe and components of engines can be damaged by the removal of ice from the aircraft surfaces [2]. Many techniques were described in the literature to reduce ice adhesion on a surface. De-icing protocols like spraying aircraft with glycol-based fluids can be used to narrow down the complications by easily removing the accumulated ice from the surface. Another protocol can be sacrificial coatings such as silicone grease that removes ice as they are shed from treated surfaces. Both applications can negatively impact the environment and require constant reapplication which can be expensive [1].

A passive way to mitigate the icing problem is to create icephobic surfaces – those that can prevent ice formation. This ability depends on the freezing time of supercooled water at the liquid-solid interface. Small droplets below the freezing point temperatures are repelled by hydrophobic surface - i.e. surfaces that are not easily wet by water [3]. For example, superhydrophobic surfaces are extremely difficult

to wet - they usually have a contact angle greater than 150°, and low contact angle hysteresis [4]. These properties make water droplets roll off from the surface easily to avoid icing. Research has also shown that hydrophobic surfaces can delay icing [5] when it is not avoidable.

Icephobicity can also mean low adhesion force between ice and a solid surface. Critical shear stress can be calculated to understand if the surface is icephobic. Icephobic surfaces can have shear strength between 150 kPa and 500 kPa and even as low as 15.6 kPa [4]. Receding contact angle on an icephobic surface, as well as the initial size of interfacial cracks, are responsible for the force needed to detach a piece of ice [6]. One interesting aspect of the research is to find out the relationship between ice adhesion strength and hydrophobicity of a surface. If the wettability of a surface is viewed from a thermodynamic viewpoint, the equilibrium work of adhesion on a solid surface can be related to the equilibrium contact angle and the surface tension of the liquid [1,7]. In previous studies, different methods have been used to measure ice adhesion, for example through a pusher test, centrifuge test, or a Blister test [8].

Metallic (stainless steel) superhydrophobic surfaces are being studied in our research group for corrosion and ice mitigation on offshore structures in harsh marine environments. A combination of surface roughening and low energy coating is used to fabricate these non-wetting surfaces. Samples in this research are roughened by sandblasting, followed by electrochemical deposition of a Zn layer and a stearic acid finishing [9]. This paper presents results of an experimental study on the ice adhesion of these surfaces.

II. EXPERIMENTAL SETUP AND METHOD

In this work, ice adhesion is tested on hydrophobic 17-4 pH stainless steel surfaces. The samples have a dimension of 30mm x 30mm x 1mm. They are fabricated with random roughness through sandblasting with Al_2O_3 particles (100 μ m or 250 μ m), followed by electrochemical deposition of Zinc for desired roughness and finished with a stearic acid layer to prevent oxidation and to provide a lower

surface energy [9]. Seven different surfaces were tested. The properties of these surfaces are summarized in Table II.

The experimental setup is shown in Fig. 1. It contains a force probe attached to a holding plate with four wheels for maneuver. The probe is connected to a computer for ice adhesion data acquisition, where the maximum stress can be recorded. The surface samples sit on a 3-D printed stand block (see Fig.2a), which was made in such a way that the steel sample does not tip while testing for the ice adhesion strength. The block has a support part to ensure it does not move which is essential for the readings of ice adhesion strength. The setup was adjusted at a certain height to ensure consistency of the probe hitting area in the mold.

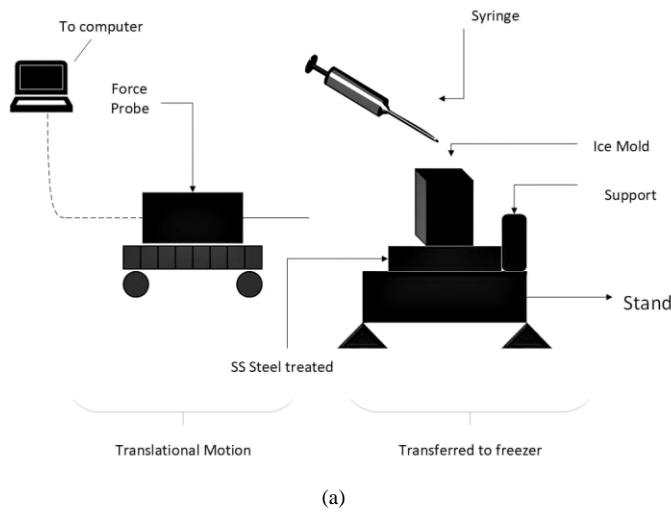


Figure 1 Experimental setup, (a) schematic of the system, also showing the way of ice formation on a test substrate, and (b) a picture showing the measurement

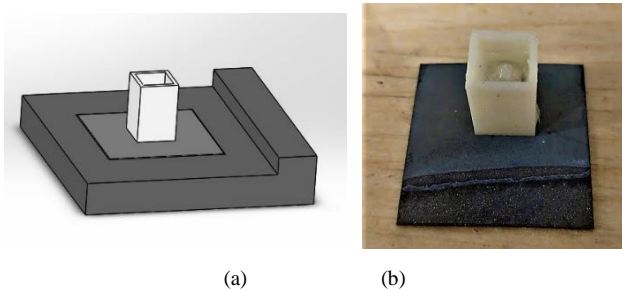


Figure 2 SolidWorks model of the sample stand block (a) and ice formation inside a mold on a coated surface (b)

Ice is made on these surfaces by using a 3D printed ice formation mold (light coloured part in Fig.2). The mold has a dimension of 8mm x 8mm and 15mm long, with hollow ends on both sides. Deionized water was injected into the mold to freeze under constant temperature (-12 °C) for the same time (90 minutes) and maintaining an overall relative humidity of 55-65% in a cold room. Humidity and temperature of the air are well maintained and monitored throughout the experiments. The water temperature is also measured with a thermocouple. To form an ice sample, A 50 ml syringe was filled with deionized water and precooled to 1 °C. Deionized water was used since water impurities could result in a large discrepancy in ice adhesion strength. The mold with water is put in a cold room where the temperature was about -12°C and left to cool down. Deionized water of 1 ml from the syringe was poured at the same rate onto the test substrate. The ice was left for 90 minutes to form before the ice adhesion experiment was carried out.

The procedure of the experiment starts with securing the test substrate onto the block stand. The wheels were attached to the force probe to control the motion, reduce friction and make the ice shedding process smooth. The force probe was aligned with the sample and set at 7.5mm above the test surface. The inclination of the surface (tilt angle) is strictly maintained to 0° (horizontal) since components of force affect the readings of ice adhesion strength. Each test surface was measured multiple times keeping all the parameters constant for reliability. Table I shows a summary of the experimental parameters.

TABLE I EXPERIMENTAL PARAMETERS

Cooler Temperature	-12 °C
Freezing Time	90 minutes
Temperature of water	1 °C
Relative humidity	55-65 %
Number of trials	6
Area of ice (m ² x 10 ³)	0.064

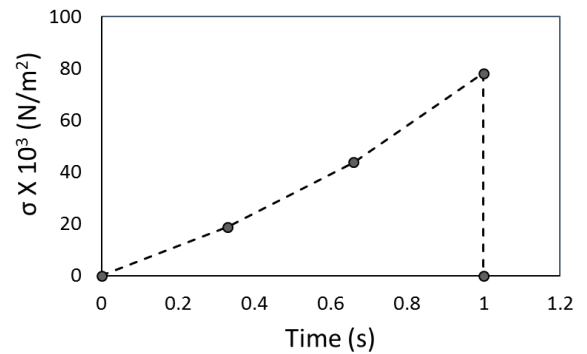


Figure 3 An example of average ice adhesion stress against time during the experiment

The measured stress σ comes from force readings from the probe divided by the surface area of ice on the stainless steel sample surface. By pushing the force probe against the

ice sample, the measured stress increases with time and eventually reach a maximum before a sudden drop to zero when the ice is shed off the surface. Figure 3 shows the variation of stress with time. The maximum stress is clearly shown when the ice is broken at the stainless-steel surface interface. The data was plotted using software which can record 30 frames per second from a video of ice shedding. The graph matched with the literature [2] and established an expected behaviour. The adhesion strength was considered as the maximum force per area observed before ice shedding. In this example, the peak stress of 78.125×10^3 N/m² can be seen from the graph.

Uncertainties were present in the experiment which may lead to some anomalous results. The uncertainty of the syringe of 50 ml can be estimated as ± 0.5 ml. The uncertainty of the force probe from Omega manual is estimated as $\pm 0.1\%$. We can also expect Vernier Caliper's uncertainty to be ± 0.02 cm. The uncertainty of digital thermometer is ± 0.1 °C and the relative humidity meter to be $\pm 0.5\%$.

III. RESULTS AND DISCUSSION

Table II shows a summary of the seven sample surfaces measured, with water dynamic contact angles (i.e., advancing angle θ_{adv} , receding angle θ_{rec}) and the contact angle hysteresis on the surfaces, as well as the measured ice adhesion stress on each surface.

TABLE II SUMMARY OF EXPERIMENTAL RESULTS

Surface (Stainless Steel)	Ice Adhesion Stress (Average) N/m ² x10 ³	θ_{adv}	θ_{rec}	Contact Angle Hysteresis (CAH)
Zinc + Stearic Acid 250 μ m Sandblasted	79.69	142.01	106.34	36.67
Zinc + Stearic Acid 100 μ m Sandblasted	82.37	146.68	123.00	23.68
100 μ m sandblasted + Stearic Acid	83.93	154.23	121.00	33.24
250 μ m sandblasted + Stearic Acid	107.37	127.97	91.94	36.03
250 μ m sandblasted	104.69	102.83	65.71	37.12
100 μ m sandblasted				
Zinc + Stearic Acid -1.55V	94.64	112.47	73.36	39.12

The magnitudes of the ice adhesion stress are comparable with the literature values. Each coated surface was tested six times keeping all the parameters constant which yields to a quantitatively meaningful data. The measured average stress of ice adhesion is plotted against two different measures of the water contact angle to allow for ready comparison with previous literature [1,2,7]. For superhydrophobic surfaces, measuring static contact angle can be difficult [1,9], so as an alternative, advancing angle, θ_{adv} and contact angle hysteresis were chosen for the analysis.

Figure 4 shows a decreasing linear relationship between average stress of ice adhesion and the advancing angle θ_{adv} . As the theory suggests, higher contact angle means higher hydrophobicity, which leads to a decrease in ice adhesion stress as indicated by the results. It is found that the advancing contact angle is a vital parameter in determining ice adhesion stress on hydrophobic surfaces.

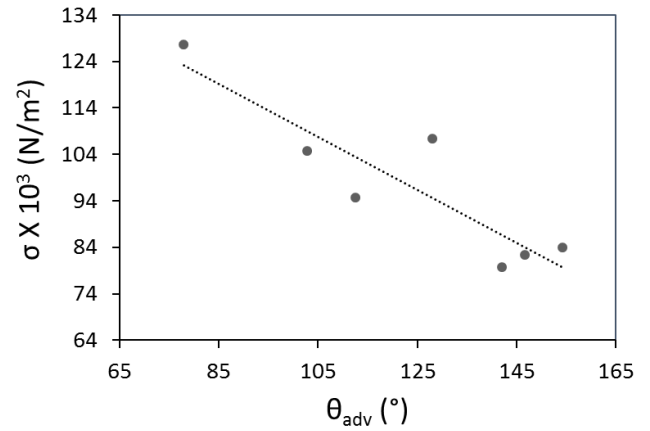


Figure 4 Average stress of ice adhesion measured at -12 °C for seven different coatings on stainless steel are plotted against the advancing contact angle (θ_{adv}) with a freezing time of 90 mins. The dotted line is the best fit for the data acquired.

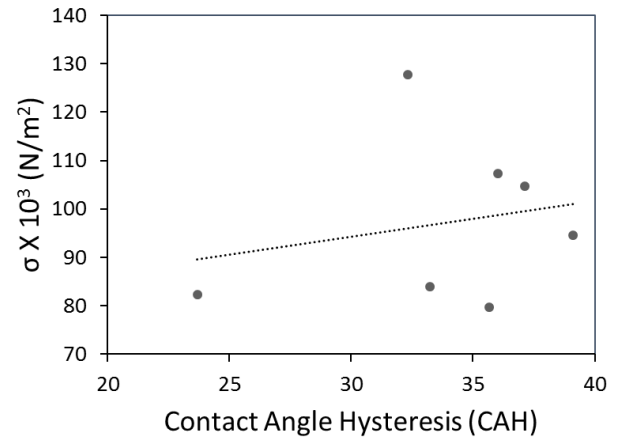


Figure 5 Average stress of ice adhesion measured at -12 °C for seven different coatings on stainless steel are plotted against Contact Angle Hysteresis (CAH) with a freezing time of 90 mins. The dotted line is the best fit for the data acquired

Improved hydrophobicity of a surface also requires a low contact angle hysteresis (CAH). Results shown in Figure 9 suggest an overall linear relationship between the ice adhesion stress and contact angle hysteresis. From this point of view, improved hydrophobicity may also lead to low ice adhesion on these surfaces. The plot also matches the linear trendline in the literature [1] for another type of repellent surfaces.

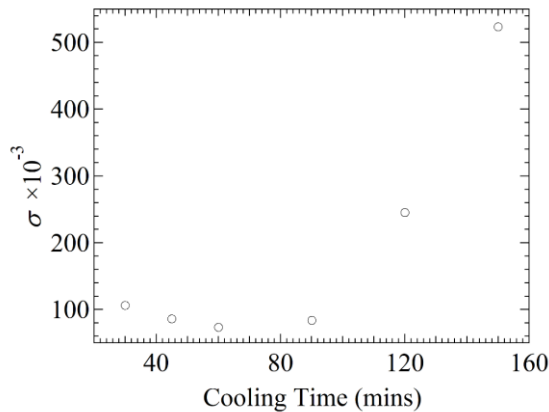


Figure 6 Ice adhesion stress after different cooling time for 100 μm sandblasted and zinc coating sample at $-12\text{ }^{\circ}\text{C}$.

Results in Figure 6 show ice adhesions on a sample with 100 μm sandblasting and coated with zinc but with different freezing times at $-12\text{ }^{\circ}\text{C}$. There is a huge overall increase of ice adhesion stress with the increase of cooling time which implies that longer ice formation may result in an increase in ice adhesion stress. However, there is a slight decrease in ice adhesion at 45 and 60 minutes which was doubtful. Most of the comparison from scholarly literature does not show the relation between ice adhesion stress and icing time, hence more detailed research is needed for a stronger conclusion.

IV. CONCLUSION

A simple experimental setup was developed for ice adhesion measurement. The average stresses of ice adhesion were measured on stainless steel samples with seven different surface structures, i.e., a combination of sandblasting, and coatings of Zinc and stearic acid. Advancing and receding contact angles are measured to investigate water wettability on these surfaces. The measured ice adhesion stresses and contact angle hysteresis were analyzed and a strong correlation was found between the ice adhesion stress, advancing angle and contact angle hysteresis. Increase in contact angle and decrease in the hysteresis reduce ice adhesion stress on the surfaces. It was demonstrated that Zinc coating, stearic acid and sandblasting relatively increase hydrophobicity on stainless steel surfaces.

Durability of the coatings can be a concern while carrying out ice adhesion experiments. Parts of the coating may be scraped off from a surface thus affecting the later experiments on the same sample. This problem may be

solved by cleaning the sample with ultrasonic cleaner to scrape off any residues.

REFERENCES

- [1] J. Meuler, J. D. Smith, K. K. Varanasi, J. M. Mabry, G. H. McKinley, and R. E. Cohen, "Relationships between water wettability and ice adhesion," *ACS applied materials & interfaces*, vol. 2, no.11, pp. 3100-3110, 2010.
- [2] P. S. R. Beeram, "Characterization of ice adhesion strength over different surfaces pertinent to aircraft anti-/de-icing", 2017, pp. 17-27
- [3] C. Cui, X. Duan, B. Collier, and K. Poduska, "Fabrication and wettability analysis of hydrophobic stainless steel surfaces with microscale structures from nanosecond laser machining," *Journal of Micro and Nano-Manufacturing*, vol. 6, no. 3, p. 031006, Jun. 2018.
- [4] V. Hejazi, K. Sobolev, and M. Nosonovsky, "From Super hydrophobicity to icephobicity: forces and interaction analysis," *Scientific Reports.*, Vol.3, pp.2194, July 12, 2013.
- [5] Y. Pan, K. Shi, X. Duan, G. F. Naterer. "Experimental investigation of water droplet impact and freezing on micropatterned stainless steel surfaces with varying wettabilities." *International Journal of Heat and Mass Transfer*, vol. 129, p.953-964, 2019.
- [6] V. Hejazi, and M. Nosonovsky, "Why superhydrophobic surfaces are not always hydrophobic," *ACS Nano*, vol. 6, no.10, pp.8488-8491, 2012.
- [7] L. Makkonen, "Ice adhesion- theory, measurements, and countermeasures," *Journal of Adhesion Science and Technology*, vol. 26, no.4-5, pp. 413-445, 24 July 2012.
- [8] H. Work Jr, "The measurement of the adhesion of glaze ice." University of Louisville, 2018.
- [9] B. Gao and K. M. Poduska, "Electrodeposited Zn for Water-Repellent Coatings," *Journal of The Electrochemical Society*, vol. 165, no. 10, pp. D472-D476, Jul. 2018.

PIV Analysis of Flow Around a Simplified Ship Hull

Jennie M. Flaig
Mechanical Engineering Department
University of Manitoba
Winnipeg, Manitoba, Canada
flaigj@umanitoba.ca

Xingjun Fang
Mechanical Engineering Department
University of Manitoba
Winnipeg, Manitoba, Canada
xingjun.fang@umanitoba.ca

Baafour Nyantekyi-Kwakye
Mechanical Engineering Department
Memorial University of Newfoundland
St. John's, NL, Canada
bnyantekyikw@mun.ca

Mark F. Tachie
Mechanical Engineering Department
University of Manitoba
Winnipeg, Manitoba, Canada
mark.tachie@umanitoba.ca

Abstract—This paper experimentally investigates the influence of Reynolds number on the flow field of a simplified ship model using a time-resolved particle image velocimetry measurement system. The tests are conducted at three different velocities resulting in three different Reynolds numbers based on the free stream velocity and ship length. All tests were performed in an open recirculating water channel. The draft of the model varies due to the different stream velocities between 0.16 and 0.23 m. The pattern of the flow field was similar for all three test conditions, but showed increasing Reynolds stresses and faster velocity recoveries at higher Reynolds numbers.

Keywords—*Reynolds number; Reynolds stress; ship; spanwise vorticity; time-resolved PIV*

I. INTRODUCTION

The flow around a ship hull is strongly dependent on the velocity and subsequently the Froude number at which it is operated. In the last two decades PIV systems have been used to measure stresses and velocities evolving around ship bodies. This delivers a greater understanding of the flow field around a ship body and will therefore help improving ship hydrodynamics. In many applications it has replaced the Pitot tube. In previous studies, both static [1] as well as moving [2] PIV systems have been tested in towing tanks.

Previous experimental studies also investigated flow around hull forms. Flow characteristics around modern commercial ship models at constant Froude numbers have been examined by [3]–[4] as well as flow patterns around ship models with rotating propellers [5]. Mean-flow and turbulence measurements in the boundary layer and wake of a ship double model have been performed in a wind tunnel [6]. Variations of Froude number have been performed to examine the structure of bow waves in a towing tank [7] and to examine the effect of drift angles on the ship flow [8]. Reference [3] studied the global force and wave pattern around container and crude oil vessels to identify the

resistance and propulsion efficiency using a resistance dynamometer, servo-needle-type wave height gauges and 5-hole pitot tubes. Reference [4] utilized PIV to take a closer look at the stern and near-wake region of a modern container hull to investigate the flow characteristics. Paik et al. performed PIV analysis of the flow around a container ship model with a rotating propeller to get a more accurate formula for the nominal wake caused by the propeller and looking for ways to enhance the effective wake. Reference [6] performed mean flow and turbulence measurements in the boundary layer and in the wake of a ship double model in a wind tunnel. They analyzed the physical aspects and complexities of the three-dimensional flow and the flow near the trailing edges and the way these evolve into the wake.

To date, only two experimental studies have investigated the influence of Froude number and thus Reynolds number on the flow behavior within the vicinity of a ship hull. However, these studies focused on effects of drift angles [8] and the structure of bow waves [7]. The present study therefore seek to investigate the influence of the Reynolds number on the mean flow field and the Reynolds stress distribution around the entire ship body. Areas of interest are especially the interaction with the water surface in the stern section and the velocity and stress development in the wake region. The velocities chosen for this studies are within a realistic range regarding the size and ship type the model is representing.

II. EXPERIMENTAL SETUP

A. Model

The model used for this experiment is based on a container ship but is strongly simplified. The container model has a design waterline length of 0.38 m, a width of 0.062 m, a draft of 0.023 m and a block coefficient of 0.61. It was made of acrylic glass to a 1/450 scale adaption of an 1850 TEU container ship design. The curvature of the waterline at the design draft was taken from

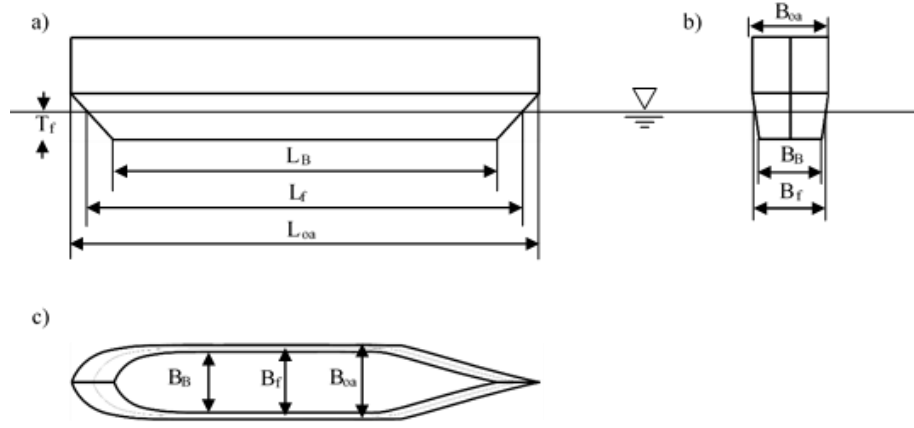


Figure 1 Schematic of model used, (a) profile view, (b) front view, and (c) top view, with dimensions explained in Table 1

the design and then scaled up by a factor of 1.09 in both length and width at a height of 0.415 m above the baseline and scaled down by a factor of 0.89 to the base. It was fastened securely by four screws that protruded out of the top of the model into a mount. This assured that the model could not move during the experiment. To prevent deflection of the laser beam and to enhance visibility the hull of the model was painted black in the near water and water regions. The model was fixed at the same height above the ground of the water tunnel throughout all test drives. As the water level changes due to the pump frequency, this led to slightly different drafts for the different test conditions. The exact dimensions for the performed test conditions can be viewed in Table 1. The Reynolds number and Froude number were calculated based on the water line length L_f . The layout of the model can be seen in Fig. 1.

B. Measuring system and experimental facility

The experiments were carried out in a recirculating open water tunnel with a length of 6.0 m, a width of 0.6 m and a depth of 0.425 m. The pump was operated at frequencies between 11 and 16 Hz resulting in free stream velocities between 0.38 and 0.56 m/s. The flow is seeded with 10-micron (d_p) silver-coated hollow glass particles that have a density of $1.41 \times 10^3 \text{ kg/m}^3$. Following [9], the slip velocity of particles was estimated as

$$U_s = \frac{d_p^2(\rho_p - \rho_f)}{18\rho_f\nu}g \quad (1)$$

where ρ_p being the density of the seeding particles and ρ_f the density of the working fluid (water). This yields a U_s value of $2.23 \times 10^{-5} \text{ m/s}$, which is much smaller than the smallest velocity scale involved in this investigation. Therefore, the seeding particles follow the flow motion. To enumerate the response

TABLE 1. Dimensions of the model based on test pump frequencies

	11 Hz	14 Hz	16 Hz
L_f [m]	0.381	0.373	0.368
B_f [m]	0.062	0.061	0.060
T_f [m]	0.023	0.019	0.016
CB [-]	0.61	0.57	0.520
U_∞ [m/s]	0.38	0.49	0.560
Re [-]	1.45×10^5	1.83×10^5	2.06×10^5
Fn [-]	0.197	0.256	0.295

time of the seeding particles to sudden changes in the flow velocity, the relaxation time τ_p was estimated from

$$\tau_p = \frac{d_p^2(\rho_p - \rho_f)}{18\rho_f\nu} \quad (2)$$

The value of τ_p is $2.23 \times 10^{-6} \text{ s}$. Assuming that the smallest temporal scale (τ_f) is $L_f/U_\infty = 0.657$ based on the highest frequency and the Stokes number ($S_k = \tau_p/\tau_f$) is approximately 0.000003, which is in its recommended range of $S_k \leq 0.05$. Based on this, it can be said that the adopted seeding particles followed the fluid motions, and their instantaneous velocities were accurate representation of the instantaneous local fluid velocities.

The PIV system used for this experiment consists of a dual head high speed Nd:YLF laser and initially three high speed 12-bit complementary metal oxide semiconductor cameras with a 60 mm Nikon lens that each have 806 frame rates at a full resolution of $2560 \text{ pixel} \times 1600 \text{ pixel}$ and a pixel pitch of $10 \mu\text{m}$. The interrogation area of the cameras was set at $32 \text{ pixels} \times 32 \text{ pixels}$ with 75% overlap corresponding to $0.788 \text{ mm} \times 0.785 \text{ mm}$. The data acquisition and vector calculation was controlled with DaVis software supplied by LaVision. An in-house MATLAB script was used to post-process the acquired data.

Measurements were performed at 16 Hz, 14 Hz and 11 Hz using one camera, but these measurements focused on the stern and wake of the model as shown in Fig. 2. For each frequency, a sample size of 12000 was acquired. The measurements of the whole model showed that the boundary layer of the bow was laminar. Therefore, this paper focusses on the wake region only.

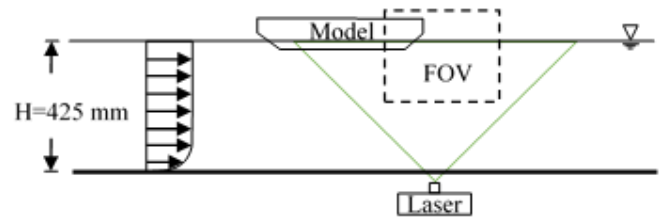


Figure 2. Experimental setup showing the model in its field of view

III. RESULTS AND DISCUSSION

A. Contour plots

The streamwise and cross-stream mean velocity contours provide a general understanding of the flow field within the vicinity of the ship model. Differences due to the Reynolds numbers can be observed and set into relation. Although the three test frequencies resulted in different lengths, L_f , (see Table 1), both x and y were normalized by the design length of 0.38 m in order to keep the plots comparable. The velocities were normalized by the corresponding free stream velocity U_∞ as specified in Table 1. The contour plots for the streamwise velocity (U/U_∞) illustrated in Fig. 3, show a qualitatively similar mean flow pattern. The normalized velocity beneath the boat for all three test conditions are almost equal to the free stream velocity. Lower values of the mean velocity existed in the wake region but it increases sharply to $0.5U_\infty$. At approximately half of the ship's length behind the lower ship edge, the mean velocity was as high as $0.8U_\infty$. The general development of velocity behind the stern is similar for all three test conditions. At ship level, it appeared that a higher Reynolds number leads to a faster recovery of the velocity to that of the free stream velocity. At approximately 10 % of the ship's length behind the lower edge of the ship a bump of a velocity lower than free stream velocity was visible, stretching into the free stream velocity field. The bump gets more distinct at higher Reynolds numbers. This observation can be attributed to the intensity of wake formed behind the stern.

Fig. 4 shows contours of the cross-stream mean velocity. Positive velocities existed beneath the ship is positive whereas negative velocities dominated behind the ship model. The positive values denote entrainment of bottom fluid into the wake region. Within the wake region, the negative section extended slightly below the base of the ship. At the lowest Reynolds number, the negative section extended gradually. For the intermediate Reynolds number the negative section extended gradually at first, but then transitions into an upward wave motion. The highest Reynolds number features an almost parabolic shape reaching the lowest y -level within the field of view compared to the other test conditions. Additionally, a lower Reynolds number seems to intensify the negative velocities in the wake region, whereas a higher Reynolds number seems to increase the positive velocity right beneath the lower edge of the ship moving downward.

The spanwise mean vorticity (shown in Fig. 5) was normalized by the free stream velocity and the design waterline length (U_∞/L). The mean vorticity displayed in Fig. 5 developed in a similar mode for all three test cases. The diffusion of vorticity is relatively small, it returns to zero within the field of view. For the test cases examined herein, the higher Reynolds numbers seem to need a longer distance to even out again. While this difference is hardly recognizable between the lower and intermediate Reynolds number, it is clearly visible for the higher Reynolds number. A higher Reynolds number also leads to a more rapid spread of the mean vorticity in the cross-stream direction. The vorticity directly behind the leading edge is almost zero for the lower Reynolds number, while the highest Reynolds number has an area of higher vorticity stretching up to $y/L=0.03$.

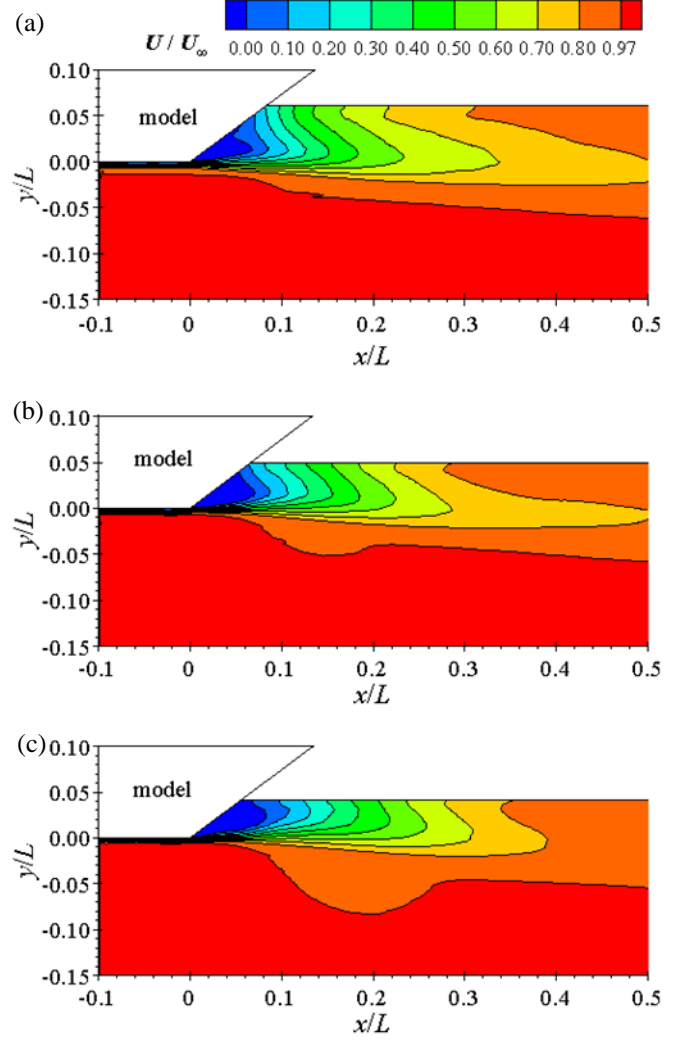


Figure 3. Contour plots of stream-wise mean velocity (U/U_∞) of Reynolds numbers (a) 1.45×10^5 , (b) 1.83×10^5 and (c) 2.06×10^5

The streamwise Reynolds normal stress shown in Fig. 6 developed similarly for all three test conditions. The region of higher stresses spreads out from the surface downward and continues to lower itself at an angle of approximately 10° , extending from the baseline of the model starting at the lower corner of the leading edge. The regions of higher stresses are slightly larger for higher Reynolds numbers, but not significantly. The general area of higher stresses is wider for the lower Reynolds number, this however seems to be related to the bigger draft.

The general area of increased cross-stream Reynolds normal stress (not shown) spreads more towards the surface as was observed for the streamwise component. However, the values were slightly lower than the streamwise normal stresses. Specifically, the peak values of the cross-stream Reynolds normal stress was approximately 25% lower than the corresponding values for the streamwise Reynolds normal stress. The levels of the streamwise and cross-stream Reynolds normal stresses implied that the flow in this region was highly

anisotropic so that turbulence models based on isotropic assumption cannot explicitly account for the anisotropic nature of this flow and will not be able to predict the flow behavior reliably. Evident for both stress distributions is that the area of noticeable stresses falls at about 10° starting from the base of the ship. This was highly visible in the vorticity and shear stress distributions. This results in a 45° angle between the lower stress regions and the stern. Likewise, the distribution of the Reynolds shear stress (not shown) revealed no significant differences between the three test conditions. A region of slightly higher shear stress was prominent for the lowest Reynolds number right beneath the ship, this effect was not noticeable for the other test conditions.

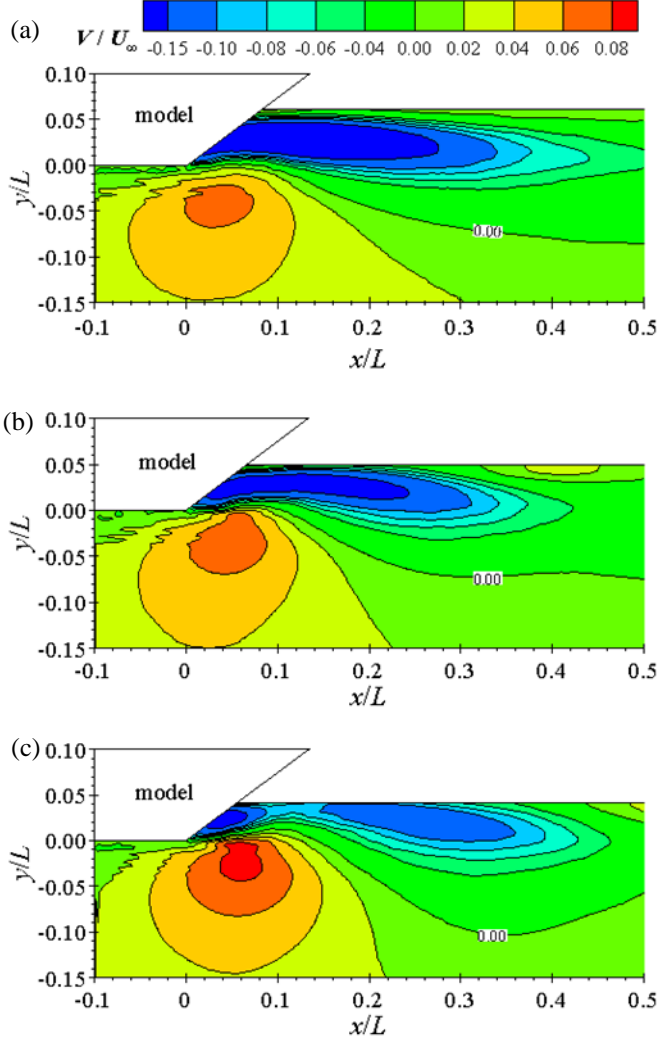


Figure 4. Contour plots of cross-stream mean velocity (V/U_∞) of Reynolds numbers (a) 1.45×10^5 , (b) 1.83×10^5 and (c) 2.06×10^5

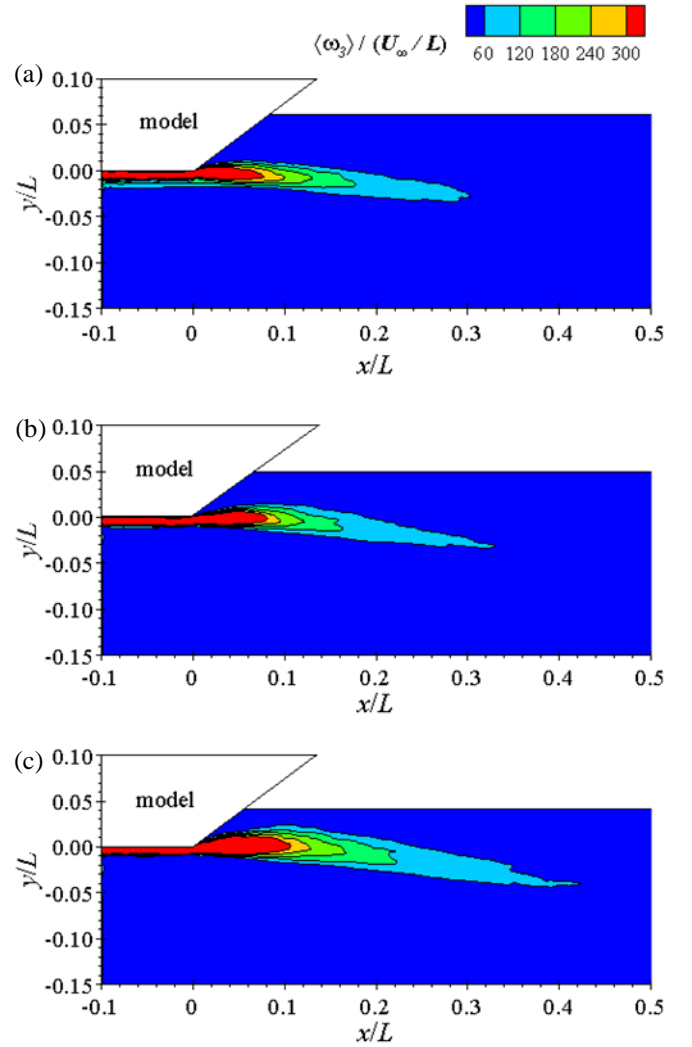


Figure 5. Contour plot of mean vorticity ($\omega_3/(U_\infty/L)$) of Reynolds numbers (a) 1.45×10^5 , (b) 1.83×10^5 and (c) 2.06×10^5

B. One-dimensional profiles of mean flow and Reynolds stresses

To quantitative explanation of the velocity field, as well as to elucidate the flow development with respect to increasing Reynolds number, one-dimensional profiles of the mean velocity, mean vorticity and Reynolds stress contours have been examined. The following profiles are each taken at the same positions in relation to the ship model, as can be seen in Fig. 7. The first and second vertical profiles were extracted at $x/L = 0.05$ and 0.25 , respectively. It should be remarked that the profiles at $x/L = 0.25$ do not interact with the ship hull and can therefore be used up to the water surface which is located at $y/L = 0.042$ ($Re = 2.06 \times 10^5$), $y/L = 0.05$ ($Re = 1.83 \times 10^5$) and $y/L = 0.061$ ($Re = 1.45 \times 10^5$). The first horizontal profile was extracted at $y/L = 0$ corresponding to the ship's baseline; while the second horizontal profile was extracted approximately halfway between the base and the waterline of the highest Reynolds number, $y/L = 0.02$.

1) Streamwise mean velocity

These four profiles show that the three test cases result in very similar stream-wise velocity patterns that support the observations of the contour plots. The first vertical profiles (Fig. 8a) feature the section right behind the ship. For all test cases,

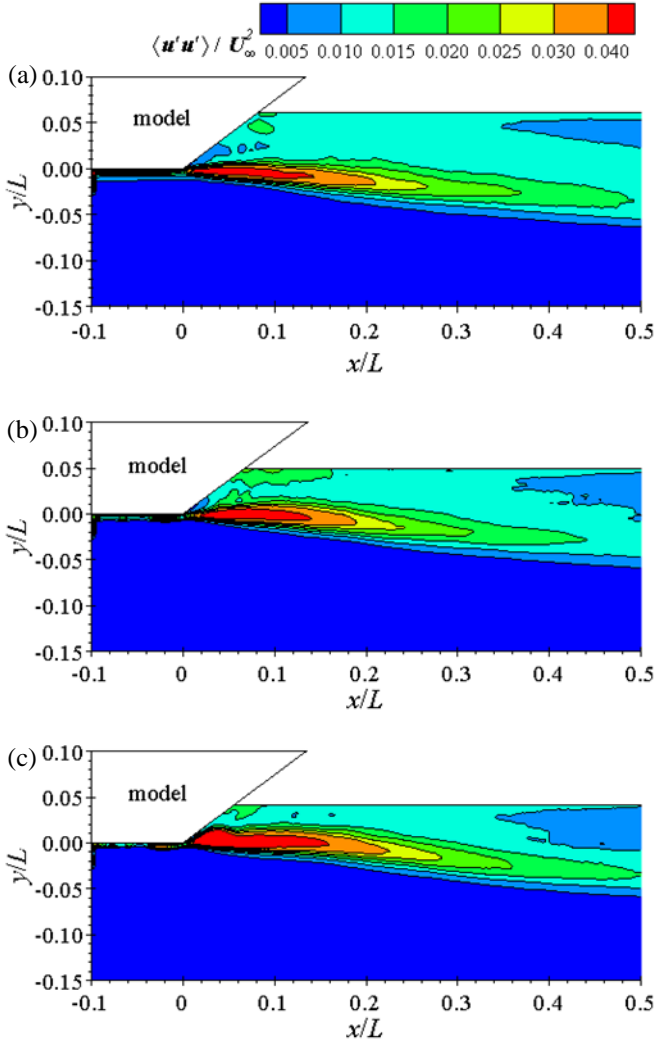


Figure 6. Contour plot of mean streamwise Reynolds normal stresses ($\langle u'u' \rangle / U_\infty^2$) of Reynolds numbers (a) 1.45×10^5 , (b) 1.83×10^5 and (c) 2.06×10^5

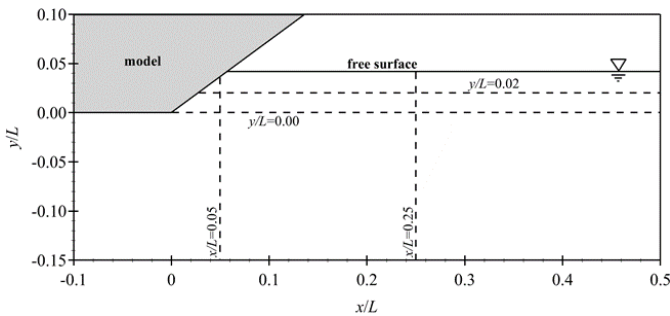


Figure 7. Illustration of locations for vertical and horizontal profiles

the velocity starts out at free-stream velocity and then dips drastically from just before the base of the ship to about $y/L = 0.03$ above the baseline where the velocity goes below zero. In the second vertical profile the velocity development has a slight decline starting from the beginning of the field of view until approximately $y/L = 0.05$ beneath the ship base. From there the slope decreases resulting in a linear fall. The velocity has its minimum at $y/L = 0$ with a value between $0.6U_\infty$ and $0.7U_\infty$, with the highest Reynolds number having the lowest value, followed surprisingly by the lowest Reynolds number. After this point the velocity increases again until the profile reaches the free surface. In the first horizontal profile (Fig. 8c) the data before $x/L = 0$ is most likely junk data due to the interaction with the ship body and are therefore not shown. Right behind the lower ship edge the velocity increases drastically from 0 to $0.4 U_\infty$ at $x/L = 0.05$ behind the lower edge and then steadies itself at a lower rate where it slowly approaches $0.9U_\infty$ at the end of the field of view, i.e. $x/L = 0.5$. The second horizontal profile is located halfway between the baseline and the free surface of the highest Reynolds number. From this point the velocity increases in a parabolic way to $0.8U_\infty$ ($Re = 1.45 \times 10^5$, $Re = 1.83 \times 10^5$) and $0.9U_\infty$ ($Re = 2.06 \times 10^5$).

2) Mean spanwise vorticity

The distribution of the mean spanwise vorticity (Fig. 9) was similar for all three test conditions. In the vertical profiles the vorticity increased shortly beneath the baseline of the ship attaining a peak value at the height of the baseline for the first profile and at $y/L = 0.2$ beneath the baseline for the second profile. While the first vertical profile shows large differences in the magnitude of the vorticity, with the highest Reynolds number being correlated to a higher vorticity, the second profile shows only marginal differences. This fact is also illustrated by the first horizontal profile (Fig. 9c) where the vorticity after having a peak just behind $x/L = 0$ evens out to zero quite quickly. The peaks of the first horizontal profile show significant differences in the magnitude between the three test cases. In the second profile only the high Reynolds number leads to a noticeable peak at $x/L = 0.1$, but then also evens out to zero shortly thereafter. This supports what has been observed earlier in the vorticity contour plots. There the regions of higher vorticity for the high Reynolds number stretched out further towards the free surface than for the other two test cases.

3) Streamwise Reynolds normal stress

The stream-wise Reynolds normal stress is depicted in Fig. 10. While the vertical profiles show any difference between the different test conditions, the horizontal profiles show distinct Reynolds number effects on the distribution and the intensity of the stresses. Specifically, the levels of the Reynolds stresses increase with increasing Reynolds number. Both, the vertical and horizontal profiles near the ship body show peak values close to $y/L = 0$ and $x/L = 0$, respectively. Moving further away from the edge leads to the peak being before $y/L = 0$ for the vertical profiles and after $x/L = 0$ for the horizontal profiles. Besides the above mentioned peaks, the streamwise Reynolds stresses stay at a relatively constant level.

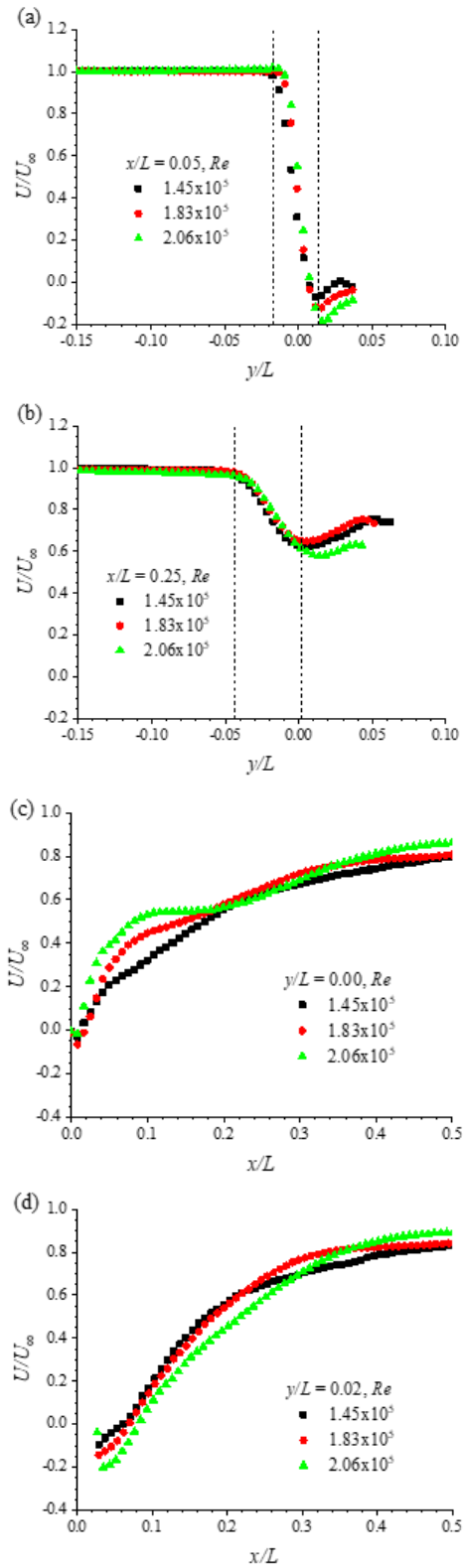


Figure 8. Streamwise velocity profiles at selected locations

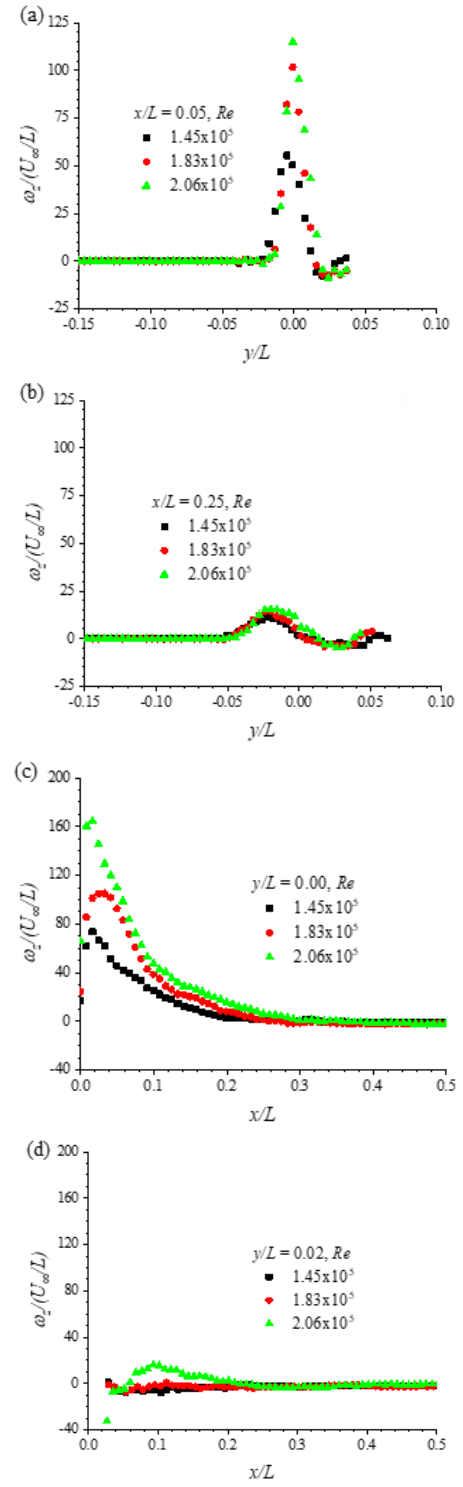


Figure 9. Profiles of mean spanwise vorticity

4) Reynolds shear stress

Fig. 11 shows the distribution of the Reynolds shear stresses. Significant in the second vertical profile is that the stresses drop down into the negative values behind the peak before they even out at zero. As with the horizontal stress profiles, there was a

difference in the levels of the shear stress. Especially near ship edge profiles, $x/L = 0.05$ (Fig. 11a) and $y/L = 0.00$ (Fig. 11c) the highest and intermediate Reynolds number produce much higher stresses than the lowest Reynolds number. Surprising for all three stress cases is the second vertical profile. The Reynolds shear stresses develop similarly with no observable Reynolds number effects. However, the location of the peak at the highest Reynolds number is delayed, and the magnitude of negative stress decreases with increasing Reynolds number and is almost absent at the highest Reynolds number. Thus, the Reynolds number seems to be quite irrelevant for the development of stress at a certain distance from the leading edge. The second horizontal profile delivers a very interesting result. There is no characteristic pattern obvious between the three test cases, besides starting to even out towards zero at approximately $x/L = 0.2$. As the Reynolds number increases, the streamwise location of the local minimum occurs sooner and so is the location where the profile changes sign from negative to positive for the intermediate and highest Reynolds numbers. This is very similar to what was observed for the second vertical profile of the cross-stream velocity.

IV. CONCLUSIONS

An experimental study was performed with a time-resolved PIV system to analyze the influence of Reynolds numbers on the flow field of a simplified container ship model. Three Reynolds numbers were tested: 1.45×10^5 , 1.83×10^5 and 2.06×10^5 . The flow field proofed to be similar for all Reynolds numbers, showing shared characteristic that mostly differed by intensity only. The streamwise velocity contour and profiles showed that the velocities develop alike, but that velocities of higher Reynolds numbers recover more quickly than those with lower Reynolds number. The cross-stream velocity distribution was negative behind the model and positive beneath (as a result of entrainment).

Unlike the stream-wise velocities, the vorticity distributions show that a higher Reynolds number will increase the area of higher vorticity behind the ship. Noticeable for the vorticity as well as the stress contours was a characteristic drop visible for the vorticity and stress areas, respectively. In all cases this drop equaled an approximate 45° angle in relation to the stern of the model. While the Reynolds number hardly had any influence on the area of increased streamwise stresses, the horizontal profile on baseline level showed, that it does have a significant impact on the absolute intensity of the stresses. The maximum stress for the highest Reynolds number was twice as high as for the lowest Reynolds number on this level. The cross-stream Reynolds stresses recover slightly faster for higher Reynolds numbers. Although there was a clear increase of stress from lower to higher Reynolds numbers, this was not as distinct as for the stream-wise Reynolds stresses. In fact, the stress increase from the intermediate to highest Reynolds number is only about 9% instead of roughly 30% in the case of the streamwise Reynolds stresses. The lowest Reynolds number is at a similar level compared to the highest Reynolds number.

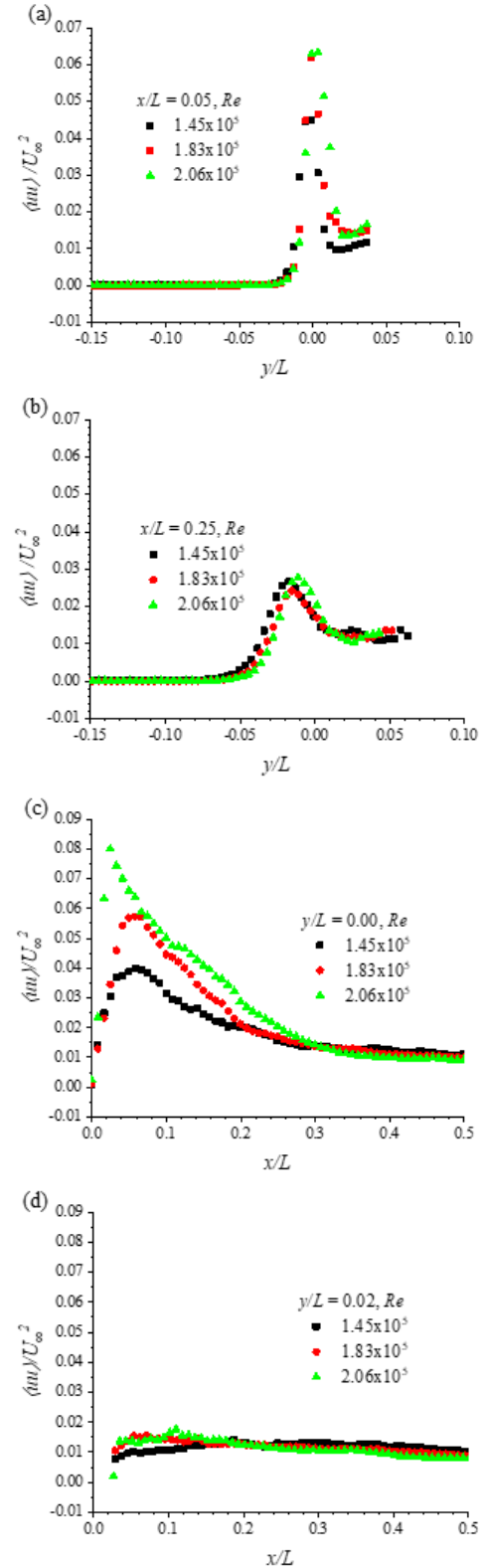


Figure 10. Profiles of streamwise Reynolds normal stress

The Reynolds shear stresses provided similar results for the maximum stress level. The maxima of the intermediate and the high Reynolds number were rather close, whereas the maximum of the lowest Reynolds number is significantly lower. There is no significant difference between the stress areas, but it was noticeable in the contours that higher Reynolds numbers led to higher shear stresses in general.

REFERENCES

- [1] L. Gui, L. Longo, and F. Stern. Towing tank PIV measurement system, data and uncertainty assessment for DTMB Model 5512. Experiments in Fluids No. 31 (2001): pp. 336-346.
- [2] J.-H. Chen, and C.-C. Chang. A moving PIV system for ship model test in towing tank. Ocean Engineering No. 33 (2006): pp. 2025-2046.
- [3] W. J. Kim, S. H. Van, and D. H. Kim. Measurement of flows around modern commercial ship models. Experiments in Fluids No. 31 (2001): pp. 567-578.
- [4] S.-J. Lee, M.-S. Koh and C.-M. Lee. PIV velocity field measurements of flow around a KRISO 3600TEU container ship model. Journal of Marine Science and Technology No. 8 (2003): pp. 76-87.
- [5] B. G. Paik, C. M. Lee, and S. J. Lee. PIV analysis of flow around a container ship model with a rotating propeller. Experiments in Fluids No. 36 (2004): pp. 833-846.
- [6] V. C. Patel and O. P. Sarda. Mean-Flow and turbulence measurements in the boundary layer and wake of a ship double model. Experiments in Fluids No. 8 (1990) pp. 319-335.
- [7] R. R. Dong, J. Katz, and T. T. Huang. On the structure of bow waves on a ship model. Journal of Fluid Mechanics Vol. 346 (1997): pp. 77-115.
- [8] J. Longo, and F. Stern. Effects of drift angle on model ship flow. Experiments in Fluids No. 32 (2002): pp. 558-569.
- [9] M. Raffel, C. E., Willert, S. Wereley, and J. Kompenhans. Particle image velocimetry: A practical guide. Springer, Second edition (2007).

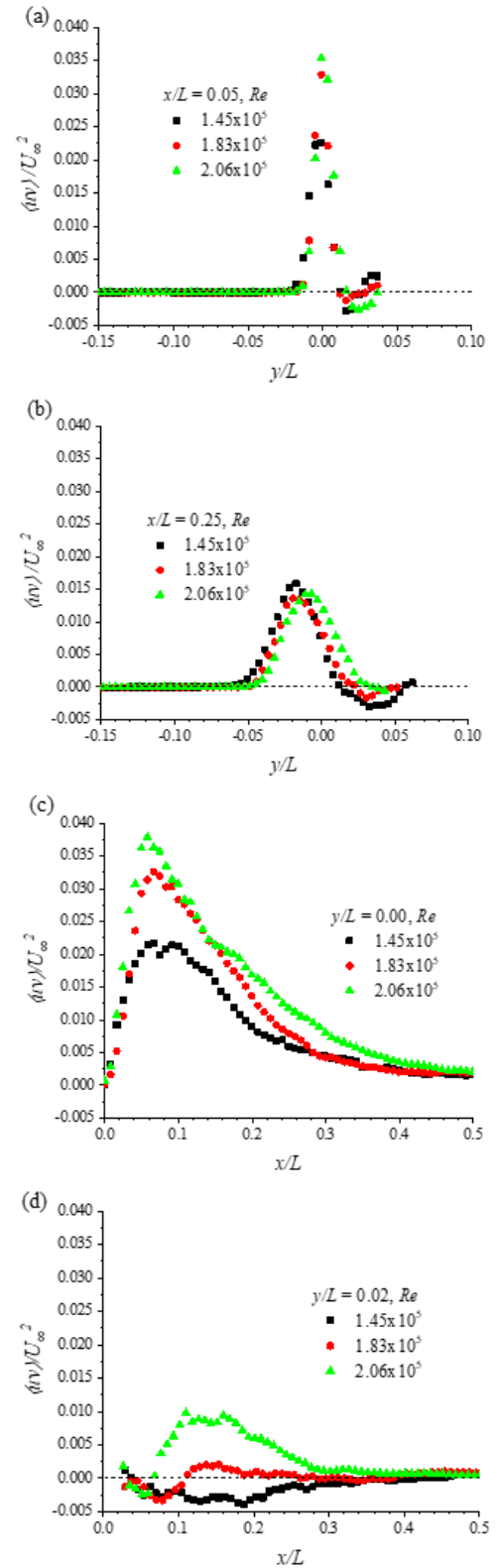


Figure 11. Profiles of Reynolds shear stress

Aspect Ratio Effects on Turbulent Flow over Forward-Backward-Facing Steps

Heath A. Chalmers
Mechanical Engineering Department
University of Manitoba
Winnipeg, Manitoba, Canada
chalmerh@myumanitoba.ca

Xingjun Fang
Mechanical Engineering Department
University of Manitoba
Winnipeg, Manitoba, Canada
xingjun.fang@umanitoba.ca

Baafour Nyantekyi-Kwakye
Mechanical Engineering Department
Memorial University of Newfoundland
St. John's, NL, Canada
bnyantekyikw@mun.ca

Mark F. Tachie
Mechanical Engineering Department
University of Manitoba
Winnipeg, Manitoba, Canada
mark.tachie@umanitoba.ca

Abstract— An experimental study was conducted to investigate the effects of streamwise aspect ratio on the characteristics of turbulent flow separation induced by forward-backward-facing steps submerged in a thick turbulent boundary layer. In these experiments, the spanwise dimension and step height, h , were kept constant while the streamwise dimensions, L , of the steps were varied to achieve the following 8 aspect ratios, $L/h = 1, 2, 3, 4, 5, 6, 8$, and 10 . The Reynolds number based on the freestream velocity and step height, was 13200 , while the ratio of the boundary layer thickness to the step height is 6.5 . A time-resolved particle image velocimetry system was used to perform detailed velocity measurements within the separation regions upstream, over and behind the steps. The results show that the reattachment length of the recirculation bubble on top of the step is not sensitive to changes in aspect ratio, however, the reattachment length of the separation bubble behind the step decreases as the aspect ratio increased from 1 to 2 and then remains constant for larger aspect ratios. The effects of aspect ratio on the mean velocities and Reynolds stresses are also examined.

Keywords-component; Turbulent flow separation; forward-backward-facing step; particle image velocimetry

I. INTRODUCTION

Flow separation caused by bluff bodies is a common phenomenon encountered in diverse engineering and industrial applications including buildings, bridges, airfoil, heat exchangers and wind turbines. The impact of flow separation in these and other applications can be either beneficial or detrimental, depending on the context. For instance, flow separation can give rise to undesirable effects such as flow-induced vibration, acoustic resonance, fatigue failure, increased drag or non-optimal performance of fluid devices. In other applications, flow separation is exploited to enhance heat transfer or improve flow-mixing properties and stabilize a combustion process.

Due to their practical importance, flow separations induced by surface-mounted bluff bodies such as forward facing and backward facing steps [1-7], hereafter referred to as FFS and BFS, respectively, have been investigated extensively over the past decades. Most of these investigations focused on understanding how changes in Reynolds number ($Re = U_e h / \nu$, where U_e is the freestream velocity, h is step height and ν is kinematic viscosity), boundary layer thickness (δ) of the approach flow relative to step height, i.e. δ/h , and upstream roughness affect the mean flow and turbulent characteristics of the separated and reattached shear layers. Reference [5] performed detailed particle image velocimetry (PIV) measurements to investigate the effects of Reynolds number and relative boundary layer thickness (δ/h) on the characteristics of turbulent flow over a FFS. Their results showed that, for a fixed δ/h , the reattachment length on top of the step increased linearly with Reynolds number up to $Re = 8500$, but remained unchanged for $Re \geq 8500$. A similar transition of the reattachment length from a Re-dependent regime to Re-independent regime has been observed in other studies except that the critical Reynolds number at which this transition occurs is not universal. For instance, [6] investigated the effects of Reynolds number on FFS submerged in both smooth-wall turbulent boundary layers and rough-wall turbulent boundary layers using a PIV. In the case of the smooth-wall turbulent boundary layers, the reattachment length of the recirculation bubble on the FFS increased linearly with Reynolds number up to 6380 beyond which no further change in the reattachment length was observed. Meanwhile, the reattachment length for the rough-wall turbulent boundary layers decreased initially with Reynolds number and attained Re independence at $Re = 4000$.

Regarding the effects of upstream roughness on the reattachment length, results presented by [6] showed that, at similar Reynolds numbers, upstream wall roughness can cause up to 40% reduction in the reattachment length on top of a FFS compared to values obtained for upstream smooth-wall cases.

More recent PIV experiments [7, 8] also reported a reduction in reattachment length over a FFS due to upstream wall roughness except that the level of reduction varies from one study to the other. For a BFS, [1] reported that the reattachment length behind the step increased from $5.8h$ to $6.2h$ when the downstream smooth wall was replaced by rough walls. Altogether, these results clearly demonstrate that the characteristics of turbulent flow separation are acutely sensitive to initial and boundary conditions.

In spite of their geometrical simplicity, turbulent flow separation induced by two-dimensional surface-mounted bluff bodies of finite streamwise extent (i.e., those with streamwise aspect ratio, L/h , much smaller than for FFS or BFS) can exhibit remarkable topological and dynamical complexity. This type of bluff bodies are often referred to as forward-backward-facing steps. Recently, [9] applied a time-resolved particle image velocimetry (TR-PIV) to examine the time-averaged turbulent velocities and unsteady characteristics of separation bubbles caused by a forward-backward-facing step with $L/h = 2.36$ that was submerged in a thick turbulent boundary layer ($\delta/h = 4.8$). They found that even though the mean flow reattached on top of the step at $1.6h$ from the leading edge, the separation bubble reattached on the step on intermittently. Furthermore, the close proximity of the separation bubbles on top of and behind the step gave rise to a strong interaction between the bubbles, and flow dynamics that is more complex than observed in isolated recirculation bubble over FFS or BFS. In spite of their interesting flow features, only two experimental studies [10, 11] have been performed thus far to examine how streamwise aspect ratio of forward-backward-facing steps affects the dynamics of the recirculation bubbles on top of and behind the step. In the first experiments [10], L/h was varied from 1 to 10, and the turbulence intensity and thickness of the approach boundary layer were respectively 0.5% and $\delta/h = 0.48$. The velocity measurements were performed using a hot wire anemometry and only reattachment length, streamwise mean velocity and turbulence intensity were reported. The second study [11] reported detailed PIV measurements over and behind steps with $0.1 \leq L/h \leq 8$, and in these experiments, $\delta/h = 1.38$ and the turbulence intensity at $y = h$ was 4%.

It is clear from the above discussion that the effects of aspect ratio on the characteristics of turbulent flow over and behind forward-backward-facing steps immersed in a thick turbulent boundary layer are still not well understood. Therefore, the objective of this study is to investigate how streamwise aspect ratio affects the mean velocity and Reynolds stresses in turbulent flows over and behind forward-backward-facing steps submerged in a turbulent boundary layer whose thickness is significantly larger than the step height and turbulence levels that exceed those investigated in previous experiments [10, 11].

II. EXPERIMENTAL PROCEDURE

A. Test facility

The experiments were conducted in a recirculating open water channel. The test section of the channel is 6000 mm long and

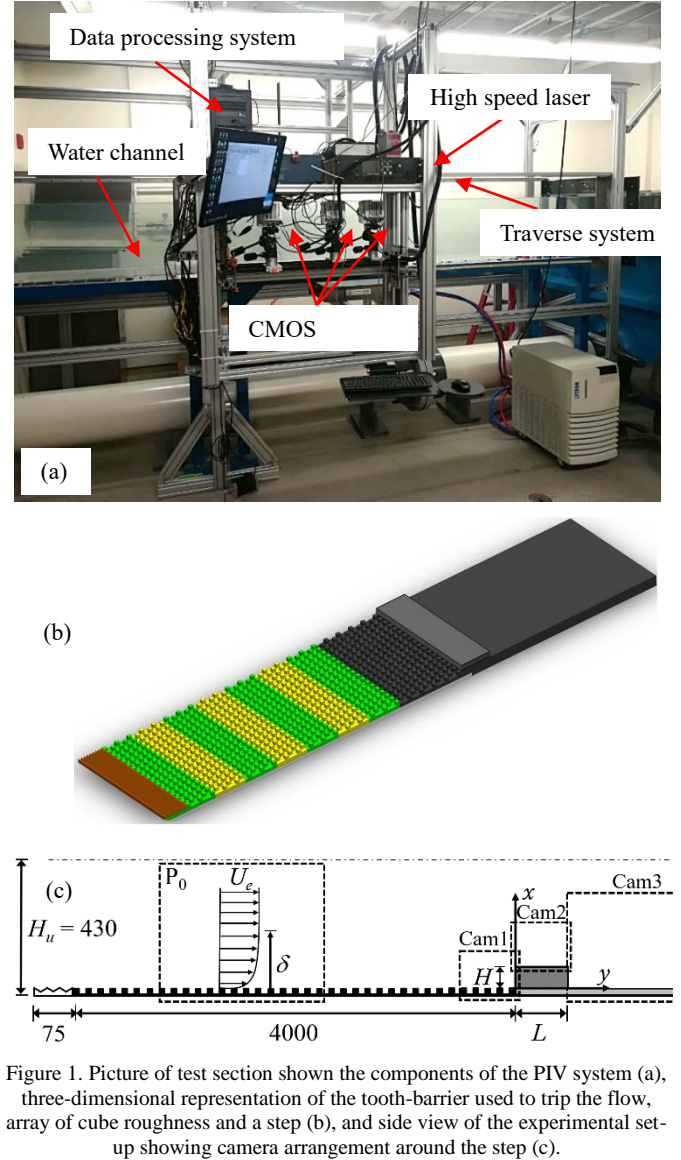


Figure 1. Picture of test section shown the components of the PIV system (a), three-dimensional representation of the tooth-barrier used to trip the flow, array of cube roughness and a step (b), and side view of the experimental set-up showing camera arrangement around the step (c).

its rectangular cross section is 450 mm deep and 600 mm wide. The two side walls and bottom wall of the test section were manufactured with 31.8 mm thick Super Abrasion Resistant transparent acrylic plates to facilitate optical access to the flow. Figure 1a shows a picture of a portion of the test section and the PIV system used in the present study. To generate the desired thick turbulent boundary layer upstream of the steps, a combination of two-dimensional toothed barrier and staggered array of cubes of height, 3 mm was used. The toothed barriers had a height of 15 mm, with triangular cut-outs at the top of pitch 15 mm and depth 12 mm. The cubes were machined from a 6 mm acrylic plate, and had centre-to-centre spacing of 6 mm in both the streamwise and spanwise directions. A three-dimensional representation of the toothed barriers, arrays of cubes and a typical step used to induce flow separation is shown in Figure 1b while a side view of the experimental set-up is

shown in Figure 1c. As shown in Figure 1c, a Cartesian coordinate system is adopted in the present study with the origin of the streamwise coordinate, x , chosen at the leading edge of the step and the origin of vertical coordinate, y , set at the bottom of the step.

A total of 8 test cases were investigated. In all cases, two-dimensional steps of fixed height $h = 30$ mm and spanwise dimension of 600 mm were used but the streamwise extents were varied to the following aspect ratios: $L/h = 1, 2, 3, 4, 5, 6, 8$, and 10 were examined. The steps were screwed onto a 6 mm thick acrylic plate which was then screwed onto the bottom floor of the channel. In order to minimize reflection of the laser, both the steps and the preceding cubed plates were painted with non-reflective black paint.

B. Measurement procedure

Detailed velocity measurements were performed in the streamwise-vertical (x - y) plane using a time-resolved particle image velocimetry (TR-PIV) technique. The PIV system consists of a high speed double-pulsed (Nd:YLF) laser which is capable of emitting green light up to a maximum energy of 30 mJ/pulse at a frequency of 1000 Hz, high speed complementary metal-oxide semiconductor (CMOS) cameras and data acquisition and processing system, all supplied by LaVision. The full resolution of the CMOS cameras was 2560 pixel \times 1600 pixel at an acquisition frequency up to 807 Hz. The water depth and freestream velocity were set to 415 mm and 0.441 m/s, respectively. Thus, the Reynolds number based on the freestream velocity and step height was 13200. The flow was seeded using 10 μ m silver coated hollow glass spheres which had specific gravity of 1.4. The slip velocity, relaxation time and Stokes number of the seeding particles were respectively 2.18×10^{-5} m/s, 2.2×10^{-6} s and 0.0014. Based on these values, it was concluded that these seeding particles follow the flow faithfully and their instantaneous velocity is an accurate representation of the local fluid velocity.

First, two sets of measurements were performed at the mid-span of the channel to characterize the state of the approach turbulent boundary layer. These measurements were performed with the steps removed from the test section. For the first and second sets of measurements, the fields of view were set to 170 mm \times 155 mm and 335 mm \times 305 mm, respectively, and was centered approximately ten step heights upstream of the step, i.e., $x/h = -10$. For the flow field around the step, a total of three CMOS cameras were used to simultaneously capture the region upstream of the step (Cam1), the recirculation region above the step (Cam2), and the region behind the step (Cam 3), as shown in Figure 1c. For these measurements, different lenses (SIGMA DG

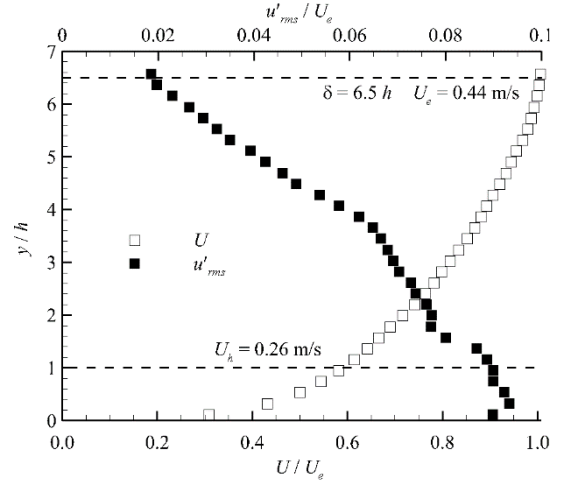


Figure 2. Vertical profiles of streamwise mean velocity and turbulence intensity in the approach turbulent boundary layer

MACRO 105mm f 2.8 D for Cam 1 and Cam 2, and AF MICRO NIKKOR 60mm f 2.8 D lens for Cam 3) were used to achieve fields of view that are appropriate to adequately resolve the flow field in different regions around the step. The sampling frequency was set to 807 Hz and up to 48000 image pairs were obtained in each of the fields of view and for all test cases. The velocity vectors were calculated from the images using a multi-pass cross-correlation algorithm. More specifically, the interrogation area (IA) was initialized as 128 pixel \times 128 pixel with 50% overlap, and finalized as 32 pixel \times 32 pixel with 75% overlap.

Figure 2 shows vertical profiles of the streamwise mean velocity and turbulence intensity in the approach turbulent boundary layer. In this figure and subsequent plots, the freestream velocity and step height are adopted, respectively, as the characteristic velocity and length scales. The boundary layer thickness, defined as the vertical distance from the wall at which the streamwise mean velocity increased to $0.99U_e$ is 195 mm, which corresponds to $6.5h$. The streamwise mean velocity and turbulence intensity at the step height ($y = h$) are respectively 0.26 m/s (approximately $0.59U_e$) and 10%. Meanwhile, the boundary layer displacement and momentum thicknesses are 39.7 mm and 25.7 mm, respectively, and the Reynolds number based on the freestream velocity and momentum thickness is $Re_\theta = 11300$.

III. RESULTS AND DISCUSSION

A. Mean velocity and reattachment length

The effects of aspect ratio on the mean flow are examined using the streamwise mean velocity and the mean reattachment length. Although measurements were obtained for 8 different aspect ratios, plots are presented for only $L/h = 1, 2$ and 5 due to space limitation. Contour plots of the mean velocity are shown in Figure 3 to facilitate whole-field visualization of the mean flow topology. Also shown in these plots are the mean

streamlines and the isopleths of 50% forward-flow fraction, which are denoted by red solid lines on the contour plots.

Irrespective of L/h , the mean flow first separated upstream of the step and reattached on the vertical face of the step. Next, the flow separated at the leading edge and either reattached onto the bottom wall of the test channel behind the step ($L/h = 1$) or onto the top surface of the step ($L/h = 2$ and 5). In the latter cases, there is a third separation at the trailing edge and a subsequent reattachment onto the bottom wall of the test channel behind the step. Thus, the flow over steps of varying L/h can be sorted into two distinct cases based on the topology of the streamwise mean velocity: (i) ‘short step’, which represents cases for which the step is not long enough for the mean flow to reattach on top of the step; and (ii) ‘long step’ which describes situations for which the streamwise extent of the step is long enough for the mean flow to reattach on top of the step. For the particular approach turbulent boundary layer condition examined herein, only $L/h = 1$ belongs to the ‘short step’ category and all other aspect ratios (i.e., $L/h \geq 2$) fall into the ‘long step’ category. It is also evident from Figure 3 that the recirculation bubble behind the step for $L/h = 1$ is significantly larger than those for $L/h \geq 2$. In the case of $L/h = 2$, the recirculation bubble on top of the step is fairly close to the recirculation bubble behind the step. This close proximity would lead to an intense interaction between these two recirculation bubbles. As the aspect ratio increases ($L/h = 5$), the distance between the two recirculation bubbles increased, thus weakening any potential interaction between the two recirculation bubbles.

The reattachment lengths of the recirculation bubbles on top of and behind the steps are used to examine the effects of aspect ratio on the size of the recirculation bubbles, and to compare the present results with those obtained in prior experiments. For the recirculation bubble over the step, the reattachment length, L_T , is defined as the distance from the leading edge of the step to the streamwise location where the mean dividing streamline reattached on top of the step. For the recirculation bubble behind the step, the recirculation length, L_B , is defined as the distance from the trailing edge to the streamwise location where the mean dividing streamline reattached onto the bottom wall of the channel. Alternatively, the isopleth of 50% forward-flow fraction can be used to estimate the reattachment lengths. Evidently, these isopleths which are shown as red solid line in the contour plots, exhibit a pattern that originates from the leading edge (or trailing edge) of the step, passes through the centroid of the recirculation bubble and terminates on top of the step (or the bottom wall of the channel behind the step). The differences between the reattachment points determined from the mean streamline and the 50% forward-flow fraction are within measurement uncertainty. For this reason, the reported reattachment lengths in this paper are those calculated from the mean dividing streamlines.

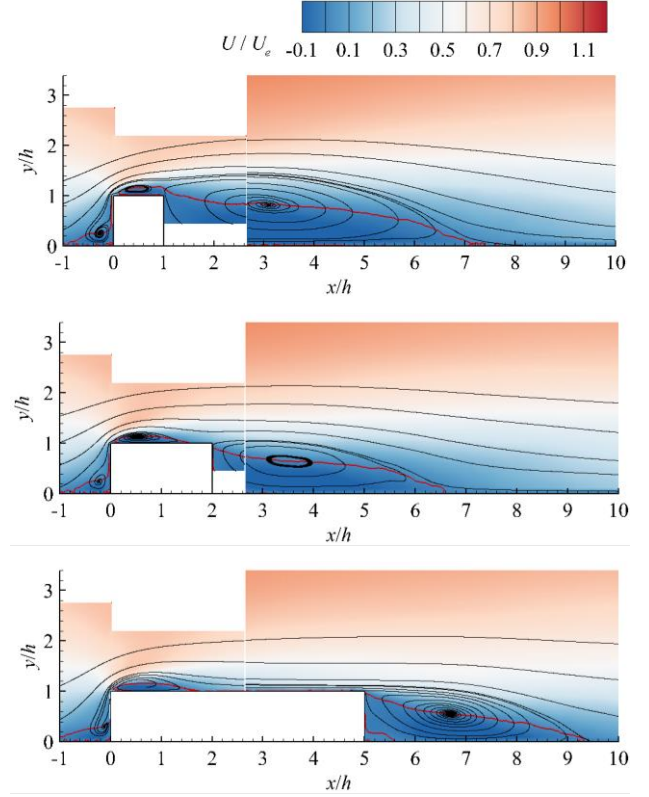


Figure 3. Contours of streamwise mean velocity with mean streamlines for $L/h = 1$ (top), 2 (middle) and 5 (bottom). The red isopleth is forward velocity fraction at 50%.

Figure 4 shows the reattachment length on top of the step (Figure 4a) and behind the step (Figure 4b) as a function of aspect ratio. The results from [10] are also shown for comparison. In Figure 4a, data are not presented for $L/h = 1$ for the present study and $L/h < 5$ for [10] because the mean flow did not reattach on the step for those aspect ratios. Figure 4a shows that reattachment length on top of the step is independent of aspect ratio, but the present values are only 42% of those reported in [10]. These significant differences among the reattachment lengths are explained by the nature of the approach turbulent boundary layer. It should be recalled that the relative boundary layer thickness are $\delta h = 6.5$ and 0.48, respectively, in the present study and [10]. Meanwhile, the turbulence intensity at $y = h$ in the present is 10% which is two orders of magnitude higher than in [5]. It has been shown that the effects of an enhanced turbulence intensity and larger relative boundary layer thickness are to shorten the reattachment length [6]. Thus, the present results are consistent with literature. Behind the step (Figure 4b), the reattachment length of the recirculation bubble decreased initially with aspect ratio but become independent of aspect ratio for $L/h > 2$ in the present study and for $L/h \geq 5$ in [10].

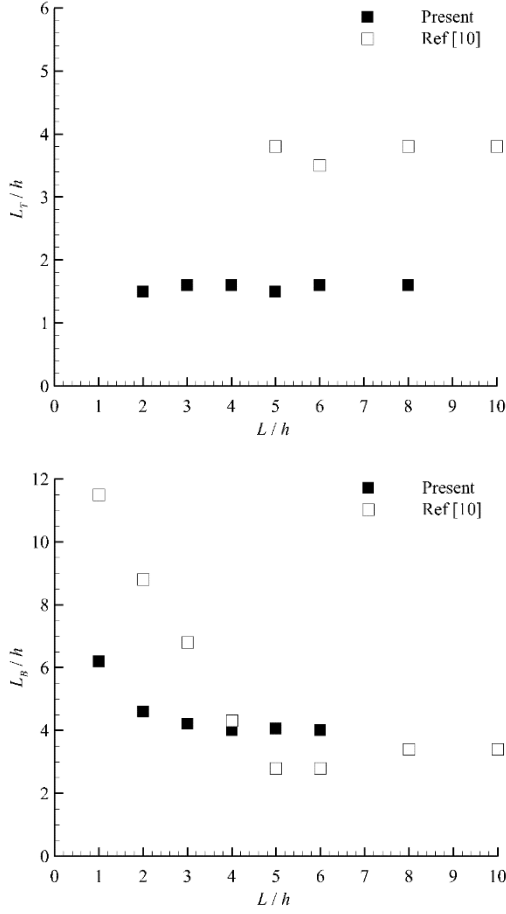


Figure 4. Variation of reattachment lengths on top of the step (top) and behind the step (bottom) with aspect ratio.

B. Reynolds stresses and eddy viscosity

Contours of the Reynolds normal stresses are shown in Figures 5 and 6. In both cases, very high turbulent levels are apparent near the leading edge of the step and decrease in the downstream direction as the shear layer spreads and turbulence diffuse into the recirculation region and the outer shear layer. The presence of high turbulence levels in the vicinity of the leading edge of the steps is remarkably different from flow separations with approach laminar or low turbulent intensity. In particular, for flow separation over a blunt body placed in a uniform flow, transition to turbulence occurs in the rear part of the separation bubble [12]. While the turbulence levels over the step are nearly independent of aspect ratio, there is a notable reduction in the magnitude of the Reynolds normal stresses behind the step for a higher aspect ratio. For the vertical Reynolds normal stress (Figure 6), there exist an additional region of elevated turbulence level above the dividing streamline. This patch of high turbulence level, which is likely advected from the approach boundary layer, has not been reported in previous studies over surface mounted steps.

The topology of the Reynolds shear stress is qualitatively similar to that of the vertical Reynolds normal stress except for

a narrow region along the dividing streamline near the leading edge where the Reynolds shear stress is negative. Previous experiments [6, 9] and numerical simulation [13] of turbulent flow over FFS also reported negative Reynolds shear stress in the vicinity of the leading edge.

In simple turbulent shear flows including the zero pressure gradient turbulent boundary layer and turbulent free jets, the Reynolds shear stress and mean strain rate are of the same sign so that the kinematic eddy viscosity which is defined as $\nu_t = -\overline{u'v'} / (\partial U / \partial y + \partial V / \partial x)$ is always positive. In view of the peculiar distribution of the Reynolds shear stress observed in Figure 7 (i.e., regions of both positive and negative values), the kinematic eddy viscosity was calculated and plotted in Figure 8. While the eddy viscosity is positive over most of the flow domain, regions of negatively-valued eddy viscosity are observed above the step and also in the front face of the step. Furthermore, the magnitude of the negative eddy viscosities are similar to the positive values. The occurrence of regions of negative eddy viscosity implies a counter-gradient diffusion phenomenon exists [6, 13]. Therefore, turbulence models that are based on the Boussinesq eddy-viscosity assumption will not be able to reliably predict the complex shear flow produced by a forward-backward-facing step submerged in a thick turbulent boundary layer.

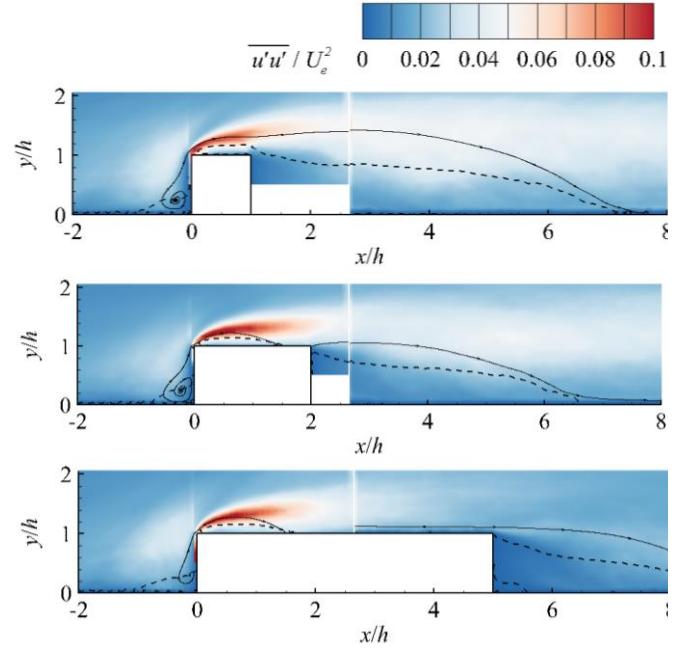


Figure 5. Contours of streamwise Reynolds normal stress, $\langle u'u' \rangle$, for $L/h = 1$ (top), 2 (middle) and 5 (bottom).

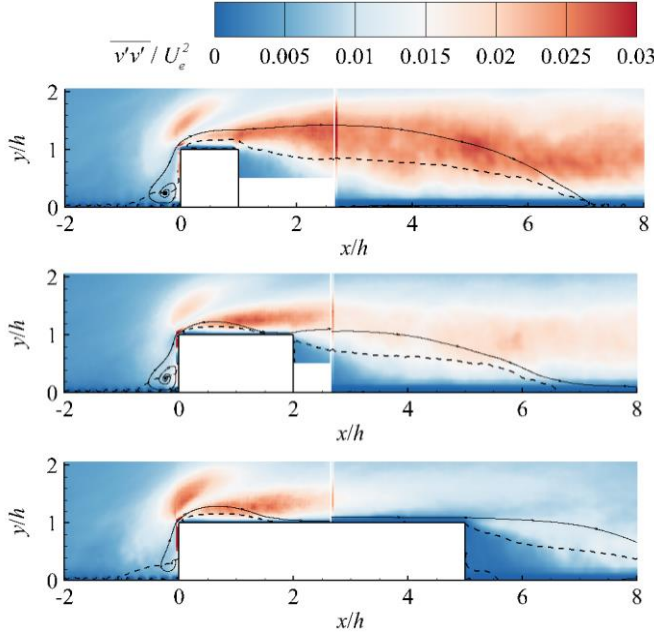


Figure 6. Contours of vertical Reynolds normal stress, $\langle v'v' \rangle$, for $L/h = 1$ (top), 2 (middle) and 5 (bottom).

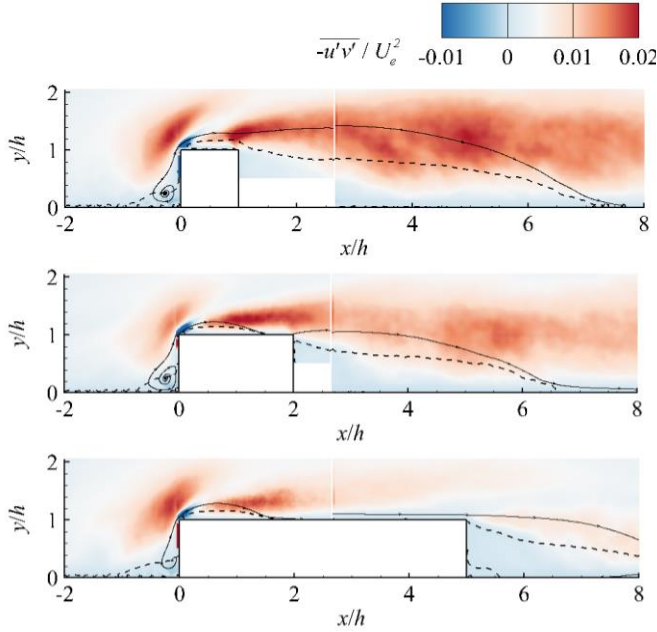


Figure 7. Contours of Reynolds shear stress, $\langle -u'v' \rangle$, for $L/h = 1$ (top), 2 (middle) and 5 (bottom).

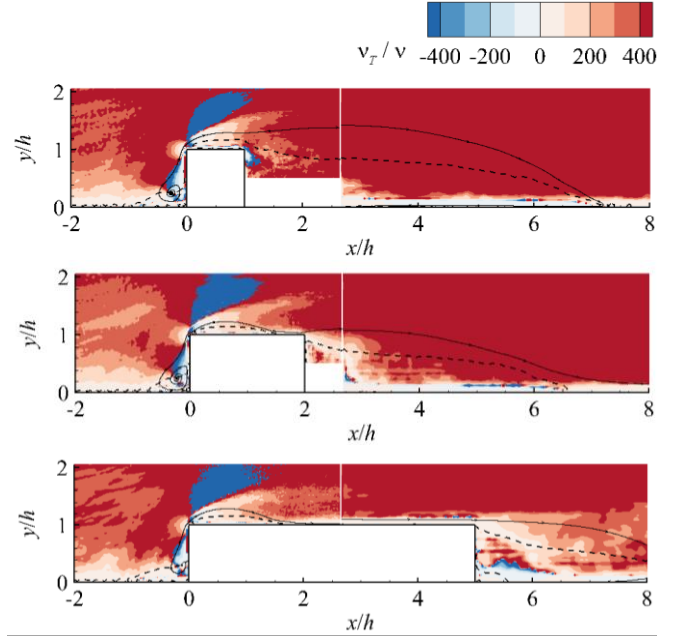


Figure 8. Contours of kinematic eddy viscosity for $L/h = 1$ (top), 2 (middle) and 5 (bottom).

C. Vertical profiles of mean velocity and Reynolds stresses

In this section, profiles of the streamwise mean velocity and Reynolds stresses at three selected streamwise locations (two on top of the steps and one behind the steps) are used to investigate the effects of aspect ratio on the mean flow and turbulent characteristics of the separated and reattached shear layers. On top of the steps, profiles were extracted $0.5h$ downstream of the leading edge and $0.1h$ upstream of the trailing edge, while those behind the steps were extracted at $2h$ downstream of the trailing edge of the step. Profiles of the mean velocity and Reynolds stresses are shown in profiles are shown in Figure 9 and Figure 10, respectively. The mean velocity profiles close to the leading edge are nearly independent of aspect ratio. Since this location is within the recirculation region for all the aspect ratios examined, the profiles show the characteristics negative velocity in the region immediately above the step. The magnitude of the maximum back flow at this location is approximately 20% of the freestream velocity for all three aspect ratios. Similarly the vertical location at which the mean velocity changes from negative to positive is also similar for the three aspect ratio. The similarity among the profiles close to the leading edge disappears close to the trailing edge and behind the step. Close to the trailing edge, only profiles for $L/h = 1$ show a region of flow reversal. This is because the separated shear layer from the leading edge did not reattach on top of this particular step while those for the larger aspect ratios reattached prior to the trailing edge and therefore do not show regions of flow reversal. Since the reattachment length for both $L/h = 2$ and 5 are the same, the profile over the longer step ($L/h = 2$) has a much longer redevelopment length on top of the step. For this reason, the mean velocity profile for $L/h = 5$ is more uniform than observed for $L/h = 2$. Behind the step, the region of flow reversal is much larger for a smaller

aspect ratio even though the maximum back flow is nearly the same for the three aspect ratios. The profiles near the leading edge and the maximum backflow are akin to those reported on a FFS at similar locations. There is also a close resemblance between profiles behind the step and those obtained in BFS as similar streamwise location.

The distributions of the streamwise Reynolds normal stress are also similar among the three aspect ratios close to the leading edge but those profiles close to the trailing edge and behind the step are quite sensitive to aspect ratio. Near the leading edge, profiles of the vertical Reynolds normal stress and Reynolds shear stress show characteristic double peaks. While the peaks close to the wall are produced by the intense shear along the dividing streamline, the outer peaks are likely an artefact of the high turbulence within the approach turbulent boundary layer. The outer peaks in both the vertical Reynolds normal stress and Reynolds shear stress have not been reported in previous turbulent flows over FFS or forward-backward-facing steps. Note, however, that the double peaks in the vertical Reynolds normal stress and Reynolds shear stress disappeared prior to the trailing edge of the step.

IV. SUMMARY AND ONCLUSION

A time-resolved PIV system was used to investigate the effects of aspect ratio on turbulent flow over and behind forward-backward-facing steps immersed in a turbulent boundary layer of thickness 6.5 step heights. Eight different streamwise aspect ratios, ranging in value from 1-10, were tested at a fixed Reynolds number of 13200. The reattachment length was used to examine the effects of aspect ratio on the size of the recirculation bubbles while contour plots and one-dimensional profiles of the streamwise mean velocities and Reynolds stresses are used to assess the effects of aspect ratio on the mean flow and turbulent characteristics. The results show that an aspect ratio of two is long enough for the separated flow to reattach on top of the step. Furthermore, the reattachment length on top of the step was measured as $1.6h$, and remained invariant with respect to aspect ratio. The reattachment length of the recirculation bubble behind the step decreases from $6.3h$ for an aspect ratio of 1 and approaches an asymptotic value of $4.2h$ for aspect ratios of 3 and larger.

High levels of Reynolds normal and shear stresses are observed along the dividing streamline near the leading edge of the step. The vertical Reynolds normal and shear stress show two unique features near the leading edge: first, distributions of these stresses exhibit double peaks with the magnitude of the outer peaks similar to the inner peaks that are formed along the mean dividing streamline; secondly, the Reynolds shear stress showed a region of intense negative values. One of the immediate implications of this unique feature of the Reynolds shear stress is the occurrence of negative eddy viscosity or counter-gradient momentum transport. As a consequence, turbulence models that are based on the overly simplistic Boussinesq eddy-viscosity assumption will not be able to adequately predict the separated shear layer on top of a two-

dimensional step submerged in a very thick approach turbulent boundary layer.

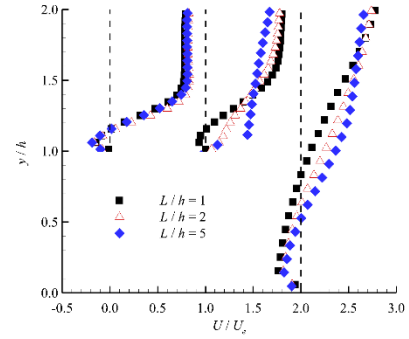


Figure 9. Vertical profiles of streamwise mean velocity at $0.5h$ downstream the leading edge (left), $0.1h$ upstream the trailing edge (middle) and $2h$ downstream the trailing edge (right) for $L/h = 1, 2$ and 5 .

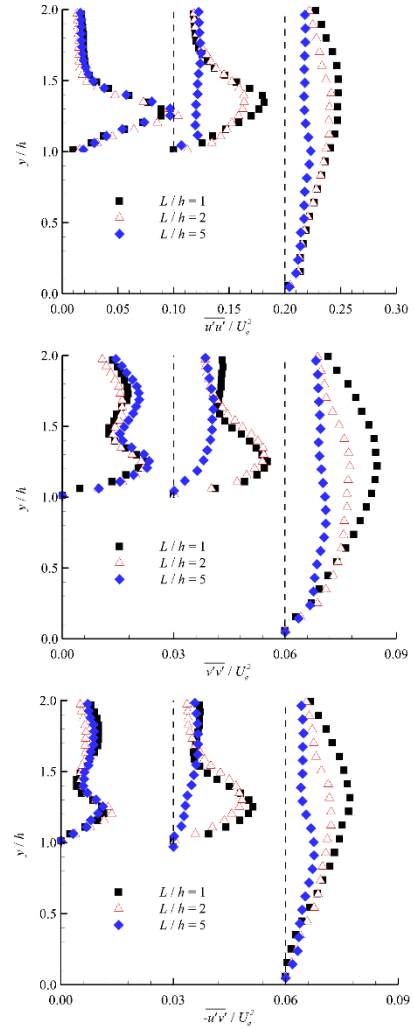


Figure 10. Vertical profiles of streamwise Reynolds normal stress (top), vertical Reynolds normal stress (middle) and Reynolds shear stress (bottom) for $L/h = 1, 2$ and 5 . In each plot, profiles are shown at $0.5h$ downstream the leading edge (left), $0.1h$ upstream the trailing edge (middle) and $2h$ downstream the trailing edge (right).

The distributions of the streamwise mean velocity and Reynolds stresses are nearly independent of aspect ratio. As the aspect ratio increases, the development length of the reattached flow over the step also increases and the mean velocity profile becomes more uniform and similar to canonical turbulent boundary layers. For this reason, the mean shear and consequently the Reynolds stresses close to the trailing edge of the step weaken considerably for a larger aspect ratio. Since the flow at the trailing edge serves as initial or upstream condition for flow separation at the trailing edge of the step, the differences observed among the mean velocity and Reynolds stresses at the trailing edge of the various aspect ratios are also manifested in the shear layer behind the step.

REFERENCES

- [1] Essel, E.E, Tachie, M.F., "Roughness effects on turbulent flow downstream of a backward facing step," *Flow Turbul Combust.* 94:125–153, 2015.
- [2] Wu, Y., Ren, H., Tang, H.: Turbulent flow over a rough backward-facing step. *Int. J. Heat Fluid Flow.* 44, 155–169 (2013)..
- [3] Jovic, S., Driver, D.: Reynolds number effect on the skin friction in separated flows behind a backward-facing step. *Exp. Fluids.* 18, 464–467 (1995).
- [4] Largeau, J., and Moriniere, V., "Wall pressure fluctuations and topology in separated flows over a forward facing step," *Exp. Fluids*, vol. 42, pp. 21–40, 2007.
- [5] Sherry, M., Lo Jacono, D., and Sheridan, J., "An experimental investigation of the recirculation zone formed downstream of a forward facing step," *J. Wind Eng. Ind. Aerod.* vol. 98, pp. 888–894, 2010.
- [6] Essel, E.E., Nematollahi, A., Thacher, E.W., and Tachie, M.F., "Effects of upstream roughness and Reynolds number on separated and reattached turbulent flow," *J. Turbul.*, vol. 16, pp. 872–899, 2015.
- [7] W. Shao and M. Agelin-Chaab, "Turbulent flows over forward facing steps with surface roughness," *ASME J. Fluids Eng.* **138**, 021103 (2016). Shao & Martin
- [8] Fang, X., and Tachie, M.F., 2019, 'On the unsteady characteristics of turbulence separations over a forward-backward-facing step'. *Journal of Fluid Mechanics*, vol. 863, pp. 994–1030.
- [9] Nematollahi, A., and Tachie, M.F., AIP, Nematollahi, A. & Tachie, M. F. 2018 Time-resolved PIV measurement of influence of upstream roughness on separated and reattached turbulent flows over a forward-facing step. *AIP Adv.* 8, 105110.
- [10] Bergeles, G., Athanassiadis, N., "The flow past a surface-mounted obstacle," *J. Fluid Eng.* 105 (4), 461–463, 1983.
- [11] Van der Kindere, J., and Ganapathisubramani, B., "Effect of length of two-dimensional obstacles on characteristics of separation and reattachment," *J. Wind Eng. Ind. Aerod.*, vol. 178, pp. 38–48, 2018.
- [12] ALAM, M. & SANDHAM, N. D. 2000 Direct numerical simulation of 'short' laminar separation bubbles with turbulent reattachment. *J. Fluid Mech.* 410, 1–28
- [13] H. Hattori and Y. Nagano, Investigation of turbulent boundary layer over forward-facing step via direct numerical simulation, *Int. J. Heat Fluid Flow.* 31 (2010), pp. 284–294

EXPERIMENTAL MEASUREMENTS OF COUGH AIRFLOW IN THE FAR-FIELD

N. Dudalski, A. Mohamed, E. Savory
Dept. Mechanical & Materials Engineering
The University of Western Ontario
London, Canada

S. Mubareka
Dept. Biological Sciences
Sunnybrook Research Institute
Toronto, Canada

Abstract— Seasonal influenza epidemics have been responsible for causing increased economic expenditures and many deaths worldwide. Particle image velocimetry (PIV) and hot-wire anemometry (HWA) measurements were conducted at 1 m away from the mouth of human subjects to develop a model for cough flow behaviour at greater distances from the mouth than were studied previously. Biological aerosol sampling was conducted to assess the risk of exposure to airborne viruses. Throughout the investigation 77 experiments were conducted from 58 different subjects. From these subjects, 21 presented with influenza like illness. Of these, 12 subjects had laboratory confirmed respiratory infections. A model was developed for the cough centreline velocity magnitude time history. The experimental results were also used to validate computational fluid dynamics (CFD) models based on the Unsteady Reynolds Averaged Navier-Stokes (URANS) method and large eddy simulation (LES). The peak velocity observed at the cough jet centre, averaged across all trials was 1.17 m/s, and an average spread angle of $\theta=24^\circ$ was measured. No differences were observed in the velocity or turbulence characteristics between coughs from sick or healthy participants.

Keywords—cough; airflow; Virus transmission; Aerosol sampling; particle image velocimetry; hot-wire anemometry

I. INTRODUCTION

Communicable respiratory diseases have demonstrated the potential to cause global pandemics, which has resulted in increased economic expenditures, and many deaths worldwide [1,2]. The significance of the airborne route of virus transmission has been a topic of recent interest in the healthcare community.

Respiratory activities, such as coughing, sneezing, breathing and talking, generate and disperse pathogen bearing aerosols [3]. Recent reviews supported the claim that droplet nuclei smaller than 5 μm behave much like a gas and are capable of remaining suspended within the air for long periods of time [4,5].

There is a widespread adoption of the “3 ft/1 m” and 6 ft/2 m rule” [6,7], which has considered such separation distances

from patients infected with respiratory viruses to be safe, without any evidence to support the claim.

The objective of the present investigation is to rigorously test the “3 ft/1 m rule,” by conducting velocity measurements and bioaerosol sampling at 1.0 m from the mouth of human subjects. Based on an average mouth opening diameter of $D=0.02$ cm [8], this region is calculated to be approximately $x=50D$. Experiments were conducted to map the flow field of human coughs at greater distances than those studied previously and to develop a model for transient cough jet behaviour. Subjects who have been naturally infected with influenza participated in experiments while they were sick and again when they had recovered. A cohort of healthy volunteers was recruited as a reference to assess any difference in the aerodynamic behaviour of coughs from the two groups. The velocity measurements were also used to validate computational fluid dynamics (CFD) models based on unsteady Reynolds-averaged Navier-Stokes (URANS) and large eddy simulation (LES) methods developed by Bi [9].

II. METHODOLOGY

A. Experimental facility and procedure

The experimental chamber [10], consisted of a 1.81 m x 1.81 m x 1.78 m enclosure, with all interior surfaces painted black, except for a glass window allowing optical access (Figure 1). The dimensions of the chamber were selected so that the cough airflow was not noticeably influenced by the chamber walls. A pear-shaped opening, with a padded head rest and chin rest, fixed the participant’s head in place, but allowed the participant to cough into the chamber with their nose and mouth unobstructed. Subjects were asked to cough 3 times each for individual trials in separate particle image velocimetry (PIV) and hot-wire anemometry (HWA) experiments with aerosol sampling. Participants waited for approximately 2 min between coughs to minimize the residual air motion within the chamber.

B. Particle image velocimetry (PIV)

PIV has commonly been used as a non-intrusive, technique to measure airflow fields due to its accuracy and spatial resolution. For PIV experiments, the chamber was seeded with aerosolized Titanium Dioxide (TiO_2) particles (ranging between 0.15 to 0.47 μm with 69% of particles between 0.34-

0.43 μm) that were dried in a vacuum oven. The particles were stored in a drum which was placed on top of a loudspeaker that vibrated to aerosolize the powder. The seeding particles were carried into the chamber from the drum by a 30 kPa air-line. Particles were illuminated by a 120 mJ per pulse, 532 nm Nd:YAG laser, which was directed through an angled mirror and a cylindrical/spherical lens assembly to create a laser sheet (~ 1 mm thickness). Each pulse has a duration of 3-5 ns, and synchronized image pairs were recorded by a CCD camera at a frequency of 15 Hz (2018 experiments) and 16.7 Hz (2014 experiments). An image separation time of $\Delta t = 100$ μs was used, and laser pulse delay and PIV exposure times of 400 μs and 405 μs were selected, respectively.

In the experiments conducted in 2014, much of the cough flow missed the small field of view in some trials, and many coughs missed completely. To contain the entire width of the cough within the imaged window and due to a necessary change in the camera used, a new field of view was selected for the experiments conducted in 2018 (Figure 2). Both windows were centred at $x = 1$ m downstream from the participant and they were translated downward to compensate for the typically downward initial cough trajectory [8]. A recursive Nyquist cross-correlation algorithm was used, in TSI Insight 4G software, to process images into velocity vector arrays. A final spot dimension of 32×32 pixels was used for the 2014 images, whereas it was increased to 64×64 pixels to compensate for the decrease in spatial resolution from 10.0 pixels/mm to 6.87 pixels/mm. Since the camera is farther from the laser sheet, increasing the field of view, a larger final spot dimension is necessary to ensure the proportion of validated vectors is over 85%. Eighty image pairs are recorded for each cough, allowing enough time for the cough to pass the field of view completely. The time at which the cough is first visible in the field of view was set to be $t = 0.0$ s in all further data analysis.

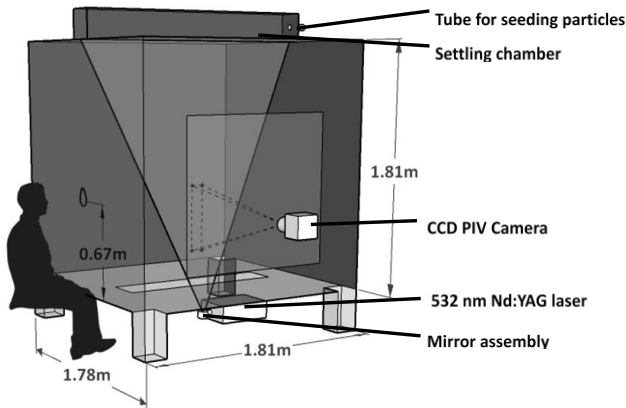


Figure 1. Schematic diagram of experimental facility

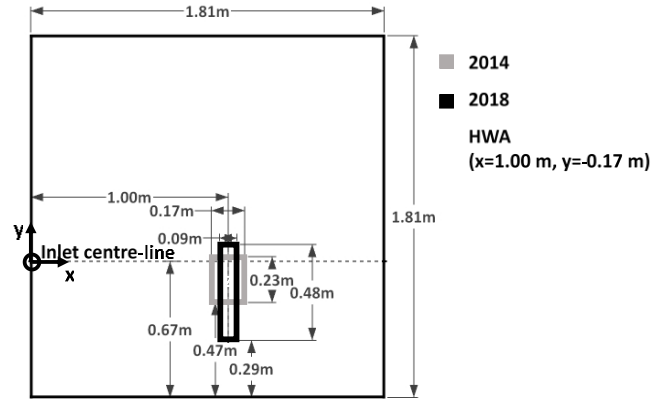


Figure 2. PIV field of view on cough chamber centre plane

Time histories for the 2D velocity magnitude, V , were computed according to Equation 1.

$$V = \sqrt{u^2 + v^2} \quad (1)$$

where: u is the streamwise velocity component in the x -direction, v is the vertical component in the y -direction

The time histories were extracted at the cough jet centre as well as at the chamber centre-line and the HWA location, $x = 1$ m away from the subject. The cough jet centre was defined as the mid-point of the cough, where the greatest velocities are present, after examining the velocity contours and vector arrays (typical cough shown in Figure 3). If no clear jet centre was evident, the cough was not used for analysis. Coughs were excluded from the analysis if, at each extraction location, the peak velocity was below 0.20 m/s, because low velocities made it impossible to distinguish the jet boundaries and locations required for this analysis. The peak velocity was calculated after subtracting the residual air motion present within the chamber prior to the arrival of the cough. This quantity was usually measured to be approximately 0.05 m/s.

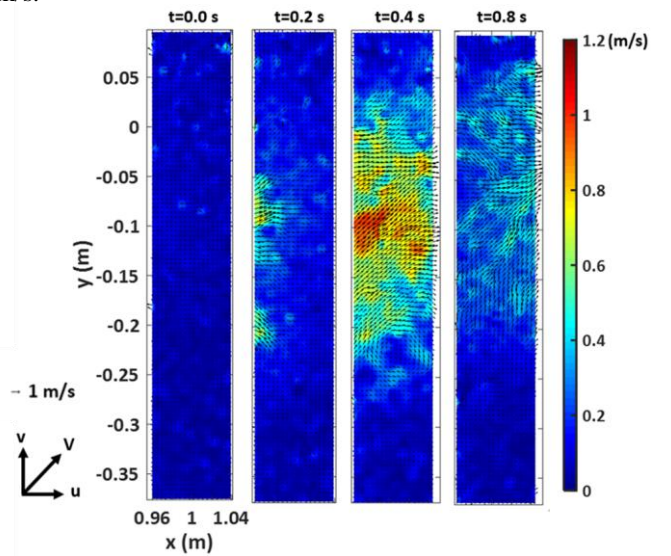


Figure 3. Instantaneous 2D velocity magnitude contour with overlaid vector arrows

Figure 4 displays an example of an instantaneous velocity time history. For further comparison, the cough jet centre velocity magnitude profiles were normalized according to Equation 2. Time was normalized by a similar method according to Equation 3.

$$V_{norm} = \frac{V(t) - V_s}{V_{peak} - V_s} \quad (2)$$

where: V_s is the residual velocity within the chamber at $t=0$, V_{peak} is the maximum velocity at the peak of the cough

$$\tau = \frac{t}{t_{peak}} \quad (3)$$

where: t_{peak} is the time at which V_{peak} is observed.

C. Hot-wire anemometry (HWA)

HWA was used to measure the instantaneous velocity magnitude at single location. A constant temperature anemometry (CTA) unit was attached to a probe containing a tungsten wire 1.25 mm long and 5 μ m in diameter), and voltage readings are recorded at a rate of 1 kHz. The system was calibrated for low airflow speeds using a specialized facility detailed in [11]. The sensor was located $x=1.00$ m downstream from the mouth of the participant, $y=-0.17$ m below the inlet centre-line, since preliminary trials had indicated that the cough jet travelled along a slight downward trajectory. A moving average filter was applied to the measurements to filter out high frequency fluctuations and noise. The size of the averaging window was selected such that further increasing the window size does not substantially alter the root mean square of velocity fluctuations about the moving average. The HWA velocity time histories were normalized by the same method described earlier.

D. Computational fluid dynamics (CFD)

Two separate CFD models have been previously developed, based on LES and URANS methods [9], in which the boundary conditions were selected based on the average volumetric flow rate profiles and mouth opening diameters determined experimentally by Gupta et al. [8]. The inlet velocity was specified to match the velocity shown in Figure 5, and a mouth diameter of $D=0.0217$ m was used for the simulation. The LES data was filtered using the same moving averaged method utilized to smooth the HWA data.

E. Biological sampling (“Sick” participants only)

To assess the presence of viral pathogens within the cough airflow, 2 polytetrafluoroethylene (PTFE) membrane filter cassettes (1 μ m pore size) were attached to low-flow (4000 ± 40 mL/min) sampling pumps (SKC Inc., Airchek 224-PCXR3). The filter cassettes were suspended within the chamber at distances of 0.5 m and 1.0 m from the mouth of the subject, along the chamber centreline. Air sampling was performed concurrently with the HWA measurements and the pumps drew air onto the filters for a period of 15 minutes after the 3 coughs were performed.

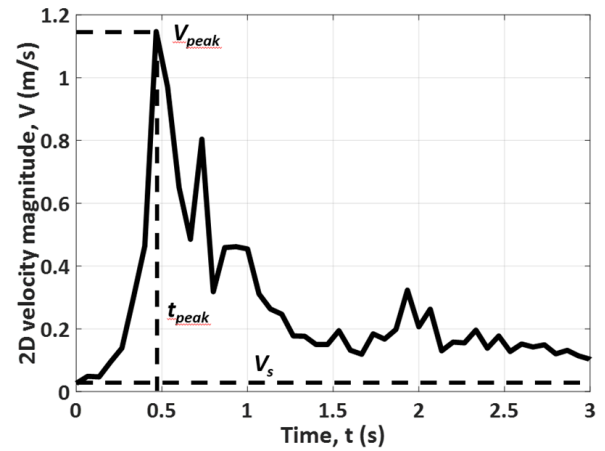


Figure 4. Example instantaneous velocity time history at cough jet centre with key quantities labeled

A mid-turbinate swab specimen was also self-collected immediately prior to the experiments to identify and verify the illness experienced by a subject. The filters and MTS were shaken for 10 s within a vial of UTM viral transport medium by a vortex shaker. The vials were stored at -80°C in a freezer until they were shipped, on dry-ice, for analysis at Sunnybrook Health Sciences Centre and Research Institute: University of Toronto by the Department of Microbiology, Division of Infectious Diseases. The identity of all pathogens was determined by a multiplex polymerase chain reaction (multiplex-PCR) from a panel of respiratory viruses. Biological sampling was not performed during “convalescent” trials, or throughout experiments with the group of “healthy” control subjects.

F. Participant recruitment

The recruitment procedures were approved by Western’s Research Ethics Board (REB approval no. 108945). Subjects with influenza-like illnesses were referred to the study after making an appointment to see a physician at Western Student Health Services. Following referral, subjects completed a questionnaire to determine if they were eligible to participate in experiments. Participants needed to be between 18-35 years of age, and, in the last 24 hours, they should have experienced a fever and cough/sore throat in the absence of any other known cause of illness (e.g. allergies). Anyone who was immunocompromised, had underlying cardiopulmonary conditions, was pregnant, or smoked was excluded from the study. Presumably-ill subjects were recruited during the flu seasons of 2013-14, 2016-17 and 2017-18. Measurements were recorded for “sick” trials, while the subjects are ill, and, when they had recovered, the subjects return for a set of “convalescent” experiments. Healthy control subjects were recruited outside of the flu season.

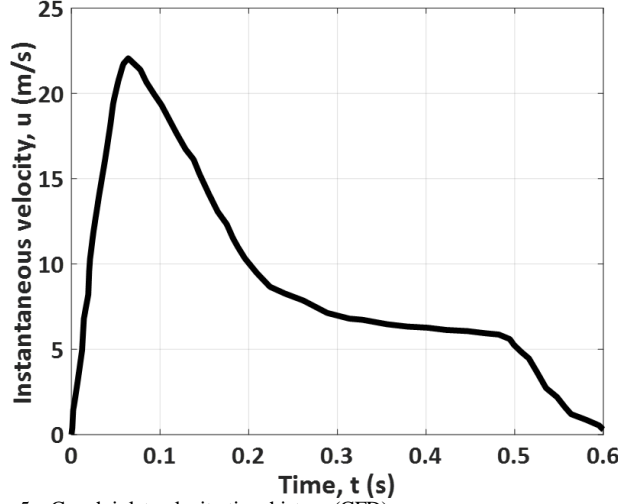


Figure 5. Cough inlet velocity time history (CFD)

III. RESULTS AND DISCUSSION

Throughout the duration of the investigation, 77 sets of experiments were conducted for 58 different subjects. Table 1 summarizes the number of measurements taken for each method. In the 2014 investigation, the HWA probe was only used as a reference, and it was not calibrated. PIV data were not available during the 2017 study period.

A. Biological samples (Filter cassettes and MTS)

Table 2 summarizes the MTS results obtained throughout the investigation. Overall, the percentage of subjects with a confirmed respiratory virus (57%) was very good considering the recruitment methods. A recruitment strategy based on self-referral may have resulted in an increased number of subjects, but it is expected that such a strategy will result in a much greater proportion of negative results. While several investigations have demonstrated the presence of viral particles within the cough aerosols of human subjects [12,13,14], no viral RNA was detected on any of the PTFE filters. It is possible that the virus-containing aerosols were directed past the sampling locations due to the strength and angle of the cough.

TABLE I. TOTAL NUMBER OF SAMPLES ACROSS ALL COHORTS

Date	Cohort	Number of subjects	Coughs sampled (PIV)	Coughs sampled (HWA)	Filter Cassette Samples	MTS Samples
2017-18	Sick	7	21	21	14	7
	Convalescent	7	21	21	-	-
	Healthy	25	75	75	-	-
2016-17	Sick	9	-	27	18	9
	Convalescent	9	-	24	-	-
2013-14	Sick	5	15	-	10	5
	Convalescent	3	9	-	-	-
	Healthy	12	36	-	-	-
Totals		77*	177	171	42	21

*77 sets of experiments from 58 different subjects

TABLE II. MTS RESULTS

Date	Number of subjects	Recovered pathogens			Recruited subjects with verified illness
		Pathogen	Subjects	Total	
2018	7	Influenza B	3	5	71%
		Adenovirus	1		
		Coronavirus NL63	1		
2017	9	Influenza A, H3N2	1	4	44%
		Coronavirus NL63	1		
		Coronavirus OC43	1		
		Respiratory syncytial virus	1		
2014	5	Influenza A, H1N1	1	3	60%
		Coronavirus NL63	1		
		Respiratory syncytial virus	1		
Totals	21	-	-	12	57%

Another important reason is that the total volume of liquid expelled as droplets was very low. Since the sampling method relied on the inertial impaction of aerosolized particles onto the filter surface, only the smallest particles could be collected. With solid impaction techniques, there is also a possibility that a virus could be destroyed upon contact with the membrane surface [15].

B. Cough direction and spread angles

The average location of the cough centre at $x=1.0$ m, was at $y=-0.122$ m (below the cough inlet centerline, or 0.548 m above the chamber floor). The total range of cough exit angles measured was between 3° (above the centre-line) to 15° (below the centreline), and an average downward angle of $\theta_c=7^\circ$ was observed. Previously, cough angles have been defined at the mouth by two downward jet angles: $\theta_1=15 \pm 5^\circ$ and $\theta_2=40 \pm 4^\circ$ (Figure 6) [8]. The average cough exit angle from the literature, determined by the midpoint between θ_1 and θ_2 , was 27.5° below the horizontal, indicating that the head and chin rests used in the investigation adequately reduced this angle so that the cough could be measured at the desired location. The jet spread angle, θ , was calculated from the maximum measured width of the jet at $x=1.0$ m. The upper and lower jet boundaries were selected to be the location at which the velocity fell below 0.05 m/s. An average spread angle of $\theta=24^\circ$ (Min. = 15° , Max. = 27° , Std. dev. = 3°) was measured, which is consistent with the spread angles previously determined [8,16]. Through a critical review and analysis of experimental and numerical investigations, Ball et al determined that steady free jets also exhibited spreading angles between 20 - 35° [17], which was consistent with the findings of the present investigation.

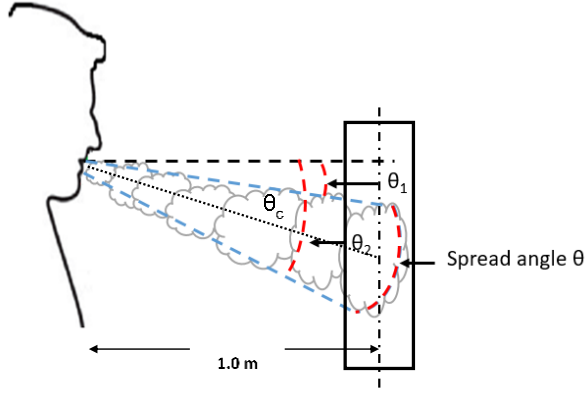


Figure 6. Cough jet angles, $\theta_1=15 \pm 5^\circ$, $\theta_2=40 \pm 4^\circ$, $\theta=25 \pm 9^\circ$, (adapted from [8])

C. Cough velocity

The peak velocity magnitude, V_{peak} , was used to characterize the strength of cough flow behaviour in the far field. A comparison of the strengths of coughs from different cohorts is displayed in Table 3. As expected, the greatest average velocities were noticed at the cough centre, followed by the HWA location. The fewest number of coughs were used for the velocity analysis at the centreline location. Very often, the cough missed this location entirely, or the peak velocity at the location was below 0.20 m/s. Also, the profiles often would not resemble the characteristic cough jets examined here, and only strong fluctuations were observed in the excluded profiles.

TABLE III. AVERAGE PEAK 2D VELOCITY MAGNITUDE POINT EXTRACTIONS (PIV)

Extraction Location	Date	Cohort	Coughs used for analysis	Average V_{peak} (m/s) by cohort	Average V_{peak} (m/s) at location	Coughs used for analysis per location
Apparent cough jet centre (PIV), (x=1.00 m, y=0.12 m) (average location)	2017-18	Sick	13	1.34	1.17	126
		Convalescent	19	1.03		
		Healthy	66	1.19		
	2013-14	Sick	9	0.75		
		Convalescent	4	0.65		
		Healthy	15	0.91		
Hot-wire location (PIV), (x=1.00 m, y=0.17 m)	2017-18	Sick	14	1.03	0.92	138
		Convalescent	20	0.93		
		Healthy	69	1.03		
	2013-14	Sick	10	0.68		
		Convalescent	4	0.59		
		Healthy	21	0.90		
Cough chamber inlet centreline (PIV), (x=1.00 m, y=0.00m)	2017-18	Sick	2	0.83	0.71	53
		Convalescent	5	0.90		
		Healthy	14	0.77		
	2013-14	Sick	11	0.62		
		Convalescent	4	0.47		
		Healthy	17	0.72		
Hot-wire anemometer (HWA), (x=1.00 m, y=0.17 m)	2017-18	Sick	20	0.78	0.82	134
		Convalescent	13	0.86		
		Healthy	68	0.75		
	2016-17	Sick	18	0.88		
		Convalescent	15	1.01		

A t-test with a significance level of $\alpha = 0.05$ did not show a statistically significant difference in the peak cough velocity magnitude, V_{peak} , between each cohort. Similar values were noticed for the average peak velocities noticed at the cough centre, HWA location and along the inlet centerline between cohorts, so it is expected that if more participants were recruited, and a greater number of coughs were analyzed, the average peak values would converge. Table 4 shows the resulting p values obtained for the comparisons between the sick cohort and the healthy group, convalescent group, and a group containing both. These values represent the probability that the average V_{peak} is the same between the groups indicated for each comparison. The obtained values are very high compared to the significance level. This indicated a strong possibility that the differences between the groups were obtained due to chance rather than a true difference in the strengths of cough between groups.

The normalized profiles were plotted for the sick, convalescent and healthy groups, and an average curve was obtained for each. These averages were plotted together (Figure 7), and the similarity in coughs between cohorts was further observed. A rapid increase in velocity up to the peak value is observed, followed by a quick decay of velocity in the region of $1 < \tau < 3$, and ending with a gradually decreasing tail.

The overall average curve representing all the cohorts can be used to develop a simple mathematical model for the normalized velocity profile, as described by the formulae in Equation 4 and shown in Figure 10. These formulae may be used to approximate the centreline velocity of a cough jet at $x = 50D$. The applications of such a model are in the development of computational cough and transient jet simulations, or in the validation of coughs produced by physical simulations.

$$V_{fit}(\tau) = \begin{cases} \tau^{2.884}, & 0 \leq \tau \leq 1 \\ \frac{0.5}{\tau - 0.5}, & 1 < \tau \end{cases} \quad (4)$$

TABLE IV. P-VALUES CALCULATED FOR COMPARISON OF MEAN VELOCITIES

Sampling location (Method)	p value for comparison between sick cohort and group indicated		
	Convalescent	Healthy	Conv + Healthy
Cough jet centre (PIV)	0.56	0.31	0.50
HWA location (PIV)	0.81	0.33	0.46
Inlet centreline (PIV)	0.71	0.49	0.50
HWA location (HWA)	0.30	0.32	0.27

To better understand the relationship between the peak velocity and the time at which it occurs, the two quantities were plotted in Figure 8, and a reciprocal fitting curve was obtained (Equation 5). Although there is scatter in the data, by using Equations 2-5, the centreline velocity time histories at $x=50D$, can be developed based on any specified peak velocity. A user chooses a value V_{peak} , which allows for τ to be computed so that Equation 4.

$$V_{peak} = \frac{k}{\tau_{peak}} \quad (5)$$

where k is the average distance between the front of the jet and the point at which the maximum velocity occurs ($k=0.55$ m)

The k values computed for all cohorts and methods were similar ($k=0.62$ m for HWA, $k=0.55$ m for cough jet centre (PIV), $k=0.53$ m for HWA location (PIV), $k=0.39$ m for centerline (PIV), $k=0.33$ m for LES, $k=0.43$ m for URANS).

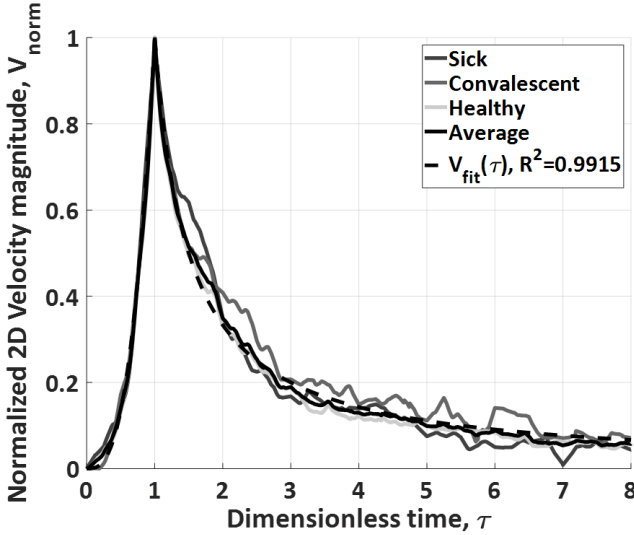


Figure 7. Mathematical modeling of average normalized time histories at cough centre ($x=1.0$ m)

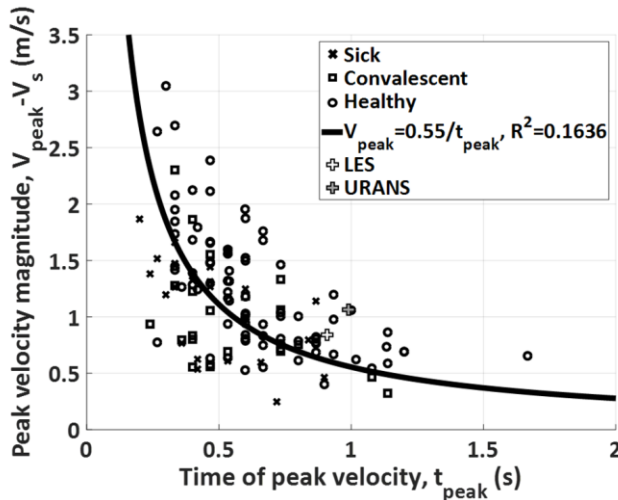


Figure 8. Variation of peak velocity with time of peak at cough centre ($x=1.0$ m, y location varies)

D. Turbulence characteristics

The fluctuations about the moving average of the HWA signal was used to characterize the turbulence present at the peak of the cough. The normalized power spectral density (PSD) was used to indicate the distribution of energy within turbulent eddies of variable size. Figure 9 displays an example of the normalized power spectral density. A slope of $-5/3$ was observed between 8-100 Hz, which is consistent with the Kolmogorov decay law, and only subtle differences are observed between coughs. The turbulence intensity was estimated from Equation 6 (I_{ua}) and from integrating the area under the PSD curve (I_{us}).

$$I_{ua} = \frac{V_{rms}}{V_{peak} - V_s} \quad (6)$$

where: V_{rms} is the root mean square of velocity fluctuations

Overall, the turbulence intensities computed using both methods were very similar. Table 5 shows the average turbulence intensities calculated for all coughs.

There are several factors that must be considered as limitations to the present work. It is possible that the type of viral infection will influence the production of mucous within the lungs and may alter the way a person will cough, such that it may not be accurate to treat all respiratory viruses the same. It is also assumed that the trajectory of a natural cough may be different from those studied here since the motion of the head is restricted in the trials. It is also assumed for this investigation that a forced cough will behave aerodynamically like a naturally occurring cough.

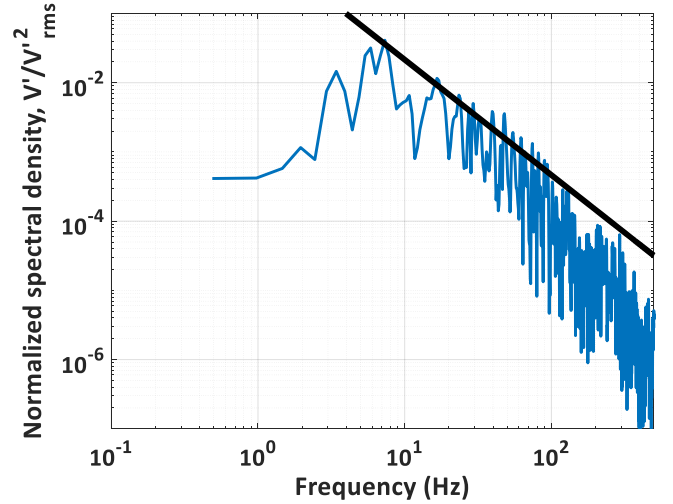


Figure 9. Normalized PSD of velocity fluctuations

TABLE V. COMPARISON OF AVERAGE TURBULENCE INTENSITIES OBTAINED FROM EACH METHOD

Quantity	I_{ua} (%) (from moving average)	I_{us} (%) (from spectrum)
Average	8.9	6.9
Std. Dev.	3.9	2.4

E. Computational fluid dynamics (CFD) model comparison and analysis

The validity of the numerical simulation was determined by comparing the normalized experimental velocity profiles with those determined computationally (Figure 10). A peak velocity of 0.83 m/s is obtained from the URANS simulation, whereas it is higher in the LES with a value of 1.07 m/s. The peak velocity observed in the LES was closer to the average value obtained from PIV experiments (1.17 m/s) for which there was no prescribed inlet velocity. This slight difference was expected, since in the analysis of experimental data, coughs with very low velocities were excluded, whereas the average value determined by Gupta et al. [8] included the lowest velocity coughs as well. Despite this, both the URANS and LES showed good agreement with the experimental results. The LES may have been a better representation of a realistic cough, but it was substantially more computationally expensive and so may not be practical for some investigations.

To gain a better understanding of the decay in V_{peak} with respect to distance from the source, the instantaneous velocity time histories were extracted along the simulated cough centreline every 0.1 m, starting at the inlet and extending to $x=1.5$ m. The normalized peak velocity, V_{peak-n} , was calculated according to Equation 7, and it was plotted with the distance from the inlet (Figure 11).

$$V_{peak-n} = \frac{V_{peak}}{V_{inlet}} \quad (7)$$

where: V_{peak} is the peak velocity magnitude observed at each location, V_{inlet} is the peak velocity magnitude observed at the inlet

As examined by Bi [9], the width of both steady and transient jets increases linearly from the virtual origin, x_0 . The cross-sectional area of the jet is therefore proportional to $(x-x_0)^2$ so a fitting curve that could be used to predict the decay of the peak velocity observed downstream was determined (Equation 8). This relationship was based on the LES simulation since it provides a better approximation of real flow behaviour.

$$V_{peak-fit}(x) = \frac{0.085}{(x+0.154)^2} \quad (8)$$

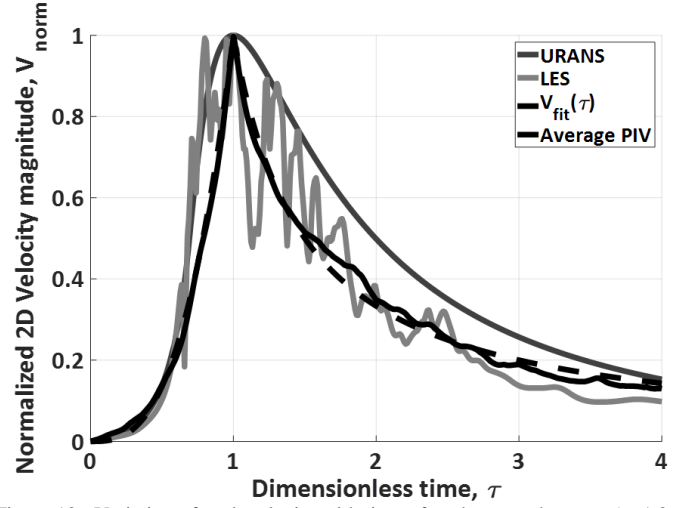


Figure 10. Variation of peak velocity with time of peak at cough centre ($x=1.0$ m, y location varies)

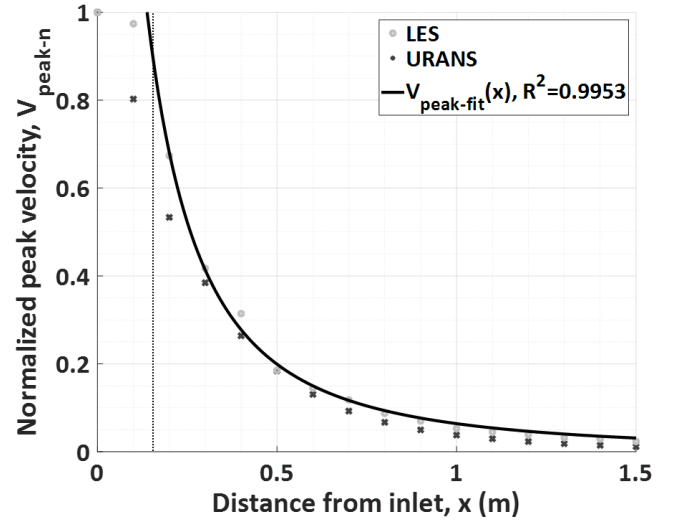


Figure 11. Variation of normalized peak velocity and distance from the inlet (CFD)

By examining Figure 11, the far field was estimated as the approximate region where $x > 0.15$ m. If the velocity time histories at each location are normalized by their respective peak and time that the peak occurs, the validity of this boundary distance is confirmed for the simulated jets examined here.

It may be considered that the present model for far-field cough flow is valid in the region $x \geq 7 D$, based on an estimated mouth opening diameter of 0.0217 m [8]. By combining numerical and experimental results, the model can be conservatively applied to approximate the cough jet behaviour at any location in the far field, although measurements of actual cough flow fields at greater distances than $x = 1.0$ m is recommended to confirm the validity.

IV. CONCLUSIONS

The objectives of the present investigation were to rigorously test the “3 feet” or “1 metre” rule as a supposed safe separation distance from a coughing person, as well as to develop an experimental model for human cough flow behaviour. This model was used to validate a computational fluid dynamics (CFD) model based on unsteady Reynolds-averaged Navier-Stokes (URANS) and large eddy simulation (LES) methods [9].

Throughout the study period, a total of 77 experiments were conducted from 58 different subjects. From these subjects, 21 presented with influenza like illness. Of these 21, 12 subjects had laboratory confirmed respiratory infections, which was a good result given the recruitment methods. In the experiments, 177 coughs were measured using particle image velocimetry (PIV), 171 using hot-wire anemometry (HWA) and 63 were sampled with polytetrafluoroethylene (PTFE) filter cassettes.

By combining the experimental measurements and data obtained from previous CFD simulations, a complete model for cough centreline velocity behaviour has been developed. Although a high degree of variability is observed between individual coughs, the average profiles are a good representation of the expiratory airflow field produced by a human cough in the region $x \geq 12 D$. This approximate location can be used to distinguish between near-field and far-field cough jet behaviour. An average peak velocity of 1.17 m/s was observed at the cough centre, 1.0 m downstream, which is comparable to the 1.07 m/s velocity determined by LES. The strong velocities measured provided evidence to dispute previously assumed safe separation distances, and therefore, there is incentive to develop more viable infection prevention methods. The average spread angle of 24° obtained experimentally was comparable to the values previously published for steady free jets, where a spreading angle between $20\text{--}35^\circ$ is typically measured [17].

No statistically significant differences were observed in the velocity or turbulence characteristics between coughs from sick or healthy participants and evidence was provided to support the claim that velocity data obtained from healthy participants in previous studies can be used to approximate the flow field of coughs from individuals who have been infected with respiratory viruses.

The biological air sampling method used in this study was unable to assess the relative significance of far field virus transmission. This could mean that that three forced coughing events do not produce a significant amount of viral droplet, or it is possible that the virus was present but not sampled due to the small area of the sampling cassettes. There is also a possibility that the air sampling method used destroyed the viral particles upon impaction and, perhaps in future investigations, more coughs should be sampled, or an alternative sampling method should be used.

ACKNOWLEDGMENTS

I would like to thank our partners at the division of infectious diseases, department of microbiology at Sunnybrook Research Institute, University of Toronto as well as the University Institute of Cardiology and Pulmonology, Laval University, and the staff at Western Student Health Services for their assistance with the recruitment of subjects for experiments

REFERENCES

- [1] de Francisco, S. N., Donadel, M., Jit, M., and R. Hutubessy. (2015). A systematic review of the social and economic burden of influenza in low- and middle-income countries. *Vaccine*. 33(48):6537-44.
- [2] Lafond, K. E., Nair, H., Rasooly, M. H., Valente, F., Booy, r., Rahman, M., Kitsutani, P., Yu, H., Guzman, G., Coulibaly, D., Armero, J., Jima, D., Howie, S., Ampofo, W., et al. (2016). Global role and burden of influenza in pediatric respiratory hospitalizations, 1982-2012: A systematic analysis. *PLOS Medicine*. 13(6):e1002060.
- [3] Morawska, L. (2006). Droplet fate in indoor environments, or can we prevent the spread of infection? *Indoor Air*. 16:335-347.
- [4] Tellier, R. (2009). Aerosol transmission of influenza A virus: a review of new studies. *J. R. Soc. Interface*. 6:6.
- [5] Ai, Z. T., Melikov, A. K. (2018). Airborne spread of expiratory droplet nuclei between the occupants of indoor environments: A review. *Indoor Air*. 28(4):500-524.
- [6] Deller, B., Stolarsky, G., and Tietjen, L. (2008). Preventing the transmission of avian or pandemic Influenza in health care facilities with limited resources: Learning resource package. *Jhpigo: An affiliate of Johns Hopkins University*. 2:5.
- [7] Kennamer, M. (2007). Infection Control for Health Care Providers. *Thomson Delmar Learning, Alabama*. 79-85.
- [8] Gupta, J. K., Lin, C. H., and Chen, Q. (2009). Flow dynamics and characterization of a cough. *Indoor Air*. 19:517-525.
- [9] Bi, R. (2018). A numerical investigation of human cough jet development and droplet dispersion. *M.E.Sc. Thesis, Dept. Mech. & Matls. Eng., University of Western Ontario, London, Canada*.
- [10] Savory, E., Lin, W. E., Blackman, K., Roberto, M. C., Cuthbertson, L. R., Scott, J. A., Mubareka, S. (2014). Western Cold and Flu (WeCoF) aerosol study preliminary results. *BMC research notes*. 7(1):563.
- [11] Mohamed, A. (2017) Experimental measurements of far-field cough airflows produced by healthy and Influenza infected human subjects. *M.E.Sc. Thesis, Dept. Mech. & Matls. Eng., University of Western Ontario, London, Canada*.
- [12] Steltzer-Braid, S., Oliver, B.G., Blazey, A. J., Argent, E., Newsome, T.P., Rawlinson, W.D., Tovey, E.R. (2009). Exhalation of respiratory viruses by breathing coughing and talking. *J Med Virol*. 81(9):1674-9
- [13] Lindsley, W. G., Blachere, F. M., Thewlis, R. E., Vishnu, A., Davis, K. A., Cao, G., Palmer, J. E., Clark, K. E., Fisher, M. A., Khakoo, R., and Beezhold, D. H. (2010) Measurements of airborne influenza virus in aerosol particles from human coughs. *PLoS ONE*. 5:e15100.
- [14] Lindsley, W. G., Blachere, F. M., Beezhold, D. H., Thewlis, R. E., Noorbakhsh, B., Othumpangat, S., Goldsmith, W. T., McMillen, C. M., Andrew, M. E., Burrell, C. N., and Noti, J. D. (2016). Viable influenza A virus in airborne particles expelled during coughs versus exhalations. *Influenza and Other Respiratory Viruses*. 10(5):404-413.
- [15] Verrault, D., Moineau, S., Duchaine, C. (2008) Methods for sampling of airborne viruses. *Microbiol Mol Biol Rev*. 72(3):413-444.
- [16] Tang, J. W., Liebner, T. J., Craven, B. A., and Settles, G. S. (2009). A Schlieren optical study of the human cough with and without wearing masks for aerosol infection control. *J. R. Soc. Interface*. 6(Suppl. 6):S727-S736.
- [17] Ball, C. G., Fellouah, H., and Pollard, A. (2012). The flow field in turbulent round free jets. *Prog. Aerospace Sci*. 50:1-2

Generation of Self Organized Bubble Train Flow

alyhafezg@gmail.com

Aly Gadallah

Department of Mechanical and Materials Engineering
Western University
London, Ontario, Canada
Department of Mechanical and Power Engineering
Tanta University
Tanta, Egypt

Kamaran Siddiqui

Department of Mechanical and Materials Engineering
Western University
London, Ontario, Canada
ksiddiq@uwo.ca

Abstract—Gas-liquid two phase flow is a widespread phenomenon occurring in both natural and industrial processes. Bubble train flow is a type of gas-liquid flow pattern that consists of elongated bubbles "Taylor bubbles" with equivalent diameter usually several times that of the channel diameter, separated by liquid slugs. The recirculation within the liquid slugs improves heat and mass transfer from liquid to wall and interfacial mass transfer from gas to liquid. Here we report on a novel capillary tube design for generating self-organized bubble train flow in the vertical circular capillary tube submerged in a stagnant liquid, through a gas supply from the lower end of the capillary tube at constant flow conditions. The design is based on the creation of a side hole in the capillary tube wall perpendicular to the tube axis and away by a certain distance from the lower end of the capillary tube. The generation of bubble train flow is based on the fact that the differential pressure between inside and outside the capillary tube in the side hole region, which resulting from the lowered density of the liquid-gas mixture inside the capillary tube as compared to the liquid outside the capillary tube, induces the surrounding liquid to enter the capillary tube through the side hole. Without any need for external energy to pump and control the liquid flow the resulting buoyant force causes the liquid pumping through the side hole. We show that the local liquid flow through the side hole acts to squeeze the gas thread inside the capillary tube in the radial direction until a visible neck is formed and eventually ruptured, resulting in the Taylor bubble formation. This work establishes the dynamics of self-organized bubble train flow generated in vertical circular capillary tube submerged in stagnant liquid. The proposed novel design provides an effective alternative to the existing monolith reactor design by providing uniform distribution of gas and liquid in each channel of the reactor to enhance its performance.

Keywords—Gas liquid; two phase flow; bubble; High-speed imaging

I. INTRODUCTION

In natural, gas-liquid two phase flow is exist in many of the phenomena's around us such as breaking waves on the ocean surface, waterfalls and strombolian volcanic. Moreover, gas-liquid flows receives great attentions in industrial processes due to its wide range of applications such as energy production (e.g., oil-gas pipelines, steam generators, cooling systems, nuclear reactors) and chemical engineering (e.g., bubble columns, airlift reactors, aeration systems, absorbers, distillation columns). Many interesting miniaturized reactors, with considerable industrial and practical applications. One of the important two phase flow patterns is bubble train flow or Taylor flow, which belongs to a class of segmented or intermittent flows. Bubble train flow is consists of gas bubbles with lengths greater than the tube diameter (Taylor bubbles) that move along the tube of capillary dimension separated from each other by liquid slugs. The Taylor bubbles almost completely fill the cross section of the tube, remaining separated from the wall by a thin film of liquid. The combination of the Taylor bubble and the liquid slug beneath it forms a unit cell and in general all unit cells have the same length along the tube. The axial segregation of the bulk liquid, the presence of bubbles in front and at the back of the slugs, significantly reduces axial dispersion, while at the same time inducing a circulating flow pattern in the liquid slugs, toroidal vortices extending the length of the slug [1, 2], which enhances radial transport. Therefore, bubble train flow offers many advantages for carrying out reactions compared with other two phase flow patterns and to single-phase laminar flow [3, 4].

The Taylor bubble formation in small channels and microchannels is strongly dependent on the geometry of the gas-liquid contacting section [5, 6] because they bring together the phases and affect the mechanism of gas bubble formation and its size which will influence the hydrodynamics and mass transfer performance of the bubble train flow [7, 8]. The geometry design of the inlet contacting section is classified into mainly two techniques: direct and indirect contacting. The indirect technique designating the case where both phases are contacted in a separate zone known as mixing chamber followed by a stepwise reduction of the mixing chamber down to the diameter of the

main channel [9, 10]. Unstable bubble train flow, characterized by different length of unite cells along the tube "channel" or the occurrence of different flow patterns at a fixed set of gas-liquid superficial velocities, are observed using the indirect contacting technique [11,12]. Whereas literatures suggest that the direct contact technique always resulted in the formation of a stable bubble train flow. Different geometric designs of direct contacting section are identified, where the inlet channels of both phases are contacted directly to the main channel which include;

- T-shaped junction where the gas and liquid inlets are perpendicular to the main channel [6];
- Pseudo "or Side" T-shaped junction where either the gas or liquid inlet is in the direction of the main channel [5];
- Y- shaped junction where the inlets for gas and liquid phases join at an angle less than 180° [15];
- Cross-shaped junction where the two phases merge in a cross where the gas is flowing in the main channel and the liquid is supplied through two orthogonal inlets [16];
- Converging shaped or flow focusing junction where the two inlets of the liquid phase in the cross-shaped junction join the third inlet of the gas phase at an angle less than 90° [14].

To the best of authors' knowledge, the formation of fully uniform "regular" bubble train flow in capillary tubes and micro channels always requires an active way to control the flow rate of the liquid phase such as Syringe pumps, digital piston pumps and peristaltic pumps.

Here, we report an experimental observation of passive, self-organized and fully uniform "regular" bubble train flow generation in vertical circular capillary tube submerged in stagnant liquid by using a novel capillary tube design. The design is based on the creation of one side hole in the capillary tube wall perpendicular to the tube axis and a certain distance away from the lower end of the capillary tube from where the gas is supplied at constant flow conditions. The diameter of this side hole is designed to be smaller than or equal to the tube inner diameter. This study also reports the experimental verification of the effective range of this novel capillary tube design for the generation of self-organized bubble train flow taking into account the influence of (i) capillary tube size, (ii) gas phase flow rate, (iii) viscosity of the liquid phase outside the capillary tube and (iv) side hole size. The range of gas flow rates that guarantee stable generation of bubble train flow in different capillary tubes size and side hole to capillary tube diameter ratio (D_h/D) are listed in Table 1. Three capillary tubes made of glass, with inner diameters (D) of 1.78, 2.36 and 3.45 mm are used in the present investigation. Air and two liquids, water and 50% (v/v) Glycerol/Water are used as test fluids. Figure 1(a) illustrates the schematic of the experimental set-up. It is also observed that the distance between the side-hole position and the outlet side of the tube, H (see Fig. 1a), should be greater than 30 mm to get a stable generation of self-organized bubble train flow for all capillary tubes.

Table 1 The dimensions of the capillary tube diameter (D), side hole/capillary tube diameter ratio (D_h/D), considered in the study and the range of gas flow rate that guarantees stable generation of bubble train flow (Q_g).

D (mm)	Dh/D	Qg (ml/min)	Ql (ml/min)	Liquid
1.78	1	2.8-173.1	6.9-14.4	Water
2.36	1	7.8-186.9	18.3-32.2	
3.45	1	55.7-431.8	85-100.4	
1.78	1	4.2-28	4.9-9.5	50% Glycerol/Water (v/v)
2.36	0.585	4.53-191.9	8.5-18.7	
1.78	0.494	3.7-42.6	8.6-16.3	Water
2.36	0.424	9.13-144	13.4-21.4	
2.36	0.585	12.9-182.6	16.5-24.6	
3.45	0.522	20-315.8	50.5-73.7	
3.45	0.725	12.5-453.3	45-75.5	

II. MECHANISM OF SELF-ORGANIZED TAYLOR BUBBLE FORMATION IN THIS NOVEL PASSIVE DESIGN

The driving force for the suction of the surrounding stagnant liquid through the side hole in the vertical capillary tube is the imbalance of the hydrostatic pressure, at the position of the side hole, between the inner capillary tube and the surrounding liquid. This hydrostatic pressure difference is due to the difference gas hold-up, i.e. the overall density of the fluid in the two zones. The mechanism of the bubble train formation in a vertical capillary tube, submerged in a stagnant liquid, is divided mainly into three stages; filling "blocking" stage, squeezing stage and pinch-off stage.

A. Filling "Blocking" stage

The cycle of Taylor bubble formation starts immediately after the previous bubble detaches from the main gas volume (see figure 1-b, image 1). The train of previously formed Taylor bubbles rising above in the tube causes the hydrostatic pressure of the liquid pockets to be lower than that of the stagnant liquid at the same height outside the tube. As a result, after a bubble detachment, the liquid from the surrounding stagnant liquid domain freely enters the capillary tube through the side hole above the rising gas volume, as shown in figure 1-b (images 2-4), until the gas volume approaches the upper edge of the side hole wherein covering the entire cross-section of the side hole and blocking the entrance of the liquid phase, as shown in figure 1-b (image 5). At this stage, the height of the liquid pocket above the moving gas volume and the previously detached bubble is set fixed, which defines the hydrostatic pressure exerted by the liquid pocket onto the moving gas front. Liquid moving through the gap between the thread and the wall of the capillary tubes subject to an increased viscous resistance, which leads to a local

pressure build-up in the liquid phase upstream of the tip of the gaseous thread at the side hole region and marks the start of the squeezing stage.

B. Squeezing stage

Due to the lower pressure on the leading edge of the gas volume compared to that of the stagnant outside liquid, as soon as the gas leading edge closes the side hole, the higher pressure outside liquid starts to squeeze the gas, as shown in figure 1-b (images 6-9). The gas-liquid interfacial tension tends to resist this external force. At the same time, during the squeezing stage, the gas stream continues to grow in the tube due to continuous supply of the gas. Note that the surface tension at the leading edge of the gas stream keeps the interface shape intact and resists any gas expansion at the interface due to the squeezing effect of the outside liquid. This is confirmed by the fact that the liquid pocket about the gas leading edge maintained the constant height during the squeezing stage. The outside liquid continues to fill

in the space created by the squeezing of the gas stream inside the capillary tube.

C. Pinch-off stage

At the end of the squeezing stage a clearly visible neck begins to develop, as shown in figure 1-b (image 1 or 9), the neck width becomes smaller and smaller until it eventually ruptures, resulting in new bubble detachment inside the vertical capillary tube, as shown in figure 1-b (image 10).

The process of the outside liquid squeezing and entering the tube is quantitatively shown in figure 1-c through the map of the liquid velocity vectors. The results clearly show the rushing in of the surrounding liquid into the tube by squeezing the gas stream. The results also show that the velocity with which the liquid enters the tube increases as the gas “neck” becomes thinner. This could be due to the reason that the thinning of the gas neck allows the liquid to fill in the space from all sides, thus, increasing the liquid volume flow rate.

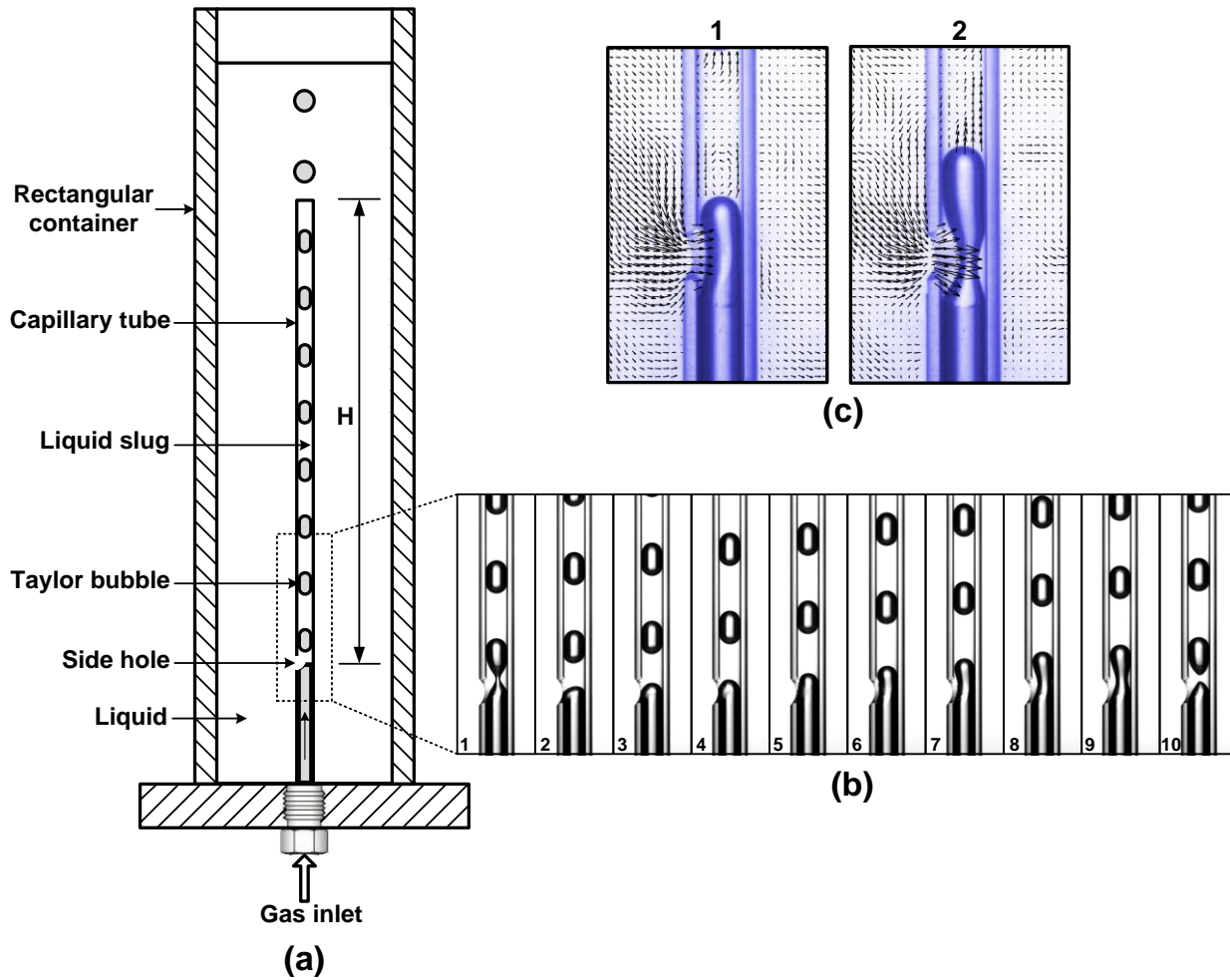


Figure 1: Bubble train flow generation in a vertical capillary tube submerged in a stagnant liquid, (a) A schematic representation of experimental set-up which consists mainly of (i) a single capillary tube, made of glass, designed with a side hole and (ii) transparent rectangular container for the liquid phase, the constant flux of air is delivered through the bottom of the capillary tube, the diameter of capillary tubes used in the present study are 1.78 mm, 2.36 mm and 3.45 mm (b) Successive images of bubble train formation mechanism inside the capillary tube showing the process of Taylor bubble generation in the side hole region which involve three stages: filling stage, squeezing stage and pinch-off stage (c) Vector maps of liquid velocity around the gaseous stream at the side hole region.

III. BUBBLE FORMATION RATE INSIDE THE CAPILLARY TUBE

The rate of Taylor bubble formation inside the capillary tube is shown in figure 2 as a function of the superficial gas velocity for various parameters. In general, it is seen that the bubble formation rate increases with an increase in the superficial gas velocity in the lower range of the superficial gas velocity. However, in the higher range of superficial gas velocity, the rate of bubble formation becomes almost independent of the superficial gas velocity. At a given superficial gas velocity, and a side hole-to-capillary tube diameter (D_h/D) ratio equal to one, the bubble formation rate decreases with an increase in the capillary tube diameter, as shown in figure 2(a). It indicates that the bubble break up process (filling stage, squeezing stage and pinch-off stage) take long time with increasing the diameters of both the side hole and the capillary tube. The sizes of the side hole and capillary tube influence the flow velocity and surface tension force. For a given superficial gas velocity, as the tube diameter decreases, both the surface tension force and the gas flow rate decrease, which allows the liquid to accelerate the liquid breakup process. However, the effect of D_h/D ratio on the bubble formation rate does not have a consistent behavior for all capillary tube diameters considered. It is seen that the bubble formation rate decreases when the D_h/D ratio decreases for the

capillary tubes size ($D = 1.78$ and 2.36 mm), as shown in figure 2 (b) and (c), while for the larger capillary tube size ($D = 3.45$ mm), the bubble formation rate increases when the D_h/D ratio decreases (see figure 2d). A decrease in the D_h/D ratio corresponds to a decrease in the liquid entrance area compared to the gas flow area. As discussed earlier, the bubble breakup process is influenced by the local fluid pressure and velocity, as well as the surface tension. For smaller capillary tube size, the gas flow rate and surface tension force are low. A decrease in the side-hole size reduces the liquid flow and hence the liquid drag to squeeze the bubble. Thus, it takes longer time to pinch off the bubble neck as a result, the bubble formation rate decreases. The results in figure 2(b) and (c) also show that for a given D_h/D ratio, the bubble formation rate decreases when the viscosity of the liquid phase increases (comparison between water and 50% Glycerol/Water (v/v)). This is likely due to the reason that an increase in liquid viscosity increases the flow resistance that slows down the bubble train movement inside the capillary tube. It is also observed that an increase in liquid viscosity also affects the stability of the bubble train generation. As results in figure 2 (b) and (c) show, the gas flow rate range that ensures stable generation of bubble train decreases to lower gas flow rates when the capillary tube diameter decreases for high viscosity liquid, which could be due to a further increase in the flow resistance.

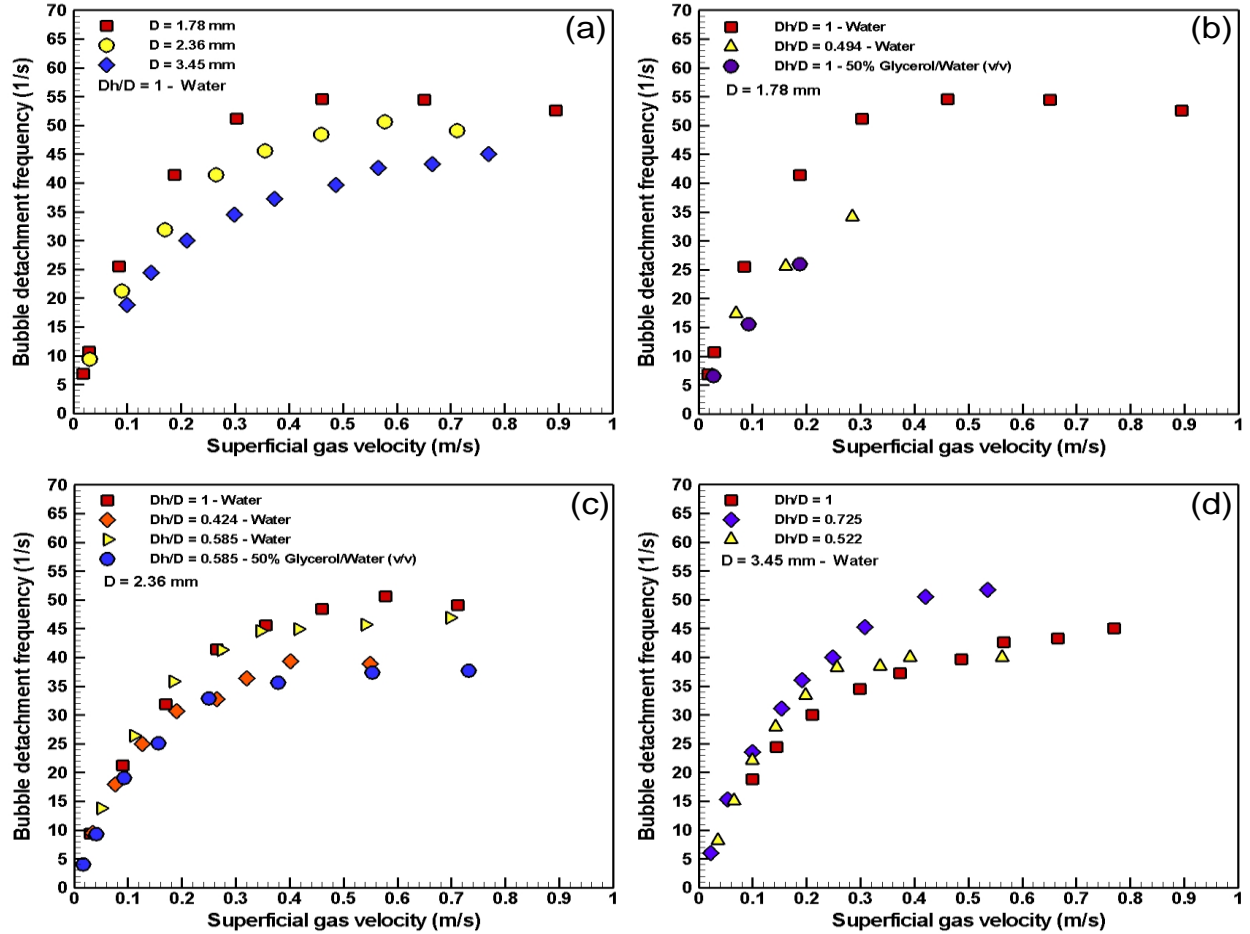


Figure 2: Bubble detachment frequency inside the capillary tube as a function of the superficial gas velocity at different diameters of capillary tube and side hole, and two liquids.

IV. THEORETICAL PREDICTION OF BUBBLE DETACHMENT FREQUENCY

The prediction of the bubble detachment frequency inside the capillary tube with a side hole and submerged in a stagnant liquid, is mainly dependent on the theoretical definition of the bubble growth time " T_g " which is the time between two consecutive bubbles or the time of formation of the unit cell. The bubble growth time can be divided into two time intervals; the liquid slug filling time " T_{fill} " and the bubble formation time " T_b ". The liquid slug filling time is the time during which the liquid can creep through the side hole under the rising gaseous stream, assuming that the liquid flow will stopped when the gas front close the side hole entrance, so T_{fill} can be defined as:

$$T_{fill} = \frac{V_{fill}}{Q_g} = \frac{\left(\frac{\pi D^2}{4}\right) \cdot D_h}{Q_g} \quad (1)$$

Where V_{fill} is the volume of the gaseous stream when it is grown up to the upper edge of the side hole. Once the gas front closes the side hole entrance, the liquid starts to squeeze the gas domain at the side hole inlet in the radial direction, at the same time the gaseous stream continues to grow in the tube, downstream of the side hole due to the continuous flow of the gas phase. During this process the bubble formation time is control by both liquid and gas flow rates [17], which can be approximately equal to:

$$T_b = \frac{L_B}{U_{TP}} \quad (2)$$

Where L_B is the bubble length and U_{TP} is the superficial two-phase velocity ($U_{TP} = U_g + U_l$, where U_g and U_l are the gas and liquid superficial velocities, respectively) and can be expressed as,

$$U_{TP} = U_l + U_g = \frac{Q_l + Q_g}{\left(\frac{\pi D^2}{4}\right)} \quad (3)$$

So the time of formation of a Taylor bubble in the capillary tube will be given as:

$$T_g = T_{fill} + T_b = \frac{\left(\frac{\pi D^2}{4}\right) \cdot D_h}{Q_g} + \frac{L_B}{U_{TP}} \quad (4)$$

The bubble detachment frequency f_B can then be calculated as:

$$f_B = \frac{1}{T_g} = \left(\frac{\left(\frac{\pi D^2}{4}\right) \cdot D_h}{Q_g} + \frac{L_B}{U_{TP}} \right)^{-1} \quad (5)$$

The above prediction model for the bubble detachment frequency was validated against the experimental data. The results are presented in figure 3 which shows a good agreement

between the experimental and theoretical results over a range of capillary tube diameters and side hole diameters.

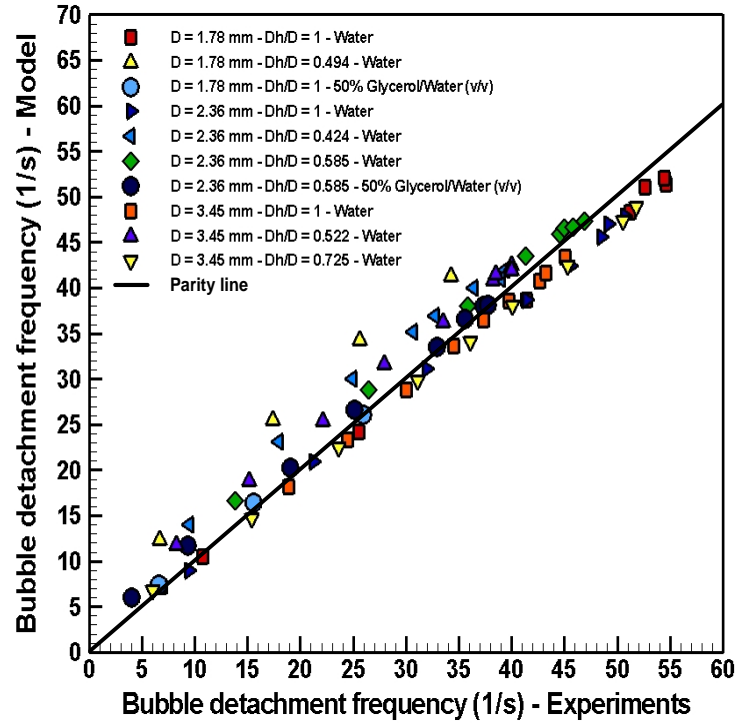


Figure 3: Validation of the prediction model Eq. (6) against the experimental values of bubble detachment frequency for all cases.

V. APPLICATION FOR THE PRESENT PHYSICAL OBSERVATION

The current design of monolith reactors faces a major challenge of uneven distributions of gas and liquid in different channels of the reactor [11, 18-21]. This phase maldistribution is the main reason for the decrease in the monolith reactor performance. The present novel design of capillary tube with a side hole provides an effective solution to this maldistribution in the monolith reactors. At the above results show, this novel design is a self-regulating mechanism to generate a train of bubbles whose size can be controlled passively via the diameters of the side hole and capillary tube, as well as, actively through the gas flow rate. The sets of these capillary tubes with side holes can be compactly "arranged" in a geometric form to create a reactor. This new reactor design would be an effective replacement for monolith reactors since this new reactor design will have a completely uniform bubble train flow inside all capillary tubes of the reactor. Furthermore, it will have the advantages of the monolith reactor which include high-transport rates of heat and mass per unit pressure drop, high surface area to volume ratios, low pressure drop, minimum axial dispersion and ease of scale-up [12].

VI. METHODS

A. Experimental set-up

The experimental set-up is shown in Fig. 1a. The experiments are conducted in a vertical column 120×120 mm in cross-section and 1.2 m in height. The column is made of acrylic to allow optical access for visualization. The column is open to atmosphere under ambient conditions. The water height is maintained at 750 mm, from the bottom of the column, for all experiments. A glass capillary tube is placed in the middle of the rectangular column. A compressed air line is used to provide pressurized air to the capillary tube from the bottom inlet. The air from the line first entered a settling chamber 0.16 m³ in volume through a pressure regulator. The purpose of the settling chamber is to dampen any pressure fluctuations from the compressed air line. The air from the settling chamber passed through a flow meter to the capillary tube via a long Teflon tube (3m) with inner diameter 0.89 mm to reduce pressure fluctuations and maintain nearly constant flow conditions. Two needle valves are attached along this line to precisely control the air flow rate. The accurate measurement of gas flow rate at low flow rates is very challenging. For this purpose, at each flow rate prior to the actual bubble measurements, the gas flow rate was adjusted by the needle valve and measured by the soap film meter where an inverted funnel was placed near the liquid surface in the column. The air bubbles were trapped and collected through the funnel and guided towards a soap flow meter. The liquids used in the experiments are water and 50% Glycerol/Water (v/v).

B. High-speed imaging

A Photron FASTCAM SA5 high-speed CMOS camera with the resolution of 1024×1024 pixels is used to record the process of bubble formation inside the capillary tube. The camera has a global electronic shutter from 16.7ms to 1 μ s independent of the frame rate. A 60 mm f 1:2.8D Nikon AF Micro-Nikkor lens is used. The Photron FASTCAM Viewer software is used to control the camera operating system including frame rate triggering, image resolution and shutter speed through a personal computer which is connected to the camera via gigabit Ethernet. The lighting source is a halogen lamp (500 watt) placed behind the test column through a diffusion screen which decreased small-scale inhomogeneity in the illumination. Each bubble was clearly projected by the back-lighting method. For each experimental run images were acquired at a rate up to 50,000 frames per second. This frame rate was found to be sufficient to resolve the transient effects of bubble formation.

VII. REFERENCES

- [1] Thulasidas, T.C., Abraham, M.A., Cerro, R.L. (1995). "Bubble-train flow in capillaries of circular and square cross section", *Chemical Engineering Science*, vol. 50, pp. 183-199
- [2] Thulasidas, T. C., Abraham, M. A., and Cerro, R. L., (1997) "Flow Patterns in Liquid Slugs During Bubble-Train Flow Inside Capillaries", *Chemical Engineering Science*, vol. 52, pp. 2947-2962.
- [3] Burns, J.R., Ramshaw, C. (2001) "The intensification of rapid reactions in multiphase systems using slug flow in capillaries", *Lab on a chip*, vol. 1, pp. 10-15
- [4] Majumder, A., Mehta, B., and Khandekar, S., (2013), "Local Nusselt number enhancement during gas-liquid Taylor bubble flow in a square mini-channel: An experimental study", *International Journal of Thermal Science*, vol. 66, pp. 8-18.
- [5] Garstecki, P., Fuerstman, M. J., Stone, H. A., and Whitesides, G. M., (2006) "Formation of Droplets and Bubbles in a Microfluidic T-Junction—Scaling and Mechanism of Break-Up", *Lab on a chip*, vol. 6, pp. 437-446.
- [6] Shao, N., Gavrilidis, A., Angeli, P., (2011) "Effect of inlet conditions on Taylor bubble length in microchannels", *Heat Transfer Engineering*, vol. 32, pp. 1117-1125
- [7] Irandoust, S., Ertlé, S., Andersson, B. (1992), "Gas-liquid mass transfer in Taylor flow through a capillary", *Canadian Journal of Chemical Engineering*, vol. 70, pp. 115-119.
- [8] Kreutzer, M., "Hydrodynamics of Taylor Flow in Capillaries and Monolith Reactors", Delft University Press, Delft, the Netherlands, 2003.
- [9] Serizawa, A., Feng, Z., Kawara, Z., 2002. Two-phase flow in microchannels. *Experimental Thermal and Fluid Science*, vol. 26, 703-714.
- [10] Kawahara, A., Chung, P.M.-Y., Kawaji, M., (2002) "Investigation of two-phase flow pattern, void fraction and pressure drop in a microchannel". *International Journal of Multiphase Flow*, vol. 28, pp.1411-1435.
- [11] Mantle, M.D., Sederman, A.J., Gladden, L.F., Raymahasay, S., Winterbottom, J.M., Stitt, E.H. (2002) "Dynamic MRI visualisation of two-phase flow in a ceramic monolith" *AIChE J*, vol. 48: pp. 909-912
- [12] Kreutzer, M.T., Bakker, J.J.W., Kapteijn, F., Moulijn, J.A., Verheijen, P.J.T., (2005), "Scaling-up multiphase monolith reactors: linking residence time distribution and feed maldistribution", *Industrial & Engineering Chemistry Research*, vol. 44, pp. 4898-4913.
- [13] Hessel, V., Ehrfeld, W., Herweck, T., Haverkamp, V., Lowe, H., Schiewe, J., Wille, Ch., (2000) "Gas/liquid microreactors: hydrodynamics and mass transfer. Proceedings, IMRET 4, Atlanta, GA, March 5-9, AIChE: New York, pp.174-186.
- [14] Haverkamp, V., Hessel, V., Löwe, H., Menges, G., Warnier, M.J.F., Rebrov, E.V., de Croon, M.H.J.M., Schouten, J.C., Liauw, M., (2006) "Hydrodynamics and mixer-induced bubble formation in microbubble columns with single and multiple channels" *Chemical Engineering & Technology*, vol. 29, pp. 1015-1026.
- [15] Kumar, V., Vashisth, S., Hoarau, Y., Nigam, K.D.P., 2007. "Slug flow in curved microreactors: hydrodynamic study". *Chemical Engineering Science*, vol. 62, pp. 7494-7504.
- [16] Cubaud, T., Tatineni, M., Zhong, X., Ho, C.M., (2005) "Bubble dispenser in microfluidic devices". *Phys. Rev. E*, vol. 72, 037302.
- [17] Ben Abdelwahed M.A., Wielhorski, Y., Bizet, L., and Breard, J., (2012) "Characterisation of bubbles formed in a cylindrical T-shaped junction device," *Chemical Engineering Science*, vol. 76, pp. 206-215
- [18] Roy S, Al-Dahhan M, (2005) "Flow distribution characteristics of a gas-liquid monolith reactor", *Catalysis Today*, vol. 105, pp. 396-400.
- [19] Bauer T, Roy S, Lange R, Al-Dahhan M. (2005) "Liquid saturation and gas liquid distribution in multiphase monolithic reactors". *Chemical Engineering Science*, vol. 60, pp. 3101-3106.
- [20] Behl M, Roy S. (2008)"Experimental investigation of gas-liquid distribution in monolith reactors". *Chemical Engineering Science*, vol. 62, pp. 7463-7470.
- [21] Sederman AJ, Heras JJ, Mantle MD, Gladden LF. (2007) "MRI strategies for characterising two-phase flow in parallel channel ceramic monoliths", *Catalysis Today*, vol. 128, pp. 3-12.

Heat Transfer

EXPERIMENTAL INVESTIGATION OF A LOOP HEAT PIPE WITH A TWO-WAY PRESSURE REGULATED VALVE

Hooman Jazebizadeh, Tarik Kaya

Mechanical and Aerospace Engineering department
Carleton University
Ottawa, Canada

Hooman.jazebizadeh@carleton.ca , tarik.kaya@carleton.ca

ABSTRACT

Loop Heat Pipe (LHP) is a two-phase heat transfer device which takes advantage of the capillary forces created in a fine porous wick. This self-regulating device utilizes the latent heat of a working fluid to transfer applied heat from a hot source to a cold sink without using any external power. This paper focuses on an LHP with a two-way pressure regulated valve (PRV) to keep the operating temperature nearly constant even when the applied power or the sink conditions are changed. The operation of the PRV is such that it degrades the LHP conductance by increasing the pressure drop in the loop. The main application of this type of LHP is the thermal management of space systems such as geostationary satellites and planetary rovers.

The LHP used for this work is the engineering prototype of a unit that has been flown to space onboard Hispasat-1E in 2010. The tests are performed in ambient conditions and a heat exchanger is installed on the radiator panel to cool the condenser by a chiller. The device is instrumented by 38 thermocouples to measure temperature of all components, ambient and sink.

During the test campaign, several tests are performed to assess the operational behavior of the LHP and its two-way PRV. These tests can be categorized into start-up, power cycling, and sink cycling. The PRV has three operation modes including closed, regulating, and open. The start-up tests examined the transition between the closed and regulating mode; and particularly the required minimum applied power for the loop start-up/transition. The PRV behaves differently when applying a low power (5 W) than a high one (40 W). Furthermore, the power cycling experiments are performed to evaluate the regulating and open modes and the transition between them. It is observed that during regulating mode, the LHP's operating temperature is varied only within 1.5 K from the regulating temperature that is set by PRV and it is independent of the sink temperature. However, the applied power for the regulating-to-open transition is detected to vary with the sink temperature. In the open mode, the operating temperature is a function of both the sink and applied power. It is concluded that the LHP in the open mode behaves as a regular LHP but with a different initial condition affecting its operation.

Moreover, high-frequency and low-amplitude oscillations of the LHP temperatures are observed as soon as the PRV loses its regulating capability. A thorough analysis is performed to find the effects of the applied power, sink and ambient temperatures on the oscillation characteristics. As a part of the analysis, the sink cycling tests are performed to investigate the effects of the sink temperature not only on the operating temperature and transition among modes but also on the amplitude, frequency and phase change of the oscillations. The results of the study can be used to optimize the design of LHPs with an PRV for future space missions such as a moon lander/rover.

Effects of Avocado Ripeness and Extraction Temperature on Polyphenolics Yield Using Subcritical Water Extraction

W. Mazyan, A. Ahmadi

University of Prince Edward Island:
Faculty of Sustainable Design
Engineering
Charlottetown, Canada

E. Martin, A. Vogt

National Research Council: Aquatic
and Crop Resource Development
Charlottetown, Canada

E. Charter

BioFoodTech: Bioscience and Food
Department
Charlottetown, Canada

ABSTRACT

The increasing demand for phenolic compounds over the past decade has led to technical advancements in subcritical water extraction [1]. Their health benefits in cancer prevention, diabetes type II prevention, and skin/hair preservation have made them admired by different pharmaceutical and cosmetic companies around the globe. Several studies investigated the extraction of polyphenolics from a variety of organic materials. For example, Karacabey *et al.* [2] extracted phenolic compounds using pressurized water and 25% ethanol as a cosolvent. They were able to enhance the extraction yield by 44% with the addition of ethanol to water. Another study conducted by Singh *et al.* [3] aimed at extracting phenolic compounds from potato peel waste using subcritical water extraction. Their results showed 2% extraction yield (by dry mass basis). Alvarez *et al.* [4] also studied the phenolic compounds extraction from potato peel using subcritical water extraction at 40 bar and 190°C. Their results showed a 20 mg/g extraction yield. The exploration of extracting phenolic compounds from additional phenolics rich foods, would provide larger amounts of phenolic compounds to the market. An excellent candidate listed under phenolic compounds rich foods is avocado. This paper investigates the feasibility of extracting phenolic compounds from avocado meat using subcritical water extraction. Extraction yields from fresh and ripe avocado meat were investigated at 105°C, 87 mL/min, and 18 bar. The proposed extraction method shows efficient yields of the phenolic compounds due to their polar chemical bonds. The extraction yield is evaluated by calculating the phenolic compounds mass to avocado's dried mass. The phenolic compounds count was evaluated using Folin-Ciocalteu reagent solvent, and under the spectrometer, the outlet phenolic compounds rate is evaluated. It is observed that subcritical water extraction provides higher phenolic compounds when extracting the components from ripe fruit. The extract from the ripe fruit shows 30% enhancement in the phenolic compounds yields. This is due to a burst in ethylene production in the fruit which softens the fruit surface and allows the subcritical water to permeate easily for the extraction purpose. Eventually, ripe avocado contains higher levels of phenolic compounds which are easily accessible due to softer surface and would be an alternative material for phenolic compounds extractions.

References:

- [1] L. P. Martínez-Padilla, L. Franke, X. Q. Xu, and P. Juliano, "Improved extraction of avocado oil by application of sono-physical processes," *Ultrason. Sonochem.*, vol. 40, no. June 2017, pp. 720–726, 2018.
- [2] E. Karacabey, G. Mazza, L. Bayindirli, and N. Artik, "Extraction of Bioactive Compounds from Milled Grape Canes (*Vitis vinifera*) Using a Pressurized Low-Polarity Water Extractor," *Food Bioprocess Technol.*, vol. 5, no. 1, pp. 359–371, 2009.
- [3] P. P. Singh and M. D. A. Saldaña, "Subcritical water extraction of phenolic compounds from potato peel," *Food Res. Int.*, vol. 44, no. 8, pp. 2452–2458, 2011.
- [4] V. H. Alvarez, J. Cahyadi, D. Xu, and M. D. A. Salda, "Optimization of phytochemicals production from potato peel using subcritical water: Experimental and dynamic modeling," *J. Supercrit. Fluids*, vol. 90, pp. 8–17, 2014.

Experimental Investigation of Phase Change Material in Concrete and Bentonite-Based Thermal Energy Storage Media

Pedram Atefrad, Hiep V. Nguyen, Reza
Daneshazarian, Aggrey Mwesigye, Seth B. Dworkin
Department of Mechanical and Industrial Engineering
Ryerson University
Toronto, Canada
Corresponding author: seth.dworkin@ryerson.ca

David Salt
McMaster Innovation Park
Capture Technologies
Hamilton, Canada

ABSTRACT

Ground-source heat pumps (GSHPs) are an economically practical option to sustainably offer space cooling and heating. There are many parameters that can affect the long-term performance of these systems such as the thermal characteristics of the ground, the profile of the building's heating and cooling loads, and the available space to drill the borefield. In the province of Ontario, Canada, the ground has an undisturbed temperature of 8-12°C; however, operation of GSHPs can change the temperature of the ground based on the rate of heat injection/extraction into/from the ground. This possible difference between injection and extraction rates (mismatched cooling and heating loads of building) can cause the ground temperature to rise or fall in the long term, and thereby reduces the efficiency of GSHPs.

To enhance the performance of GSHP systems, they can be coupled with thermal energy storage so that the temperature fluctuations of the heat source/sink are minimized. Minimizing these fluctuations will increase the system efficiency. To improve the performance of the thermal energy storage, different phase change materials (PCM) can be used. One low-cost PCM is paraffin-based, which like other PCMs, can be a solution to the thermal imbalance of the ground, since PCMs act as thermal capacitors and help to alleviate the aforementioned problem of thermal imbalance. This is due to high energy storage density of PCMs.

In order to study the viability of an underground thermal storage medium employing PCM, three experiments were performed in this study. The first one was with a test-bench cylindrical thermal storage tank with a diameter of 1 m and a height of 1.5 m. Sensors were embedded in the tank to analyze the performance of the system and measure the temperature distribution. A bentonite/sand mixture with and without PCM was studied. In the second experiment, a u-shaped tube placed in a flat cavity domain with overall dimensions of 15 cm by 11 cm was used to characterize the thermal property of different PCMs. The third experiment was investigating the performance of PCM in a concrete cylinder 1 m in diameter and 1 m high with twelve PCM and fluid tubes. The PCM was used in different configurations. The attained results reveal that there is potential in improving the thermal balance of the medium by employing PCM. An added benefit is that a smaller volume of heat/sink is then required as compared to using the natural ground/soil, which may permit larger geo-exchange installations with a smaller footprint.

Thermal Characteristics of Steel and Composite Risers for Flow Assurance

Samar M. Aly

Graduate Student

Department of Mechanical Engineering
 Memorial University of Newfoundland
 St. John's, NL, Canada
 smaly@mun.ca

Samer Nakhla

Assistant Professor

Mechanical Engineering, Faculty of Engineering
 Emergency Medicine, Faculty of Medicine
 Memorial University of Newfoundland
 St. John's, NL, Canada
 satnakhla@mun.ca

ABSTRACT

In oil and gas production, flow assurance guarantees a successful and economical flow of the fluids from the reservoir to a designated processing facility. Flow assurance is one of the biggest challenges that pipeline designer faces, especially under deep water where the temperature is low and the pressure is high. These deep water conditions favor the formation of solid deposits, which leads to blocking the flow line, reducing oil production, and ultimately may cause the well to shut down. To avoid this problem, the flow line temperature must always be kept above the solid deposit formation temperature. This necessitates accurate analysis of pipelines thermal properties to choose the best material suitable under these harsh conditions.

For many years, steel catenary risers (SCR) were considered as the best solution for under deep water operations due to their lower cost and significant dynamic resistance. However, they suffered from many disadvantages including heavy weight, low corrosion resistance, and high thermal conductivity. Recently, many designers proposed to use composite catenary risers (CCR) instead of SCR. Composite materials are consider superior over its metallic counterparts because of their better thermal characteristics, high strength, low density and light weight, high corrosion resistance, excellent fatigue properties, and lower coefficient of thermal expansion (CTE) [1]. These favorable properties privileged using composite materials in catenary risers.

This paper presents a comparison between traditional steel catenary risers (SCR) and Composite Catenary Risers (CCR) based on their thermal characteristics for flow assurance purposes. The comparison is based on predicting the fluid flow temperature along the pipeline to show which material will keep the temperature above the solid deposit formation temperature, which is set in this paper to be 20 °C [2]. To predict the fluid flow temperature, a mathematical model based on the exact energy conservation equation for fluid transport in a pipeline is used in the steady state fluid flow. Nominal homogenized mechanical and heat transfer properties are used for composite and steel pipelines of the same thickness and diameter. Initial comparison between riser materials without insulation provided in Figure 1 shows expected exponential heat loss in the SCR whereas the CCR demonstrated significantly superior performance.

To establish rational comparisons between SCR and CCR, other aspects of their performance must be considered. Performance aspects regarding material strength, expected life and minimal weight design constitute the most essential minimal set for these comparisons. Comprehensive investigations are conducted and conclusions are extracted to quantify overall performance aspects of SCR and CCR.

Keywords: Flow assurance, Composite material, Thermal conductivity.

References

- [1] Jones, R. (1999). *Mechanics of Composite Materials*, 2nd Edition, Taylor & Francis Group, ISBN 1-56032-712-X
- [2] Su, J., et al. (2012). "Thermal design of multi-layered composite pipelines for deep water oil and gas production." *Int. Journal on Comput. Appl. Technology* 43(3): 248-259.
- [3] Bai, Y. And Q. Bai (2005). *Subsea Pipelines and Risers*, Elsevier Science, ISBN 0-080-4456-67
- [4] The Canadian Association of Petroleum Producers (CAPP) (2017). "Use of Reinforced Composite Pipe (Non-Metallic Pipelines)".

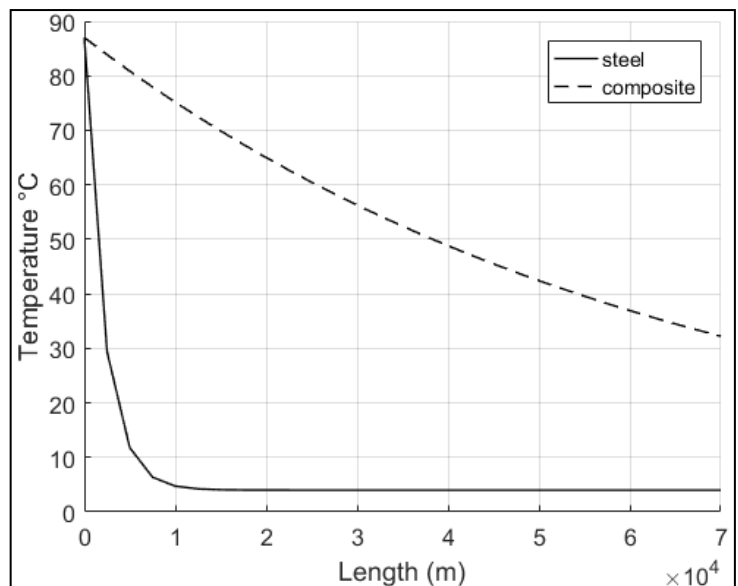


Figure-1: Comparison of SCR vs CCR

Thermal management of the hotspot on the spray cooled surface using jet impingement

Yakang Xia, Ri Li*

School of Engineering
 The University of British Columbia
 Kelowna, Canada
sunny.li@ubc.ca

ABSTRACT

Although spray cooling is able to provide a uniform cooling and has been widely investigated as one of the potential cooling method for compact electronic products. However, a hotspot exists in the central region of the spray cooled surface. The high temperature in the center location and the large temperature gradients degrade the performance, reliability, and lifetime of the device severely. In this work, we propose and test a combined jet and spray cooling method to lower the hotspot temperature and the overall temperature gradient.

A thin film heater (20mm×20mm) on an aluminum nitride substrate (AlN) with high heat flux and uniform temperature was designed and fabricated. The temperature distribution on the lower surface of the substrate, which has been painted to be black with an emissivity of 0.96, is acquired by using an infrared (IR) camera (SC660, FLIR). The optical resolution of the IR images is around 20.7pixels/mm², which provides more detailed and reliable temperature than local temperature measurement by thermocouples and the resistance temperature detector.

As observed from the detailed temperature distribution across the surface of the heated substrate with the aid of the infrared photography, a hotspot appears in the center of the heated surface during the spray cooling because of the velocity and volume flux characteristics of the spray itself, as shown in Fig. 1(a). To remove the hotspot and lower the temperature on the surface, an auxiliary jet is utilized to directly impinge onto the central region, as illustrated in Fig. 1(b), where H is defined as the distance from the spray nozzle tip to the upper surface of the AlN substrate, and α is the inclination angle of the spray relative to the vertical direction.

The combination of the spray and a jet lowers the maximum temperature located in the central region. At the same time, the average temperature and the temperature gradient are decreased as well compared to the case of spray only cooling, as shown in Fig.1 (c).

$$\delta R = \frac{1}{2Aq''} \int ((T_s - T_{avg}) + |T_s - T_{avg}|) dA \quad (1)$$

Non-uniformity thermal resistance was defined as Eq. (1), where q'' is the heat flux (W/m²), T_s is the local surface temperature (K), T_{avg} is the average surface temperature (K) and A is the area of the heater (m²). Results indicate that the heat transfer coefficient h is higher than the spray only cooling case with the same water flow rate and the combined jet and spray cooling technique gives a more uniform temperature distribution and much higher CHF.

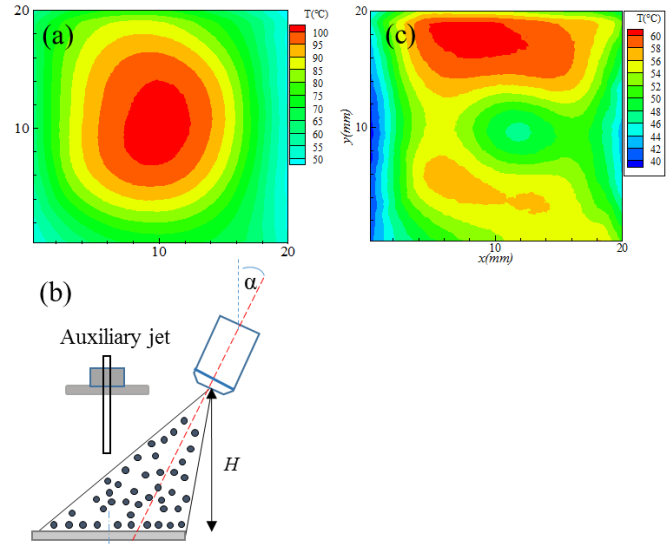


Figure 1. (a) Temperature distribution of spray cooling only, (b) proposed jet and spray combination cooling method, (c) temperature distribution of combined jet and spray cooling

A FAST TWO-WAY COUPLED INTEGRAL APPROACH FOR IMPINGING JET HEAT TRANSFER

Yunpeng Wang, Roger E. Khayat
 Mechanical & Materials Engineering
 Western University
 London, Canada
 ywan988@uwo.ca

Abstract— A non-iterative two-way coupled semi-analytical method was designed to evaluate the performance of liquid jet impingement cooling of solid surface heated with constant heat flux. With a modified version of the integral approach, we are able to account for the variation of fluid viscosity with temperature without the need of iterative schemes. Maintaining the same of level of accuracy, the computational time is much shortened, and the danger of divergence induced by numerical nonlinearity is also successfully avoided. The effects of two-way coupling on the hydrodynamic and thermal boundary layer, the wall temperature and the Nusselt number are reported and discussed. We also validate our method by comparing the current results with existing experiment and iterative solutions, and find reasonably good agreements.

Keywords—liquid jet; impingement cooling; two-way coupled method; integral approach

I. INTRODUCTION

Liquid jet impingements are widely used and studied as it offers very low thermal resistances for cooling purposes and is relatively simple to implement. Its industrial applications include cooling of combustor wall and turbine blades, quenching of steel, rapid cooling involved in glass manufacturing and surface cleaning, etc. The first major contribution in impinging jet cooling is made by Chaudhury [1]. He solved the energy equation for the flow field of impingement jet immediately after Watson [2] solved the momentum equation following Watson's similarity approach. It should be noted that Chaudhury's similarity method is only applicable for the constant wall temperature condition. However, most practical applications are operated under heat flux condition (see Fig. 1). Indeed, temperature condition is more artificial as it requires feed-back control system. It is also important to note that Chaudhury's similarity solution assumes constant fluid properties as most of later theoretical studies do.

We note that while the density, heat capacity and thermal conductivity for most liquids remains in the same order of magnitude (i.e. small variation) for most liquids, it's a fact that

viscosity of liquids decrease moderately or even significantly with temperature as a result of the decrease of the cohesive forces between liquid molecules [3]. Being the mostly used liquid, for instance, water has a kinematic viscosity of 1.79 cSt near the ice point but drops to 0.29 cSt when reaching the boiling temperature. The viscosity of other non-metallic liquids can have even larger variations of multiple orders of magnitude [4].

Though most theoretical studies and lots of numerical works neglected the temperature dependence of liquid viscosity, Liu and Lienhard [5] adopted a coupled integral approach to solve the energy equation. They obtained the thermal boundary layer thickness using existing result for velocity boundary layer thickness [6]. To account for the change of viscosity, the designed a delicate iterative algorithm and achieved good agreements with their experiments.

To the author's knowledge, the non-iterative two-way coupling is still very rare is there are any. In this context, we developed a coupled integral approach which is capable of accounting for the variation of fluid viscosity and also free from iteration. This approach inherently eliminates the potential divergence with iterative schemes and maintains the same level of accuracy as we shall see.

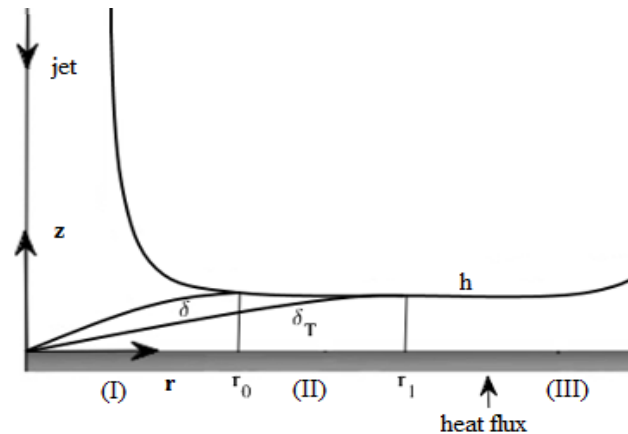


Figure 1. Illustration of circular hydraulic jump.

II. PHYSICAL DOMAIN AND PROBLEM STATEMENT

As illustrated in Fig. 1, we consider a laminar jet of radius a and temperature T_f impinging at a volume flow rate Q on a circular disk, which is heated with uniform heat flux q . We note here dimensionless variables and parameters are used in Fig. 1. The length and velocity scales are taken to be jet radius and average velocity of the incoming jet. The heat flux is scaled by kT_f/a and temperature by T_f . The hydraulic jump region is not of consideration in the current work, and the hydrostatic pressure caused by gravity is neglected as the variation of the liquid film is assumed to be gradual.

In the current problem, we identify three distinct flow regions: a developing boundary-layer region (I), where both the velocity boundary layer and the thermal boundary layer are below the free surface and grow until the former one reaches the film surface at r_0 , and a fully viscous region (II), where the viscous effect has already reached the free surface but the thermal boundary layer is still below it. Then region (III) will be identified when the thermal boundary layer reaches the free surface at r_1 . We note that the hydrodynamic boundary layer grows faster than the thermal boundary layer since the kinematic viscosity, for most non-metallic liquids, is larger than the thermal diffusivity (i.e. Prandtl greater than unity). The velocity and temperature outside the corresponding boundary layers remain almost constant, equal to that of the impinging jet. As convection is dominant, we neglect heat losses at the free surface.

Throughout this study, the stagnation region under the impinging jet will not be considered since the distance from the stagnation point for the inviscid flow to reach uniform horizontal velocity is small [2]. In this case, we assume that both the hydrodynamic and thermal boundary start to develop from the center of the disk. To account for the change of viscosity with temperature. We need to choose a proper viscosity model. Among varieties of viscosity models available in the literature, the classical one proposed by Fulcher [7] was commonly recommended as it gives fairly satisfying fits with experimental data [8] except at very low temperature [9]. Fulcher's model is commonly known as Vogel–Tamman–Fulcher (VTF) or Vogel–Fulcher–Tamman (VFT) equation. The kinematic for kinematic viscosity using VFT equation can be written in dimensionless form (scaled by a reference viscosity ν_0 at a certain temperature) as

$$\nu = A \times 10^{B/(T-C)}, \quad (1)$$

with A , B and C being the fitting parameters from experiments. The VFT equation will be directly introduced to our formulation.

III. THE TWO-WAY COUPLED APPROACH

As the length scale liquid layer in the horizontal direction is much larger than the scale in the vertical direction, we can approximate the problem using boundary layer equations [10]. Consequently, the dimensionless governing equations for mass, radial momentum and energy equations are simplified to

$$\frac{\partial u}{\partial r} + \frac{u}{r} + \frac{\partial w}{\partial z} = 0, \quad (2)$$

$$Re \left(u \frac{\partial u}{\partial r} + w \frac{\partial u}{\partial z} \right) = \frac{\partial}{\partial z} \left(A \times 10^{B/(T-C)} \frac{\partial u}{\partial z} \right), \quad (3)$$

where $Re = Q/(\pi a \nu_0)$ with ν_0 being a reference kinematic viscosity at a certain temperature (20°C) and

$$Pe \left(u \frac{\partial T}{\partial r} + w \frac{\partial T}{\partial z} \right) = \frac{\partial^2 T}{\partial z^2}, \quad (4)$$

where $Pe = Q \rho C_p / (\pi a k)$ with ρ being density, C_p being heat capacity and k being thermal conductivity. In addition, as the total volume flow rate of liquid is conserved, we have

$$\int_0^h u dz = \frac{1}{2r}. \quad (5)$$

In region (I), the velocity and temperature profiles are given by

$$u = \frac{3}{2\delta} z - \frac{1}{2\delta^3} z^3, \quad (6)$$

$$T = \frac{2q\delta T}{3} + 1 - qz + \frac{q}{3\delta^2} z^3. \quad (7)$$

Here (6) is designed in a way such that the no-slip, no-shear condition are satisfied at the bottom and the free surface. In addition, another condition is applied so that the velocity at the boundary layer edge is the same as the free stream. Similarly, profile (7) is designed to satisfy the heat flux condition at the wall as well as the zero heat loss and matching condition at the thermal boundary layer edge. By integrating (3) and (4) between the disk surface and the upper edge of the corresponding boundary layers, and using (5), (6) and (7), we obtain

$$h - \frac{3}{8}\delta = \frac{1}{2r}, \quad (8)$$

$$\frac{d\delta}{dr} = \frac{140A}{13 Re \delta} 10^{\frac{B}{\frac{2q\delta T}{3} + 1 - C}} - \frac{\delta}{r}, \quad (9)$$

and

$$\frac{d\delta_T}{dr} = \frac{Pe\delta_T^3 r \left(14\delta_T^2 - 3\delta_T^2\right) \frac{d\delta}{dr}}{Pe\delta_T^2 \delta r \left(42\delta_T^2 - 5\delta_T^2\right)} + \frac{-14Pe\delta_T^3 \delta_T^3 + Pe\delta_T^5 + 140\delta_T^4 r}{Pe\delta_T^2 \delta r \left(42\delta_T^2 - 5\delta_T^2\right)}. \quad (10)$$

Here (8), (9) and (10) will be solved simultaneously to obtain the boundary layers and liquid thickness. The transition location from region (I) to region (II) is found by equating the hydrodynamic boundary layer thickness to the total liquid height.

In region (II), the free surface velocity U start to change with radial distance, and we assume u/U takes the same form as in the boundary layer region, so that

$$u = U \left(\frac{3}{2h} z - \frac{1}{2h^3} z^3 \right). \quad (11)$$

where U is determined from (5). Here the temperature profile stays the same as (7) since the thermal boundary layer is yet below the free surface. Similarly, the integral equations in region (II) are given by

$$\frac{dh}{dr} = \frac{525Ar}{136Re} 10^{\frac{2q\delta_T}{3} + 1 - C} - \frac{h}{r}. \quad (12)$$

$$\frac{d\delta_T}{dr} = \frac{4(\delta_T^3 - 7h^2\delta_T)}{5h\delta_T^2 - 42h^3} \frac{dh}{dr} + \frac{175rh^4}{Pe\delta_T^2(5\delta_T^2 - 42h^2)}. \quad (13)$$

Similarly, the transition location between region (II) and (III) by equating the thermal boundary layer thickness to the liquid height.

In region (III), both the velocity boundary layer and the thermal boundary layer have reached the free surface. The velocity profile is unchanged, whereas the temperature profile now becomes

$$T = \frac{2qh}{3} + T_s - qz + \frac{q}{3h^2} z^3. \quad (14)$$

In this case, the pair of integral equations in region (III) are given by

$$\frac{dh}{dr} = \frac{525Ar}{136Re} 10^{\frac{2qh}{3} + T_s - C} - \frac{h}{r}. \quad (15)$$

$$\frac{dT_s}{dr} = -\frac{26q}{175} \frac{dh}{dr} + \frac{2qr}{Pe}. \quad (16)$$

Throughout the current work, the Nusselt number, which reflects the strength of convection over conduction is defined, using dimensionless variables by

$$Nu = \frac{q}{T_w - 1}. \quad (17)$$

IV. RESULTS AND DISCUSSIONS

Fig. 2 illustrates the influence of wall heat flux on the boundary layer height and film thickness. As we can see, the thermal boundary layer seems to be always able to reach the free surface. This seems contradictory to the finding of Liu and Lienhard [5], which finds that the thermal boundary layer would never reach the free surface when the Prandtl number is above 4.859. However, this is not truly a contradictory finding as the critical Prandtl number 4.859 is based on constant viscosity, whereas as the viscosity in fact varies significantly with temperature. Therefore, this critical value obtained using constant Prandtl number is not always practical especially when the heat flux is relatively large. This finding is also reasonable in the sense that the temperature of the liquid will continue to increase as it travels downstream. So that the thermal boundary layer has to eventually reach the free surface. It is noteworthy, however, that this behavior is only valid in the absence of hydraulic jump. As the hydraulic jump may occur before the thermal boundary layer reaches the free surface, more detailed treatments are necessary at the jump in that case, which is not included in our present work.

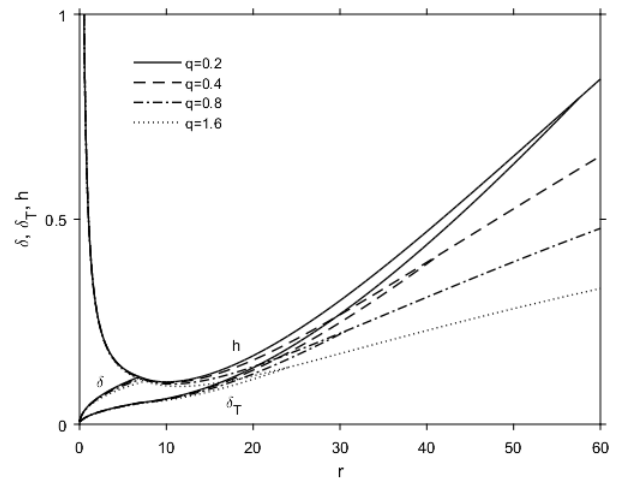


Figure 2. Physical domain and grid configuration ($Re=4000$).

In any practical applications with a constant heat load, it is the temperature at the solid surface that needs to be controlled and monitored, thus it's crucial to investigate the influence of heat flux on the local wall temperature as shown in Fig. 3. We observe that the wall temperature is very close to that of the incoming jet near the impingement point and gradually increases with radial distance as a result of the continuous heat absorption from the solid wall. It is therefore a clear indication that cooling performance far away from the impingement point is not as good as in the near region. The poor cooling efficiency downstream can be also explained as rise of the liquid temperature with radial distance due to heat accumulation illustrated in Fig. 4, where the surface temperature of the liquid is plotted against the radial distance under different wall heat fluxes. We observe that the free surface temperature starts to increase after the thermal boundary layer reaches the air-liquid interface. Also, for a higher liquid temperature, the wall temperature has to increase to maintain the same level of temperature gradient when the flux heat flux stays constant.

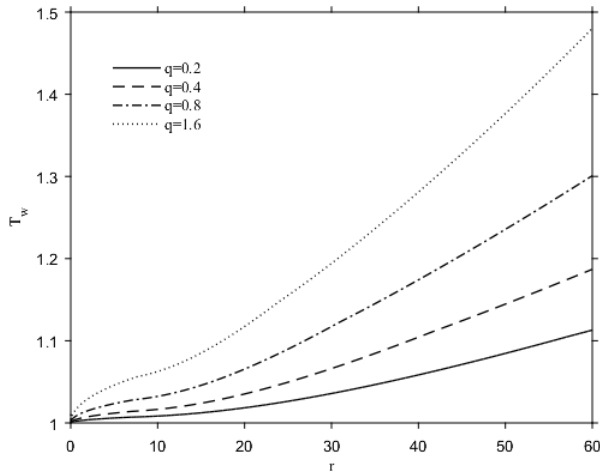


Figure 3. The influence of wall heat flux on the wall temperature ($Re=4000$).

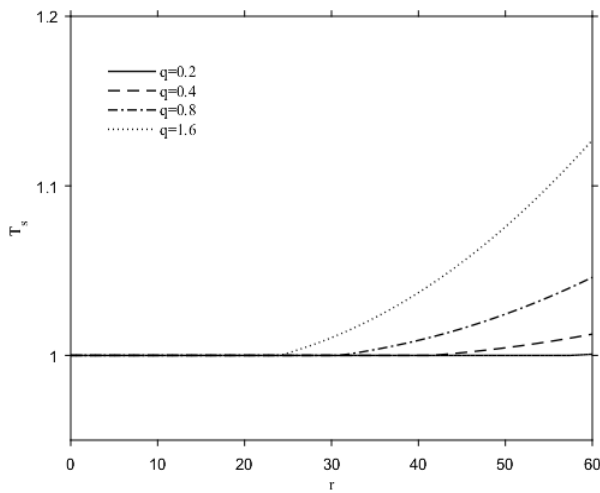


Figure 4. Influence of the wall heat flux on the free surface temperature ($Re=4000$).

It was traditionally understood that convective and conductive heat transfer co-exist in impingement cooling while the former prevails with high inertia. As inertia can dramatically decrease with radial distance as the result of accumulating viscous effect, the Nusselt number, which reflects the strength of convection over conduction, behaves in a similar manner demonstrated in Fig. 5. As can be seen, the convection is much stronger upstream whereas conduction gradually takes over as the flow decelerates. Interestingly, the convective heat transfer is enhanced by wall heat flux when Re stays the same, which is largely due to the decrease of viscosity with temperature (i.e. liquid flows faster).

Fig. 6 depicts the influence of Reynolds number on the Nusselt number. Also shown in the figure is the corresponding experiment [5] under the same condition. The value of applied dimensional heat flux is 12378 W/m^2 . We can see that the agreements become worse for high Reynolds number. Apart from the inaccuracy in the cubic profile, another cause of the disagreement with high Reynolds number might be the initial development of turbulence whereas the modeling is based on laminar flow assumption.

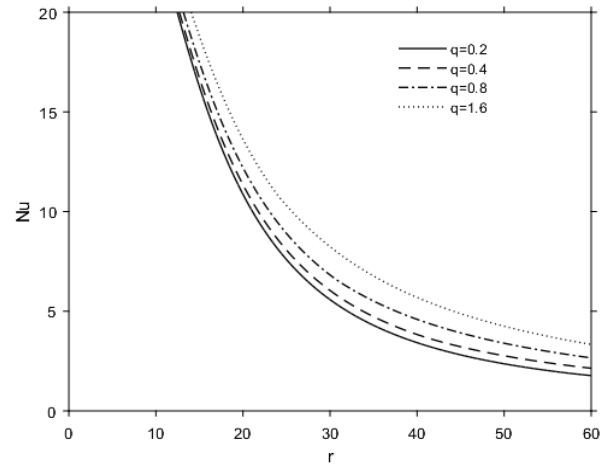


Figure 5. Influence of wall heat flux on the Nusselt number ($Re=4000$).

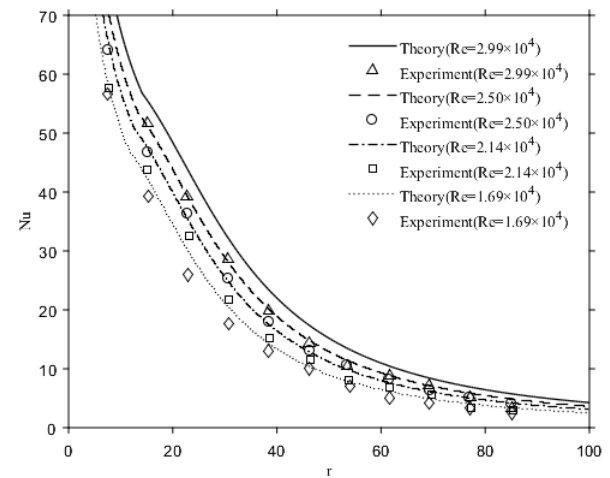


Figure 6. Influence of Reynolds number on the Nusselt number.

V. CONCLUDING REMARKS

In this theoretical study, we use an two-way coupled integral approach to solve the heat transfer problem of a circular liquid jet impinging onto a uniformly heated disk by incorporating the temperature-dependent viscosity model into the momentum equation. With this approach, we can avoid the potential risk of divergence during the numerical scheme in literature. We find the previously documented critical Prandtl number above which thermal boundary layer will never reach the free surface does not always provide us with practical use due to the constant viscosity assumption. According to our results, the thermal boundary layer will always reach the free surface regardless of any parameter involved in the case of constant wall heat flux. The disk surface temperature and the free surface temperature both increase with wall heat flux as well as radial distance. The local Nusselt number are calculated and reasonable agreements are achieved when comparing with existing experiments. We emphasize that though this method may not be as accurate as iterative approach, the high computational efficiency should be desirable when quick evaluations are needed especially when large temperature variation is present.

REFERENCES

- [1] Z. H. Chaudhury, "Heat transfer in a radial liquid jet," *J. Fluid Mech.*, vol. 20, pp. 501-511, 1964.
- [2] E. Watson, "The spread of a liquid jet over a horizontal plane," *J. Fluid Mech.*, vol. 20, pp. 481-499, 1964.
- [3] P. K. Kundu, I. M. Cohen, and D. R. Bowling, *Fluid Mechanics*. London, UK: Elsevier, 2016.
- [4] C. J. Seeton, "Viscosity-temperature correlation for liquids," *Tribology Letters*, vol. 22, No.1, pp. 67-78, 2006.
- [5] X. Liu and J. Lienhard, "Liquid jet impingement heat transfer on a uniform flux surface," *ASME Heat Transfer Phenomena in Radiation, Combustion, and Fires*, vol. 106, pp. 523-530, 1989.
- [6] A. Sharan, "Jet-disc boiling: burnout predictions and application to solar receivers," Master's Thesis, University of Houston, 1984.
- [7] G. S. Fulcher, "Analysis of recent measurements of the viscosity of glasses," *J. Am. Ceram. Soc.*, vol. 8, No. 6, pp. 339-355, 1925.
- [8] J. C. Mauro, Y. Yue, A. J. Ellison, P. K. Gupta, and D. C. Allan, "Viscosity of glass-forming liquids," *PNAS*, vol. 106, No. 47, pp. 19780-19784, 2009.
- [9] G. W. Scherer, "Editorial comments on a paper by Gordon S. Fulcher," *J Am Ceram Soc.*, vol. 75, pp. 1060-1062, 1992.
- [10] H. Schlichting, *Boundary-layer theory*. New York: Springer, 2000.

DEVELOPMENT OF A NUMERICAL MODEL FOR HYDROTHERMAL CONVERSION OF BIOMASS

Mohammad Heidari, Animesh Dutta

School of Engineering,
University of Guelph,
Guelph, Canada
mheidari@uoguelph.ca
adutta@uoguelph.ca

Bishnu Acharya

School of Sustainable Design Engineering,
University of Prince Edward Island,
Charlottetown, Canada
bacharya@upei.ca

ABSTRACT

Hydrothermal carbonization (HTC) is an up and coming approach to produce a value-added solid product (hydrochar) from various biomass feedstocks, especially when it comes to biomass that has a high moisture content. Although many research works have been performed for HTC, there is still a need to replace this process from laboratory setups to industrial scales. To date, there is no reliable numerical model that can consider the heat transfer, kinetics, and the reactions of the porous structure of biomass and water due to complex reaction mechanisms. This study aims to produce such a numerical model by obtaining required experimental data and by considering the heat and mass transfer in the porous media and a representative reaction of the whole process. A batch reactor setup is utilized as to perform HTC experiments on sawdust. The temperature inside the reactor and power consumption were monitored by OMEGA data logger and FLUKE energy analyzer respectively. The obtained products from HTC experiment were characterized to find required data which can represent the chemical reactions. Based on the data found by experiments and from literature, a comprehensive heat and mass transfer model was developed using chemistry, heat transfer, and transport of diluted species in porous media modules in COMSOL Multiphysics software. The comparison of the results from the model with experimental data shows that the model can be used for predicting the mass yield, the real energy consumption during the process, and time and amount of the power released by exothermic reactions. Furthermore, the simulation can be used to assess the effects of varying different parameters to find out the best operational conditions for performing HTC.

THE EFFECT OF GEOMETRY ON THE STRENGTH, STORAGE AND COMBUSTION CHARACTERISTICS OF BIOCHAR AND HYDROCHAR PELLETS

Trishan Deb Abhi, Animesh Dutta

School of Engineering
University of Guelph
Guelph, Canada
tabhi@uoguelph.ca
adutta@uoguelph.ca

ABSTRACT

With a view to mitigating greenhouse gases (GHGs) and other toxic emissions, lignocellulosic biomass has emerged as a popular option among other renewable energy sources due to its potential to produce environmentally sustainable clean energy. However, the inferior physicochemical properties of raw biomass make it not suitable as a direct replacement of coal. In this study, two different biomass thermochemical conversion processes namely hydrothermal carbonization (HTC), and torrefaction, followed by densification process had been performed. The effect of the different geometry of densified pellets had been studied for determining their feasibility in terms of their strength, storage, transportation, and combustion characteristics and for finding out the most suitable pellet geometry for energy storage purpose. Although, both thermochemical processes gave satisfactory results for upgrading biomass properties, the HTC pellets showed better physicochemical properties than the raw and torrefied pellets. Moreover, the mass density, the volumetric energy density, hydrophobicity, durability, and the carbon content of the HTC pellets increased significantly. Results also showed that the ash and the oxygen content of HTC pellets decreased dramatically, improving the combustion characteristics of the biomass. In addition, the durability, compression strength, and bulk density increased with decreasing the pellet size. In summary, this study revealed that HTC pellets can be used as a potential substitute of coal in coal-fired industries without further modification in the existing infrastructure.

NUMERICAL INVESTIGATION OF THE EFFECTS OF COKE ON THE TRANSPORT PROPERTIES IN AN ETHANOL FUEL CELL REFORMER

Precious Arku, Animesh Dutta

School of Engineering
University of Guelph
Guelph, Canada
parku@uoguelph.ca

ABSTRACT

The use of solid oxide fuel cells to replace internal combustion engines in automobiles has increased in recent years. However, the performance and stability of catalysts in the presence of carbon is key for the commercial success of fuel cell reformers. In this paper, finite element method was used to study the effect of coke deposition on heat and mass transfer during the catalytic partial oxidation of ethanol in a packed bed reactor. A mechanism for the formation of syngas from ethanol was proposed, while kinetic parameters were taken from the literature. The model showed that an oxygen/carbon (O/C) ratio of 0.25 resulted in a conversion of less than 36 % and this conversion increased with increase in O/C ratio until a maximum of 100 % was achieved at O/C = 1.0. The properties of Ni/Al₂O₃ catalyst bed were investigated after subjection to several hours of carbon buildup. Bed permeability, porosity and temperature distribution were significantly affected after just 1500 s of reaction time. It was also observed that void fraction and permeability became non-uniform across the bed. After 12000 s on stream, the void fraction at the reactor inlet reduced from an initial value of 0.5 to 0.1 while that at outlet reduced to 0.32. The porosity and temperature decreased with axial position and the difference became more pronounced with time. A decrease in bed porosity reduced the bed temperature due to an increase in effective thermal conductivity and ethanol conversion and hydrogen selectivity decreased as a result. Thus, it was concluded that heat transfer becomes a limiting factor in reforming reactions in the presence of carbon. Production distribution prior to deactivation was also studied and it was observed that a maximum ethanol conversion of 100% was achieved at 600 °C and C/O ratio of 1.0. Finally, results from the reactions were compared with that of a different study to validate the reaction mechanism and similar results were found in the literature. It was inferred that the proposed reaction mechanism in this model is able to predict the conversion and selectivity of the main products at high temperatures. However, a limitation is that it is not able to account for the formation of acetaldehyde at low temperatures. Thus, it is possible that the average deviation of selectivity values will be significant at lower reaction temperatures. Nonetheless, results from this model further clarify why catalyst performance decreases with time in a reformer.

The impacts of exterior surface convective heat transfer coefficients on high-rise building energy consumptions

Meseret T. Kahsay

Department of Civil & Environmental
Eng./WindEEE Research Inst.
University of Western Ontario
London, ON, Canada
mkahsay@uwo.ca

Girma T. Bitsuamlak

Department of Civil & Environmental
Eng./ WindEEE Research Inst.
University of Western Ontario
London, ON, Canada
gbitsuam@uwo.ca

Fitsum Tariku

Building Science center of Excellence
British Columbia Institute of Technology
Burnaby, BC, Canada
Fitsum_tariku@bcit.ca

ABSTRACT

Energy analysis of low-rise buildings can be managed easily by many Building Energy Simulation (BES) tools; however, high-rise buildings pose some fundamental challenges for the current state of the art BES tool. Some of the challenges are the scale of the building, the microclimate changes with altitude, and the uncertainty of the existing convective heat transfer coefficient (*CHTC*) correlations. To tackle some of these challenges, integrated Computational Fluid Dynamics (CFD) and BES tools such as EnergyPlus is explored. First, the micro-climate interaction with the study building is investigated. Specifically, the wind effect variation with altitude is studied using CFD to develop new-*CHTC* correlations. Then these new-*CHTCs* and the existing correlations are compared to quantify the impact of the environmental factor on building annual energy consumptions using the BES tool. A typical 100 m height isolated building having 100% WWR located in Boston, MA, is considered as a case study. The result shows that on average the annual heating energy consumption would have been overestimated in the order of 10% and the annual cooling energy consumption overestimated in the order of 5%.

Keywords: *High-rise building, computational fluid dynamics, convective heat transfer coefficient, building energy simulation, EnergyPlus, turbulence*

CHARACTERIZATION OF TRANSIENT FLOW VELOCITIES DURING MELTING OF PCM IN A UNIFORMLY HEATED CIRCULAR CAVITY

Kyle Teather, Kamran Siddiqui
Department of Mechanical and Materials Engineering
Western University
London, Canada

ABSTRACT

According to projections, it is predicted that by 2040 global energy usage will increase by one-third of the current energy demand [1]. With technology advancing at the current rate and without changes in government policies, it is projected that about 75% of that energy produced in 2040 will be supplied by fossil fuels [1]. In order to minimize damage to the environment and ensure sustainability, it is necessary to increase energy conservation as well as decrease reliance on fossil fuels. An attractive solution that can be applied to both necessities is energy storage in the form of heat; for example, heat that would otherwise be wasted can be conserved for use, or energy could be stored from renewable sources in times of availability until times of high demand to help remedy their intermittent nature of production.

Methods of storing energy using sensible heat have been implemented, but thermal storage devices that make use of the latent heat of phase change are more effective than sensible-heat based devices. Using latent heat greatly increases the thermal capacity of the device due to the exploitation of the enthalpy of phase change, as well as allows charging and discharging of heat within a narrow temperature range since temperature remains constant during phase change. The materials used in latent heat storage are known as *phase change materials* (PCM). The challenge with PCM is its inherent resistance to heat transfer: these materials have low thermal conductivity, which is compounded by solidification at heat transfer surfaces during discharge. Many effective methods of increasing heat transfer to PCM involve maximizing heat transfer area. One of these methods involves using a PCM within a foam made of a material with a high thermal conductivity. The exact mechanisms of heat transfer at the pore scale, however, are not well understood; this inhibits the ability to efficiently design systems that make use of this material. There is another common method of implementing PCM called microencapsulation, which involves packaging the PCM in a small sphere to create a simple, self-contained thermal storage unit. Since these capsules share similar geometry with the porous media composite described above, their functionality could also be optimized with detailed knowledge of phase change behavior inside the encapsulation. Therefore, there is a need to characterize the phase change behaviour of a PCM in this geometry to improve numerical models as well as to improve system design.

This study aims to characterize these mechanisms of heat transfer in a circular geometry. A detailed experimental investigation was conducted to examine the transient melting behaviour of a PCM in a thin circular cavity to emulate the cross-section of a sphere. This thin circular cavity filled with PCM was heated uniformly through the walls via water jacket that provided uniform heating, and the front and back surfaces were insulated. Rubitherm RT26 was the PCM used, a commercial PCM with a melting temperature of 26 degrees Celsius. Thermocouples and a thermal camera were employed to measure heat loss from the apparatus in order to quantify the heat flux to the PCM. Particle image velocimetry (PIV), an optical technique used to capture velocity data with high spatial resolution, was used to capture the flow velocities in the melted region at different times during the phase change. Three different tests were conducted at different magnitudes of applied heat flux.

Results show that melting proceeds in stages. In the beginning, the PCM is entirely solid and heat transfer is by pure conduction. A uniform layer of melted PCM forms around the boundary, until it is thick enough that the buoyancy forces begin to overcome the viscosity of the fluid. At this point, vortices are formed at the left and right sides that enhance the heat transport from the wall to the PCM. These vortices induce a sloped region to form in the upper region of the PCM. Once the top region of the PCM has melted, leaving an elliptical region of solid PCM isolated in the center, the flow appears to become stratified. The flow velocities slow down, and the strength of natural convection weakens. The shape of the melt front is consistent as heat flux increases; however, the flow becomes more energetic and melting proceeds to completion in less time. Further work is planned to capture the temperature distribution in the volume during melting. More detailed results including temperature fields within the PCM will be presented and discussed.

[1] BP (British Petroleum), 2018, "BP Energy Outlook: 2018 Edition", pp. 15,68

EXPERIMENTAL STUDY OF NATURAL CONVECTION IN A HORIZONTAL UNIFORMLY HEATED CYLINDER

Kyle Teather, Kamran Siddiqui
Department of Mechanical and Materials Engineering
Western University
London, Canada

ABSTRACT

Fluid-filled horizontal cylinders subjected to heat transfer through the pipe wall have applications in nuclear and food processing industries. Heat transfer in such geometries is driven by buoyancy-induced convection. Very limited research work has been reported in the literature in this area where the heat transfer was investigated under different thermal boundary conditions that varied from differential heating and cooling, an isothermal boundary to constant heat flux. The experimental studies were mainly restricted to the temperature measurements, while numerical and analytical studies used simplified models to approximate heat transfer. There is a scarcity of studies providing a detailed structure of temperature and velocity fields. Characterizing the flow behaviour in such geometry would serve to further the understanding of the fundamental physics of convective flow and heat transfer in this configuration, as well as offer improved knowledge to optimize the associated industrial processes.

A detailed experimental investigation was conducted with the objective of characterizing the convective flow behaviour in a circular geometry subjected to isothermal wall heating. A thin circular cavity was constructed from acrylic, which served as the control volume once filled with water. A water jacket surrounded the radial walls of the cavity, and the front and back surfaces were insulated. Hot water was pumped through the water jacket from a reservoir fitted with a temperature control system. Thermocouples were fitted to monitor the inlet and outlet temperatures of the water jacket, as well as the bulk water temperature near the cavity surface and the surface temperature at the top and bottom. A thermal camera was employed to monitor the surface temperature of the common front face shared by the cavity and water jacket to collect data about the thermal gradient. Particle image velocimetry (PIV), an optical, high-resolution velocity measurement technique was used to capture the flow velocities once the system reached steady state. The water inside the cavity was seeded with near-neutrally buoyant hollow glass spheres with a mean diameter of 15 μm . This seed particle was then illuminated with laser light, and the instantaneous position of the seed particles was captured using a synchronized camera. These images, taken with a known frequency, were then analyzed using a cross-correlation technique to obtain velocity fields. Experiments were conducted at four different inlet temperatures spanning 30 to 60 degrees Celsius.

The analysis of acquired PIV data shows the cavity can be separated into three regions based on flow behaviour. In the first region, adjacent to the wall, the flow is dominated by convection due to heat addition from the boundaries. The fluid rises along the wall toward the top with relatively high velocity. The second region lies in the lower central section of the cavity, where the fluid flow seems to be driven by the shear from the fluid passing along the boundaries. This induces a symmetric pair of counter-rotating vortices – which enhance the transportation of heat. The third region corresponds to the upper region of the core where the warm fluid accumulates, and the stable stratification dampens buoyancy-induced flow leading to a significant reduction in the flow velocities. As a result, conduction becomes the dominant mode in this region. Future work includes the combined analysis of the velocity and temperature data within the cavity to obtain a detailed insight into the underlying physical process to provide a more accurate explanation of these observed trends. Further investigations will also be conducted to characterize the thermo-fluid behaviour in the cavity during cooling. Detailed results will be presented and discussed at the conference.

THE EFFECT OF CHANGING THE MAIN COMPRESSOR INLET TEMPERATURE ON THE OVERALL CYCLE EFFICIENCY OF A 10 MW S-CO₂ GAS TURBINE POWER PLANT

Ibrahim Ali

Department of Mechanical and Aerospace Engineering
 Carleton University
 Ottawa-Ontario, Canada
 e-mail: ibrahimali@cmail.carleton.ca

Henry Saari

Department of Mechanical and Aerospace Engineering
 Carleton University
 Ottawa-Ontario, Canada
 e-mail: henrysaari@cunet.carleton.ca

Abstract—The unique properties of supercritical carbon dioxide (S – CO₂) make it very adaptable, and capable of behaving like gas or liquid. As a result, it is difficult to measure the physical and thermodynamic properties, as they will behave differently depending on the density of the supercritical gas. Increasing the density of the gas to the critical point appears to increase the compressor's efficiency by reducing the work. This paper analyses the effect of the main compressor inlet temperature on the performance of the 10 MW S – CO₂ gas turbine power plant. The design point of the main compressor inlet temperature is 308.15K, the mass flow rate is 70.3kg/m³ and the compressor inlet pressure is 8.55 MPa. However, over a range of main compressor inlet temperatures it is evident that even a slight change in the main inlet temperature significantly affects the cycle performance. The temperature is very sensitive, and any change can have a huge impact on the density of the S – CO₂, which, in turn, affects the performance of the cycle. From 308.15 K to 309.15 K, at a constant pressure of 8.55 MPa the overall cycle efficiency decreases by 0.57%. Since the inlet temperature depends on the performance of the cooler for cycle heat rejection, it is important to monitor the cooling tower performance to deliver ideal water temperature to the cooler.

Keywords- Supercritical CO₂ Brayton cycle; main compressor inlet temperature; turbine; density.

I. INTRODUCTION

Studying carbon dioxide CO₂ is considered very important, due to its interchangeable properties at different temperatures and pressures. As a result, CO₂ was selected because of its adaptable properties, as well as effective thermal stability, low corrosion levels with the appropriate material, availability and low cost; the physical and thermodynamic properties are well researched. Also, the cycle must be operated beyond the critical pressure of CO₂ [1]. This study demonstrated that for 400 MW power plant an increase of 1% in the cycle efficiency raise the electricity product to 86.4 MWh a day and led to a savings of approximately \$0.79 million per year, or \$31.54 million over the plant lifetime (40 years) [2]. To achieve high efficiency in

Brayton cycles, the compressor inlet temperature must be relatively low. Companies such as General Electric, Sandia National Laboratories, research institutes and universities such as Carleton have been conducting studies to develop and improve the S – CO₂ cycles used for generating power. Carleton University is working on a project that involves all the aspects or properties of S – CO₂ gas turbine design, known as the Carleton University Brayton Cycle Loop (CU-BCL), as shown in Figure 1; the CU-BCL design was based on the recompression cycle (US Department of Energy, 2015) [3] [4]. Since it is necessary to adjust the compressor inlet temperature to its design point, the performance of the cooling tower that delivers the correct temperature to the cooler for cycle rejection (i.e. T = 295.15K) is also essential. The latter can significantly affect the performance of the plant's mechanical system [5]. In a Brayton cycle using ideal gas, the lower the compressor inlet temperature the higher the cycle efficiency [3] [6]. It is important to analyze the parameters affecting the performance of the S – CO₂ gas turbine, including the interchangeable turbine inlet temperature, the minimum cycle temperature, maximum pressure and the temperature of the environment, to ensure cycle stability and performance.

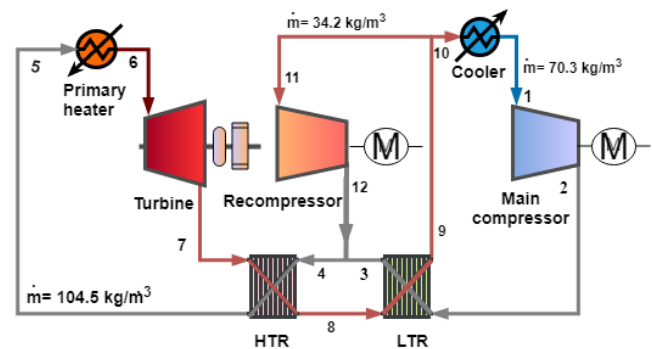


Figure 1. 10 MW S-CO₂ Carleton University Brayton Cycle Loop (CU-BCL) layout.

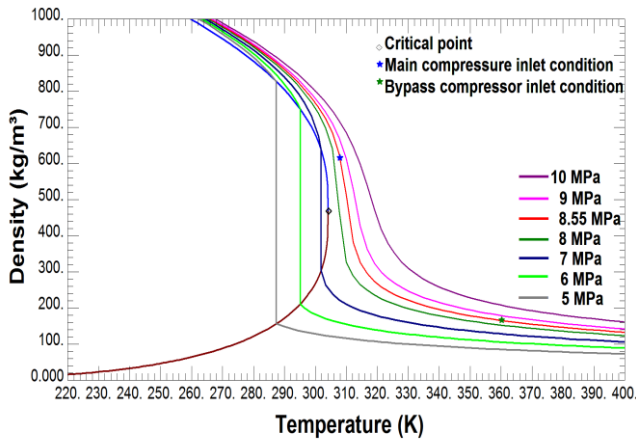


Figure 2. The density rises more rapidly near the critical point with decrease in the temperature

Dostal [6] discusses the importance of parameters in cycle operations, such as the compressor inlet temperature, cycle pressure ratio, compressor outlet pressure and turbine inlet temperature. This paper will further explain the effects of changing the compressor inlet temperature on the cycle operation. As the cycle is re-compression based, the transient code must be able to recognize both compressors operating at the same time. Since the compressors operate differently, it is important to check and monitor the change in CO₂ properties continually. As shown in Figure 2, the main compressor (shown in blue) is approaching the critical point, and the re-compressor, or bypass compressor (shown in green) is approaching ideal gas behavior; thus, the main compressor inlet condition requires more attention. It is obvious that the change in temperature, particularly near the critical point at which the specific heat reaches a peak for the given pressure $P = 8.55$ MPa, also it is difficult to product the properties of supercritical condition due to its fast change at the critical point,

II. 10 MW CU-BCL SUPERCRITICAL CO₂ BRAYTON CYCLE LAYOUT

The data in this paper is from 10MW re-compression cycles (US Department of Energy, 2015) [7]. The compression cycles are designed very precisely, with two compressors functioning identically but with different thermodynamic properties and flow rates; the main compressor mass flow is $\dot{m} = 70.3$ kg/m³ and the bypass compressor is $\dot{m} = 34.2$ kg/m³. Figure 3 displays the 10 MW S-CO₂ CU-BCL T – S diagram and shows that once the fluid passes through the primary heater where it expands upon reaching the turbine inlet as shown in steps 6 and 7, it will produce energy. Subsequently, the heat produced by the hot fluid will be absorbed by the high temperature recopercator (HTR) and the low temperature recopercator (LTR), as shown in steps 7 and 8. The fluid then passes through the LTR and is split into two different streams, as shown in steps 1 and 10. One stream passes through the cooler for heat rejection and from the cooler to the main compressor, as shown in steps 1 and 2. The other stream passes through the bypass compressor where the

fluid compressed up to $P = 23.99$ MPa, as shown in steps 11 – 12. The fluid pressure is increased by the main compressor to preheat the fluid in LTR 2 and 3. Fluid streams at points 3 and 12 are then merged at point 4 before entering the HTR 4 – 5. The energy from the primary heat source is transferred at a high temperature of 1089.15 K to the SCO₂ working fluid, as shown in steps 5 and 6. The main compressor is analyzed more than the re-compressor, as the bypass compressor operates at a temperature and pressure beyond the critical point ($T_{\text{recomp}} = 361.15$ K $P_{\text{recomp}} = 8.69$ MPa) and is easier to design and monitor. To avoid recompressor overheating, the main compressor recycle line can be tied upstream of the cooler, allowing continuous recycling of gas without overheating. For optimal efficiency, the temperature of the main compressor inlet should be as close to the critical point as possible. In addition, the pressure of the entire cycle can be measured to the pressure of the main compressor outlet since it has the highest pressure in the cycle ($P = 24.13$ MPa), and the compressor inlet temperature can be identified as the lowest temperature of the cycle $T = 308.15$ K [2] [8]. Preliminary compressor geometry must first be determined in the design mode, to obtain the best starting point for further evaluation in the analysis mode. A smart condition monitoring system is also very important during transient and steady state operation. The density at the supercritical region, as shown in Figure 2, could affect the tip radius of the compressor, thus the hub/tip diameter ratio must be designed in a moderate way. In his work, Dostal focused on stress due to high power density that could damage the machine and advised that the compressor inlet temperature should be monitored at all times.

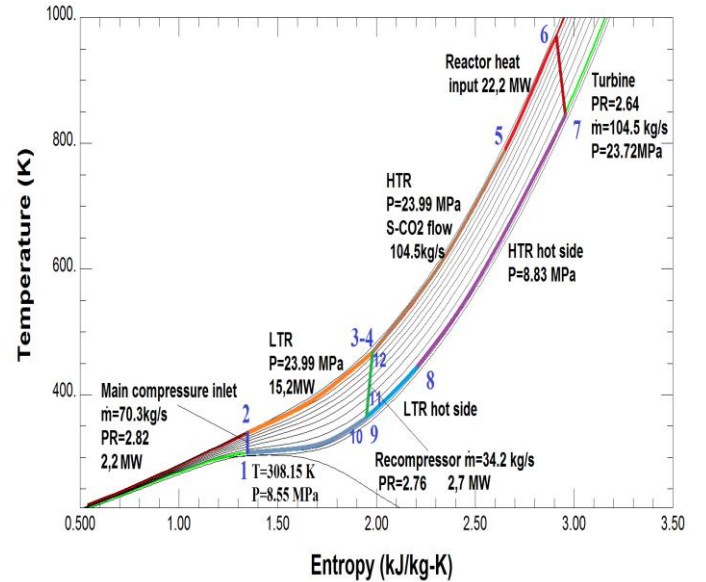


Figure 3. T-s diagram for the 10 MW S-CO₂ gas turbine Brayton cycle.

The Peng-Robinson and Span & Wagner equations of state (EOS) are most commonly applied, and the latter is widely used as it is the most accurate over the years, several EOS have been applied to capture the real gas properties of S-CO₂, and each

has certain limitations for carbon dioxide [9] [10] [11]. Zhao [12] studied six EOS and concluded that the (SW) EOS has the highest precision when calculating the properties of S-CO₂ in the supercritical region. The precise EOS is the basis of REFPROP for CO₂. REFPROP is a commercial tool available from the US National Institute of Standards and Technology (NIST). REFPROP version 10 is used to calculate the CO₂ properties in this paper and is accurate to within 0.03% of the density near critical point with maximum error of 0.2% for the working region of the cycle [2].

III. NUMERICAL ANALYSIS

The specific enthalpy (h_1) and entropy (S_1) of the fluid entering the turbomachine can be determined, since the main compressor isentropic efficiency, inlet pressure and temperature are known: $\eta = 82\%$, $P = 8.55$ MPa and $T = 308.15$ K respectively. The specific enthalpy (h_{s2}) can be calculated at the outlet of the turbomachine if it is isentropically compressed, which is dependent on the inlet specific entropy and the outlet pressure. The isentropic specific work (w_i) can be calculated according to:

$$w_{isent} = (h_1 - h_{s2}) \quad (1)$$

$$W = \dot{m} * w_{isent} \quad (2)$$

And the definition of the isentropic efficiency, the actual specific work (w) for the compressor can be calculated as:

$$w_{comp} = \frac{w_{isent}}{\eta_{isent}} \quad (3)$$

$$w_{main-comp} = \dot{m}(h_2 - h_1) \quad (4)$$

$$w_{re-comp} = \dot{m}(h_{12} - h_{11}) \quad (5)$$

and turbine

$$w_{turbine} = w_{isent} * \eta_{isent} \quad (6)$$

$$w_{turbine} = \dot{m}(h_6 - h_7) \quad (7)$$

$$w_{net} = w_{turbine} - w_{main-comp} - w_{re-comp} \quad (8)$$

The specific enthalpy h_{s2} of the fluid located at the outlet of the adiabatic turbomachine is calculated according to the following energy balance equation:

$$h_{s2} = h_1 - w \quad (9)$$

And the thermodynamic state can be determined at the outlet of the turbomachine from the calculated outlet enthalpy and known outlet pressure [10]. Thus, the integration process is calculated as the final outlet enthalpy h_2 , and used as shown below to calculate the equivalent isentropic efficiency:

$$\eta_{main-comp} = \frac{(h_{2s} - h_1)}{(h_2 - h_1)} \quad (10)$$

$$\eta_{re-comp} = \frac{(h_{12s} - h_{11})}{(h_{12} - h_{11})} \quad (11)$$

$$LTR = \dot{m}_2(h_3 - h_2) = \dot{m}_8(h_8 - h_9) \quad (12)$$

$$HTR = \dot{m}_5(h_5 - h_4) = \dot{m}_7(h_7 - h_8) \quad (13)$$

$$Q_{core} = \dot{m}_5(h_6 - h_5) \quad (14)$$

$$\eta_{turbine} = \frac{(h_6 - h_7)}{(h_6 - h_{7s})} \quad (15)$$

$$Q_{pre-cooler} = \dot{m}_1(h_9 - h_1) \quad (16)$$

$$\eta_{total} = \frac{w_{net}}{Q_{in}} \quad (17)$$

IV. ANALYSIS AND DISCUSSION

When the main compressor is active, the carbon dioxide near the critical point does not follow the ideal gas law, due to the extreme variability properties of S-CO₂ and the non-constant heat capacity (C_p) [13]. As well, the ideal gases have low density and high compressibility, while the S-CO₂ has high density and small compressibility, viscosity and thermal conductivity. Figure 2 shows the significant change in inlet density from 619.06 kg / m³ to 649.55 kg / m³ when the temperature decreases from 308.15 K to 307.15 K. However, if the temperature goes from 308.15 K to 309.15 K, the corresponding decrease in density is 619.06 kg / m³ to 578.7 kg / m³, which can be significant as it will affect the fluid velocity and, consequently, the difference in the outlet temperature becomes 334.87 K to 326.48 K, with the flow rate remaining at 70.3 kg / m³ through the compressor. As the compressor inlet temperature decreases, the system power output increases. The higher density at the compressor inlet results in a higher system flow rate, which increases the power output by the turbine generator by $\dot{m} = \rho * v * A$, where ρ is the density, v the velocity and A is the area. Due to the radical temperature and pressure dependency of specific heat, the temperature difference between the hot and cold fluid varies widely in the recuperator. The specific heat of the cold side flow for LTR is 2.30 kJ/kg-K, which twice that of the hot side flow of 1.15 kJ/kg-K, and this maximizes the 10 MW S-CO₂ gas turbine efficiency. Since the HTR operates with equal mass flow rates on both sides, the minimum value of the temperature difference is not always achieved in the recuperators, which reduces recuperator effectiveness and lowers cycle efficiency [14]. The heat flow noted as Q is relatively proportional to the mass flow rate of CO₂ and the specific heat capacity, provided the pressure and temperature are constant. Figure 4 shows the S-CO₂ property as C_p increases slowly, then abruptly $T = 300$ K the $C_p = 3.537$ kJ/kg-K. However, at $T = 308.15$ K it is equal to $C_p = 8.38$ kJ/kg-K, and at $T = 310$ K C_p reaches a maximum value at $C_p = 15.49$ kJ/kg-K.

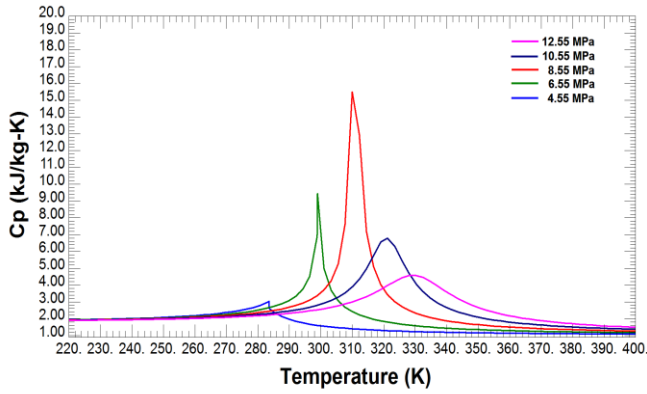


Figure 4. The variation of the specific heat and temperature correspond to the S-CO₂ 10 MW gas turbine analyses.

Then there is a sharp decrease of $C_p = 2.01 \text{ kJ/kg} \cdot \text{K}$ at a temperature of $T = 250 \text{ K}$. In this zone, the flow rate due to a temperature variation of one-unit changes with the CO₂ temperature. This behaviour is apparent when the temperature rises from 250 K to 310 K, and then drops back to 250 K. Such severe fluctuations present significant challenges for modelling thermodynamic properties, and thus must be closely monitored. Figure 5 shows the change in enthalpy near the critical point, before it becomes a plateau. However, with the exception of pressure, none of the other dynamic properties peak close to critical pressure. In addition, the CO₂ conditions can vary from supercritical to subcritical liquid, supercritical to subcritical gas and two-phase liquid gas mixture, all according to the temperature and flow of CO₂ at the main compressor inlet. For this paper, the design point of the compressor outlet pressure is $P = 24.13 \text{ MPa}$ at a temperature of $T = 351.15 \text{ K}$ and a turbine inlet temperature of $T = 973.15 \text{ K}$.

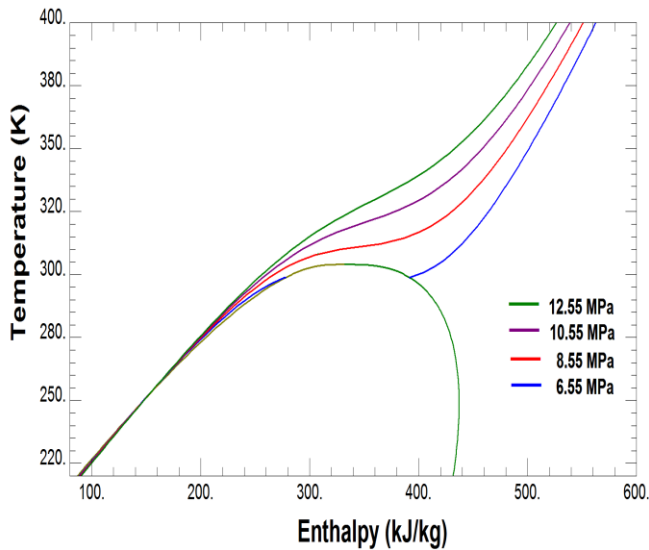


Figure 5. Enthalpy vs temperature for 10 MW S-CO₂ gas turbine case study.

It is noted that the temperature of the main compressor inlet has a major impact on overall cycle operations and compressor performance. Thus, the compression cycle performance is enhanced by maintaining the main compressor inlet temperature at the design point, as higher design point temperatures increase the compression power, which can significantly affect overall cycle efficiency. Conversely, if the main compressor inlet temperature goes below the design point, two phase maybe produced which could cause problem. As shown in Figure 6, the compressor inlet temperature affects the workload demands of the compressors. This is quantifiable, since for every 1 K rise in the main compressor inlet temperature there is a 1% increase in compressor workload, which decreases the effectiveness of the turbine and the generator. The analysis shows that increasing compressor temperature by 1 K leads to a drop in cycle efficiency of 0.57%. Thus, it is important that the pre-cooler design and recirculation of the heat sink is carefully executed to avoid undesirable cycle inlet temperatures and, more importantly, to optimize the economics by maximizing output against operational costs.

The significance of the cooling tower in the cycle is clear, since it increases the performance of the cycle and overall energy efficiency, particularly in warmer weather. Therefore, it is also important to monitor the cooling tower performance as it optimizes the heat transfer in the cycle, as this helps to increase efficiency and delivery of the ideal temperature to the cooler. To maximize compressor efficiency and control the compressor inlet temperature at a set point of $T = 308.15 \text{ K}$ in order to achieve the desired outcome, the flow rate and temperature of the cooling water that flows into the pre-cooler must be controlled. Thus, for maximum efficiency the design of the gas turbine must include high reliability and availability, which is quantified as the percentage of time the turbine is available to generate power in any given period [14] [15]. Analysis highlights the importance of adding a standby circulation pump to the cycle, to ensure optimal water flow and heat transfer. Pumps can fail due to loose connections, damaged bearings, cavitation, clogged strainers or other reasons. The result is reduced water flow to the cooler, decreased efficiency and the subsequent loss of reliability and availability.

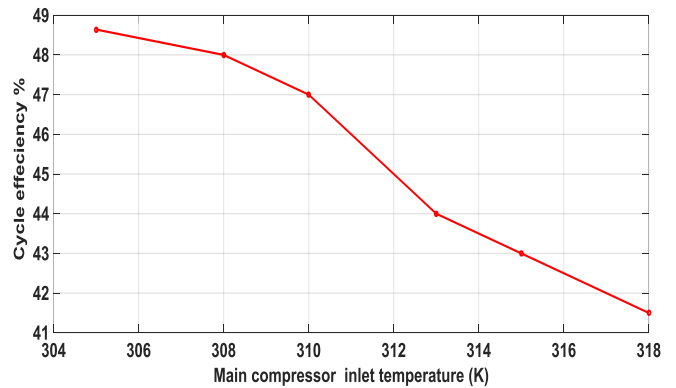


Figure 6. The effect of main compressor inlet temperature on the cycle performance.

V. CONCLUSIONS

Maintaining high cycle efficiency can be achieved by taking advantage of the remarkable property variations of CO₂ near the critical point, which can significantly reduce the workload in the main compressor. The main compressor inlet temperature should be maintained at the design point and at operational transients close to the design point that are above the critical temperature, to prevent two-phase flow inside the main compressor. As cycle efficiency is very sensitive to the compressor inlet temperature, the temperature should be monitored at all times during operations. Figure 2 shows the significant impacts on the density as the inlet temperature changes, and how it can affect cycle efficiency. Thus, by monitoring and maintaining the inlet temperature at the design point $P = 8.55 \text{ MPa}$ $T = 308.15 \text{ K}$ cycle performance will be at highest capacity. After applying equations 1 to 17 to calculate the main compressor inlet temperature effectiveness on overall cycle efficiency, it is clear that any slight change in inlet temperature significantly affects cycle performance. As shown in Figure 6, if the temperature increases from 308.15 K to 310.15 K at a constant pressure of 8.55 MPa, the overall cycle efficiency decreases by 1%. Special attention needs to be paid in cold weather operating conditions that are under the saturation line, as this can cause the CO₂ to condense. The problem is more serious in hot climates and can result in reduced cycle output and investment return compared to much cooler climates for the same level of capital investment and initial operating costs. This can be an unexpected factor in a cooling system during turbine operations.

REFERENCES

- [1] M. Persichilli, A. Kacludis, E. Zdankiewicz, and T. Held, "Supercritical co₂ power cycle developments and commercialization: why s-co₂ can displace steam," Power-Gen India & Central Asia, 2012.
- [2] A. Moiseyev, "Passive load follow analysis of the star-1m and star-h2 systems," Ph.D. dissertation, Texas A&M University, 2004.
- [3] V. Dostal, M. J. Driscoll, and P. Hejzlar, "A supercritical carbon dioxide cycle for next generation nuclear reactors," Ph.D. dissertation, Massachusetts Institute of Technology, Department of Nuclear Engineering, 2004.
- [4] J. R. Elliott and C. T. Lira, *Introductory chemical engineering thermodynamics*. Prentice Hall PTR Upper Saddle River, NJ, 1999, vol. 184.
- [5] B. Buecker, "Proper cooling tower maintenance increases power plant efficiency," Energy-Tech, Tech. Rep., 2008.
- [6] V. Dostal, M. Driscoll, P. Hejzlar, and N. Todreas, *CO₂ Brayton Cycle Design Optimization*. Center for Advanced Nuclear Energy Systems, Massachusetts Institute of Technology, 2002.
- [7] N. W. R. D. R. Ames, *Fundamentals and Applications of Supercritical Carbon Dioxide (sCO₂) Based Power Cycles*, Fossil energy 12. Elsevier Ltd, 2017.
- [8] Y. Wang, G. Guenette, P. Hejzlar, and M. Driscoll, "Compressor design for the supercritical co₂ brayton cycle," in 2nd International Energy Conversion Engineering Conference, 2004, p. 5722.
- [9] R. Span and W. Wagner, "A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1100 K at pressures up to 800 MPa," *Journal of physical and chemical reference data*, vol. 25, no. 6, pp. 1509–1596, 1996.
- [10] J. J. Dyreby, "Modeling the supercritical carbon dioxide brayton cycle with recompression," Ph.D. dissertation, The University of Wisconsin-Madison, 2014.
- [11] F. Jacob, A. Rolt, J. Sebastianpillai, V. Sethi, M. Belmonte, and P. Cobas, "Performance of a supercritical co₂ bottoming cycle for aero applications," *Applied Sciences*, vol. 7, no. 3, p. 255, 2017.
- [12] Q. Zhao, M. Mecheri, T. Neveux, R. Privat, and J.-N. Jaubert, "Thermodynamic model investigation for supercritical co₂ brayton cycle for coal-fired power plant application," in Fifth International Supercritical CO₂ Power Cycles Symposium, San Antonio, TX, Mar, 2016, pp. 29–31.
- [13] C. I. Krau, D. Kuhn, and T. Schulenberg, "Heat transfer phenomena of supercritical fluids," 2008.
- [14] M. P. Boyce, *Gas turbine engineering handbook*. Elsevier, 2011.
- [15] A. Gad-Briggs, T. Nikolaidis, and P. Pilidis, "Analyses of the effect of cycle inlet temperature on the precooler and plant efficiency of the simple and intercooled helium gas turbine cycles for generation iv nuclear power plants," *Applied Sciences*, vol. 7, no. 4, p. 319, 2017.

COMPARING ARTIFICIAL NEURAL NETWORKS (ANN) WITH REFPROP SOFTWARE IN PREDICTING THE CARBON DIOXIDE (CO₂) PROPERTIES AS A WORKING FLUID FOR THE 10MW GAS TURBINE POWER PLANT

Ibrahim Ali

Department of Mechanical and Aerospace Engineering
 Carleton University
 Ottawa-Ontario, Canada
 e-mail: ibrahimali@cmail.carleton.ca

Henry Saari

Department of Mechanical and Aerospace Engineering
 Carleton University
 Ottawa-Ontario, Canada
 e-mail: henrysaari@cunet.carleton.ca

Abstract—The unique properties of supercritical carbon dioxide (S – CO₂) make it very adaptable, and capable of functioning as a gas or liquid. This makes it difficult to measure the physical and thermodynamic properties, since they will vary depending on the density of the supercritical gas. S-CO₂ power cycles have great potential to provide alternative power generation systems that can achieve higher plant efficiency, and one of the most important factors to consider is the purity of the working fluid. Since much of S – CO₂ cycle performance improvement is derived from the physical properties of supercritical CO₂, the cycle efficiency will decrease as the CO₂ concentration decreases. Since the carbon dioxide near the critical point does not follow the ideal gas law, due to the extreme variability properties of S-CO₂, the need of the developed modeling framework provides the consistent performance predictions that are required for further application-specific analyses during the turbine operations. This paper proposes a novel approach that combines an artificial neural network (ANN) and REFPROP software to calculate the working fluid properties of CO₂. To accomplish this, the main compressor inlet temperature is fixed at 308.15K with the isoperity pressure from 0.1 to 20MPa used as input, and the density target measurements taken at 200 points. At T=308.15 K and P= 8.55 MPa, the density of pure CO₂ is $\rho = 619.06\text{kg/m}^3$, a maximum reduction of CO₂ density at T = 308.15K at different pressure values due to impurity is clear: CO₂/Ar, CO₂/O₂, CO₂/CO and CO₂/N₂ are 376.82kg/m³, 347.40kg/m³, 337.50kg/m³ and 326.92kg/m³ respectively. With only 3% impure nitrogen, the density dropped significantly by almost 50%. ANN model is subsequently applied to predict the behaviour of the CO₂, and the data is compared to the REFPROP software data with corresponding ANN predictions. The results are promising, and it is proposed that ANN can be used for an online monitoring system as it displays detection of small changes rapidly and detects changes even at the supercritical region.

Keywords- *artificial neural networks ANN, CO₂ properties, density*

I. INTRODUCTION

Eustace and Merrington (1995) described a neural network as a diagrammatic representation of a mathematical equation that receives values (inputs) and provides results (outputs) [1]. As a black-box methodology, artificial neural networks (ANNs) have become appropriate and powerful tools for data processing, modeling and the control of highly nonlinear systems such as gas turbines. In addition, due to the high demands of the electricity market, power producers support continuous investigation of new methods to optimize the design, manufacturing, control and maintenance of gas turbines. ANNs have previously solved complex problems in the field of gas turbines successfully, and additional work in this area will lead to the optimal design and manufacture of gas turbines with minimal supervision and energy consumption [2]. The National Institute of Standards and Technology (NIST) distributes the Thermodynamic and Transport Properties Database (REFPROP). REFPROP is a program that can calculate the thermodynamic and transport properties of fluids and mixtures. The properties of carbon dioxide are calculated by REFPROP according to the Helmholtz free energy equation of state; more details are available in the literature [3]. The Span-Wanger (SW) equation of state (EoS) has the highest accuracy when calculating the properties of S-CO₂ [4] [5]. Version 10 of REFPROP will be used to generate the data for this study. Artificial neural networks is one of the techniques that have a key role in system identification and modelling industrial systems, since they are capable of capturing dynamics of systems without prior knowledge of their intricate dynamic equations. Due to the advanced and nonlinear dynamic behaviour of gas turbines, they require constant monitoring of the dynamics of the systems to reveal unknown factors that could contribute to issues during S-CO₂ turbine operations [6]. A demonstration case using a Siemens gas turbine was conducted for the proposed method by comparing two ANN models: one trained with operational data and the other with simulated data. The comparison showed that the trends are very similar, and the study had promising results that support further research in this field [7]. Nikpey [8] developed an ANN-based model to monitor the combined heat and power for micro gas turbines by using the data collected from a modified micro gas

turbine on a test rig. His prediction results showed that the final optimum ANN model can serve as an accurate baseline model for monitoring applications. Further, Lazzaretto [9] used ANN to investigate zero-dimensional design and off-design modeling of a single-shaft gas turbine, and found out that reliability of reproducing the relationships among the most important thermodynamic parameters may be key to their use in the simulation of complex thermal systems. Arriagada [10] used ANN for fault diagnosis of a single-shaft industrial gas turbine, and the results showed that an ANN-based fault diagnosis system is not only capable of fault isolation and identification with high reliability, but can also identify many fault types before they are fully developed and become problematic. Using simulation data allows the integration of ANN-based condition monitoring in the control system of new gas turbines. The thermodynamic parameters of the supercritical CO₂ power cycle can be predicted with high accuracy by using an ANN under variable working conditions to reach optimum performance [11]. This could benefit plant owners through increased availability, and original equipment manufacturers (OEM) by savings in maintenance [7]. ANN is now recognized as one of several successful approaches that can disclose nonlinear behavior of such complicated systems [2]. ANN based models can be created directly from the operational data of an actual supercritical CO₂ gas turbine, or by using the simulated data when operational data is unavailable. An ANN is useful for predicting gas turbine performance if it can be suitably trained with operational data. This is demonstrated by the very high prediction accuracy of the developed ANN model which can be used for several purposes, including offline simulation of gas turbine performance, online condition monitoring of the turbine for early detection of faults or degradation and sensor validation [12] [13]. The major objective of this study is to test the ability of an ANN to track and monitor CO₂ as a working fluid for a 10MW S – CO₂ gas turbine in the most unpredictable and concerned region at main compressor inlet condition.

II. 10 MW SUPERCRITICAL CO₂ BRAYTON CYCLE LAYOUT

The data in this paper was obtained from 10MWe recompression cycle (US Department of Energy, 2015) [14]. Compression cycles are designed very precisely, with two compressors functioning equivalently but with different thermodynamic properties and flow rates; the main compressor mass flow is $\dot{m} = 70.3 \text{ Kg/s}$ and the bypass compressor is $\dot{m} = 34.2 \text{ Kg/s}$. As shown in Figure 1; T – S, once the fluid passed the primary heater, it expands through the turbine and produces energy, as shown in steps 6 and 7. The heat produced by the fluid is absorbed by the high temperature recopercator HTR and the low temperature recopercator LTR, as shown in steps 7 and 8. The fluid flows through the LTR and is split into two different streams, as shown in steps 10 and 1. One stream passes through the cooler for heat rejection then moves to the main compressor, as shown in steps 1 and 2. The other stream passes through the bypass compressor where it is compressed up to $P = 23.99 \text{ MPa}$, as shown in steps 11 and 12.

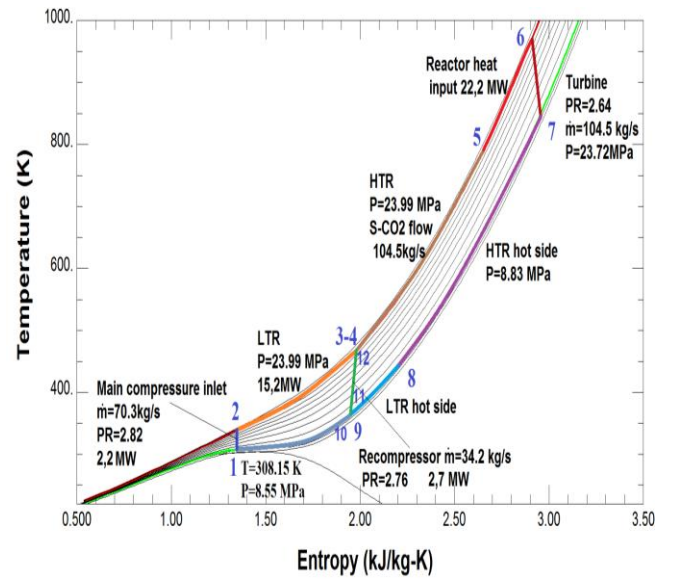


Figure 1. T-s diagram for the 10 MW S-CO₂ gas turbine Bryon cycle.

The main compressor to preheat the fluid in LTR 2 and 3 increases the fluid pressure. At points 3 and 12 the streams are merged at point 4, before entering the HTR 4 – 5 where the fluid gains heat up to 806.15 K, then it passes through the primary heat source, and heats up to 973.15 K as shown in steps 5 and 6. The unpredictable behavior of S-CO₂ at the main compressor inlet conditions affects the inter-cycle performance. Therefore, the implemented method to tack and monitor the fluid at this region is significant.

III. EFFECTS OF IMPURITIES ON THE DENSITY

The working fluid density one of the most important properties of an S-CO₂ gas turbine, and the effects of impurities on CO₂ density are analyzed using REFPROP software 10. The density of CO₂ with several gaseous impurities was measured at the main compressor inlet, with a fixed temperature of $T = 308.15 \text{ K}$ and a range of pressures. The non-condensable components N₂, O₂, CO and Ar are common impurities in all CO₂ capture processes, with a mass fraction of 3% of impurities in a rich CO₂ binary system of 97%. The findings indicate a significant reduction of CO₂ density in the presence of the mentioned impurities. The ability of the S-CO₂ cycle to reach high efficiency is due to the reduced compressor workload as the compressor inlet conditions approach the critical point of CO₂. As shown by equation 1, due to the density decreases, the mass flow rate into the cycle is reduced, and, consequently, the gas turbine output.

$$\dot{m} = \rho * A * c * Ma \quad (1)$$

Mass flow (\dot{m}) through a fixed geometry is a function of the axial cross section area (A), the axial Mach number (Ma), the speed of sound (c) and the density (ρ) [15] [16] [17] [18] [19] [20].

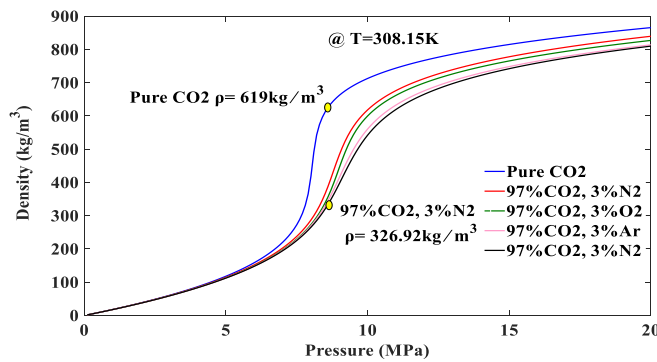


Figure 2. Effect of impurity on the density.

However, the density is highest at pure CO₂ with 619.06 kg/m³, and impurities in the CO₂ stream can potentially affect the density in supercritical regions, which can reduce cycle efficiency. Figure 2 shows the reduction of CO₂ density with the four impurities of O₂, N₂, CO and Ar, with 97% CO₂ and 3% for each impurity. A maximum reduction of CO₂ density at T=308.15 K, and different pressure values from 0.1 to 20 MPa due to impurities, is apparent at P= 8.55 MPa CO₂/Ar, CO₂/O₂, CO₂/CO and CO₂/N₂, at 376.82kg/m³ 347.40kg/m³, 337.50kg/m³ and 326.92kg/m³ respectively. For example, with only 3% impure N₂ the density dropped almost 50%. As shown in Figure 3, non-condensable O₂, Ar, N₂ and CO impurities significantly reduce the speed of sound of the CO₂ stream. This change is slightly higher above the saturated line near the critical point, as the pressure increases it becomes less significant away from the critical point of the mixture.

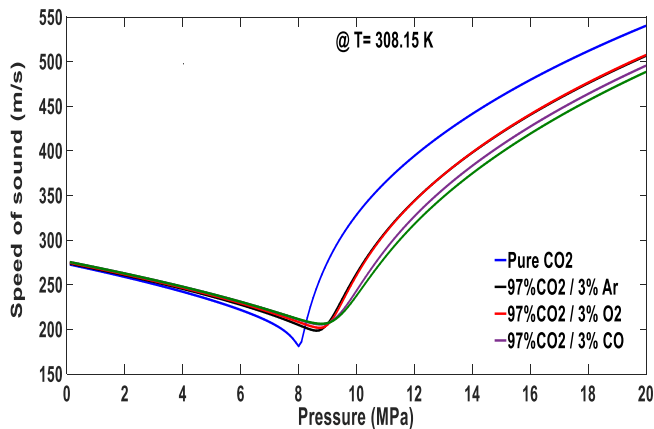


Figure 3. Reduction of the speed of sound due to impurities.

IV. STRUCTURES OF ANN

ANNs are collections of mathematical models that emulate some of the properties of biological nervous systems, and draw on analogies of adaptive biological learning. The key element of an ANN model is the unique structure of the information processing system, which is composed of many highly interconnected processing elements that are analogous to neurons, tied together with weighted connections that are analogous to synapses. Thus, a typical neuronal model comprises weighted connectors, an adder and a transfer function. Figure 4 shows the single-input neuron structure (the multi-input neuron is shown in Figure 5). The scalar input p is multiplied by the scalar weight w to form wp equations 2 and 3, one of the terms that is sent to the sum. The other input (1) is multiplied by a bias b and then passed on to the sum. The sum output (n), often referred to as the net input, goes into a transfer function f which produces the scalar neuron output [21]. The main limitations of every ANN are typically due to weak selection of network weights and parameters, or the number of hidden neurons [22] [6] [23] [24]. However, despite the limitations, using ANN can still produce remarkable enhancements in the process of industrial system modeling and control.

$$n = wp + b \quad (2)$$

where a is the network output signal, w is the weight of the input signal, p is the input signal, b is the neuron specific bias, F is the transfer/activation function and n is the induced local field or activation potential [20].

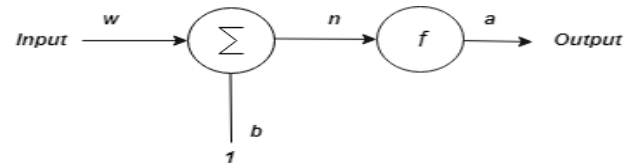


Figure 4. Neuron structure single-input.

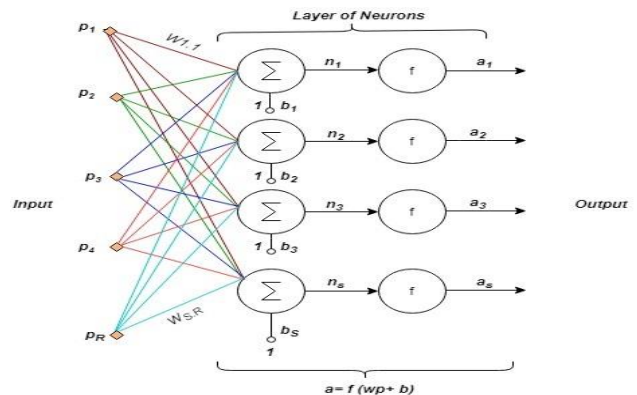


Figure 5. Multiple-input neuron.

V. EVALUATING THE ANN WITH THE REFPROP GENERATING DATA

The purpose of this study is to investigate how an ANN can be successfully employed to predict the properties of the working fluid S-CO₂ gas turbine using REFPROP data. The five inputs are pure CO₂, and the CO₂/O₂, CO₂/N₂, CO₂/CO and CO₂/Ar impurities, with 97% CO₂ and 3% for each impurity obtained from REFPROP at a fixed temperature and different pressure values where the output is density. This test ensures that an ANN is the ideal method to monitor the properties of the working fluid. The results of the training were recorded, and the performances were evaluated and compared in terms of their mean squared errors (MSE). The most accurate multilayer perceptron (MLP) with the minimum MSE was selected, as it has equivalent generalization characteristics of the model. The results from the model for different densities of S-CO₂ gas turbine values (predicted values) were compared with the values of the generated data from the REFPROP software. The best resulting network based on the average performance of the trained structure is depicted in Figure 6, and it also shows that 31 was the iteration in which the validation performance error was lowest. The mean squared error of the performance at this point is also quite low. The training continued for ten more iterations before stopping. Figure 7 shows the regression plot of the relationship between outputs of the network and outputs of the system (targets). The R value is an indication of the relationship between the outputs (ANN) and the targets (REFPROP). The R values for all the graphs are 1, which means the results of training, validation and test data sets indicate an excellent fit. The resulting neural network based model can predict reactions of the system to changes in input parameters with high accuracy, and is capable of system identification with high reliability.

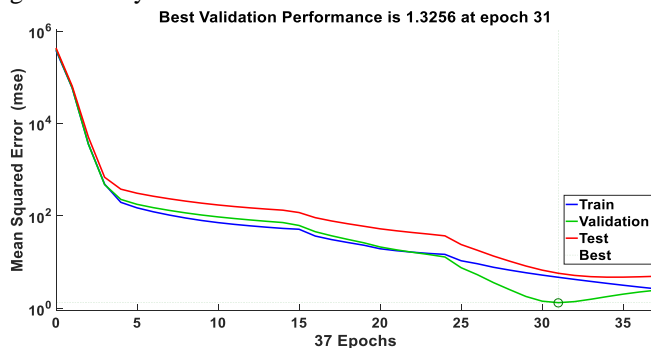


Figure 6. Performance of the resulting MLP model.

VI. RESULT AND DISCUSSION

Due to the economic benefits of CO₂ as an inexpensive and non-toxic source of carbon, many studies have focused on capturing it from different streams. The complexity of predicting the behaviour of CO₂ at supercritical region and

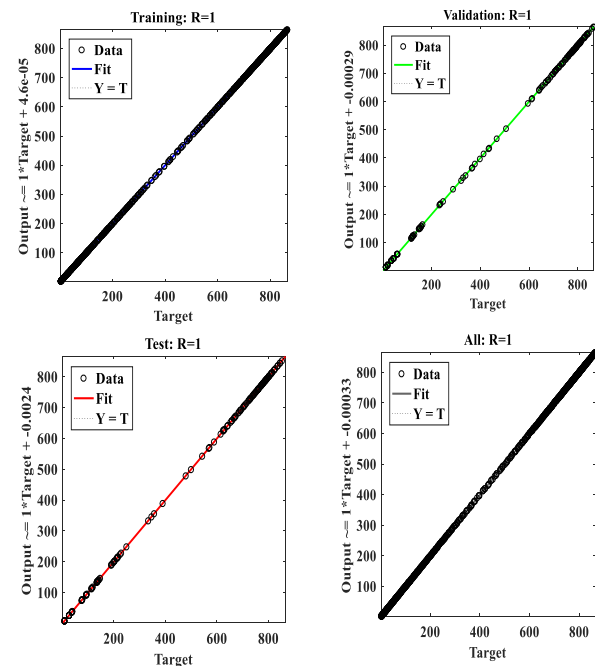


Figure 7. Regression of the resulting MLP model.

implementing a model to monitor the process can be difficult. ANNs are nonlinear mathematical models that relate inputs to outputs; a total of 200 data for pure CO₂ and the four impurities were used. The network was trained, validated and tested using 70%, 15% and 15% of the total data with one hidden layer by a hyperbolic tangent sigmoid transfer function. An optimization procedure based on a genetic algorithm was performed to select the best ANN architecture. According to the collected performances, the MSE was considered the right parameter to select the best model; the best observed performance belonged to the model with one hidden layer and ten neurons. The MSE for Training, validation and testing are 1.16×10^{-4} , 3.23×10^{-4} and 1.41×10^{-4} respectively. Figure 6 shows performance improvement during the training, as well as the validation and testing processes. As the epochs increased, the performance of the training, validation and testing improved until epoch 31, where the performance increased only slightly. The following regression plots display network outputs with regard to targets for training, validation and test sets. For a perfect fit the data should fall along a 45-degree line, which indicates that the network outputs are equal to the targets. The fit is excellent for all data sets, with the coefficient of determination R values for the graphs 1 (Figure 7). Thus, the results for training, validation and test data indicate an excellent fit. As shown in Figure 8, the output density of REFPROP software and ANN-based models are compared. The outputs of the ANN model followed the targets precisely, and showed that the resulting neural network-based model can predict the reaction of the working fluid with high accuracy, and is capable of highly reliable system identification [6] [25] [26].

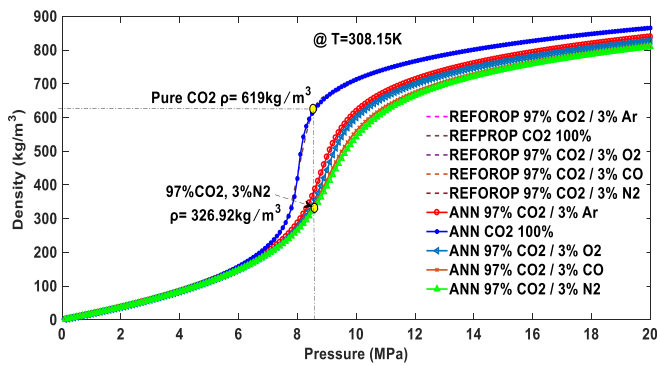


Figure 8. Comparison between outputs of the REFPROP and ANN models CO2 tested data.

VII. CONCLUSION

The thermophysical properties of supercritical fluids change swiftly close the pseudocritical point. This introduces difficulty in predicting the behavior of supercritical fluid. For this purpose, many experiments have been conducted to develop correlations for predicting fluid behavior. In this paper the results show that the predicted values agree closely with the REFPROP software data, and the maximum absolute error deviation for the best predictor was no more than 3.44%. The resulting model showed that the ANN-based method can also be applied to predict the S-CO₂ behavior at the critical point. The ability to estimate and predict CO₂ property behavior through a wide range of temperature and pressure in different regions is a major advantage of using ANN modeling, compared to the other thermodynamic models.

REFERENCES

- [1] V. Palade and C. D. Bocaniala, Computational intelligence in fault diagnosis. Springer, 2006.
- [2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] H. Asgari, X. Q. Chen, M. B. Menhaj, and R. Sainudiin, “Ann-based system identification, modelling and control of gas turbines—a review,” in Advanced Materials Research, vol. 622. Trans Tech Publ, 2013, pp. 611–617.
- [4] R. Span and W. Wagner, “A new equation of state for carbon dioxide covering the fluid region from the triplepoint temperature to 1100 k at pressures up to 800 mpa,” Journal of physical and chemical reference data, vol. 25,
- [5] Q. Zhao, M. Mecheri, T. Neveux, R. Privat, and J.-N. Jaubert, “Thermodynamic model investigation for supercritical co₂ brayton cycle for coal-fired power plant application,” in Fifth International Supercritical CO₂ Power Cycles Symposium, San Antonio, TX, Mar, 2016, pp. 29–31.
- [6] J. J. Dyreby, “Modeling the supercritical carbon dioxide brayton cycle with recompression,” Ph.D. dissertation, The University of Wisconsin-Madison, 2014.
- [7] H. Asgari, “Modelling, simulation and control of gas turbines using artificial neural networks,” 2014.
- [8] M. Fast, M. Assadi, and S. De, “Condition based maintenance of gas turbines using simulation data and artificial neural network: A demonstration of feasibility,” in ASME Turbo Expo 2008: Power for Land, Sea, and Air. American Society of Mechanical Engineers, 2008, pp. 153–161.
- [9] H. Nikpey, M. Assadi, and P. Breuhaus, “Development of an optimized artificial neural network model for combined heat and power micro gas turbines,” Applied Energy, vol. 108, pp. 137–148, 2013.
- [10] A. Lazzaretto and A. Toffolo, “Analytical and neural network models for gas turbine design and off-design simulation,” International Journal of Thermodynamics, vol. 4, no. 4, pp. 173–182, 2001.
- [11] J. Arriagada, M. Genrup, A. Loberg, and M. Assadi, “Fault diagnosis system for an industrial gas turbine by means of neural networks,” in Proceedings of the International Gas Turbine Congress, 2003, pp. 2–7.
- [12] J. Wang, Z. Sun, Y. Dai, and S. Ma, “Parametric optimization design for supercritical co₂ power cycle using genetic algorithm and artificial neural network,” Applied Energy, vol. 87, no. 4, pp. 1317–1324, 2010.
- [13] M. Fast, M. Assadi, and S. De, “Development and multiutility of an ann model for an industrial gas turbine,” Applied Energy, vol. 86, no. 1, pp. 9–17, 2009.
- [14] M. Fast and T. Palme, “Application of artificial neural networks to the condition monitoring and diagnosis of a combined heat and power plant,” Energy, vol. 35, no. 2, pp. 1114–1120, 2010.
- [15] N. W. R. D. R. Ames, Fundamentals and Applications of Supercritical Carbon Dioxide (sCO₂) Based Power Cycles, Fossil energy 12. Elsevier Ltd, 2017.
- [16] A. Dahlquist, “Conceptual thermodynamic cycle and aerodynamic gas turbine design,” Lund University, Sweden, 2016.
- [17] A. Kather, “Co₂ quality and other relevant issues,” in 2nd working group meeting on CO₂ quality and other relevant issues, Cottbus, Germany, 2009.
- [18] E. S. Rubin, H. Mantripragada, A. Marks, P. Versteeg, and J. Kitchin, “The outlook for improved carbon capture technology,” Progress in Energy and Combustion Science, vol. 38, no. 5, pp. 630–671, 2012.
- [19] M. Anheden, A. Andersson, C. Bernstone, S. Eriksson, J. Yan, S. Liljemark, and C. Wall, “-co₂ quality requirement for a system with co₂ capture, transport and storage,” in Greenhouse Gas Control Technologies 7. Elsevier, 2005, pp. 2559–2564.
- [20] E. De Visser, C. Hendriks, M. Barrio, M. J. Mølnvik, G. de Koeijer, S. Liljemark, and Y. Le Gallo, “Dynamis co₂ quality recommendations,” International journal of greenhouse gas control, vol. 2, no. 4, pp. 478–484, 2008.
- [21] I. Al-Siyabi et al., “Effect of impurities on co₂ stream properties,” Ph.D. dissertation, Heriot-Watt University, 2013.
- [22] M. T. Hagan, H. B. Demuth, and O. D. Jes’us, “An introduction to the use of neural networks in control systems,” International Journal of Robust and Nonlinear Control, vol. 12, no. 11, pp. 959–985, 2002.
- [23] A. Khoukhi and M. H. Khalid, “Hybrid computing techniques for fault detection and isolation, a review,” Computers & Electrical Engineering, vol. 43, pp. 17–32, 2015.
- [24] H. A. Talebi, F. Abdollahi, R. V. Patel, and K. Khorasani, Neural network-based state estimation of nonlinear systems: application to fault detection and isolation. Springer, 2009, vol. 395.
- [25] Q.-J. Zhang, K. C. Gupta, and V. K. Devabhaktuni, “Artificial neural networks for rf and microwave design—from theory to practice,” IEEE transactions on microwave theory and techniques, vol. 51, no. 4, pp. 1339–1350, 2003.
- [26] Z. Ahmadloo, “Predictability of carbon dioxide and ethane solubility in ionic liquids: A simulation approach,” Ph.D. dissertation, Faculty of Graduate Studies and Research, University of Regina, 2016.
- [27] V. Ochkov, K. Orlov, V. Voloshchuk, and N. Rogalev, Thermal Engineering Studies with Excel, Mathcad and Internet. Springer, 2016.

HEAT FLUX ANALYSIS OF A LI-ION CELL DURING CHARGING AND DISCHARGING PERIODS

A. R. M. Siddique
 School of Engineering
 University of Guelph
 Guelph, Canada
 Email: asiddi04@uoguelph.ca

S. Mahmud, S. A. Gadsden, and B. V. Heyst
 School of Engineering
 University of Guelph
 Guelph, Canada
 Email: smahmud@uoguelph.ca

Abstract— Due to recent trends in environmental uncertainties and polluted climates, electric vehicles have quickly become the main driving technology to replace gasoline and diesel-powered vehicles. The main power source and storage device for electric vehicles is the rechargeable batteries e.g., Lithium-ion (Li-ion) cell. Excessive heat generation in Li-ion cells during charging and discharging periods reduces the operational lifetime and efficiency of the cell. Extending the operational lifetime of these batteries is critical for the improvement and wide-spread acceptance of electric vehicles. In this article, a series of experimental studies were conducted to analyze the heat flux behavior of Li-ion electric vehicle batteries. Three different temperature-controlled environments were studied: natural convection (in air), inside an enclosed glass jar, and in a liquid (water). Results demonstrate that heat flux changes significantly depending on the environment. The best results are achieved while operating in a controlled environment deployed by a TEC cooling system which also improves the battery operation time. In this study, the heat flux varied between -100 W/m^2 and 300 W/m^2 at 20°C to 23°C for 9,000 seconds under a 0.5A discharge current.

Keywords—charging cycle; discharging cycle; heat flux; Li-ion cell; thermoelectric cooler (TEC)

I. INTRODUCTION

Electric vehicles have become a reliable mode of transportation, and are a viable alternative to gasoline and diesel-driven vehicles. Since fossil fuels are becoming less abundant and causing environmental pollution, there is a growing demand for smart electric vehicle technology [1]. The electric battery is the key system found within electric vehicles. The performance of the battery depends on many factors, such as electrochemical reactions, heat generation in the cell, surrounding temperature, state of charge and health, operational hours, and so forth [2].

Among various determining factors, heat generation within the cell is also important to study for Li-ion batteries. Different battery thermal management systems (BTMS) have been studied in order to manage the excessive heat generation

during high charge/discharge cycle [4-6]. This high amount of generated heat inside the cell causes to reduce battery operation time and efficiency as well. BTMS such as air-cooled system [7], using fins [8], phase change materials (PCM) based cooling system [9], and thermoelectric cooler (TEC) [10] based cooling system are widely used for EVs. Air and water have been using as a suitable option for any thermal management system especially for BTMS over the last decades for EV/HEV [11, 12]. However, Siddique et al. [13] surveyed a comprehensive literature on PCM and thermoelectric based battery thermal management system from where it was found that there is still scope to work with TE system with any other active/passive cooling system in order to increase the efficiency of the overall system.

The aim of this article is to study the heat flux behavior of the Li-ion cell during charging/discharging periods in order to identify how much power per unit area does the cell lose during this period. In order to analyze this characteristic, a series of experiments were performed under three different situations that include in natural convection (in air), inside an enclosed glass jar, and inside the water flow with control temperature. A thermoelectric cooler (TEC) was used with a temperature controller to control the water temperature. Surface temperature and voltage profile of the Li-ion cell were also examined.

II. HEAT GENERATION IN LI-ION CELL

Li-ion cell is a chemical power source which mainly linked to the temperature. A complex electrochemical reaction takes place inside the cell during charging/discharging cycle [14]. Heat is generated during the chemical reaction and excessive heat interrupts the normal operation of the Li-ion cell. A general formula of heat generation from a Li-ion battery was developed by Bernardi et al. [15] is given in Eq. (1).

$$q = I(U - V) - I \left(T \frac{\partial U}{\partial T} \right) \quad (1)$$

where q is the generated heat by the cell, I is the operating cell current, U and V are the open circuits and operating

potential of the cell; receptively, and T is the operating temperate of the cell. The first term on the right side of Eq. (1) represents the ohmic loss and the second term presents the irreversible heat generation term [16].

During charging/discharging cycle many complex mechanisms including electrochemical reactions electricity and heat transfer are involved, and are changed due to the state of charge (SOC), state of health (SOH), temperature, time etc. [16]. The first term is much bigger than the second term in terms of high charge/discharge current rate. Therefore, the second term is negligible for electric vehicles (EVs). The electrochemical reactions and heat generation by a Li-ion cell varies with different environments [15].

III. EXPERIMENTAL TESTS AND RESULTS

A. Experimental et-up

In order to analyze the heat flux and surface temperature during charge/discharge of Li-ion cell at different environmental conditions i.e., in natural air flow, inside a vacuum glass jar, and inside liquid flow, a series of experimental tests were performed. Figure 1 illustrates the experimental set-up for charging a Li-ion cell. A charge/discharge battery tester (Model: iMax B6AC v2, supplier: SKYRC) was used to charge the Li-ion cell (Model: BRC 18650 2200mAh, 3.7 V rated, supplier: Ultrafire) during charging cycle. The voltage of the cell during charging/discharging period was recorded by a data acquisition device (Model: NI USB-6221, supplier: National Instruments). Two K-type thermocouples were used in order to measure the surface and surrounding temperature of the cell by Omega-HH374 thermometer. Heat flux generated by the cell during charging/discharging cycle was measured by a PHFS-01e heat flux sensor (supplier: FluxTeq) which was tightly attached to the surface of the cell. Thermal conductive paste (OT-201-2, OMEGA) was used between the heat flux sensor and the cell to make a good contact.

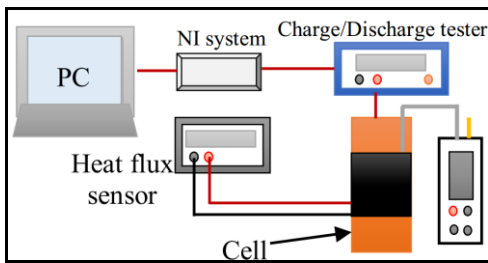


Figure 1. Schematic presentation of the experimental set-up for charging Li-ion cell.

Different environmental situations were considered during the discharging cycle of the cell. A schematic diagram was shown in Fig. 2 to present the experimental set-up in natural air and inside a vacuum glass jar. At first, the Li-ion cell was discharged in the air in normal condition (see Fig. 2(a)). A fixed load (6.75 Ω , which draws around 0.5 A) was used in all cases to discharge the cell. After that, the experimental tests were performed inside a sealed glass jar (see Fig. 2(b)). The

bottom part of the glass jar was sealed so that no air goes inside or coming out to make a vacuum environment.

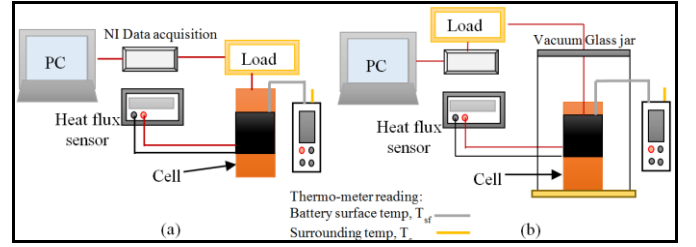


Figure 2. Schematic diagram of the experimental set-up (a) in natural air flow and (b) inside a vacuum glass jar during the discharge period.

Lastly, in order to control the surrounding temperature of the cell, the Li-ion cell was placed inside a cylindrical container full of water as shown in Fig. 3. A thermoelectric cooler (TEC) was incorporated to control the water temperature. TEC is the key component of the cooling module. The cooling module includes a TEC (4 cm \times 4 cm), fins, a cooling DC brushless fan (12 V, 0.3 A) and aluminum water block (4 cm \times 4 cm) which was attached to the cooled side of the TEC. A DC brushless pump (4.5 W, Model: DC 30A-123) was used to circulate the water through the container with connecting tubes. The flow rate of the pump was 1852 mL/min. Though it's dangerous to keep the battery without any protection inside the water; however, at first both ends of the battery was soldered and then covered with electrical tape, and then covered by glue so that water cannot go inside.

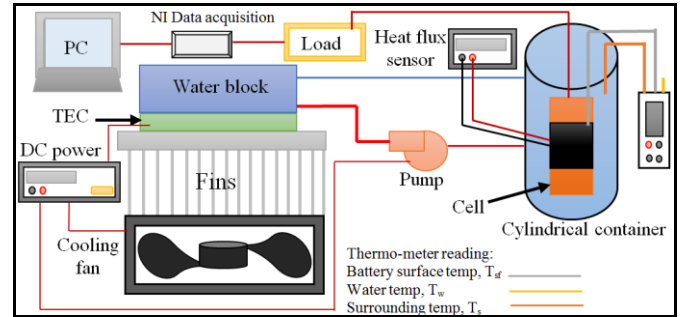


Figure 3. Pictorial diagram of the experimental set-up with the TEC system.

B. Charging cycle test

Li-ion cell was charged using the charge/discharge tester using three different current inputs i.e., 1 A, 1.5 A, and 2 A. At first the cell was discharge up to its safety level (2.9-3.2 V) and then it was charged up to nearly 4.1 V. A voltage profile with respect to time was presented in Fig. 4. It is seen that the charging voltage rapidly increases with time. However, there is a sharp rise in the voltage profile starting from 3.8 V to 4.1 V during 2 A and 1.5 A current input whereas it takes more time to go to 4.1 V. As times passes the voltage becomes a steady state.

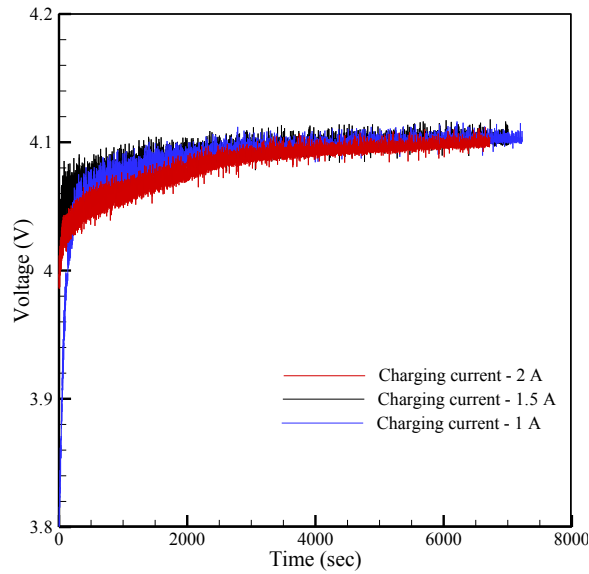


Figure 4. Voltage profile along with time during different charging cycles.

Figure 5 shows the heat flux profile with time under three different current inputs. It is seen that the heat flux and surface temperature increase gradually with time when there is a jump in voltage from 3.8 V to around 4 V. Moreover, when the voltage becomes stabilize (battery was generating less amount of heat) then the heat flux and surface temperature starts to decrease until the cell was fully charged. The heat flux starts from a negative value and as times passes the value varies between 0 W/m² to 75 W/m² under different charging current. Moreover, it is noticed that the decreasing trends of the heat flux are almost the same for different current input as the cell is generating almost the same heat due to the steady voltage level.

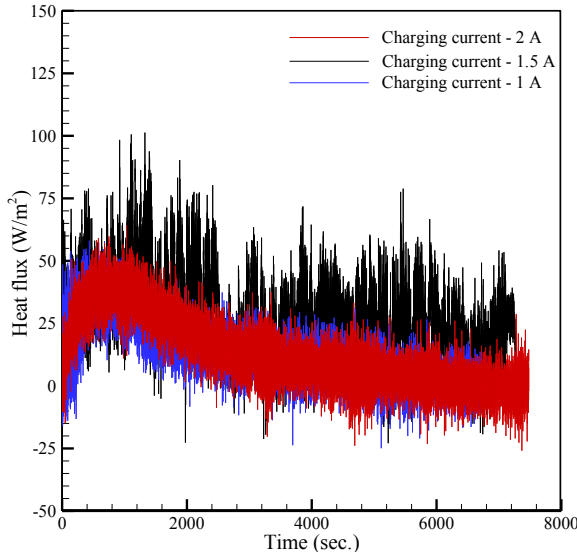


Figure 5. Heat flux profile along with time during different charging cycles.

C. Discharging cycle test

The environmental influence on the cell was tested under the discharging cycle at three environmental conditions. A

fixed load of 6.75 Ω (0.5 A) was used to discharge Li-ion cell in this experiments just to identify the characteristics of the cell discharge voltage, the heat flux, and surface temperature at different environmental situations. The surrounding temperature effects on voltage and heat flux was clearly presented in Fig. 6 and Fig. 7. Figure 6 shows the voltage profile during 0.5A discharge current. Though the cell was fully charged; however, when it was connected to the load it started to discharge it's voltage from around 3.9 V and gradually decreases from 3.9 V to 0.6V with time (In total 9000 seconds).

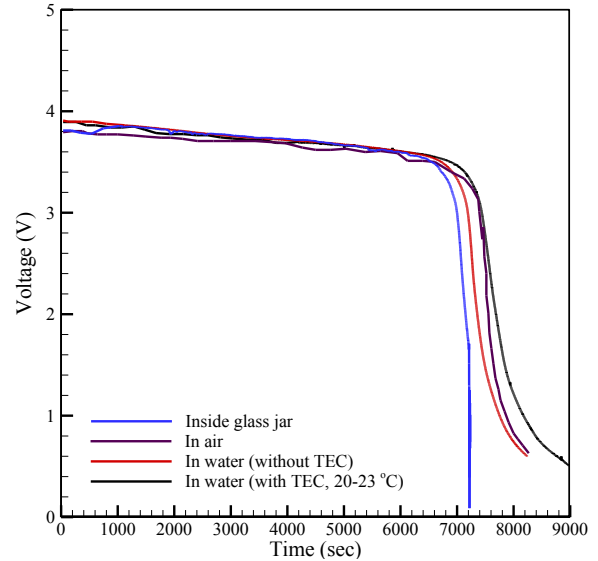


Figure 6. Discharge voltage profile with time under 0.5 A discharge current.

The experimental tests show that it takes longer time to discharge the voltage in the colder environment. The voltage decreases rapidly and takes less time to discharge under vacuum situation due to the effect of the surrounding environment. Since the place was closed, so there was no air flow in or out which means the surrounding temperature was almost fixed with time under normal situation. However, the cell was continuously generating heat during the discharging period which affected the surrounding temperature to heat up. The cell could not radiate heat smoothly as it could do under normal air flow condition. It means the cell heats up itself more and slows down the speed of the chemical reaction which reduces the operation time. In order to control the surrounding temperature, a TEC based cooling system was used with a temperature controller (supplier: INKBIRD). In this experiment, the surround the temperature was controlled between 20 °C to 23 °C.

Figure 7 shows the heat flux behavior of Li-ion cell under the natural convection, vacuum place, and inside the water with a controlled environment. The trend of the heat flux increases as the surrounding temperature decreases. Almost similar trend is noticed under natural air, inside the vacuum glass jar, and inside water at 0.5A discharge current. However, this heat flux behavior is completely different for the cell due to the cooling system in a controlled environment. The cell is

generating less heat under a cold environment which reduces the heat flux as well.

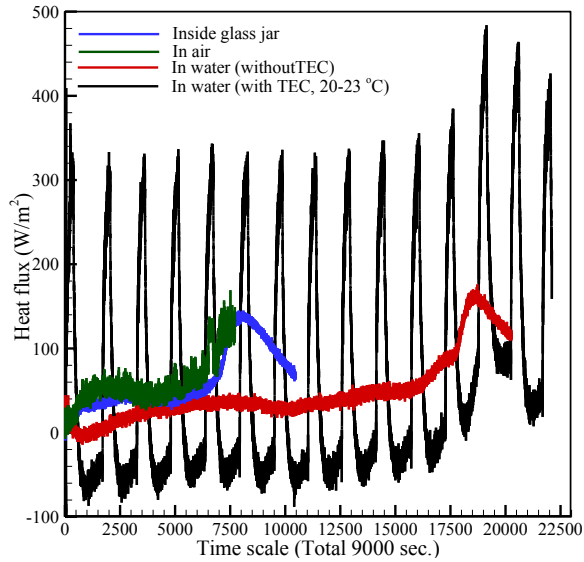


Figure 7. The characteristics of Li-ion cell in terms of generating heat flux with time under different environmental conditions.

Figure 8 presents the temperature profile under different environmental conditions at 0.5A discharge current. Similar trends are noticed under natural air flow and inside the glass jar whereas it is different underwater and in a controlled environment. The surface temperature under natural convection and enclosed glass jar increase gradually; however, it increases rapidly at the end when the battery has does not have enough power to supply. The temperature gradually increases with time without any distortion under water because water has specific heat which helps to keep a steady temperature rise in the cell. Furthermore, the surface temperature of the Li-ion cell fluctuates. The environment temperature is also fluctuating during the test which is seen from the graph by the green line.

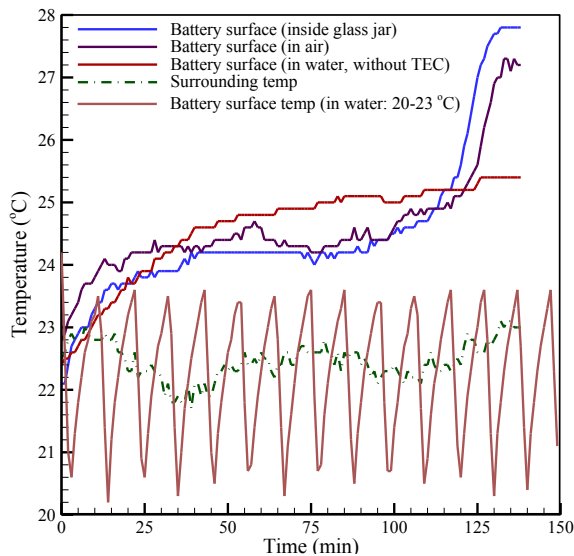


Figure 8. Temperature profile with time under different environmental conditions.

IV. CONCLUSION

In this paper, the heat flux and the surface temperature behavior of the Li-ion 18650 cell was experimentally studied. Three different environmental conditions (natural air flow, inside the enclosed glass jar, and inside water) were considered to analyze different behavior under charging and discharging cycles. Three different charging current levels (1 A, 1.5 A, and 2 A) were examined to characterize the voltage, heat flux, and surface temperature profiles of the Li-ion cell.

Best results were achieved under the controlled-environment with a TEC cooling module. In other words, fixing the water temperature between 20 °C to 23 °C in terms of voltage, heat flux, and surface temperature under 0.5 A discharge current. Moreover, the operational time of the battery increased due to the colder environment. However, heat flux changed more from -100 W/m² to 300 W/m² than other environmental conditions. Future work will focus on the battery packs of Li-ion cells and characterize its behavior under different controlled-environment temperatures. The TEC based cooling system is a suitable fit for this application and will be utilized and explored further.

REFERENCES

- [1] F.V. Conte, "Battery and battery management for hybrid electric vehicles: a review," *e & i Elektrotechnik und Informationstechnik*, vol. 123(10), pp. 424-431, 2006.
- [2] S. Al-Hallaj, R. Kizilel, A. Lateef, R. Sabbah, M. Farid, and J.R. Selman, "Passive thermal management using phase change material (PCM) for EV and HEV Li-ion batteries", *IEEE Veh. Power Propuls. Conf.*, Chicago, IL, USA, 2005.
- [3] R. Liu, J. Chen, J. Xun, K. Jiao, and Q. Du, "Numerical investigation of thermal behaviors in lithium-ion battery stack discharge", *Appl. Energy*, vol. 132, pp. 288-297, 2014.
- [4] S.A. Gadsden and S.R. Habibi, "Model-Based Fault Detection of a Battery System in a Hybrid Electric Vehicle", *Journal of Energy and Power Engineering*, vol. 7(7), pp. 1344-1351, 2013.
- [5] S.A., Gadsden, M. Al-Shabi and S.R. Habibi, "Estimation Strategies for the Condition Monitoring of a Battery System in a Hybrid Electric Vehicle", *ISRN Signal Processing*, pp. 1-17, 2011.
- [6] R. Bustos, A.R.M. Siddique, T. Cheema, S.A. Gadsden, and S. Mahmud, "State of Charge and Parameter Estimation of Electric Vehicle Batteries", *2018 Canadian Society for Mechanical Engineering (CSME) International Congress*, Toronto, Ontario, 2018.
- [7] W. Wu, X. Yang, G. Zhang, K. Chen, and S. Wang, "Experimental investigation on the thermal performance of heat pipe-assisted phase change material based battery thermal management system", *Energy Convers. Manag.* doi:10.1016/j.enconman.2017.02.022, 2017.
- [8] Y. Berthou, "Étude de parois de bâtiments passifs associant un Matériau à Changement de Phase (MCP) et une super isolation transparents", *phdthesis École nationale supérieure des Mines de Paris*, 2011.
- [9] S.Y. Kim, K. Lee, S. Park, and J. Kim, "Thermal design analysis and performance test of a 1kW thermoelectric battery cooler", In: *Thermomechanical Phenom. Electron. Syst. -Proceedings Intersoc. Conf.*, doi:10.1109/ITHERM.2014.6892446, 2014.
- [10] N. Zhoujian, J. Li, D. Yong, D. Chao, and L. Xuejiao, "A Review on Lithium-ion Power Battery Thermal Management Technologies and Thermal Safety", *J. Therm. Sci.*, vol. 26(5), pp. 391-412, 2017.
- [11] C. Alaoui, "Solid-state thermal management for lithium-ion EV batteries," *IEEE Trans. Veh. Technol.*, 2013.
- [12] Y. Liu, S. Yang, B. Guo, and C. Deng, "Numerical Analysis and Design of Thermal Management System for Lithium Ion Battery Pack Using Thermoelectric Coolers," *Adv. Mech. Eng.*, 2014.
- [13] A. R. M. Siddique, S. Mahmud, B. V. Heyst, "A Comprehensive Review

- on a Passive (Phase Change Materials) and an active (Thermoelectric Cooler) Battery Thermal Management System and Their Limitations,” *Journal of Power Source*, vol. 401, pp. 224-237, 2018
- [14] D. Bernardi, E. Pawlikowski, and J. Newman, “A general energy balance for battery systems [J]”, *Journal of the Electrochemical Society*, vol. 132(1), pp. 5–12, 1985.
- [15] R. Zhao, S. Zhang, J. Liu, and J. Gu, “A review of thermal performance improving methods of lithium ion battery: Electrode modification and thermal management”, *Journal of Power Sources*, vol. 299, pp. 557-577, 2015.
- [16] G. Xia, L. Cao, and G. Bi, “A review on battery thermal management in electric vehicle application”, *Journal of Power Sources*, vol. 367, pp. 90-105, 2017.

Numerical simulation of heat transfer and insulation performance of honeycomb structure

Liu, Hao; Tasnim, Syeda; Lubitz, William*

School of Engineering
University of Guelph
Guelph, Ontario, Canada
*wlubitz@uoguelph.ca

Abstract— For many years, heat transfer researchers have been exploring new material or new method which can apply to the solar collector or other device which requires a transparent cover with a function of heat insulation. In the current work, the honeycomb structure for heat insulation is studied using ANSYS fluent software. We use borosilicate glass as the solid material and air, argon, and vacuum as the heat transfer (HT) fluid inside. The performance of each fluid at different temperatures (200°C, 600°C, 900°C) is evaluated. In addition, the fluid motion inside the model is also studied. The results show that vacuum is the best HT fluid for heat insulation purpose. And argon has a relatively better performance than dry air. Identical patterns of fluid motion are observed for all the cases. The moving fluid has a little impact on the heat insulation since the velocity is considerably small.

Keywords- *heat transfer; numerical simulation; honeycomb; transparent heat transfer fluid; fluid motion*

I. INTRODUCTION

Honeycomb structure has been a hot spot for the heat transfer researchers ever since it was proposed. Due to its special structure of dividing the broad space into countless small cells, it can be used into many areas such as heat storage systems, heat exchangers, and more often heat insulation. The formed small cells can prevent the heat passing from one side to the other side.

Applying the honeycomb structure into the Transparent Insulation Material (TIM) can be a huge step forward for the solar collectors or other devices which need an upper layer which can have a high solar transmittance and an excellent heat insulation performance at the same time. This idea was first proposed by Francia (1961). Francia used small glass tubes to make a honeycomb structure and then put it on the upper surface of a flat plate solar collector. After that several research works considering TIM to reduce the heat loss of solar collectors have observed.

In this paper the heat insulation performance of the transparent honey comb structure with different heat transfer medium (air, argon, vacuum) is evaluated. The solid material for the model is borosilicate glass which has 92% transmittance, which is a typical value for borosilicate glass. By changing the

temperature from 200°C to 900°C, we would like to observe how thermophysical properties as well as natural convection, radiation and conduction heat transfer effect the heat transfer performance of the honey comb structure. Particularly, in the present work, we would like to study how the fluid moves inside the structure as the temperature changes. As the fluid motion pattern may differ as temperature changes.

Hollands et al. (1992) used FEP (fluorinated ethylene propylene) to build a honeycomb insulation layer. The cells that the authors used were small enough, so the structure can suppress the natural convection of the air in the cavities. The authors found that at sufficiently low temperature differences across the honeycomb, the convection was completely suppressed. Convection effects were observed as temperature increased.

Goetzberger et al. (1992), Platzter (1992), Rommel and Wagner (1992), and Schweiger (1997) all used traditional glass tubes as TIM since glass has good resistance to high temperature. However, glass is heavy, brittle and expensive. Therefore, researcher starts to seek for new transparent materials. Kessentini et al. (2011) used plastic TIM which has lower cost and lower weight. The authors added an overheating protection system to resist high temperatures.

Swann and Pittman (1961) proposed a semi-empirical formula based on numerical simulation which has always been a standard to calculate the equivalent thermal conductivity (ETC) of honeycomb structure panels. Daryabeigi (2002) used a finite difference method to calculate ETC by coupling heat conduction and radiation. Li, Wu et al. (2008) used finite volume methods to calculate ETC of a hollow clay brick in three dimensions. Rafidi and Blasiak (2005) used a two-dimensional model to simulate the thermal performance of a honeycomb heat regenerator made of two composite materials.

Zheng et al. (2013) studied the performance of the metallic honeycomb core structure in the high temperature environment. Although these authors used metal rather than TIM, the procedure and the results of the experiment and simulation in this work are very valuable to use for reference. The results showed that radiation has the largest proportion in the total heat transfer. Therefore, reducing the amount of radiation can be an effective way to reduce heat transfer.

Boukendil et al. (2012) used two-dimensional steady state model using finite volume method to simulate the coupled heat transfer through a double honeycomb wall. The authors found that overall heat flux increased almost linearly with temperature difference since conduction is the predominant mode of heat transfer.

Bhourri et al. (2011) and Li et al. (2018) used honeycomb structures to store thermal energy. Li et al. (2018) used ceramic honeycomb to store energy and used air as the heat transfer fluid in a one-dimensional model for thermal energy storage.

There are a lot of researchers who study on how different parameters effect the heat insulation performance of the honeycomb structure. Hollands (1965) proposed that the small cells with hydraulic diameters of millimeters can effectively suppress the natural convection in the cells. Cane et al. (1977) and Smart et al. (1980) studied the effect of angle of inclination on heat transfer. Asako et al. (1991) determined that the heat transfer is primarily by conduction when the cell aspect ratio is 4 and Nusselt Number is less than 1.14.

The objective of the current study was to simulate the transient heat transfer through honeycomb panel made of borosilicate glass which has a high optical transmittance of 92% and high temperature resistance up to 820°C. The data regarding the honeycomb structure is taken from Zheng et al. (2013) to reproduce their results and was used for all the simulation cases.

II. SIMULATION PROCESS

The model of Zheng et al. (2013) is taken as a reference in the simulation so that we can compare the results. As shown in Fig.1, the model takes into account the actual manufacturing process. The thickness of the upper and lower sides of the hexagon is 0.076 mm and the thickness of the rest of the four sides are 0.038 mm. The diameter of the circumscribed circle of the hexagon is 6 mm. The height of the honeycomb structure is 7.5 mm and the thickness of the upper and lower panels is 0.16 mm. So the total thickness of the board is 7.82 mm.

As shown in Fig.1, two different steels are used in the board. The honeycomb structure uses GH536 and the two panels on the side use GH3039. Considering the board consists of a countless number of identical hexagonal prism shells, we can simulate only one honeycomb unit as shown in Fig. 2. The heat transfer and fluid motion within the cell was simulated using ANSYS Fluent 19.2. Meshing was completed keeping other parameters at default settings, we set the Max Face Size and Max Tet Size both 0.1 mm as shown in Fig. 2.

Transient 300 second simulations were performed. Gravity is in the direction from the upper hot panel to the lower cold panel. Surface to surface (S2S) radiation applies to all walls, and air was selected as an incompressible ideal gas. Internal emissivity is 0.8, external emissivity is 0.54, heat transfer coefficient is 3 W/(m²K). The free stream temperature and external radiation temperature are both 30°C with a temperature boundary condition for the front surface. The front surface temperature starts from 30°C, linearly increases to 200°C, 600°C, and 900 °C in 60 s, and then remains at those temperatures depending on the boundary condition for 240 s. For side surfaces, symmetry boundary condition is applied. For rear

surface, mixed natural convection and radiation heat transfer has been considered with an external emissivity of rear surface as 0.54. Three different mesh sizes are considered with 1 mm, 0.1 mm, and 0.05 mm. The results of 0.1 mm and 0.05 mm are close enough and this is the reason 0.1 mm mesh size is considered for all the simulation cases. The modeling results are shown in Fig.3 for two different mesh sizes of 1 mm and 0.1 mm. The trend of the temperature curve with time is roughly the same as Fig. 9 in Zheng et al. (2013). The temperature vs. time graph for the 0.1 mm mesh cell size case is closer to the results in the reference paper than the 1 mm mesh cell size case. Figure 4 shows the comparison of current rear surface temperature variation with time to those of Zheng et al. (2013)'s experimental results. This comparison proves the credibility and validity of the present work. The reason for the variation of the present results to those with Zheng et al. (2013)'s experimental results is due to the assumption used in modeling such as constant emissivity of surfaces, note that emissivity of surfaces will vary with temperatures.

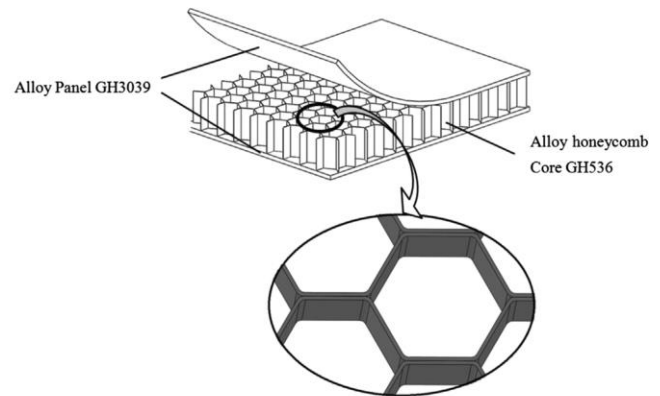


Figure 1. Honeycomb panel structure. From Zheng et al. 2013.

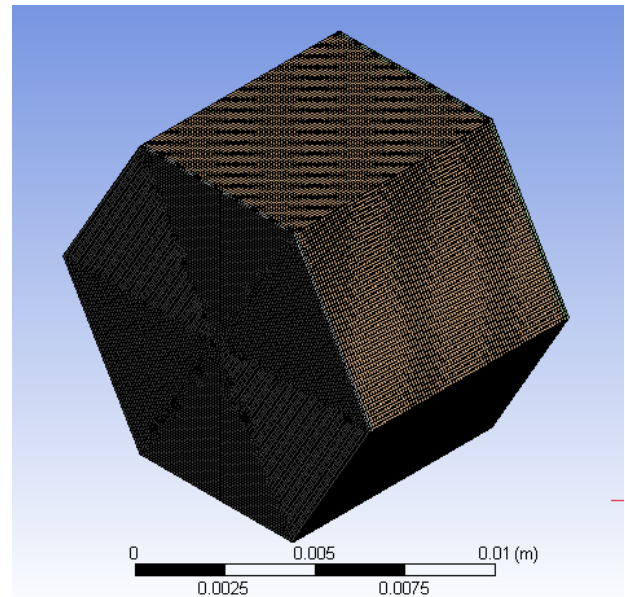


Figure 2. Honeycomb structural unit with 0.1 mm mesh.

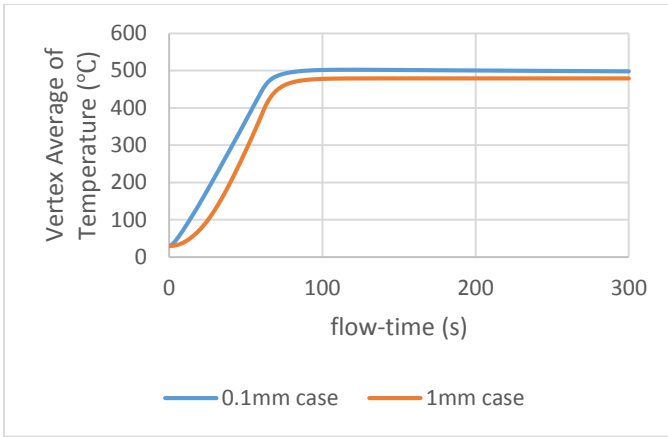


Figure 3. Rear surface temperature for two different mesh sizes.

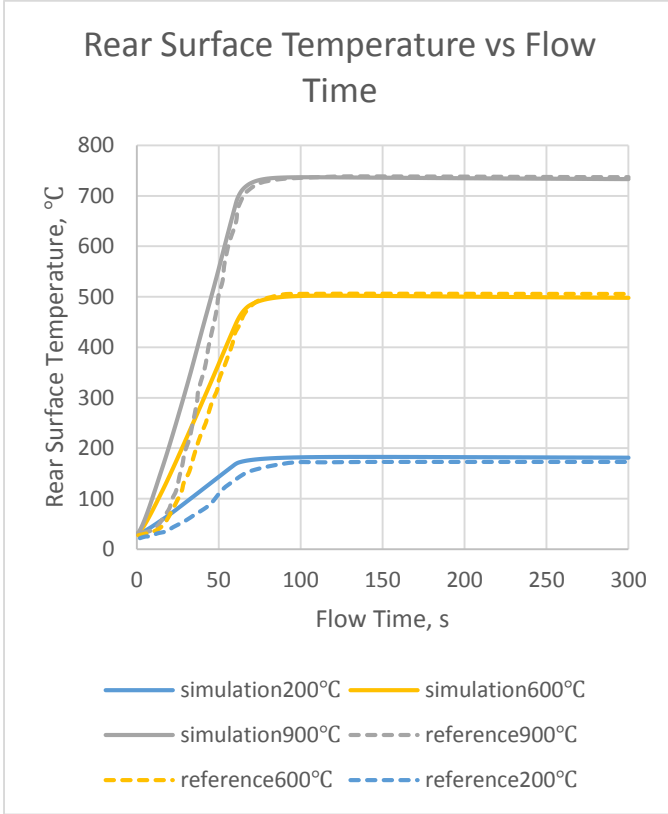


Figure 4. Rear surface temperature vs. flow time for verification cases.

The study of the motion of air inside the cell is really interesting since Zheng et al. (2013) did not discuss this aspect. The authors assumed a small size cell and neglected the convection and so does the motion of air in the structure (Zheng et al., 2013).

In the current research, we study the motion of air inside the cell even though it is expected to be negligible. The motion of air is clearly shown in Fig.5 even though it is completely different from the result in Hollands et al. (1992) shown in Fig. 6. As we can see, the flow of the air inside is very active even though the highest speed is only 6.97×10^{-5} m/s.

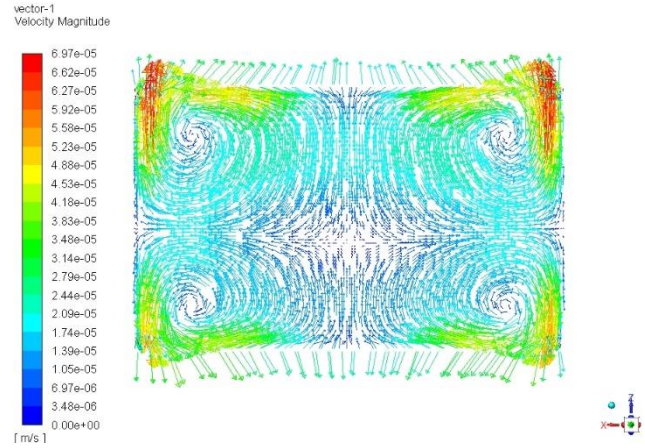


Figure 5. Velocity distribution through honeycomb cross-section, from 0.1 mm mesh simulation.

In order to find the difference in velocity results between the current work to that of Hollands et.al. (1992), we performed another simulation. In this case, the gravity vector was changed to be parallel to the front and rear surface, which is exactly the same way it was configured in the Hollands et al. (1992) work as shown in Fig.6. The simulation result with change in the direction of gravity is shown in Fig. 7. As we can see, the velocity vectors show exactly the same pattern as in Fig.6. This suggests that even though the magnitude of velocity of air inside the cell is small, it may still be relevant to consider it, particularly since the orientation of the cell with respect to gravity can greatly influence the motion pattern.

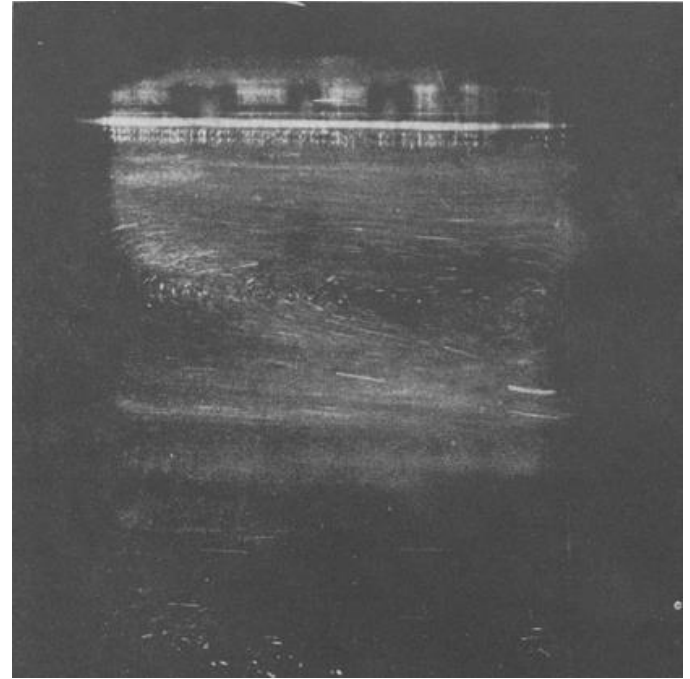


Figure 6. Visualization of fluid motion inside a cell. From Hollands et al. 1992.

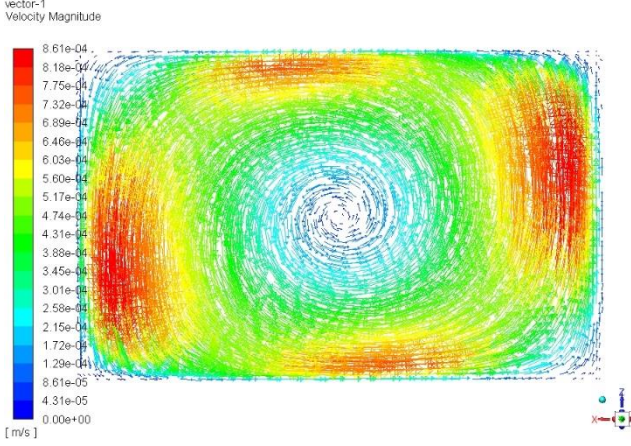


Figure 7. Air velocity distribution in honeycomb cross-section, with gravity vector parallel with the two surfaces.

III. RESULTS

Comparisons of the numerical values for the rear surface temperature from the present simulation to those of [Zheng et al. \(2013\)](#)'s experimental results are presented in Table 1.

Table 1. Temperature difference between reference case and simulations.

Front surface temperature	Rear surface temperature in the reference	Rear surface temperature in the simulation	Temperature difference
200°C	173.142°C	178.972°C	3.36%
600°C	505.801°C	497.903°C	-1.56%
900°C	737.008°C	733.256°C	0.509%

IV. TEST CASES

A total of 15 simulation cases were completed by considering incompressible ideal air, actual dry air, argon and vacuum as HT fluid and three different temperatures (200°C, 600°C, 900°C) for the front surface. Figure 8 shows the cross sectional distribution of temperature for actual dry air. The straight parallel lines for temperature distribution shows that very near the top surface heat is primarily transferred by conduction. After that heat transfer is due to combined effect of conduction convection infra-red radiation. Figure 9 shows the comparison of rear surface temperature using actual dry air, argon and vacuum as HT fluid inside the structure for three different front surface temperatures.

The final results are shown in Table 2 and Table 3. Considering the average temperature of rear surface, the thermal properties of the solid material can affect the heat insulation performance of the model. The material with low thermal conductivity can prevent more heat transferring from one side to another. And the thermal properties of the inside fluid does not have a significant influence on heat transfer. Applying a vacuum inside the honeycomb structure can effectively lower the overall heat

transfer and improve the heat insulation performance, especially in a low temperature environment.

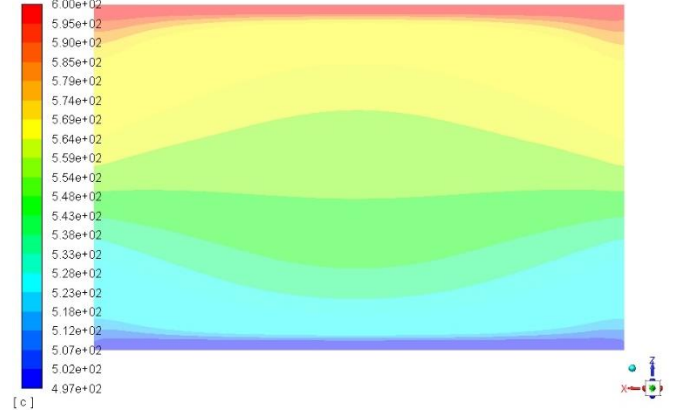


Figure 8. Temperature distribution through the honeycomb cross-section.

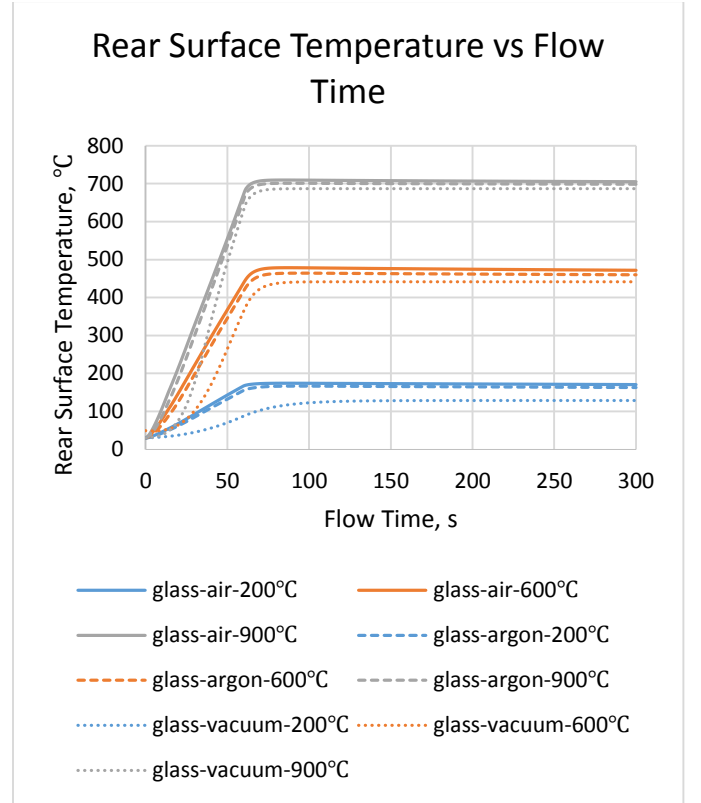


Figure 9. Rear surface temperature vs. flow time for test cases.

Considering fluid velocity within the cell, all the fluids have a very low circulation velocity within the cell. The circulating patterns of each fluid for different temperatures are similar and actual dry air are shown in Fig.10, Fig.11 and Fig. 12. All of them have four primary circulation cells within the honeycomb cell. When the front surface temperature goes high, the upper circulating air becomes stronger and pushes the smaller lower air parts. The viscosity and density of the fluids can affect the velocity. But they appear to have little effect on the temperature trend since the velocity of the trapped fluid is very low (except

for the vacuum). The circulation patterns for argon show the same trend and are not shown.

The overall heat transfer coefficient across the cell can be calculated as:

$$Q = UA(T_{rear} - T_{front}) \quad (1)$$

$$q = U(T_{rear} - T_{front}) \quad (2)$$

where Q is the total rate of heat transfer between the front surface and rear surface, in watts, and q is the rate of heat transfer per unit area. The overall heat transfer coefficient of the model is U , typically with units of W/m^2-K , and A is the surface area of the front surface or rear surface, in m^2 . The average temperature of the front and rear surfaces are T_{front} and T_{rear} respectively.

The results in Table 4 show that the overall heat transfer coefficient decreases when the solid material is changed from steel to borosilicate glass, and the fluid is changed from dry air to argon, and then to vacuum. The borosilicate glass has a better insulation performance than steel, and argon is a better insulator, resulting in lower heat transfer, than dry air. The borosilicate glass-vacuum shows the best performance on heat insulation since there is no convection in the honeycomb structure. However, there is no absolute vacuum in reality. We can evacuate the air in the cell as much as possible to create a vacuum. In this way, the convection component of heat transfer can be reduced, however, radiative heat transport across the cell, and conduction along the cell walls, will still occur.

V. CONCLUSION

Numerical study of combined heat transfer by natural convection, conduction and infra-red radiation through a honeycomb structure has been conducted. The transient results have been validated with experimental results from existing literature. The following conclusions can be drawn based on the analyses:

- The thermal properties of the solid material and inside fluid can affect the heat insulation performance of a honeycomb structure.
- Argon and air have similar heat insulation performance, argon shows better performance than air because of its low thermal conductivity.
- Applying vacuum inside the honeycomb structure effectively improves the heat insulation performance, especially at a low temperature environment.
- The circulating pattern of the inside fluid may change as the direction of gravity changes.
- The fluid has a very low speed circulating inside, and it does not have a considerable effect on the heat transfer process. A combination of borosilicate glass and vacuum shows the best heat insulating performance.

Future studies will include development of an experimental facility to conduct the experiment with borosilicate and vacuum

and compare the numerical results from present work to the experimental work. In addition experimental study will be conducted using air and argon as the HT fluids to shed some light on the fluid flow patterns and the relationship between fluid flow patterns and gravity.

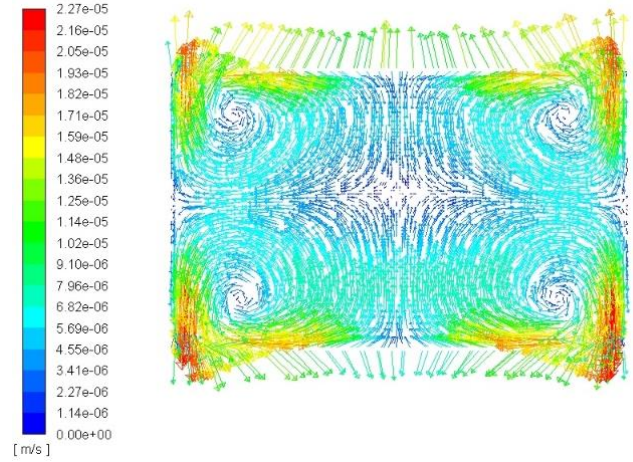


Figure 10. 200°C velocity vectors for dry air case.

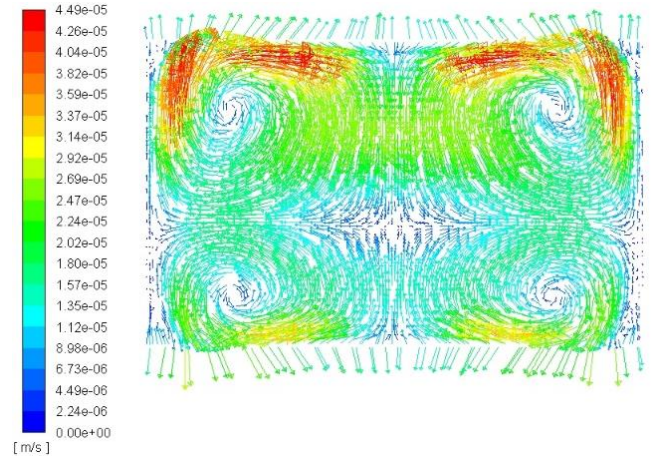


Figure 11. 600°C velocity vectors for dry air case.

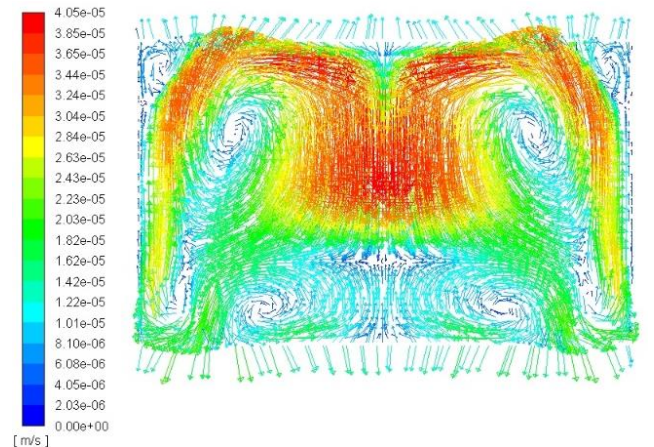


Figure 11. 900°C velocity vectors for dry air case.

Table 2. Temperatures of rear surface for different cases.

Temperature of front surface, °C	Temperature of rear surface of ideal air case, °C	Temperature of rear surface of actual dry air case, °C	Temperature of rear surface of glass and actual dry air case, °C	Temperature of rear surface of glass and argon case, °C	Temperature of rear surface of vacuum case, °C
200	181	177	168	161	127
600	497	491	468	457	439
900	730	727	701	695	683

Table 3. Velocity of interior fluid for different cases.

Temperature of front surface, °C	Velocity of inside ideal air, 10^{-5} m/s	Velocity of inside actual dry air, 10^{-5} m/s	Velocity of inside actual dry air of glass case, 10^{-5} m/s	Velocity of inside argon case, 10^{-5} m/s	Velocity of inside vacuum case, 10^{-5} m/s
200	2.77	2.88	2.27	3.36	0
600	6.97	6.25	4.49	4.75	0
900	7.18	5.14	4.05	4.55	0

Table 4. Overall heat transfer coefficient for each case.

Front surface temperature, °C	Overall HT coefficient of Steel-ideal air case, W/m ² K	Overall HT coefficient of Steel-dry air case, W/m ² K	Overall HT coefficient of Glass-dry air case, W/m ² K	Overall HT coefficient of Glass-argon case, W/m ² K	Overall HT coefficient of Glass-vacuum case, W/m ² K
200°C	78.11	59.36	43.34	32.09	10.75
600°C	116.44	108.05	81.71	70.05	56.06
900°C	199.10	193.03	152.14	143.49	129.86

REFERENCES

- [1] Asako, Y., et al. (1991). "Three-dimensional laminar natural convection in an inclined air slot with hexagonal honeycomb core." *Journal of Heat Transfer* 113(4): 906-911.
- [2] Bhourri, M., et al. (2011). "Honeycomb metallic structure for improving heat exchange in hydrogen storage system." *international journal of hydrogen energy* 36(11): 6723-6738.
- [3] Boukendil, M., et al. (2012). "Detailed numerical simulation of coupled heat transfer by conduction, natural convection and radiation through double honeycomb walls." *Building Simulation* 5(4): 337-344.
- [4] Cane, R., et al. (1977). "Free convection heat transfer across inclined honeycomb panels." *Journal of Heat Transfer* 99(1): 86-91.
- [5] Daryabeigi, K. (2002). "Heat transfer in adhesively bonded honeycomb core panels." *Journal of thermophysics and heat transfer* 16(2): 217-221.
- [6] Francia, G. (1961). *Un nouveau collecteur de l'énergie rayonnante solaire: theorie et verifications experimentales*. United Nations conference on new sources of energy, Rome.
- [7] Goetzberger, A., et al. (1992). "A new transparently insulated, bifacially irradiated solar flat-plate collector." *Solar Energy* 49(5): 403-411.
- [8] Hollands, K. G. T. (1965). "Honeycomb devices in flat-plate solar collectors." *Solar Energy* 9(3): 159-164.
- [9] Hollands, K. G. T., et al. (1992). "Manufacture, solar transmission, and heat transfer characteristics of large-celled honeycomb transparent insulation." *Solar Energy* 49(5): 381-385.
- [10] Kessentini, H., et al. (2011). Numerical and experimental study of a flat plate solar collector with transparent insulation and overheating protection system. *Proceedings of ISES Solar World Congress 2011*.
- [11] Li, L. P., et al. (2008). "Optimization of the configuration of 290×140×90 hollow clay bricks with 3-D numerical simulation by finite volume method." *Energy and Buildings* 40(10): 1790-1798.
- [12] Li, Q., et al. (2018). "Dynamic simulations of a honeycomb ceramic thermal energy storage in a solar thermal power plant using air as the heat transfer fluid." *Applied Thermal Engineering* 129: 636-645.
- [13] Platzter, W. (1992). "Calculation procedure for collectors with a honeycomb cover of rectangular cross section." *Solar Energy* 48(6): 381-393.
- [14] Rafidi, N. and W. Blasiak (2005). "Thermal performance analysis on a two composite material honeycomb heat regenerators used for HiTAC burners." *Applied Thermal Engineering* 25(17): 2966-2982.
- [15] Rommel, M., Wagner, A., 1992. Application of transparent insulation materials in improved flat-plate collectors and integrated collectors storages. *Solar Energy* 49 (5), 371–380.
- [16] Schweiger, H., 1997. Optimization of solar thermal absorber elements with transparent insulation. PhD thesis, Universitat Politecnica de Catalunya.
- [17] Smart, D., et al. (1980). "Free convection heat transfer across rectangular-celled diathermanous honeycombs." *Journal of Heat Transfer* 102(1): 75-80.
- [18] Swann, R. T. and C. M. Pittman (1961). Analysis of effective thermal conductivities of honeycomb-core and corrugated-core sandwich panels, National Aeronautics and Space Administration.
- [19] Zheng, L., et al. (2013). "Experimental investigation and numerical simulation of heat-transfer properties of metallic honeycomb core structure up to 900 °C." *Applied Thermal Engineering* 60(1): 379-386.

PERFORMANCE EVALUATION OF SIDING MATERIALS SUBJECTED TO RADIANT HEAT LOADS

Eder A. Villa Coronel
Department of Mechanical Engineering
University of Alberta
Edmonton, Alberta, Canada
ederalai@ualberta.ca

Razim Refai
Wildfire Scientist
FPInnovations
Edmonton, Alberta, Canada
razim.refai@fpinnovations.ca

André G. McDonald
Department of Mechanical Engineering
University of Alberta
Edmonton, Alberta, Canada
andre2@ualberta.ca

Abstract— A methodology was developed to evaluate the fire-resistance of different commercially available siding materials when exposed to a high heat load condition (HHC) of 50 kW/m^2 and a low heat flux condition (LHC) of 20 kW/m^2 . The siding materials that were selected for evaluation were engineered wood, fibre cement board, cedar siding, and vinyl siding. Oriented strand board (OSB) was selected as a control material for these experiments. Prototypes consisting of a siding material, a weather barrier (building paper), and (OSB) were fabricated and exposed to the radiant heat fluxes emitted from an electrically-powered radiant heater. Time to ignition and surface temperature data gathered from the burn tests of the siding material prototypes were the main metrics that served to gauge the effect of the heat load on the prototypes and to establish their failure point. Tests were terminated after 30 minutes or if a prototype failed; whichever occurred earlier. The failure criteria was defined as the time to ignition (TTI) of the prototype, where ignition could be due to flaming ignition of the siding material, or flaming or glowing ignition of the OSB behind the siding material. The results of these tests suggest that it was possible to differentiate the performance of different siding materials subjected to radiant heat loads by evaluating both its TTI and the temperature variation of its surfaces (thermal response). A variance analysis confirmed that some of the siding materials were statistically different. The tests showed that cedar siding was the least ignition resistant material while fibre cement board was the most ignition resistant material under both radiant heat loads, in spite of that, the siding of this prototype did not ignite. Under the HHC, the results suggested that the use of engineered wood material represents no advantage when compared to bare OSB. Under the LHC, independently of the failure mechanism of engineered wood and vinyl siding, there was no significant difference between their times to reach the failure point. The temperature differences between the exposed and interface surfaces allowed to better understand the thermal behaviour of the siding materials.

Keywords: *Ignition, Radiant Heat Flux, Temperature Distribution, Time to Ignition, Siding Materials.*

I. INTRODUCTION

As part of the structural fire protection solutions for the Wildland Urban Interface (WUI), siding materials can help to withstand the adverse effects of direct flame contact or radiant heat originating from a fire to prevent structural ignition and to controlling better the spread of fire among buildings. However, there is limited research available that has assessed quantitatively the failure point of siding materials. Studies have analyzed the performance of supplementary materials such as intumescent coatings, wrapping materials, and wildfire chemicals under various fire conditions [1, 2]. Other studies focused on ignition and flammability tests to validate and calibrate ignition assessment models [3] and semi-quantitative inspections after prescribed burns [4–7] where focus of the analyses were to determine vulnerabilities from ember exposures, firebrand production, and flame propagation. Furthermore, siding materials that do not withstand radiant heat loads are still being considered as suitable options for the WUI [8]. Conversely, studies and agency reports [4, 9, 10] do not recommend vinyl siding, as it is likely to fail when it is exposed to high heat from fires.

During a fire, thermal radiation occurs prior to the arrival of a flame front. This radiant heat from the flames can raise the temperature of the exterior of a building to the point of ignition [11]; therefore, incident radiant heat flux loads have been used to evaluate the fire resistance of construction materials [1, 2, 12, 13]. Heat flux on surfaces provides a value that is easy to relate to the specific amount of energy to which an object will be exposed [14]. Heat fluxes around 20 kW/m^2 are sufficient to cause flashovers [15]. Silvani and Morandi [16] measured irradiance levels ahead of a fire front from 18 to 50 kW/m^2 , recommending these levels be used in thermal degradation studies in order to provide reliable data relevant to wildfires. The National Institute of Standards and Technology (NIST) to assessed the structure separation distance in multi-structure fires, where heat fluxes reported were to be between 20 and 70 kW/m^2 , with a value of 50 kW/m^2 for much of the fire event [17]. Many design choices in the built environment are based upon knowledge of fire behavior and its effects on risk [18].

In this study, the aim was to provide quantitative information on the failure point of different commercially

available siding materials by evaluating their performance when subjected to high and low radiant heat fluxes that are representative to both wildland fires and fires in the proximity of structures. Results from this study will help frame recommendations for choosing best siding materials for structures in the WUI.

II. EXPERIMENTAL METHOD

A. Materials and equipment

The siding materials that were used for testing were vinyl siding (Double 4.5 inch Dutchlap, Ply Gem Canada, Calgary, Alberta, Canada), cedar siding (Western red cedar, Windsor Plywood, Westlock, Alberta, Canada), engineered wood (fibre substrate, 76 Series Cedar Lap Siding, LP SmartSide, Louisiana-Pacific Corporation, Nashville, Tennessee, USA), and fibre cement board (HZ5 Hardie Plank lap siding, Hardie Board, James Hardie, Fontana, California, USA). OSB was used as the control material for the experiments. The dimensions of the composite wall (referred to as ‘prototype’) were $533 \times 533 \text{ mm}^2$ ($21 \times 21 \text{ inch}^2$). The prototype was comprised of the siding material, a moisture barrier (Tygar building wrap, Old Hickory, Tennessee, USA), and a 11 mm (0.4375 inch)-thick untreated OSB (APA Wood, Tacoma, Washington, USA). No insulation was considered for the construction of the prototypes. Insulation decreases the heat loss through the back of the prototype, causing prototypes to fail at faster rates. With both faster and potentially coincident failure, distinguishing between the performances of siding materials would be challenging.

Figure 1 shows the experimental assembly. An electric-powered radiant heater (Omegalux QH-121260, Omega Engineering, Inc., Laval, QC, Canada) was used to generate radiant energy that was incident on the siding materials. A multi-channel data acquisition unit (34970A Data Acquisition/Data Logger Switch Unit, Keysight Technologies Inc., Santa Rosa, California, USA) and stainless steel-braided J-Type 24 gauge thermocouples (Omega Engineering Inc., Laval, QC, Canada) were used to measure temperatures and collect data from:

- The exposed surface of the siding material (Figure 2(a), surface A)
- The interface surface; that is, the building paper-siding material interface (Figure 2(a), surface B)
- The unexposed or backside surface of the OSB (Figure 2(a), surface C)

A metal skid was used to mount the prototype steel holder and the radiant heater panel. The steel holder was used to hold in place either the prototypes or a Schmidt-Boelter heat flux sensor (Model No.: 64-10T-10R (CaF2W)-20898, MEDTHERM Co., Huntsville, Alabama, USA) that was used to measure the heat fluxes.

Five thermocouples were used in order to have one measuring point arranged along the vertical and horizontal centerlines. Figure 2 (b), 2(c), and 2(d) show a schematic of the thermocouple location. A total of fifteen thermocouples were

used for the siding material prototypes while only ten were required for the OSB (control material).

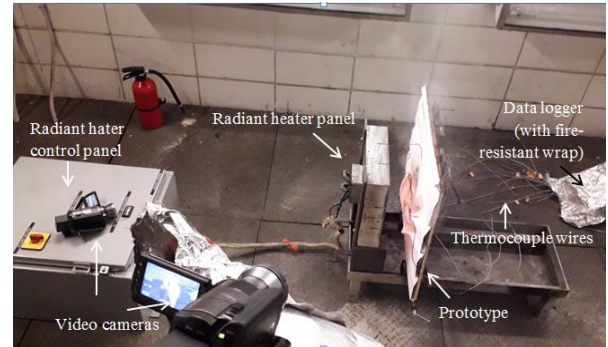


Figure 1. Experiment assembly.

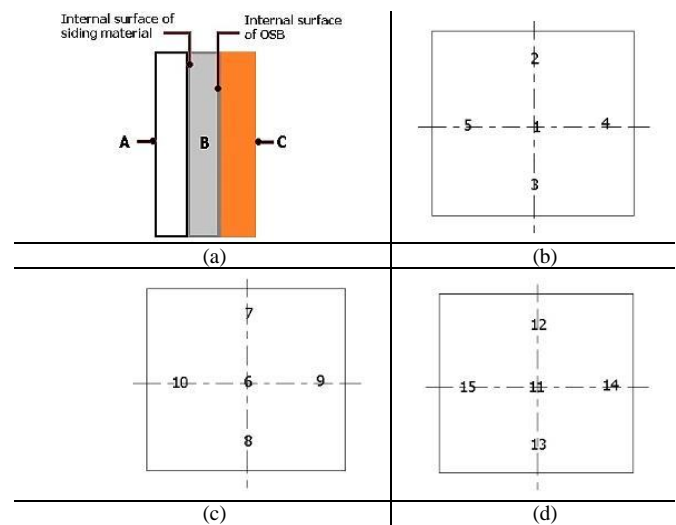


Figure 2. Schematic of (a) the composite wall comprised of the siding material exposed surface (A), the siding material-OSB interface (B), and the OSB unexposed or back surface (C) and (b) the thermocouple arrangement on each surface.

B. Burn tests

The burn tests for the present study were conducted at the Protective Clothing and Equipment Research Facility (PCERF) at the University of Alberta.

In order to expose the prototypes to a “constant” incident heat flux condition, a 2.5 inches-thick insulation board (shutter) (M-board, Industrial Insulation Group, Augusta, GA, USA) was used to shield the prototypes from incident radiation during the period in which the heat flux from the panel was increasing to a steady state value. Upon achieving steady state conditions of heat flux, the M-board was removed. This was considered as the starting point of the test (at $t = 0$).

Tests were terminated after 30 minutes or if a prototype failed; whichever occurred earlier. The failure criteria was defined as the time to ignition (TTI) of the prototype, where ignition could be due to flaming ignition of the siding material, or flaming or glowing ignition of the OSB behind the siding

material. Use of the TTI of exclusively the siding materials as the absolute test termination criterion was not suitable for these tests since the siding materials that were evaluated were both combustible materials (such as OSB and cedar siding) and non-combustible materials (such as vinyl siding).

In order to have quantitative evidence of the prototype failure, the transient temperature profiles from the thermocouples were used to both determine the failure point and to analyze the temperature variation of the surfaces (thermal response) of the prototypes. To generate sufficient data for statistical consistency, a minimum of three burn tests of each type of prototype was performed ($n = 3$) for the heat flux condition.

III. RESULTS AND ANALYSIS

A. Heat flux conditions

Two main parameters were used to obtain the desired heat flux conditions: the distance between the radiant heater panel and the samples (stand-off distance) and the heater set point temperature (HST). For the high heat flux condition (HHC), at a stand-off distance of 114.3 mm (4.5 inches) and a HST of 746°C (1375°F), a heat flux of $56.99 \pm 0.36 \text{ kW/m}^2$ was obtained. For the low heat flux condition (LHC), at the same stand-off distance and a HST of 520°C (968°F), a value of $20.3 \pm 0.31 \text{ kW/m}^2$ was measured. These conditions are representative of heat fluxes in both wildland fires and fires in the proximity of structures [15–17].

B. Time to ignition

1) Oriented strand board (OSB)

Oriented strand board (OSB) was used as the control material. No siding material or building paper was considered for these tests. Only ten thermocouples were used since no siding material was required. Under the HHC, the ignition occurred after $t = 433 \pm 66.70 \text{ s}$ of exposure to the incident radiant heat flux. An indication of the OSB ignition can be observed in Figure 3, which showed a sudden increase in the transient exposed surface temperatures at this time. Under the LHC, the tests were terminated when glowing ignition was observed on the exposure surface, which occurred after $t = 544 \pm 33.45 \text{ s}$ of exposure to the incident radiant heat flux. Under this condition, no flaming ignition was expected or observed, which was in accordance with results obtained from [13, 19].

2) Vinyl siding (VS)

Observations from the video footage showed that the vinyl siding (VS) melted after approximately 5-10 seconds of exposure under the HHC. Figure 4 shows the transient interface surface temperature behaviour of a vinyl burn test. A sudden increase in the temperatures occurred after $\bar{t} = 69 \pm 10.39 \text{ s}$ of exposure to the incident radiant heat flux, at this time, the ignition event occurred. Figure 5 show the ignition of the VS prototype. The melting of the vinyl siding took slightly longer under LHC in comparison to HHC, based on video evidence from the footage captured. It was also confirmed that no ignition occurred under this heat load

condition for either of the VS or the OSB that was part of this prototype. After an exposure time of $t = 1162 \pm 303.42 \text{ s}$, glowing ignition was visible at the interface surface.

3) Engineered wood (EW)

Flaming ignition was observed under both exposure conditions. Figure 6 shows the resulting temperature profile for the exposed surfaces under the HHC. The occurrence of ignition was visible by way of a nearly singular, rapid increase in temperature of the exposed surface; hence, under the HHC the TTI was $424 \pm 3.46 \text{ s}$.

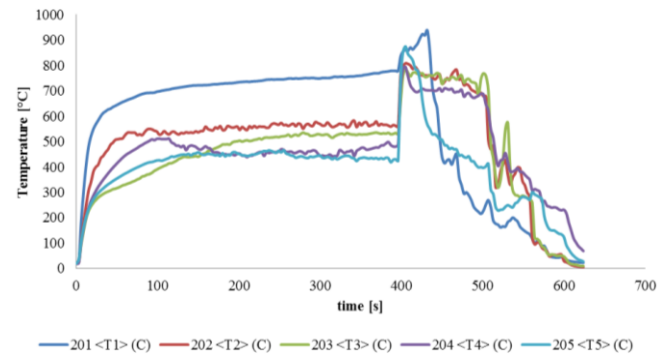


Figure 3. Exposed surface temperature behaviour of OSB under the HHC.

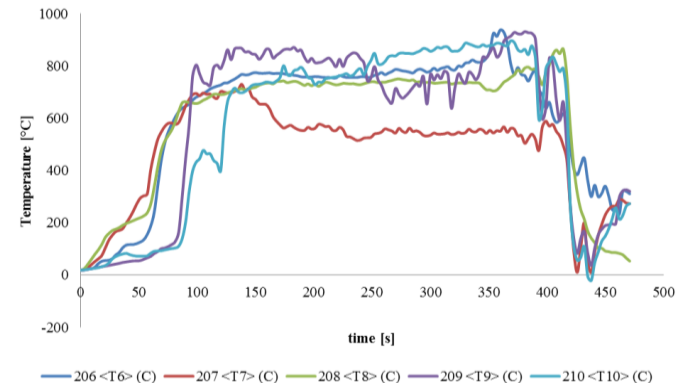


Figure 4. Interface surface temperature behaviour of VS under the HHC.

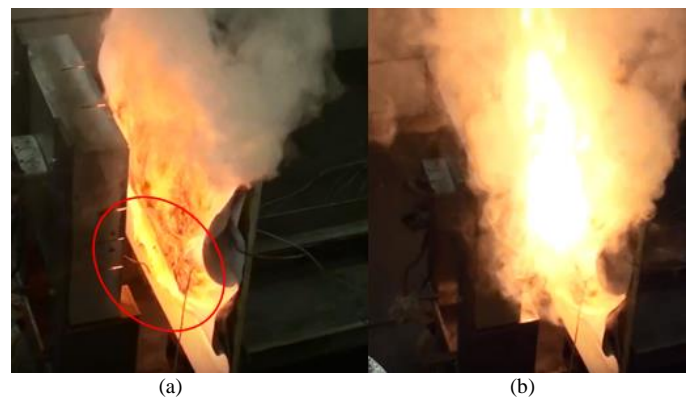


Figure 5. Images of (a) deformed VS prior to ignition and (b) VS prototype igniting.

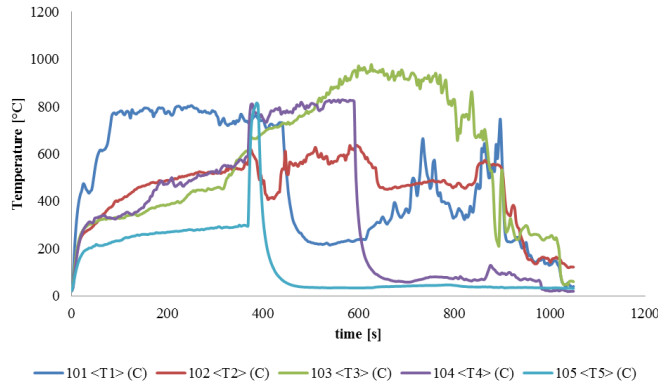


Figure 6. Exposed surface temperature behaviour of EW under the HHC.

Under the LHC, it was observed that flaming ignition process was initiated on the internal surface of the OSB at the interface surface rather than at the exposed surface of the prototype. This was confirmed by the rapid increase in the transient temperatures of the exposed and interface surfaces at the time when ignition occurred, as shown in Figure 7. It can be assumed that the continuous siding material loss that occurred as a result of the slower thermal degradation process eventually left the OSB directly exposed to the incident heat load. The resulting TTI under this condition was 1437 ± 71.38 s. Figure 8 shows a comparison between the HHC and LHC ignition processes. In Figure 8(a), it can be observed that the fuel gases that were generated from the EW ignited, whereas in Figure 8(b), flames were generated on the internal surface of the OSB at the interface of the prototype.

4) Fibre cement board (FCB)

Under both the HHC and the LHC, fibre cement board (FCB) siding did not ignite. However, under the HHC, flaming ignition occurred on the internal surface of the OSB at the interface. Under the HHC, the TTI was 1197 ± 197.45 s. Figure 9 shows the exposed and interface surface temperatures of the FCB for one burn test under the HHC, where a rapid increase in the interface surface temperature was indicative of the ignition event. Under the LHC, the tests were terminated at $t = 1800$ s where lower temperatures and no rapid increases occurred in the transient temperature profiles. Figure 10 shows the ignition of the FCB prototype.

5) Cedar siding (CS)

For the cedar siding (CS) material, flaming ignition occurred under both the HHC and LHC. Under the HHC and the LHC, the TTI was $t = 19 \pm 1.73$ s and 666 ± 374.17 , respectively. Figure 11 shows the time to ignition for the siding and the interface surfaces in a CS burn test.

6) Comparison between TTIs of siding materials

The time to ignition (TTI) of the different prototypes was compared using statistical software – R Studio. A variance analysis (ANOVA) with a confidence interval of 95% (or P value, $p < 0.05$) was performed to test the differences between the siding materials and their respective TTIs. Mean separation was done using Least Square Differences (LSD)

test with a confidence interval of 95%. The results from this analysis showed that some siding materials were significantly different.

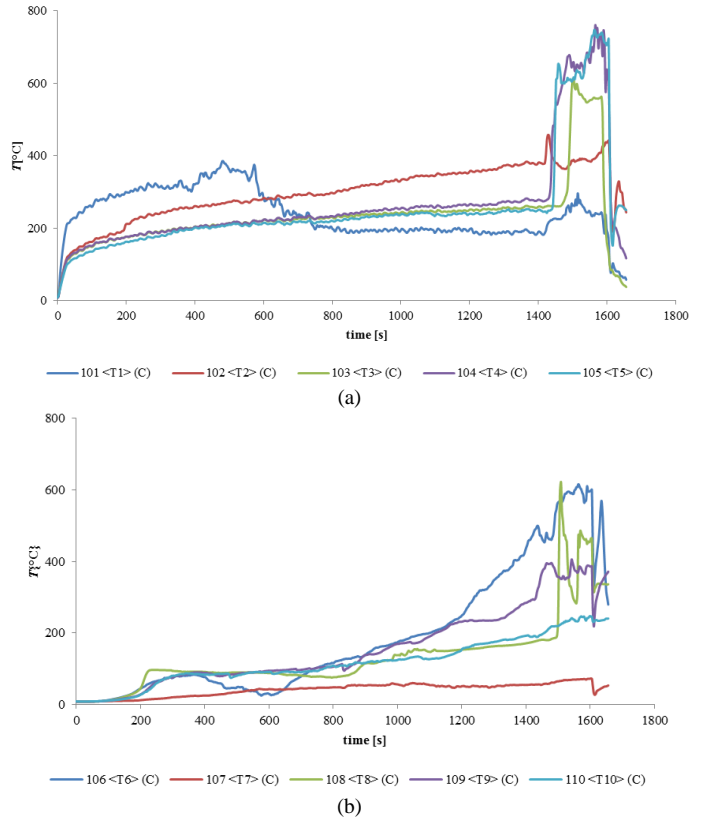


Figure 7. Rapid increase in temperature at (a) the exposed and (b) the interface surface of EW under LHC.



Figure 8. Comparison of the ignition process of the EW prototype under (a) the HHC and (b) the LHC.

Figure 12 shows the average TTI of different prototypes under HHC. The tests showed that cedar and vinyl prototypes failed the fastest while the fibre cement board took the longest time to fail. Statistical analysis showed that the difference in TTI was not significant between cedar and vinyl as a pair, owing to their similar TTIs and standard deviations. Likewise, statistical differences in TTI for engineered wood and bare OSB were also not found. Statistical differences in TTI were found for

fiber cement board compared to all other products despite its high standard deviation. These results suggest that fibre cement board offers the most resistance to ignition under high radiant heat loads while vinyl and cedar the least resistance under high radiant heat loads. Figure 13 shows the average TTIs of different prototypes under LHC. No TTI was shown for fire cement board since the prototype did not fail in the burn tests. Statistical analysis showed that cedar siding and bare OSB did not have significant differences in TTI, suggesting that the presence of cedar siding offers no real advantage to protect the OSB in its prototype from radiant heat.

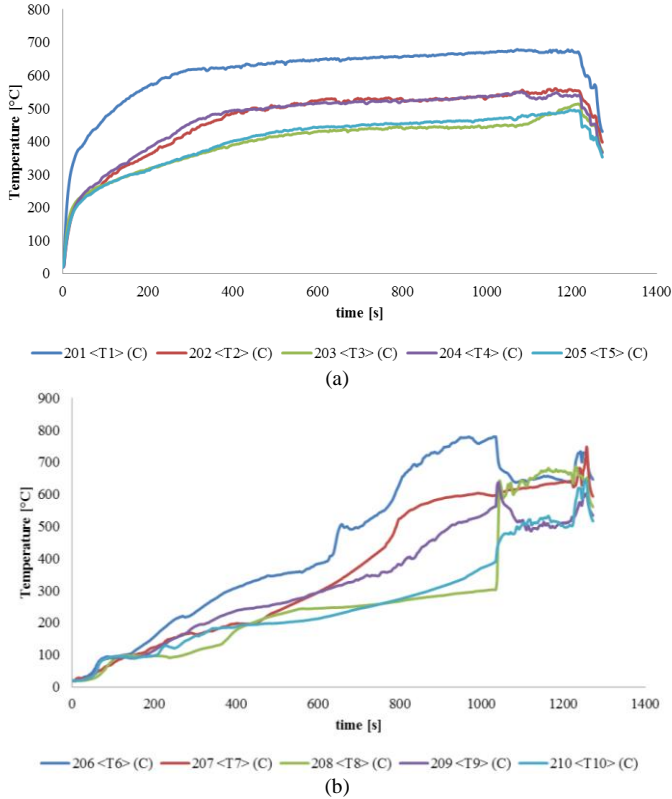


Figure 9. Temperature profile of the FCB during a burn test under HHC for the (a) exposed and (b) interface surface.

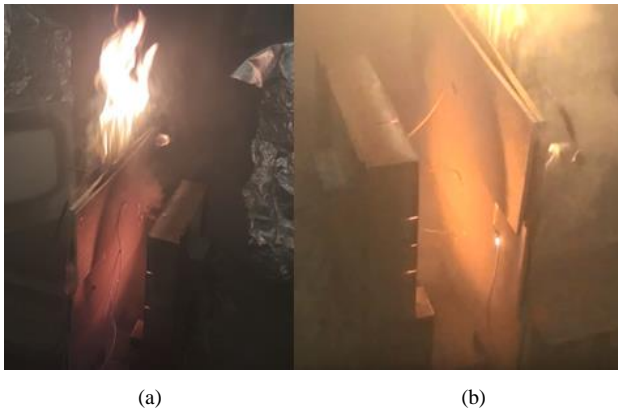


Figure 10. Ignition of the OSB during a FCB burn test from (a) left and (b) right view.

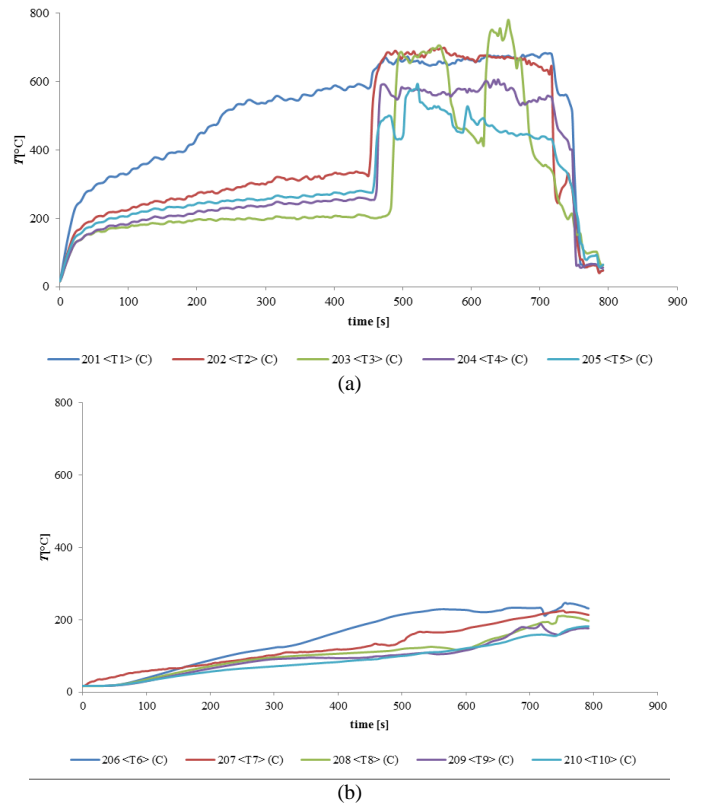


Figure 11. Temperature profiles of the CS under LHC for the (a) exposed and (b) interface surface.

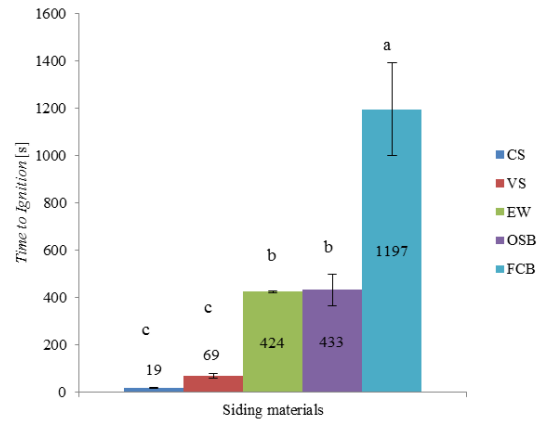


Figure 12. Time to Ignition of the siding material prototypes under the HHC. Treatments with the same letter (*a*, *b*, *c*) are not significantly different according to LSD test ($p < 0.05$). Vertical lines represent the standard deviation.

Also from Figure 13, it was observed that vinyl siding prototype took more than twice the time to fail when compared to the bare OSB. The engineered wood prototype took the longer time to fail; however, its resistance was not significantly different to the vinyl siding prototype. The large standard deviations of the cedar and the vinyl siding prototype can be attributed to the complex nature of the wood composition that is found in the cedar and to the resins and nature, composition, and orientation of the strands in the OSB.

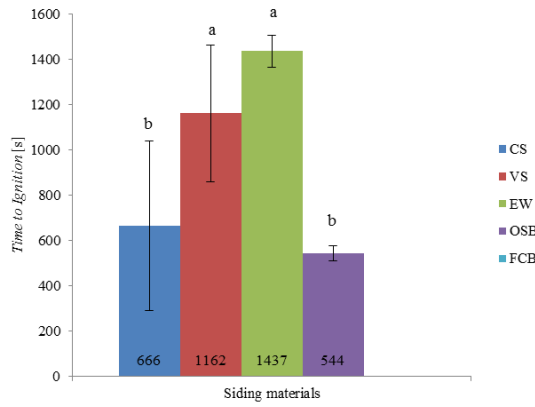


Figure 13. Time to Ignition of the siding material prototypes under the LHC. Treatments with the same letter (*a*, *b*, *c*) are not significantly different according to LSD test ($p < 0.05$). Vertical lines represent the standard deviation.

C. Ignition temperatures

Table I and II show summaries of the temperatures measured from the burn tests of the siding material prototypes. The temperatures exhibited by the siding materials previous to the flaming ignition event were defined as ignition temperatures (IT). Conversely, when no ignition occurred, the temperatures were defined as maximum material temperatures (MT).

In Table I, only the exposed surface of the fibre cement board (FCB) described a MT since the siding material did not ignite; the rest of the materials described ITs on both the exposed and the interface surfaces. In Table II, FCB, vinyl siding, and OSB described MTs since only glowing ignition or no ignition was observed.

TABLE I. SURFACE TEMPERATURES UNDER THE HHC

Materials	Exposed surface	Interface surface
Fibre cement board	633.95 ± 79.47°C <i>a</i>	742.25 ± 20.56°C <i>a</i>
Vinyl siding	N/A	448.04 ± 161.92°C <i>b</i>
Cedar siding	441.45 ± 55.67°C <i>b</i>	31.35 ± 23.11°C <i>d</i>
Engineered wood	628.14 ± 82.47°C <i>a</i>	244.87 ± 88.97°C <i>c</i>
OSB	749.03 ± 27.48°C <i>a</i>	277.98 ± 77.91°C (Unexposed surface)

N/A: Not applicable. Treatments with the same letter (*a*, *b*, *c*) are not significantly different according to LSD test ($p < 0.05$).

TABLE II. SURFACE TEMPERATURES UNDER THE LHC

Materials	Exposed surface	Interface surface
Fibre cement board	350.53 ± 21.95°C <i>a</i>	274.17 ± 24.2°C <i>b</i>
Vinyl siding	N/A	503.99 ± 113.3°C <i>a</i>
Cedar siding	340.01 ± 38.77°C <i>a</i>	277.01 ± 157.19°C <i>b</i>
Engineered wood	366.82 ± 59.78°C <i>a</i>	474.38 ± 99.26°C <i>a</i>
OSB	449.01 ± 33.08°C <i>a</i>	100.97 ± 9.58°C (Unexposed surface)

N/A: Not applicable. Treatments with the same letter (*a*, *b*, *c*) are not significantly different according to LSD test ($p < 0.05$).

Biswas, *et al.* [3] proposed different exposure conditions such as (1) short-duration (60 s), high intensity radiant heating between 40 and 50 kW/m², (2) short-duration, direct flame contact (100 kW propane flame), and (3) long-exposure (900 s), coupled radiant heating at 25 kW/m² and direct flame contact. Each test was chosen to simulate likely scenarios in the WUI. That research evaluated the response of a 1.2 x 1.2 m² fibre cement board under conditions (2) and (3). The measured temperatures reported for the short-duration, direct flame contact condition rose to about 300°C. Although the experimental method is not comparable with the present study, there is agreement with the exposed surface temperatures of the FCB under the LHC (Table II). Surface temperatures for the coupled condition were not reported to compare them with the HHC results for the FCB of the present study.

Biswas, *et al.* [3] also measured an “interior” vinyl siding temperature range, which varied from 350 to 500°C when the siding material was exposed to a 40 kW/m² radiant heat load. In the present study, the ignition temperature was 448°C (Table I), which falls within that reported range.

From a 40 kW/m² radiant heating test, cedar siding ignition temperatures were approximately 350°C [3]. In the present study, under the HHC and LHC, the temperatures at ignition of the CS were 441°C and 340°C, respectively. Although there is close agreement with the ignition temperature obtained under the LHC, there is a difference of approximately 100°C between the temperatures obtained under the HHC and the 40 kW/m² test. The additional radiant energy (20 kW/m² higher than the condition used by Biswas, *et al.* [3]) impinging on the thermocouples may have heated them at a faster rate, affecting the temperature measurements.

1) Comparison between ignition temperature of siding materials

In Table I, a large difference was observed between the exposed and the interface surface temperatures of the cedar siding. This temperature difference was due to the thermal inertia through the cedar siding. Thermal inertia is defined as the product of material thermal conductivity, k , density, ρ , and specific heat capacity, c ($k\rho c$). The lower the thermal inertia, the faster the surface temperature will rise [15, 19]. Table III shows the thermal inertia of the siding materials analyzed in the present study. It can be observed that the thermal inertia of the cedar siding corresponds to the lowest value of these siding materials.

Under the HHC (Table I) cedar siding had both the lowest exposed and interface surface temperatures. Results from Table I and 3, suggest that the use of cedar siding is more unfavourable as compared to the use of bare OSB. From the surface temperatures of the fibre cement board (FCB) prototype, it was observed that the interface surface temperature was in accordance with such from the exposed surface of the bare OSB at the onset of the ignition event. Comparing the surface temperatures from the FCB and the engineered wood (EW) it can be noted that the exposed surface temperatures were similar; nevertheless, the FCB prototype

ignited after 1200 seconds whereas the EW prototype ignited approximately three times faster. This results suggested that FCB siding shielded and resisted the heat more effectively than the EW siding. Regarding the vinyl siding prototype, an important decrease in the ignition temperature of the internal surface of the OSB at the interface was observed when compared with such from the FCB and bare OSB. Further analysis will be required to understand the occurrence of this temperature variation caused by the thermal degradation of both the vinyl siding and OSB.

Under the LHC (Table II) the temperature difference between the exposed and the interface surface of the cedar siding was much lower as compare to such from the bare OSB at their corresponding failure points; yet flaming ignition occurred in the former, while only glowing ignition occurred in the latter. Similarly as for the HHC, under the LHC, FCB and EW prototypes exhibited approximately the same exposed surface temperatures with the difference that the FCB siding shielded effectively the prototype from the heat (no flaming or glowing ignition ensued on the OSB at the interface of the prototype) while the EW siding degraded to the point of a direct exposure of the OSB that leaded to flaming ignition. The vinyl siding melted at an early stage of the burn test as occurred under the HHC; though, its interface surface temperature was in agreement with the exposed surface of the bare OSB, suggesting that, at this heat flux condition, the thermal degradation of the vinyl siding did not affected the thermal response of the OSB.

TABLE III. THERMAL INERTIA OF THE SIDING MATERIALS ANALYZED

Materials	c [J/kg·K]	k [W/m·K]	ρ [kg/m ³]	Thermal inertia [J ² /m ² ·K ⁴ ·s]
Fibre cement board [20]	840	0.245	1380	2.840×10^5
Engineered wood [21]	1300	0.14	800	1.456×10^5
Vinyl siding [21]	840	0.1	1470	1.235×10^5
OSB [20]	1880	0.092	650	1.124×10^5
Cedar siding [20]	1880	0.085	336	5.369×10^4

IV. CONCLUSIONS

The test methodology provided useful information to evaluate the performance of siding materials subjected to two radiant heat loads, a high heat flux condition (HHC) of 57 kW/m² and a low heat flux condition (LHC) of 20.3 kW/m². A failure criterion was established as the time to ignition of the prototype, where ignition could be due to flaming ignition of the siding material, or flaming or glowing ignition or the OSB behind the siding material. The transient temperature profiles from the thermocouples that were affixed to the surfaces of the prototypes were used to determine the failure point and to analyze the temperature variation of the surfaces (thermal response) of the prototypes. A variance analysis (ANOVA) confirmed that some siding materials were significantly

different when comparing their time to ignition (TTI) and surface temperatures.

Under the HHC, it was found that the cedar and vinyl prototypes were the fastest prototypes to fail while the fibre cement board (FCB) prototype had the longest TTI. The FCB siding did not ignite; in contrast, flaming ignition occurred on the internal surface of the OSB at the interface of this prototype, this further illustrated the ignition resistance of the fibre cement board. No significant differences were obtained between the bare OSB and the engineered wood prototype, suggesting that using engineered wood to protect the sheathing materials of a composite will not represent any advantage.

Under the LHC, the FCB prototype did not failed. Cedar siding had the least TTI. The vinyl and engineered wood prototypes were found to have statistically similar time to ignitions; this result suggests that regardless of the different failure mechanism that these two materials experienced (flaming ignition or melting of the siding material); the time to reach the failure point was the same.

It was noted that between the LHC and the HHC established, a critical heating rate exists at which the ignition mechanism of the OSB will bypass glowing ignition process to flaming ignition directly. From the burn test of the OSB, it was confirmed that the main thermal decomposition process under a heat flux of 20 kW/m² will be glowing ignition.

Further research will be needed to determine the OSB and vinyl siding by-products that were derived from the combustion of the materials under stoichiometric conditions for comparison with measurements of the composition of the combustion by-products.

V. REFERENCES

- [1] B. Bahrani, V. Hemmati, A. Zhou, and S. L. Quarles, "Effects of natural weathering on the fire properties of intumescent fire-retardant coatings," *Fire and Materials*, vol. 42, no. 4, 2018, 413–423.
- [2] A. Zhou, "Performance evaluation of ignition-resistant materials for structure fire protection in the WUI," in *Proceedings of the 13th Int. Conference and Exhibition on Fire and Materials*, San Francisco, CA, 2013, 355–366.
- [3] K. Biswas, D. Werth, and N. Gupta, "A Home Ignition Assessment Model Applied to Structures in the Wildland-Urban Interface," p. 10.
- [4] Insurance Institute for Business and Home Safety, "Wildfire Demonstration (2011) - IBHS," [Online]. Available: <https://disastersafety.org/ibhs/research-center-demo-wildfire-2011/>. [Accessed: 11-Jan-2019].
- [5] S. L. Manzello, S. Suzuki, and Y. Hayashi, "Exposing siding treatments, walls fitted with eaves, and glazing assemblies to firebrand showers," *Fire Safety Journal*, vol. 50, 2012, 25–34.
- [6] S. Suzuki and S. L. Manzello, "Firebrand production from building components fitted with siding treatments," *Fire Safety Journal*, vol. 80, 2016, 64–70.
- [7] S. L. Manzello, S. Suzuki, and D. Nii, "Full-Scale Experimental Investigation to Quantify Building Component Ignition Vulnerability from Mulch Beds Attacked by Firebrand Showers," *Fire Technology*, vol. 53, no. 2, 2017, 535–551.
- [8] M. Dobson, "Vinyl Siding: Fire-safe and right for Fort McMurray," *Vinyl Institute of Canada*. [Online]. Available: <http://vinylinstituteofcanada.com/vinyl-siding-fire-safe-and-right-for-fort-mcmurray/>. [Accessed: 11-Jan-2019].
- [9] S. Walkinshaw, D. Schroeder, and S. Hvenegaard, "Evaluating the effectiveness of FireSmart Priority Zones for Structure Protection | FireSmart Canada", 2012.

- [10] U.S Department of Homeland Security, “Home Builder’s Guide to Construction in Wildfire Zones”, 2008.
- [11] C. K. Randall, “Fire in the Wildland-Urban Interface: Understanding fire behavior.” University of Florida IFAS Extension, USDA Forest Service.
- [12] F. Takahashi, T. M. Murray, A. Abbott, S.-Y. Hsu, J. S. T’ien, and S. L. Olson, “Performance evaluation criteria of fire blanket materials for structure protection in wildland-urban interface fires,” presented at the Fall Meeting of the Eastern States Section of the Combustion Institute 2009, 2009, 645–650.
- [13] S. Hirle and K. Balog, “The Effect of the Heat Flux on the Self-Ignition of Oriented Strand Board,” *Research Papers Faculty of Materials Science and Technology Slovak University of Technology*, vol. 25, no. 40, 2017, 123–129.
- [14] S. Sudheer, L. Kumar, B. S. Manjunath, A. Pasi, G. Meenakshi, and S. V. Prabhu, “Fire safety distances for open pool fires,” *Infrared Physics & Technology*, vol. 61, 2013, 265–273.
- [15] N. N. Daeid, “Fire Investigation,” 1st Ed., CRC Press, 2005, 6–12.
- [16] X. Silvani and F. Morandini, “Fire spread experiments in the field: Temperature and heat fluxes measurements,” *Fire Safety Journal*, vol. 44, no. 2, 2009, 279–285.
- [17] A. Maranghides and E. J. Johnsson, “NIST Technical Note 1600: Residential Structure Separation Fire Experiments.” U.S Department of Commerce, National Institute of Standards and Technology, 2008.
- [18] S. E. Caton, R. S. P. Hakes, D. J. Gorham, A. Zhou, and M. J. Gollner, “Review of Pathways for Building Fire Spread in the Wildland Urban Interface Part I: Exposure Conditions,” *Fire Technology*, vol. 53, no. 2, 2017, 429–473.
- [19] J. G. Quintiere, “Principles of Fire Behavior,” 2nd. Edition., FL, USA: Taylor and Francis Group, 2017.
- [20] S. V. Glass, V. Kochkin, S. C. Drumheller, and L. Barta, “Moisture Performance of Energy-Efficient and Conventional Wood-Frame Wall Assemblies in a Mixed-Humid Climate,” *Buildings*, vol. 5, no. 3, 2015, 759–782.
- [21] Y. A. Cengel and A. Ghajar, *Heat and mass transfer fundamentals and applications*, Fifth Edition. McGraw-Hill Education, 2015.

INVERSE HEAT TRANSFER APPLICATION IN A SCALED MODEL OF A POWER TRANSMISSION LINE TOWER FOUNDATION

Daqian Zhang^{1,2}, Guangjun Wang², Xili Duan¹

¹Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, A1B 3X5 Canada

²School of Energy and Power Engineering, Chongqing University, Chongqing, 400044 China
 daqianz@mun.ca, wangguangjun@cqu.edu.cn, xduan@mun.ca

Abstract—In many northern regions, the metal footings of power transmission line towers are buried in permafrost. With a high thermal conductivity, the tower footing has a significant thermal effect on the local permafrost around the foundation. In previous work, heat transfer in the tower foundation was analyzed by modeling the tower footing as a line heat source. However, the distribution of heat strength for this line source is not clear. In this paper, an inverse heat transfer method based on dynamic matrix control (DMC) is used to develop a predictive model for estimation of the heat source strength distribution. A finite volume method (FVM) is used to develop a two-dimensional numerical model for ground heat transfer around the tower footing. Then, the temperatures at several measurement points in a scaled model are used as the input in the inverse heat transfer analysis to estimate the time-space-varying distribution of the heat source. Finally, the estimated heat source strength is used to calculate the transient temperatures in the tower foundation. This method could provide a more accurate heat transfer analysis for the design or maintenance of the tower foundations.

Keywords—ground heat transfer; tower footing; line heat source; inverse heat transfer

I. INTRODUCTION

Detailed analysis of ground heat transfer is necessary for the design and maintenance of power transmission line foundations that are built in the northern regions [1]. In these regions, the footing structures of the power transmission towers are buried in permafrost, the temperature of which remains at or below 0 °C at a certain depth for two or more years [2]. The excellent mechanical strength of the permafrost provides a good foundation for the buried tower structures. Above the permafrost is the active layer, which usually experiences seasonal thawing and freezing over seasonal changes, but it does not have the same mechanical properties as the permafrost. The structure of a power transmission line tower foundation is illustrated in Fig. 1. The tower footing is fixed firmly in permafrost, and the tower has an energy exchange with solar radiation Q_r , convection Q_c , and heat conduction Q_g . However, one of the problems for the integrity of the foundation is the risk from the gradual local degradation of

permafrost caused by the buried metal structures [3]. With the high thermal conductivity of the metal tower, heat is conducted through the tower downwards into the ground or upwards into the surrounding during seasonal weather changes [4].

An appropriate model to analyze heat transfer in the power transmission tower foundation is critical for structural integrity estimation and maintenance of transmission lines in these permafrost regions. In the past few decades, many valuable studies were conducted on ground heat transfer within a foundation of a structure. Several analytical models were developed [5-6], in which the ground was approximated as a semi-infinite medium. Deng and Fedler [7] used a two-dimensional model of unsteady heat conduction to predict the heat transfer in multi-layered soils with vertical ground-coupled heat pumps; Zeng et al. [8] presented an analytical line-source model of boreholes in geothermal heat exchangers. However, these works only considered the heat source as a line source with constant or uniform strength. In the past few years, Duan and Natterer [9-10] developed a Variable Heating strength (VHS) line heat source model to predict the transient heat conduction in the ground foundation with a buried power transmission line tower, and the model has a higher accuracy in that situation.

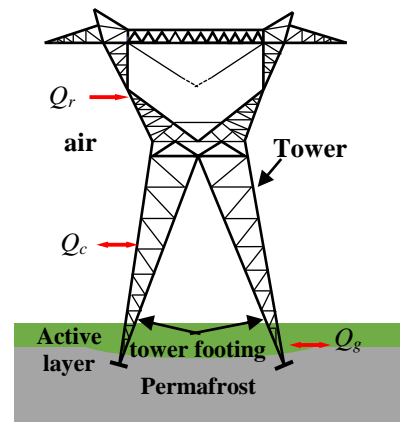


Figure 1. Schematic of Power transmission line tower foundation

Unfortunately, these models have some defects for practical engineering applications. Firstly, the quantity of the heat conducted into the tower footing does not always remain as a constant or a fixed form over time. When the tower footing is modeled as a line heat source, its heating strength actually varies with time and space. In this case, estimating the unknown heat source distribution is the key for the heat transfer analysis in the ground.

Inverse Heat Transfer Method (IHTM) is an effective way to determine the unknown properties (in this case the strength of the line heat source) using temperature measurements in a heat transfer system. In the past few decades, many useful IHTMs have been developed. Wang et al. [11] developed a double decentralized fuzzy inference (DDFI) method with a temporal-spatial decoupling characteristic for an inverse heat conduction problem to estimate the time and space-dependent thermal boundary condition. Yang et al. [12] used the conjugated gradient method (CGM) and the discrepancy principle to estimate the unknown time-dependent heat generation at the interface of cylindrical bars during rotary friction welding process from the knowledge of temperature measurements taken within the stationary bar. Samadi and Kowsary [13] applied the sequential function specification method (SFSM) to estimate the transient heat flux imposed on the rake face of a cutting tool during the cutting. These applications showed great success of the inverse heat transfer methods.

This paper aims to estimate the distribution of the heat strength of a finite line source representing the tower footing in a scaled power transmission line foundation model by using an inverse heat transfer method. The ground temperatures are then reconstructed with the estimated heat source strength. First, a two-dimensional direct numerical model is established for ground heat transfer around the tower footing. Following that, a mathematic model of the IHTM based on DMC is developed. Then, a series of numerical experiments are conducted on a scaled test cell, and the temperatures of measurement points are used to estimate the time-space-varying distribution of the heat strength of the line source. Finally, the inversed heat strength is used as the input to reconstruct the ground temperatures in the tower foundation.

II. DIRECT NUMERICAL MODEL

In order to simplify the heat transfer analysis, the tower footing is approximated as a single metal rod with a diameter of D in the ground. The simplified physical model and the positions of temperature measurements P_n are shown in Fig. 2. The simplified tower footing is buried with a depth of H_g ; the main exchanged heat from the tower footing is modeled as a line heat source with strength $q(z,t)$ [4]. The ground is expressed as a $2R \times Z$ two-dimensional symmetrical space in the heat transfer system. For higher calculation efficiency, only half of the space is considered in the analysis. The left, right and bottom sides of the domain are set as adiabatic boundary conditions. The top side is subjected to a uniform temperature T_g and it is also measured by thermocouples. Five temperature measurement points are placed on the right side of the tower footing.

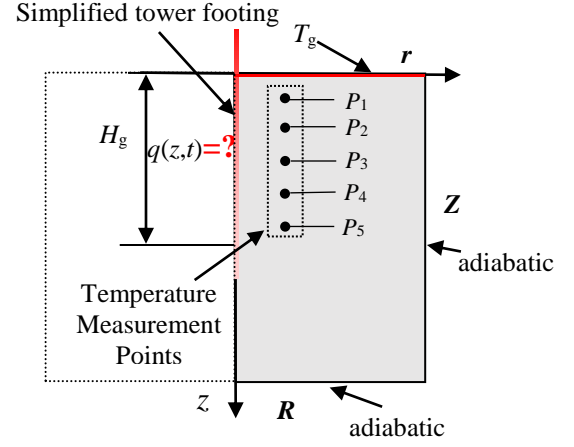


Figure. 2 Simplified scaled physical model for the heat transfer in a power transmission line tower foundation

The governing equation, initial condition and boundary conditions for the transient heat conduction in the ground are expressed as:

$$\frac{1}{\alpha_g} \frac{\partial T_g(r, z, t)}{\partial t} = \frac{\partial^2 T_g(r, z, t)}{\partial r^2} + \frac{1}{r} \frac{\partial T_g(r, z, t)}{\partial r} + \frac{\partial^2 T_g(r, z, t)}{\partial z^2} + Q_g \quad (1a)$$

where

$$Q_g = \begin{cases} q(r_0, z, t) & r = r_0, -H_g \leq z \leq 0 \\ 0 & \text{else} \end{cases} \quad (1b)$$

$$T_g|_{t=0} = T_i \quad (1b)$$

$$T_g = T_g(t) \quad r \geq r_0, z = 0 \quad (1c)$$

$$-\lambda \frac{\partial T_g(r, z, t)}{\partial r} = 0 \quad r = r_0, z \leq 0 \quad (1d)$$

$$-\lambda \frac{\partial T_g(r, z, t)}{\partial r} = 0 \quad r = R, z \leq 0 \quad (1e)$$

$$-\lambda \frac{\partial T_g(r, z, t)}{\partial z} = 0 \quad z = Z, r \geq 0 \quad (1f)$$

Here, T_g is the objective temperature function of the ground at location r, z and time t ; T_i is the initial temperature of the ground; $\alpha_g = \lambda/\rho c$ is the thermal diffusivity of the ground; λ, ρ and c are the corresponding thermal conductivity, density and specific heat capacity, respectively.

The boundary temperature $T_g(t)$ and heat strength $q(z,t)$ are both set as the following sinusoidal forms in Eq.(2) and Eq.(3a-3b), which are used to represent the annual variations [9]:

$$T_g(t) = T_0 + A_g \sin(w_g t / T_{total} + \phi_g) \quad (2)$$

$$q(z, t) = q_0(z) + A_q(z) \sin(w_q t / T_{total} + \phi_q) \quad (3a)$$

$$q_0(z) = Q_g / H_g \text{ or } q_0(z) = 2Q_g \times (1 - |z| / H_g) \quad (3b)$$

where T_0 and q_0 are the average values during the period, A_g and A_q are the amplitudes of the variation, w_g and w_q are the frequencies of variation, ϕ_g and ϕ_q are the phase lags. Q_g is the total input heat energy. The total period T_{total} is set as 146 minutes for the scaled model.

The transient temperature field $T_g(z, r, t)$ of the heat transfer system defined by Eq. (1a-1f) can be determined by a tri-diagonal matrix alternating direction implicit algorithm [14] with the finite volumes methods (FVM). $(T_m^k)_{exa}$ is the temperature of the measurement point m at the time step k . The simulated “measured” temperatures are generated by adding Gaussian white noises to the “exact” temperatures at points of measurement, as follows:

$$Y_m^k = (T_m^k)_{exa} + \omega \sigma \quad (4)$$

where ω is a random variable of normal distribution with zero mean in the interval $[-2.576, 2.576]$, σ is the standard deviation.

III. INVERSE PROBLEMS AND INVERSE METHOD

In practical situations, the time-space-varying heat strength $q(z, t)$ is an unknown quantity in Eq. (1a-f). The inverse method based on dynamic matrix control (DMC) [15] is applied to determine the $q(z, t)$ by utilizing the temperatures at the measurement points.

A. Formulation of a predictive model for the inverse problem

The heat strength $q(z, t)$ can be discretized into a set of components $q_s^k, (s=1, 2, \dots, S; k=1, 2, \dots, T)$, which has some quantitative relationship with the temperature $Y_m^k (m=1, 2, \dots, M; k=1, 2, \dots, T)$ of the measurement points, where s is the discrete time serial number, k is the discrete space serial number and m is the measurement point serial number.

Each component q_s^k can be determined through the future temperature information with the time at $t=k, k+1, \dots, k+r-1$, where r is the future time step. The component q_s^k is estimated with time.

First, the matrix equation of the predictive model is established according to Eq. (5),

$$\mathbf{T} = \tilde{\mathbf{T}}|_{q=0} + \mathbf{A} \mathbf{Q}_g^k \quad (5)$$

where \mathbf{T} is the predictive temperature matrix of the points of measurement. $\tilde{\mathbf{T}}|_{q=0}$ is the predictive temperature matrix when the heat strength q^k is zero at $t = k$. \mathbf{A} is the relationship matrix of sensitivity coefficients. \mathbf{Q}_g^k is the heat strength matrix. where

$$\mathbf{T} = [\mathbf{T}^k, \mathbf{T}^{k+1}, \dots, \mathbf{T}^{k+r-1}]^T, \mathbf{T}^{k+n} = [T_1^{k+n/k}, T_2^{k+n/k}, \dots, T_M^{k+n/k}]^T,$$

$$\tilde{\mathbf{T}}|_{q=0} = [\tilde{\mathbf{T}}^k, \tilde{\mathbf{T}}^{k+1}, \dots, \tilde{\mathbf{T}}^{k+r-1}]^T, \\ \tilde{\mathbf{T}}^{k+n} = [\tilde{T}_1^{k+n/k}, \tilde{T}_2^{k+n/k}, \dots, \tilde{T}_M^{k+n/k}]^T,$$

$$\mathbf{A} = \begin{bmatrix} \Delta\phi^0 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \Delta\phi^1 & \Delta\phi^0 & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta\phi^{r-1} & \Delta\phi^{r-2} & \Delta\phi^{r-3} & \dots & \Delta\phi^0 \end{bmatrix},$$

$$\Delta\phi^n = \begin{bmatrix} \Delta\phi_{1,1}^n & \Delta\phi_{1,2}^n & \dots & \Delta\phi_{1,S}^n \\ \Delta\phi_{2,1}^n & \Delta\phi_{2,2}^n & \dots & \Delta\phi_{2,S}^n \\ \vdots & \vdots & \ddots & \vdots \\ \Delta\phi_{M,1}^n & \Delta\phi_{M,1}^n & \dots & \Delta\phi_{M,S}^n \end{bmatrix},$$

$$\mathbf{Q}_g = [q^k, q^{k+1}, \dots, q^{k+r-1}]^T, \mathbf{q}^{k+n} = [q_1^{k+n}, q_2^{k+n}, \dots, q_M^{k+n}]^T, \\ (n=0, 1, \dots, r-1; m=1, 2, \dots, M; s=1, 2, \dots, S).$$

where $\tilde{T}_m^{k+n/k}$ is the predicted temperature when the heat strength q_s^{k+n} is zero at $t = k+n$. $\Delta\phi_{m,s}^{n/k}$ is the relationship sensitivity coefficients.

B. Step response coefficients

The sensitivity coefficient $\Delta\phi_{m,s}^n$ represents the response degree of point m concerning the discrete heat flux q_s^{k+n} during the time interval $[k, k+r-1]$, which is calculated by Eq. (6).

$$\Delta\phi_{m,s}^n = \frac{\partial T_m^{k+n/k}}{\partial q_s^k} \approx \frac{T_m^{k+n/k}(q_s^{k,0} + \Delta q_s^k) - T_m^{k+n/k}(q_s^{k,0})}{\Delta q_s^k} \quad (6) \\ n = 0, 1, 2, \dots, r-1$$

where $\Delta q_s^k = 1, q_s^{k,0} = 0$.

C. Rolling optimization objective function

Differentiating Eq. (5) with regard to \mathbf{Q}_g^k and making the optimal estimation of heat strength \mathbf{Q}_g^k according to the rolling optimization strategy:

$$\mathbf{q}^k = \Phi[(\mathbf{A}^T \mathbf{A} + \alpha \mathbf{E})^{-1} \mathbf{A}^T (\mathbf{Y} - \mathbf{T})] \quad (7)$$

where

$$\mathbf{Y} = [\mathbf{Y}^k, \mathbf{Y}^{k+1}, \dots, \mathbf{Y}^{k+r-1}]^T, \mathbf{Y}^{k+n} = [\mathbf{Y}_1^{k+n}, \mathbf{Y}_2^{k+n}, \dots, \mathbf{Y}_M^{k+n}]^T,$$

$$\Phi = [\mathbf{E}_s, \mathbf{0}, \dots, \mathbf{0}]_{r \times s}.$$

IV. RESULTS AND DISCUSSION

For the heat transfer system in Fig. 2, we used the properties based on a scaled test case [10] to carry out the simulation. In this case, the initial temperature is set as $T_i = 0$ °C, the radius of the medium and height of the medium are set as $R = 100$ mm and $Z = 200$ mm, respectively. The thermal properties of the medium are as follows: the density ρ is 780 Kg/m³, the thermal diffusivity α is 1.6×10^{-7} kJ/kg·K, thermal conductivity λ is 0.24 W/m·k, and the total input heat power Q_g is set as 1.13W.

The geometry domain is meshed into 50 and 100 equal parts along the r - and z -direction separately; the time step size $\Delta t = 5$ s. The future step $r = 3$. The positions of temperature measurement points are placed at $r = 7$ mm in R direction and $z_1 = -6.7$ mm, $z_2 = -20.1$ mm, $z_3 = -33.5$ mm, $z_4 = -47$ mm, $z_5 = -60.5$ mm in the Z direction.

A. Ground transient heat conduction with a line heat source

When the heat strength $q(t)$ is uniform, the temperature response at the measurement points over time is shown in Fig. 3. The temperature response varies with sinusoidal forms throughout the period.

B. Inverse of the Heat source

The heat strength may have different spatial distribution, and for this research, two typical distributions, i.e., uniform and linear forms, are applied in the simulation. The standard deviations σ of the measurement temperature errors are set as 0.01, 0.05, 0.1, respectively.

The corresponding exact distributions and inversed heat source distributions with different standard deviations σ are shown in Fig. 4 and Fig. 5. The simulation results show that there is some degree of deviations with the existence of measurement noises with both distribution forms. It is found that the inversed method could provide stable and accurate results when the $\sigma = 0.01$. With the increasing of σ , the inversed results show some fluctuations and they become more serious with the further increasing of σ . The main reason for this phenomenon is that the random measurement noises are becoming equal or even larger than the temperature change during each time step. However, the trends of temperature variation can still be tracked in a relatively stable way, which can reflect the exact distribution with space and time.

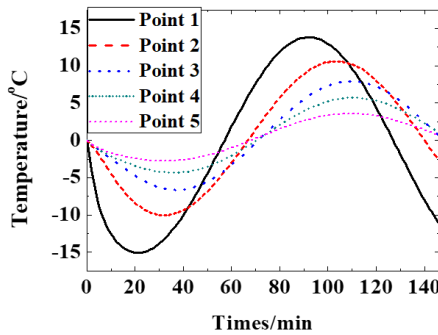


Figure 3. Temperature response at the points of measurement

C. Reconstruction of temperatures in the ground

The whole temperature field in the foundation can be reconstructed based on the inversed heat source strength in nearly real time. As an example, the inversed heat strength is considered as a linear distribution and the standard deviation is set as $\sigma = 0.1$. Fig. 6 (a, b) show the foundation temperature distributions at $t = 10$ min (the simulated coldest time in “winter”) and $t = 83$ min (the simulated warmest time in “summer”), respectively. The results indicate that the temperatures near the tower footing of the ground are obviously higher than that of other positions at the same level during the summer time, and the situation during winter time is just the opposite.

V. CONCLUSION

In this study, an inverse heat transfer method based on DMC is developed to estimate distributions of heat strength in the tower footing of a scaled test case. The inversed method is validated in the application of a power transmission line tower foundation, and then the reconstructed temperatures based on the inversed heat source are shown at two typical time points. The following conclusions can be drawn:

- (1) The inverse heat transfer method shows a very good effectiveness when the standard deviation σ is smaller than 0.05;
- (2) Bigger measurement errors lead to bigger deviations and fluctuations;
- (3) The inverse method still provides relatively stable estimation even in a bigger standard deviation $\sigma = 0.1$;
- (4) The temperature fields are reconstructed with very small errors.

In later work, the inverse heat transfer method will be further developed to consider phase change heat transfer for other applications.

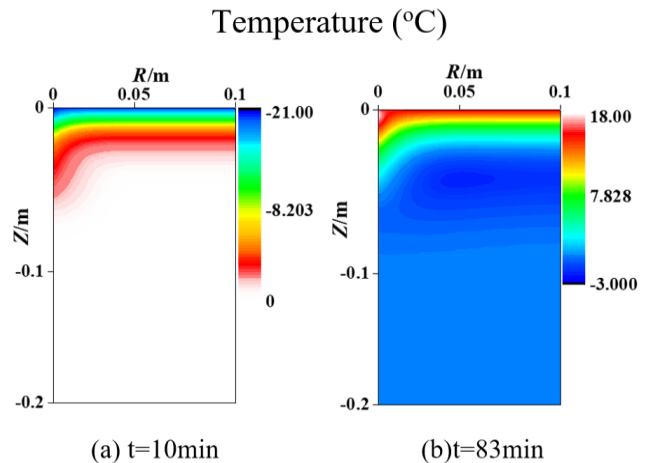
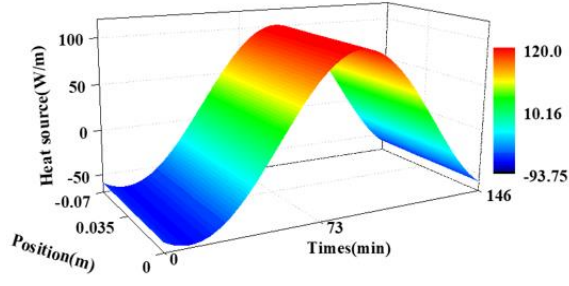
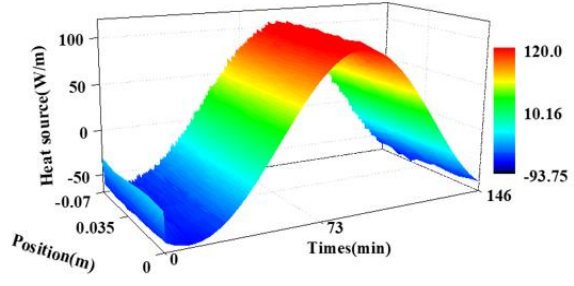


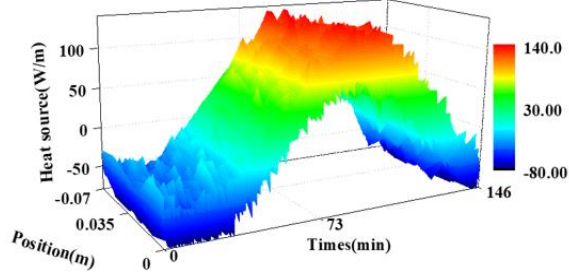
Figure. 6 Reconstruction of temperature fields in the ground: (a) $t=10$ min (b) $t=83$ min



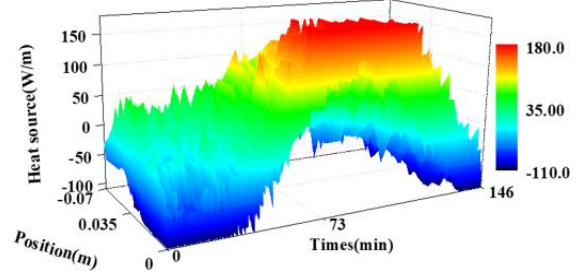
(a) Exact value



(b) Inversed heat source($\sigma=0.01$)

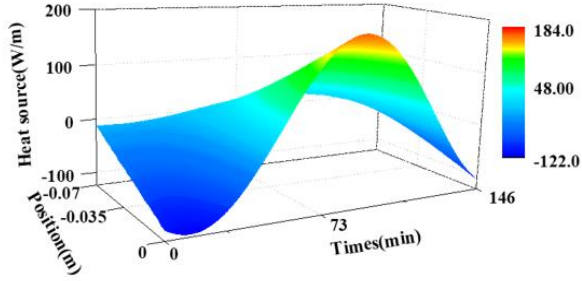


(c) Inversed heat source($\sigma=0.05$)

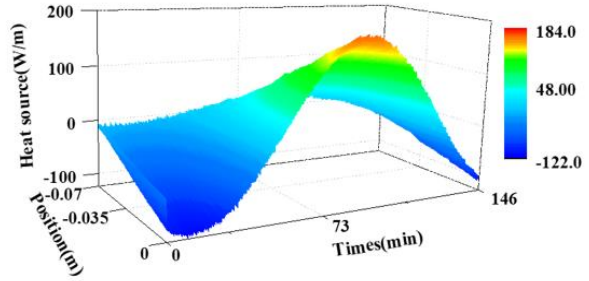


(d) Inversed heat source($\sigma=0.1$)

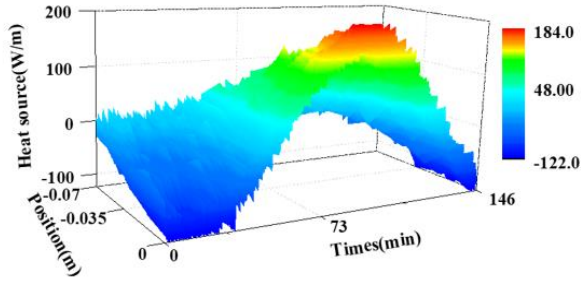
Figure 4. Comparison of Exact and Inversed heat source distribution with different standard deviation (uniform distribution)



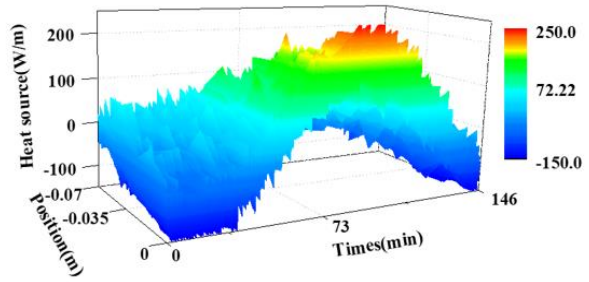
(a) Exact value



(b) Inversed heat source($\sigma=0.01$)



(c) Inversed heat source($\sigma=0.05$)



(d) Inversed heat source($\sigma=0.1$)

Figure 5. Comparison of Exact and Inversed heat source distribution with different standard deviation (linear distribution)

ACKNOWLEDGMENT

The work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. The first author(D. Zhang) also wish to acknowledge financial support from China Scholarship Council.

REFERENCES

- [1] Streletskiy D A, Shiklomanov N I, “Nelson F E. Permafrost, infrastructure, and climate change: a GIS-based landscape approach to geotechnical modeling,” *Arctic, antarctic, and alpine research*, vol. 44, no. 3, pp. 368-380, 2012.
- [2] Andersland O B, Ladanyi B, *An Introduction to Frozen Ground Engineering*, Springer US, 1994.
- [3] Mu Y, Wang G, Yu Q, et al, “Thermal performance of a combined cooling method of thermosyphons and insulation boards for tower foundation soils along the Qinghai–Tibet Power Transmission Line,” *Cold Regions Science and Technology*, vol. 121, pp. 226-236, 2016.
- [4] Duan, X., Naterer, G. F., Lu, M., and Mueller, W., “Transient heat conduction from a vertical rod buried in a semi-infinite medium with variable heating strength,” *Heat and mass transfer*, vol. 43, no. 6, pp. 547, 2007.
- [5] Sodha, M.S., R.L. Sawhney and A. Sengupta, “Shape factor for an underground vertical infinite cylindrical structure,” *International Journal of Energy Research*, vol. 18, no. 4, pp. 431-436, 1994.
- [6] Sodha, M.S., R.L. Sawhney and B.C. Jayashankar, “Estimation of steady state ground losses from earth coupled structures by simulation,” *International Journal of Energy Research*, vol. 14 no. 5, pp. 563–571, 1990.
- [7] Deng, Y. and C.B. Fedler, “Multi-layered soil effects on vertical ground-coupled heat pump design,” *Transactions of the Asae*, vol. 35, no. 2, pp. 687-694, 1992.
- [8] Zeng, H.Y., N.R. Diao and Z.H. Fang, “A finite line - source model for boreholes in geothermal heat exchangers,” *Heat Transfe-Asian Research*, Vol. 31 no. 7, pp. 558-567, 2002.
- [9] Duan, X., and G. F. Naterer, “Ground heat transfer from a varying line source with seasonal temperature fluctuations,” *Journal of Heat Transfer* vol. 130, no. 11, pp. 111302, 2008.
- [10] Duan, X., and G. F. Naterer, “Heat transfer in a tower foundation with ground surface insulation and periodic freezing and thawing,” *International Journal of Heat and Mass Transfer*, vol. 53, no. (11-12), pp. 2369-2376, 2010.
- [11] Wang, G., Wan, S., Chen, H., Lv, C., and Zhang, D, “A double decentralized fuzzy inference method for estimating the time and space-dependent thermal boundary condition,” *International Journal of Heat & Mass Transfer*, vol. 109, pp. 302-311, 2017.
- [12] Yang Y C, Chen W L, Lee H L, “A Nonlinear Inverse Problem in Estimating the Heat Generation in Rotary Friction Welding,” *Numerical Heat Transfer*, vol. 59 no. 2, pp. 130-149, 2011.
- [13] Samadi, F., Kowsary, F., Sarchami, A., “Estimation of heat flux imposed on the rake face of a cutting tool: a nonlinear, complex geometry inverse heat conduction case study,” *International Communications in Heat and Mass Transfer*, vol. 39, no. 2, pp. 298-303, 2012.
- [14] Tao, W., *Numerical heat transfer*. Vol. 5. Xi'an Jiaotong University, 2001. [陶文铨. 数值传热学. Vol. 5. 西安交通大学出版社, 2001].
- [15] Li Y, Wang G, Hong C, “Simultaneously estimation for surface heat fluxes of steel slab in a reheating furnace based on DMC predictive control,” *Applied Thermal Engineering*, vol. 80, pp. 396-403, 2015

Materials Engineering

Additive Manufacturing of Aluminum Alloys: Opportunities for Extreme Environments

Mohsen Mohammadi

Marine Additive Manufacturing Centre of Excellence (MAMCE)
University of New Brunswick, Fredericton, NB, E3B 5A1, Canada

Metal additive manufacturing techniques including selective laser melting, direct energy deposition, and electron beam melting have received tremendous attentions in last couple of years. Of particular interest is additive manufacturing of aluminum alloys for extreme situations, where complex loading conditions including compression and tension, impact, and cyclic forces often in corrosive environments are present. These complex loading scenarios are common for marine, aerospace, defence, and energy applications. Employing different electron microscopy techniques, the microstructure of a selective laser melted AlSi10Mg alloy system was investigated. A hierarchical type of microstructure including dislocation networks, nanoscale Si precipitates, cell walls and boundaries, grain structures, and melt pools was identified. The fabricated samples were then subjected to different loading conditions including quasi-static tension and high strain rate impacts using a split Hopkinson pressure bar apparatus. The mechanism of deformation for the alloys was then identified using multiscale advanced electron microscopy techniques. A strengthening model accounting for different deformation sources was finally proposed to study the effects of various phenomena active in the deformation of SLM-AlSi10Mg.

KEYWORDS: Laser additive manufacturing; AlSi10Mg; hierarchical microstructure; hardening law.

Novel welding processes and joint evaluation techniques in pipeline construction

Adrian P. Gerlich
Associate Professor,
University of Waterloo

New welding processes have continually emerged which have aimed to increase the productivity and reliability of the welding process. These features are of particular importance in large scale construction of infrastructure such as transmission pipelines, where considerable volumes of weld metal are required, and welding must be performed in harsh conditions. Utilizing a new process is often hampered by qualification of welding procedures and acceptance of these processes in existing codes and standards. Some of the latest fusion based welding processes will be reviewed which combine multiple heat sources such as arc and laser in more efficient configurations to increase production speed and quality. New methods for evaluating joint quality and properties for both procedure qualification and quality control in the field will be discussed, such as instrumented indentation and online thermal monitoring. A perspective on future welding processes such as so-called 'one-shot' welding techniques that involve drastically faster thermal cycles.

CHARACTERIZATION OF LONG-TERM MECHANICAL PERFORMANCE OF POLYETHYLENE (PE) AND ITS PIPE

Ben Jar*, Na Tan, Yi Zhang, and Souvenir Muhammad

Department of Mechanical Engineering
University of Alberta
Edmonton, Canada
ben.jar@ualberta.ca

ABSTRACT

This presentation covers a long journey of research on polyethylene (PE), from a surprising discovery of elastic modulus “degradation” of PE plaque to design of short-term tests to detect deformation transition in PE pipe. Our studies have come up with an explanation for both the elastic modulus degradation and the deformation transition, which serves as an example for the close relevance between the fundamental, curiosity-driven studies and the practical applications. Key outcomes from these studies are a series of new test schemes that can be used to evaluate mechanical performance of plastics which cannot be quantified using standard tests.

The presentation will start from a problem we faced when applying finite element (FE) simulation to mimicking the necking behavior of PE in tension. The FE simulation produced clear evidence on the significance of viscous component on the overall stress response to PE deformation. By removing the viscous component from the overall stress response to deformation, the elastic modulus degradation can then be characterized based on a damage concept. Such a concept was later given a physical meaning for PE based on data from wide-angle X-ray scattering (WAXS), which shows the complexity of deformation in the crystalline phase of PE lamellae before they are disintegrated to a fibrillary form. The results were then verified using a different approach. Understanding the deformation mechanisms has allowed us to design loading conditions to target deformation in a particular phase in PE (either amorphous or crystalline phase). The same approach was then applied to PE pipe, to determine the critical stress level for the deformation transition and time for its occurrence. Overall objective of these studies was to design mechanical testing schemes so that deformation introduced at the macroscopic level can be used to characterize influence of mechanisms that occur at the micro- or even nano-scaled levels, on the mechanical properties for PE.

In addition to the above results that have been obtained when this abstract is prepared, this presentation will provide results from an on-going research work which is to explore the possibility of using the above test schemes to investigate deformation resistance of a particular phase at the micro- or nano-scale. Potential applications for such information will also be discussed.

Moisture Effect on Mechanical Performance of Out-of-Autoclave Composite Material at Different Void Level

Afshin Bayatpour, Mehdi Hojjati
Concordia Centre for Composites
Concordia University
Montreal, Quebec, Canada
afshin.bayatpour@gmail.com, mehdi.hojjati@concordia.ca

ABSTRACT

Composite materials and especially carbon fiber reinforced polymers (CFRPs) have been broadly used due to high specific strength and stiffness ratio over metallic materials in the past few decades. CFRPs are widely utilized in aerospace sectors, civil infrastructure and marine sectors under harsh environment. As a result, there exists a strong demand to observe the mechanical behavior of composite materials while being exposed to moisture environment to avoid unpredictable effects. To reach better mechanical properties, composite materials normally cure in the autoclave which is less cost-efficient compared to out-of-autoclave process - especially for large structures. However, there exists a higher possibility of void formation in the out-of-autoclave process due to lack of higher pressure during manufacturing. This study aims, primarily, to observe the moisture absorption response of composite samples. Moreover, we study the sensitivity of CFRPs durability in harsh environment. By changing vacuum pressure during the manufacturing process, three different unidirectional laminates at three levels of void content have been manufactured. After immersing the samples to the warm water at 60°C, moisture absorption behavior of all laminates was observed by monitoring the weights of samples. Furthermore, mechanical behavior of these laminates has been studied at four different moisture levels by performing dynamic mechanical analysis (DMA) and Interlaminar shear strength (ILSS) testes. Empirical results indicate that different properties including the interlaminar shear strength, glass transition temperature, and storage modulus reduce while exposing the samples with different void contents to the moisture environment.

Discrete-element modeling of nacre-like materials: statistical variation, nonlinear deformations and fracture

Najmul Abid

Department of Mechanical Engineering
McGill University
Montreal, Canada
najmul.abid@mail.mcgill.ca

Francois Barthelat

Department of Mechanical Engineering
McGill University
Montreal, Canada
francois.barthelat@mcgill.ca

ABSTRACT

Natural materials such as nacre, collagen and spider silk are composed of staggered stiff and strong inclusions in a softer matrix. This specific type of hybrid microstructure generates remarkable combinations of stiffness, strength and toughness. Here we use discrete element modeling (DEM) to capture the mechanical response of large statistical volume elements (SVEs) of staggered composites which incorporate statistical variations. We find that even the slightest statistical perturbations precipitate strain localization and decrease ductility by a large fraction, while these negative effects can be offset by strain hardening and large strain at the interfaces. We also present DEM fracture models with +100,000 inclusions that capture crack deflection, crack bridging, inelastic process zone and residual deformations in the wake of the crack. The fracture of staggered composite is strongly affected by the arrangement of the inclusions and statistical variations. These findings suggest new design guidelines for bio-inspired composites with superior performance.

Quantifying the Effect of Hydrogen Charging on the Microstructure of Steel using Computed Tomography Imaging

Liam S. Morrissey, Stephen M. Handrigan, Sam Nakhla¹

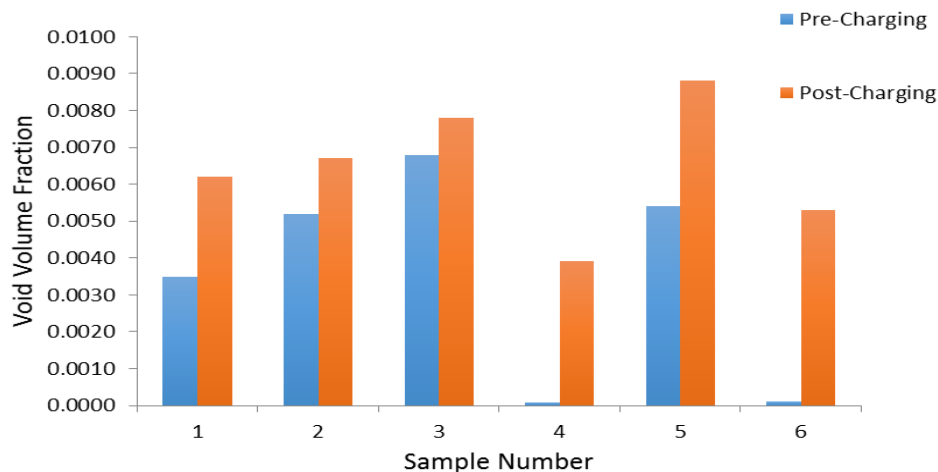
Department of Mechanical Engineering
 Memorial University of Newfoundland
 St. John's, Canada

¹ Department of Mechanical Engineering/ Department of Emergency Medicine

ABSTRACT

Hydrogen diffusion into the microstructure is a key first step for both hydrogen embrittlement and hydrogen blistering. In blistering, hydrogen recombines at voids and internal pressures induce plastic deformation of the metal lattice. For embrittlement, research has suggested that hydrogen may influence initial void fraction, which, in turn, has a direct influence on behaviour under loading. Moreover, despite the clear importance of these microscopic voids, observing and quantifying them in metals is difficult. For example, despite the importance of hydrogen accumulation at voids for both blistering and embrittlement, no study has actually quantified the effect of hydrogen on void formation. Therefore, the objective of the current study was to quantitatively determine the role hydrogen charging alone has on the void volume fraction of steel samples.

Six 10mm diameter cylindrical 13-Cr steel samples were cut from a rod and imaged using a CT scanner before and after hydrogen charging. CT imaging involves taking several images while rotating the sample, and then reconstructing these projections into a 3D model. To identify voids a global thresholding method was used that segments the model based on a single gray threshold value. Individual void volumes were then summed to produce an initial void fraction of steel. Results indicated that average initial void fraction was 0.0035, with a large range between 0.0001 and 0.0068. After charging the void fraction increased by an average of 18 times with an average void fraction of 0.0065. Therefore, hydrogen charging alone, without loading, was shown to increase the void fraction in steel. One potential explanation for the source of the void growth is a micro void nucleation theory which posits a nucleation mechanism in which hydrogen introduces super abundant micro vacancies into the metal. Then, under the presence of hydrogen, these micro vacancies migrate and aggregate, allowing for the formation of a hydrogen molecule. This hydrogen molecule then introduces further stresses into the material and leads to the formation of a larger vacancy cluster. Therefore, the growth of voids being imaged in the current study is likely the aggregation of microvoids as atomic hydrogen recombines into hydrogen gas, a preliminary step that could lead to a surface blister.. The current work provides strong evidence that hydrogen recombination in voids can cause significant changes to the microstructure, even without loading, and obvious surface blisters.



Molecular Dynamics Investigation on the Effect of Vacancies on Young's Modulus of Pure Iron

Stephen M. Handrigan*, Liam S. Morrissey, and Sam Nakhla¹

1. Faculty of Engineering and Applied Science (Mechanical) and Faculty of Medicine (Emergency Medicine)

Faculty of Engineering and Applied Science (Mechanical)

Memorial University of Newfoundland

St. John's, Canada

Corresponding author: smh113@mun.ca

ABSTRACT

The utilization of Molecular Dynamics (MD) to investigate and understand material properties as a result of atomistic processes is continually increasing. Macroscale mechanical properties vary sample to sample and are averaged amongst many samples. The reason for variance in seemingly identical samples is a result of the arrangement of atoms, point defects, dislocations, and more all occurring on the atomistic and molecular scale within the material. Therefore, it is of critical importance to fully understand the processes occurring on the atomistic and molecular scales to fully comprehend the macroscale mechanical properties.

One such atomistic process that can affect mechanical properties is the presence of point defects, such as vacancies. To observe the effect of vacancies on Young's modulus, room temperature uniaxial tension was applied to a $\langle 100 \rangle$ structure of single-crystal pure iron with vacancies ranging from 0 – 5% using two different many-body force fields – Modified Embedded Atom Method (MEAM) and Tersoff. Deformation was applied in the (100) crystal direction. The stress-strain curve was then plotted and the slope of the stress-strain curve provided Young's modulus. Stress was the virial stress and strain was engineering strain. Strains were kept to less than 3% to ensure linear elastic deformation for utilization of Hooke's law. Simulations were repeated for both force fields at each vacancy percentage. All Young's moduli were normalized to each respective force field's zero vacancy predicted modulus.

It was observed that the effect of vacancies to reduce the Young's modulus of pure iron was highly dependent on the selected force field (Figure 1). The Tersoff force field was able to capture the trend of experimental data for the entire vacancy range, whereas MEAM was only able to predict an accurate reduction for less than 4% vacancies. As such, it is clear that the selection of force field is highly influential on the results obtained for mechanical properties and that coupling MD with a correctly selected force field can result in highly accurate data comparable to macroscale experimental results.

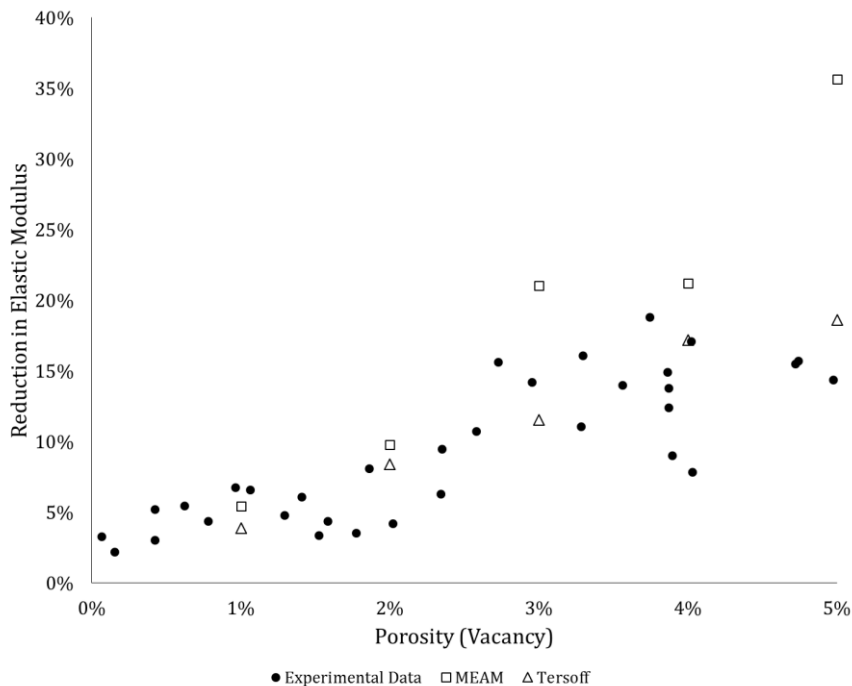


Figure 1 – Comparison of Reduction in Elastic Modulus for MEAM and Tersoff Force Fields to Experimental Data

Nanoindentation of WC-Co cemented carbides fabricated by Selective Laser Sintering

Joseph Agyapong^{*1}, Alex Czekanski² and Solomon Boakye-Yiadom³

^{1, 2, 3} Mechanical Engineering Department,
Lassonde School of Engineering, York University
Toronto, ON, Canada, M3J 1P3

ABSTRACT

Tungsten Carbide-Cobalt (WC-Co) cemented carbides are widely used as cutting, forming and machining tools. They have been employed in the mining, medical and manufacturing industries for their applications because of their high hardness and strength and excellent wear resistance properties. In addition, the WC-Co cemented carbides are also used for high temperature applications because of the ability to retain their bulk properties at high temperatures. This has been reported to be attributed to their complex composite structure of a hard, brittle carbide phase, usually WC or W₂C, and a tough metallic binder, usually Co, which has tungsten (W) and carbon (C) dissolved in it. The latter imparts some ductility/fracture toughness to the material while the former provides the high hardness and wear resistance. This makes WC-Co one of the important materials that has great potential in different application sectors. Recently, optimized 3D printing parameters have been used to process Tungsten Carbide-Cobalt (WC-Co) cemented carbides. This process induces non-equilibrium phases in the microstructure even before application under service conditions because of the irregular temperature changes and cooling patterns occurring at the same time during the processing stage. Disregarding this might affect the integrity of the adoption of the new manufacturing technique for rapid tooling even though it has desirable properties. In this study, optimized processing parameters were used to produce WC-Co (87vol%-17vol%) cemented carbides using Selective Laser Sintering (SLS). The processed specimens were heat treated to understand the effect of heat treatment on the microstructural integrity of the specimens. In order to ascertain the properties of the different constituents, nanoindentation using Atomic Force Microscopy is adopted, to be able to derive the contribution to hardness and strength of the material by the different constituent phases formed in the as-printed WC-Co material. This provides the knowledge needed to explain further how temperature and service conditions affect the structural integrity and performance of 3D printed WC-Co cemented carbides.

EFFECT OF GRAPHENE AND ZIRCONIA ON WEAR AND FRACTURE BEHAVIOUR OF ALUMINA MATRIX NANOCOMPOSITE

Solomon Hanson Duntu^{*1}, Solomon Boakye-Yiadom³

Mechanical Engineering Department, Lassonde School of Engineering,
York University
Toronto, ON, Canada, M3J 1P3

Mohammad Islam²

Advanced Manufacturing Institute,
King Saud University,
Riyahd 11421, Saudi Arabia

ABSTRACT

Advanced ceramic alumina presents relatively superior mechanical properties such as low density, high refractoriness, high specific strength and high damping capacity than monolithic alumina. However, their high temperature structural applications such as aircraft engine parts and rocket materials are limited due to their brittle nature. To improve the fracture toughness and wear properties of nanocrystalline alumina, graphene(0.5vol%), and zirconia (4vol% and 10vol%) are added to the alumina matrix via colloidal mixing followed by conventional hot-pressing technique. The fabricated specimens were subjected to detailed microstructural and mechanical characterization techniques to evaluate their microstructural integrity and mechanical properties such as densities, hardness, surface roughness (Sa), wear properties and fracture toughness. Mechanical tests such as Vickers micro-indentation and tribological tests were used to determine the hardness and frictional wear characteristics of the nanocomposite. Crack extensions from Vickers indentation were measured and used to characterize the fracture toughness of the nanocomposites. The wear track surfaces were also analyzed using the SEM and a 3D optical profilometer to determine the wear mechanisms and 3D profile of wear tracks. Microstructural characterization of the fabricated nanocomposites using the optical and scanning electron microscopy (SEM) revealed a relatively refined microstructure with the addition of 0.5vol% graphene and 10vol% zirconia than monolithic alumina, which contributed to their improved mechanical properties. The refined structure was attributed to the homogeneous dispersion of the graphene and zirconia phases within the matrix. The inclusion of 4vol% zirconia gave a rather more coarse microstructure with depleted mechanical properties than pure alumina due to insufficient dispersion of zirconia phase which led to alumina grain growth. Also, there was a relation between the spatial and volume distribution of zirconia phases on the wear resistant properties and fracture behavior of the alumina-zirconia nanocomposites which was discussed in details in this work.

MICROSTRUCTURAL EVOLUTION OF ADIABATIC SHEAR BANDS IN AZ31 MAGNESIUM ALLOY DURING HIGH STRAIN RATE IMPACT

Francis Tetteh * ¹

Mechanical Engineering Department, Lassonde School of
Engineering
York University
Toronto, Canada
ftetteh@yorku.ca

Solomon Boakyey-Yiadom ²

Mechanical Engineering Department, Lassonde School of
Engineering
York University
Toronto, Canada
sboakyey@yorku

ABSTRACT

Structural materials such as steel, Aluminum (Al) and Magnesium (Mg) are easily susceptible to strain localization, which results in the formation of Adiabatic Shear Bands (ASBs) during high strain rate impact. It is generally agreed that the initiation of strain localization and formation of ASBs are manifestations of damage in metallic materials subjected to high strain rates and large strains of deformation and may lead to catastrophic failure instantaneously. However, Strain localization and formation of ASBs is a complex problem that is still not fully understood in structural materials under high strain rate loading conditions. This is attributed to the narrow nature of the ASBs when they are formed and the rapid rates of deformation which makes it almost impossible to observe their evolution and mechanism of formation during impact. In this study, it is hypothesized that a systematic study of the microstructure of a material prior-to, and after impact can be used to track the microstructural changes that occur during the evolution of ASBs. This research is initiated to systematically study the microstructure of AZ31 Mg Alloys prior to impact and after impact to determine the effect of the pre-deformation microstructure on the nucleation and initiation of ASBs, and the mechanism of evolution of ASBs during impact. The AZ31 Mg alloys were heat treated at 400°C for 2 hours followed by high strain rate impact using the direct impact Hopkinson pressure bar (DIHPB) at different impact momentum. Strain localization, formation of adiabatic shear bands and initiation of cracks are studied using the optical microscope, scanning electron microscope and microhardness tests. Extensive grain refinement was observed within evolved ASBs along propagating cracks with the ASBs regions having high hardness compared to regions away from the ASBs. It was observed that, despite the brittle nature of the Mg alloy, ASBs evolved with cracks along the path of the ASBs. In addition, twins and micro twins were observed around evolved ASBs and away from the ASBs.

Influence of strength and trapping characteristics on material susceptibility to hydrogen Embrittlement (HE) based on experimental investigations and finite element analyses (FEA)

Tuhin Das

*Department of Mining and Materials Engineering
McGill University
Montreal, QC, Canada
tuhin.das@mail.mcgill.ca*

Salim V. Brahim

*Department of Mining and Materials Engineering
McGill University
Montreal, QC, Canada*

Jun Song

*Department of Mining and Materials Engineering
McGill University
Montreal, QC, Canada*

Stephen Yue

*Department of Mining and Materials Engineering
McGill University
Montreal, QC, Canada*

A number of materials including high strength steels suffer from premature failures due to hydrogen embrittlement (HE). The failure of these materials could be catastrophic based on the choice of applications. Therefore, in the present study, a combined approach based on experimental and finite element (FE) analyses has been adopted to understand HE failure in relation to material susceptibility. Hardness is an important material property that could be accessed easily, influences material susceptibility to HE. However, hardness is an outcome of the strength and microstructure of the material. The effect of hardness on material susceptibility has been studied extensively by employing simplified ASTM F1624 standard test method. It was observed that the HE susceptibility of the material increases with the increase in hardness.

FE simulations have been carried out to develop fundamental understanding of the problem related to various material parameters and properties. Hydrogen trapping behavior is also included into the model to point out its impact on material susceptibility considering the different cases of hydrogen charging. Finally, the calculations obtained from the FEA model are corroborated with experimental observations.

Thermal and Dynamic Mechanical Analysis of Biochar Filled Glassfiber Polymer Composite and Hemp-fiber Polymer Composite

Raj Kumar Dahal
School of Engineering
University of Guelph
Guelph, Canada
rdahal@uoguelph.ca

Bishnu Acharya
Faculty of Sustainable Design Engineering
University of Prince Edward Island
Charlottetown, Canada
bacharya@upei.ca

ABSTRACT

Agro-processing is a major industry in Canada. The processing of an agricultural product generates an enormous amount solid waste. The appropriate management of these waste streams has become a challenging issue. Biochar from spruce pellets pyrolysis at 450°C, when used in powdered form as a filler material in composite materials, will improve the mechanical and thermal properties of a composite material. On the other hand, hemp fiber as a reinforcement adds sustainability in the composites. This paper presents the comparison study of mechanical and thermal properties of biochar filled Glassfiber reinforced polymer composite with the hemp fiber reinforced polymer composite. Glassfibre and hemp fiber reinforced polymer composites were prepared at different biochar concentration. Samples were vacuum infused for curing. Thermal properties are analysed with DSC and TGA. Flame retardancy property of the cured composites was analysed by performing UL 94V-0 test. Results from 3-point bend test for mechanical properties was performed in Q800 DMA. The flame test suggested that higher biochar concentration in the composites retards the flame better. DMA tests showed increase in storage moduli thus stiffness of composites with higher biochar in the composites. Biochar caused lower damping in the composites.

Crystal orientation effect on the incipient plasticity of the thin film nanoindentation

Mahdi Bagheripoor

Mechanical and Materials Engineering Department
The University of Western Ontario
London, Canada
mbagher6@uwo.ca

Robert Klassen

Mechanical and Materials Engineering Department
The University of Western Ontario
London, Canada
rjklasse@uwo.ca

ABSTRACT

The present study aims to investigate the anisotropic behavior of crystallographic gold substrate in (001), (011) and (111) surface orientations during nanoindentation. Molecular dynamics simulation is used to compensate for the limitations of experimental nanoindentation. The homogenous defect nucleation and evolution in Au single crystal of these three crystal orientations is studied and a tight connection is made between the force-displacement curve and defect structure. Various forms of dislocation propagation are observed in different crystal orientations the elastic-plastic transition point appears later for the (111) oriented surface than the ones for (001), (011) oriented surfaces. In the early stages of the indentation, deformation is dominated by the Shockley partial dislocations (SPDs), nucleated under the indenter. Formation of stair-rod dislocations and their reaction with the current SPDs will result in the formation of new Shockley partial dislocations. The further reaction between SPDs and annihilation of them will release prismatic loops. The crystal's sensitivity to the presence of internal structural defects is also considered and it was found to be strongly dependent on its crystallographic orientation. A defect-based model can elucidate the stochastic pop-in loads in nanoindentation, as internal defects in materials are almost inevitable.

Effect of Heat Treatment on the Microstructural Evolution and Mechanism of Wear of 3D Printed Ti6Al4V Alloys

Niyousha Azgomi^{*1}, and Solomon Boakye-Yiadom^{1,2} Mechanical Engineering Department,
Lassonde School of Engineering, York University
Toronto, ON, Canada, M3J 1P3

ABSTRACT

Titanium and its alloys especially medical grade Ti6Al4V have a wide application in biomedical industry such as knee and hip implant due to its biocompatibility, high strength, and good corrosion resistance. Recently, advancements in metal 3D printing has resulted in the processing of personalized medical implants with complex geometries. However, it is challenging to produce parts from metal powders with known and predictable properties because processing parameters produce microstructural heterogeneities which influence the properties and behavior of the processed parts. After 3D printing, the parts are heat treated to improve the microstructural integrity/properties and relieve the processed parts of high residual stresses. The purpose of this study is to understand the effect of heat treatment on the microstructure, hardness and wear mechanism of 3D printed Ti6Al4V. During the heat treatment, two different cooling speed (rates) were used; water quenching and air cooling, respectively. Solutioning temperature was set at 1080°C above the β transus temperature (980°C). Specimens were held at this temperature for 4 hours then were either air cooled, or water quenched prior to aging. Three different temperatures 200°C, 500°C and 800°C were used to age both the air-cooled and water quenched specimens for 4 hours followed by air cooling. After heat treatment, all the specimens were mounted, ground, polished and etched by Kroll's reagent and studied using the optical microscope, scanning electron microscope and microhardness tests. To access the mechanical integrity of the 3D printed and heat treated Ti6Al4V alloys, reciprocal wear test based on the ASTM G133 was done using Bucker's Universal Mechanical Tester (UMT) Tribolab at room temperature to understand the types and mechanism of wear. It is demonstrated that the microstructural integrity of the Ti6Al4V alloys is a function of both the manufacturing process and the temperature/duration for heat treatment.

Acoustic topological insulators in a flow-free symmetry-broken resonator system

Xingxing Liu

Department of Mechanical and Materials Engineering,
The University of Western Ontario, London, Ontario N6A
5B9, Canada
xliu659@uwo.ca

Jun Yang

Department of Mechanical and Materials Engineering,
The University of Western Ontario, London, Ontario N6A
5B9, Canada
jyang@eng.uwo.ca

ABSTRACT

The development of topological insulators in condensed matter systems that break time-reversal symmetry using magnetic bias has inspired a quest for similar effects in classical acoustic waves. The acoustic topological insulator is a revolutionary design to control acoustics in isolation and broadband unidirectional transmission, which is topologically robust and immune to structural disorders or defects. Currently, these fascinating properties have been investigated through fan-induced moving media or acoustic capacitance adjustment. However, most of them are still associated with disadvantages including extra noise, bulky volume, and limited dynamic controlling performance. In this study, we propose an approach which could possibly overcome these limitations by introducing a modulation scheme of the acoustic metamaterial in a lattice of resonators, which demonstrates that the Floquet topological insulators with structure control can realize topologically robust and nonreciprocal acoustic propagation. This controlling strategy provides an alternative platform to conduct acoustic topological applications, especially for noiseless and miniaturized airborne acoustics. The use of structure modulation could potentially provide a platform for miniaturizing topologically insulating devices for airborne acoustics.

Kinetics of Detwinning Arising from Annealing in Chemical Vapor Deposited Nickel

Hao Sun

Mechanical and Industrial Engineering
University of Toronto
Toronto, Canada
haoge.sun@mail.utoronto.ca

Chandra Veer Singh

Mechanical and Industrial Engineering
University of Toronto
Toronto, Canada
chandraveer.singh@utoronto.ca

ABSTRACT

Nickel carbonyl chemical vapor deposition (CVD) is a high-efficiency process used to produce nickel shell molds with high yield strength and reasonable ductility. Such advantageous mechanical properties arise from its bimodal grain structure, which consists of a nanocrystalline matrix embedded with large columnar grains filled with nanotwins. The nanocrystalline matrix enhances the strength while the nanotwins contribute to the high ductility. Additionally, during deformation, dislocations interact with coherent twin boundaries (CTBs), disordering the boundary structure, thereby increasing the resistance for subsequent dislocation glide. Thus, the strain hardening capacity of CVD nickel is also improved. Accordingly, the fracture toughness of CVD nickel is superior to that of both polycrystalline and nanocrystalline nickel, allowing many design options not available with other mold-making technologies.

Despite the advantageous mechanical properties, the nanotwins do not persist at high temperatures, restricting the working temperature of CVD nickel and causing thermal degradation of its strength. In this work, the detwinning process of CVD nickel was studied by experimental examinations and molecular dynamics simulations. First, TEM examinations show that detwinning is realized by dislocation generation and incoherent twin boundary (ITB) migration. Then, the necessity of dislocations for the formation of detwinning nuclei is explained based on theoretical analysis: without dislocations, detwinning decreases the energy of atoms merely on and close to CTBs, while for the other atoms, only their crystal orientation changes. Such rotation requires a long-range driving force, which must be “perceived” by each atom inside a twin lamella but cannot be provided merely by CTBs. However, if there were plentiful dislocations inside a twin lamella, each atom would “sense” a driving force to eliminate these dislocations, making it possible to form a stable embryonic detwinning nucleus.

Next, molecular dynamics simulations were conducted, and the results showed that the dislocations generated from CTBs are intrinsic grain boundary dislocations (IGBDs). They detach from the CTB plane due to creep, driven by the internal tensile stress originating from grain growth in the nanocrystalline regime of CVD nickel. Hence, the dislocation density inside the twin lamella jumps. Subsequently, the deformation energy that is stored by these dislocations triggers the generation of Shockley partial dislocations at the intersection of CTBs and grain boundaries. Finally, these newly formed Shockley partial dislocations rearrange into an ITB, which removes the nanotwin connecting to it by its subsequent migration. Overall, unlike grain growth, mechanical stress is a necessary condition for detwinning. Due to its correlation with grain growth, such stress in CVD nickel could be eliminated by the removal of the nanocrystalline structures. Compared with the bimodal structure notwithstanding having a slightly reduced strength, a monolithic nanotwin structure of CVD nickel has enhanced thermal stability and retains high corrosion resistance. Both properties are more important considerations in the mold industry than the slightly diminished yield strength.

Work of fracture in non-crimp fabric carbon fiber reinforced epoxy composites

Aaditya Suratkar*

Mechanical and Materials Engineering
University of Western Ontario
London, Canada
asuratka@uwo.ca

Ying Fan

Mechanical and Materials Engineering
University of Western Ontario
London, Canada
yfan4@uwo.ca

John Montesano

Mechanical and Mechatronics Engineering
University of Waterloo
Waterloo, Canada
john.montesano@uwaterloo.ca

Jeffrey Wood

Mechanical and Materials Engineering
University of Western Ontario
London, Canada
jtwood@uwo.ca

The energy absorption mechanisms in advanced polymer matrix composites are investigated in this study. The materials under investigation are uni-directional heavy tow (50K) non-crimp fabric carbon fiber reinforced epoxy composites, manufactured using High Pressure-Resin Transfer Molding (HP-RTM) process. The study aims at correlating the experimental data acquired from quasi-static tensile tests and fracture surface morphologies to the work done to catastrophic final fracture in unidirectional (UD; $[0]_7$), quasi-isotropic ($[0/\pm 45/90]_s$) and multi-directional ($[\pm 45/0_2]_s$) composites. Scanning Electron Microscopy (SEM) was employed to identify the distinguishing features of the fractured surfaces. The fractured surfaces revealed various toughening mechanisms like fiber pull-out, plastic deformation in matrix, secondary cracks, etc. An appraisal of the energetic contributions of these mechanisms to fracture is performed in this study. Influence of the stacking sequences is also discussed.

On Fabrication and Durability of Superhydrophobic Electrodeposited Zn Coating on Stainless Steel: the Effect of Substrate's Surface Roughness

Mona Amiriafshar*, Xili Duan, Ali Nasiri

Faculty of Engineering and Applied Science, Memorial University of Newfoundland St. John's, NL, A1B 3X5

*Correspondence: mamiriafshar@mun.ca

ABSTRACT

With the aim to fabricate superhydrophobic 17-4 PH stainless steel surfaces for harsh environment applications, the impact of substrate's surface roughness on fabrication and durability of a superhydrophobic coating, comprised of a Zn electrodeposited layer capped with 0.05 M stearic acid, was investigated. Various micron and sub-micron scale finished surfaces with different surface roughness, namely as-received ($R_a \sim 11.93 \mu\text{m}$), sandblasted ($R_a \sim 4.62 \mu\text{m}$), and ground ($R_a \sim 0.39 \mu\text{m}$) surfaces were fabricated, followed by applying Zn electrodeposition coating to improve the water repellent properties of the stainless steel substrate. Additionally, electrodeposited Zn layer was then capped with 0.05 M stearic acid to further reduce the surface energy of the Zn coating and also to diminish the surface oxidation. The capped Zn electrodeposited coating showed an impressive degree of superhydrophobicity with the contact angle of larger than 150° .

Regardless of the initial surface roughness of the 17-4 PH stainless steel substrate, the obtained surface roughness after applying the superhydrophobic coating was found to be approximately the same ($R_a \sim 19.8 \pm 3.7 \mu\text{m}$) for all the samples. However, the compositional analysis of the coated surfaces confirmed that the higher surface roughness of the substrate is beneficial to obtain a uniformly deposited Zn coating, resulting in the improved uniformity of the coating in the order of as-received > sandblasted > ground surfaces.

To study the durability and electrochemical stability of the deposited superhydrophobic coating, electrochemical impedance spectroscopy (EIS) and contact angle measurements were performed up to 72 h of immersion time in an aerated 3.5 wt.% NaCl solution that mimics the seawater environment. The EIS results highlighted a significantly greater adhesion for the coating on the as-received substrate correlated to its higher surface roughness. Moreover, the contact angle measurement after 24 h (115°) and 72 h (107°) of immersion time indicated that the as-received sample retains some degree of hydrophobicity at longer immersion times due to the better adhesion of the coating. On the other hand, the coated sandblasted and ground surfaces indicated that the electrodeposited Zn layer due to its poor adhesion to the stainless steel substrate was locally peeled off shortly after the initial immersion, resulting in a significant decrease in the contact angle after 72 h of immersion time, *i.e.*, 77° and 66° for coated sandblasted and ground surfaces, respectively. Therefore, the substrate's surface roughness was found to be a dominating factor in controlling the adhesion and durability of the superhydrophobic electrodeposited Zn coating on the 17-4 PH stainless steel surface.

Determining the Liquid Window of Al-Si Coated 22MnB5 during Austenitization using Reflectivity Measurements

Cameron Klassen

Department of Mechanical and Mechatronics Engineering
 University of Waterloo
 Waterloo, Canada
 c2klasse@edu.uwaterloo.ca

Kyle J. Daun

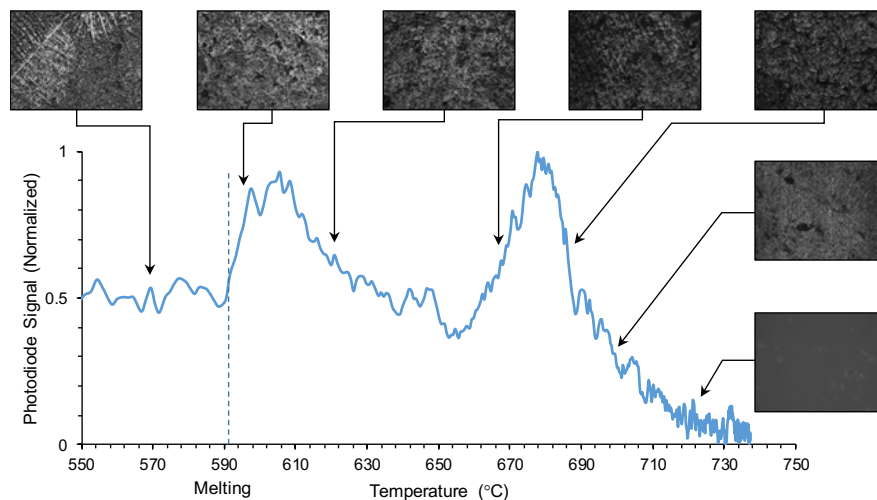
Department of Mechanical and Mechatronics Engineering
 University of Waterloo
 Waterloo, Canada
 kjdaun@uwaterloo.ca

ABSTRACT

Aluminum-Silicon coated 22MnB5 steel is a mainstay material for making high strength body-in-white automotive parts through hot stamping. The Al-Si coating prevents scale formation and decarburization during furnace-based austenitization ($\sim 950^\circ\text{C}$), and reacts with the iron in the substrate steel to form an intermetallic coating that provides long term corrosion resistance. However, the coating melts at $\sim 570^\circ\text{C}$, before it transforms into the solid intermetallic at higher temperatures. This liquid phase causes surface scoring and relocation of the blanks, which complicates automated transfer of the blanks into the die. Moreover, the molten coating impregnates the ceramic rollers, causing them to fail. The capital costs for roller replacement, coupled with the downtime required for corrective furnace maintenance, represent a considerable expense that reduces the profitability of this process. Accordingly, industry seeks to find the process settings that minimize the duration over which the coating is liquid. This, in turn, requires a technique for inferring the instantaneous steel surface state during furnace heating.

This study pioneers the use of laser reflectivity measurements to infer the surface state of Al-Si coated steel coupons (Usibor[®] 1500P AS150) in situ, to better understand the influence of heating rate on the melting of the Al-Si coating during austenitization. Light from a green laser (520 nm) was reflected off coupons as they are heated within a muffle furnace and measured to infer the instantaneous spectral reflectance. The reflected light from the coupons displayed a “speckle” pattern that evolved throughout the heating process, which indicates the surface roughness. Samples were retrieved from the furnace and quenched at different times to analyze their microstructure, surface roughness, and spectral reflectance through ex situ analysis.

Time-resolved spectral reflectance reveals two distinct peaks. The first peak corresponds to the increase in reflectance coinciding with the melting of the Al-Si coating, whereas the origin of the second peak remains unknown. Moreover, while the initiation of the first peak is insensitive to heating rate, as the setpoint of the furnace was increased, the later points of interest in the signal shifted towards higher temperatures. However, since the samples are being heated more rapidly at higher heating rates, the peaks occur quicker or with a shorter process window. This leads to the idea that the liquid window of the coating can be decreased with an increase in heating rate.



Reflectance of Al-Si coated 22MnB5 during heating, and optical micrographs made on samples extracted and quenched.

Assessing fibre metal laminate materials for their impact characteristics

Nick Tepylo, Bob De Snoo and Xiao Huang
 Department of Mechanical and Aerospace
 Engineering
 Carleton University, Ottawa, Canada

Calvin Rans
 Faculty of Aerospace Engineering
 Delft University of Technology, Delft,
 Netherlands

Abstract— Weight reduction of transportation vehicles can be achieved with the use of composite and fibre metal laminate (FML) materials. An important safety characteristic of these lightweight materials is their impact resistance compared to that of traditionally used automotive materials: steel and aluminum. In this study, low-velocity impact (LVI) was conducted on thin composite and FML panels to assess their applicability in impact prone applications. The results showed that the impact characteristics of monolithic aluminum 2024-T3 sheet performed better than carbon/epoxy, carbon/nylon, CARAL 5 2/1-0.3, and GLARE 5 2/1-0.3 lightweight panels. Due to the strain rate strengthening effects of glass fibres, GLARE was found to be the best alternative to the aluminum alloy.

Keywords- *Fibre metal laminates, low velocity impact, GLARE, impact tower, specific energy absorption*

Introduction

Fibre metal laminates are defined as a hybrid material consisting of alternating layers of monolithic metallic sheet and pre-impregnated (prepreg) fibre layers [1]. Traditional variants of FMLs were developed at the Delft University of Technology (TU Delft) in the 1980's and included aramid, carbon, or E-glass fibre layers with either aluminum 2024-T3 or 7075-T6 metallic layers [2]. At present, a FML known as GLARE (glass reinforced aluminum) has become an effective material option in aerospace applications driven by damage tolerance requirements [3] due to its improved fatigue and impact resistance over monolithic aluminum 2024-T3 and fiber reinforced polymer (FRP) panels respectively [4, 5]. Standard grades of GLARE consist of unidirectional S2 glass and FM94 epoxy layers mixed with layers of aluminum 2024-T3 as shown in Figure 1.

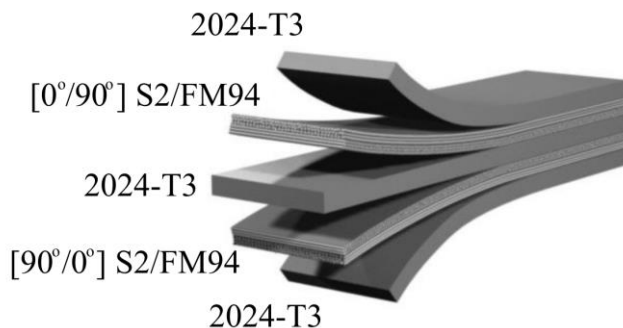


Figure 1: Lay-up of a standard cross-ply GLARE FML [6].

The advantage of layering FRPs with metallic sheet is the freedom to tailor the hybrid laminate for specific applications. Although primarily considered in design for its

excellent fatigue resistance, GLARE and other forms of FMLs are now being optimized for impact prone applications such as aircraft cargo floor panels [4]. Signs of a good impact material are large plastic deformation zones (metallic layers) and high failure strains (metallic and composite layers). For these reasons aluminum 2024-T3 is preferred over the stronger and more brittle aluminum 7075-T6 and glass fibres are preferred over the stiffer and lower strain-to-failure carbon fibres for impact prone structures [8, 9, 10]. Further impact resistance can be gained in FMLs by increasing the volume fraction of metallic sheet at the expense of also increasing the density [11].

In general, a composite material's response to an impact event is different to that of a metal. Metallic structures absorb low and intermediate impact energies through elastic and plastic deformation for which permanent deformation usually has a small effect on the load carrying capacity [12, 13]. Even at very high impact energies that produce penetration of the metallic structure, its load carrying capacity can be predicted from fracture mechanics principles allowing for safe design practices [14]. Fibre reinforced composite structures are limited in their ability to deform plastically and kinetic impact energy is dissipated through flexure, interlaminar shear deformations, fracture, and delamination of the laminate [7]. This ultimately causes large damage areas and results in reduced strength and stiffness [15, 16]. It is therefore more difficult to design for an impact event on a composite structure compared to a metallic one.

While the inclusion of metallic layers in FMLs improves the impact resistance of sandwiched composite layers, the progression of both metallic and composite failure mechanisms during an impact event limits the use of prediction models for designers [4]. To understand the impact performance of composite and FML materials in service, impact tests are often performed in the laboratories by various researchers [11, 17-20]. A general categorization of impact tests is based upon whether the impact is at low, intermediate, high/ballistic, or hyper velocity [19]. It is generally accepted that low velocity impacts are performed with large masses while high velocity impacts are performed with small masses; however, there is no clear transition between categories based on kinetic energies or velocities reached [21]. Instead of velocity limits, another designation has been suggested in which low velocity impacts (LVI) are defined as impacts where through thickness fracture or complete penetration of the laminate does not occur [22].

In this study, a series of low-velocity impact tests were conducted on several lightweight composite and FML materials to rank their suitability for impact resistant applications. Four different variations of lightweight materials were tested, each having approximately the same total thickness (1.2 mm). These consisted of carbon fibre reinforced thermoset and thermoplastic composites as well as carbon fibre and glass fibre FMLs, all of which were compared to monolithic aluminum 2024-T3.

I. MATERIALS AND EXPERIMENTAL PROCEDURES

A. Materials

The materials tested include thermoset (epoxy) and thermoplastic (nylon) reinforced carbon fibre in addition to two different variations of FMLs, all of which were manufactured at TU Delft. The mechanical properties of the laminate constituents are given in Table 1. Both glass and carbon fibre reinforced FML were manufactured in house. The glass FML is designated as GLARE 5-2/1-0.3 which means it has a $[0^\circ/90^\circ/90^\circ/0^\circ]$ S2-glass arrangement between 0.3 mm thick clad aluminum 2024-T3 layers where the 0° direction coincides with the rolling direction of the aluminum. The other FML is denoted as CARAL (carbon reinforced aluminum) 5-2/1-0.3 which is similar to the GLARE setup, except the glass fibre reinforced epoxy layers are replaced with carbon fibre reinforced epoxy. Panels of monolithic, non-clad aluminum 2024-T3 were also impact tested to gain a full spectrum of composite, FML, and metallic impact samples. As a baseline, aluminum sheet commonly used for autobody, was also tested in this study. All materials tested had a thickness of approximately 1.2 mm to establish comparability of impact results as suggested in ASTM D3763. A summary of the laminates is given in Table 2 along with their respective layups and physical properties. Thermoset based materials were cured in an autoclave while the thermoplastic panel was consolidated in a hot press.

Table 1: Mechanical properties of laminate constituents

Property	Aluminum 2024-T3 [29]	UD M30SC/DT120 prepreg [24]	UD S2 glass/FM94 prepreg [24]	*Carbon fibre reinforced nylon 6,6 [25]
E_{11} (GPa)	70.15	155	48.9	24.13
E_{22} (GPa)	70.15	7.8	5.5	N/A
G_{12} (GPa)	27.6	5.5	5.5	N/A
σ_{ult} (MPa)	480	1800	2640	193
ν_{12}	0.33	0.27	0.33	N/A
ϵ_{break} (%)	15-18	1.6	4.5	1.5-2.5
ρ (g/cm ³)	2.78	1.76	1.98	N/A
t (mm)	0.3	0.156	0.133	0.155

*Property data not available from manufacturer Tencate so reference [25] used.

Table 2: Configuration of materials tested

Material	Layup	Dimensions (mm \times mm)	Panel Thickness (mm)	Areal Weight (kg/m ²)	MVF*
Carbon/Epoxy	$[0^\circ / 90^\circ / 0^\circ / 90^\circ]_s$	110 \times 110	1.26	1.84	0
Carbon/Nylon 6,6	$[0^\circ / 90^\circ / 0^\circ / 90^\circ]_s$	110 \times 110	1.23	1.81	0
CARAL 5-2/1-0.3	$[0^\circ / 90^\circ / 90^\circ / 0^\circ]$	110 \times 110	1.23	2.6	0.49
GLARE 5-2/1-0.3	$[0^\circ / 90^\circ / 90^\circ / 0^\circ]$	110 \times 110	1.15	2.73	0.53
Aluminum 2024-T3	N/A	110 \times 110	1.22	3.42	1

*MVF- metal volume fraction

B. Low Velocity Impact Test

Low-velocity impact testing was completed on an Instron Dynatup 8200 instrumented drop weight impact tower complimented with Instron Dynatup Impulse impact analysis software at Carleton University. Impact energies can be set by either altering the height of the impactor or by placing individual 1 kg lead blocks on the impactor carriage. The velocity of the impactor was captured prior to impact by the passing of a steel velocity flag through a magnetic sensor which also triggers a pneumatic braking system on rebound to avoid multiple impacts on each sample. The contact force-time history between impactor and sample was measured by a 178 kN load cell. Velocity, displacement, and impact energy-time histories were then determined by the Impulse software through numerical integration of the force signal. ASTM standard D3763 was used as a guide for analyzing the impact data and for designing the clamping fixture to ensure multi-axial deformation. All samples were cut to square dimensions of 110×110 mm and clamped between parallel steel plates that had a circular opening of 80 mm diameter as shown in Figure 2. The reason for utilizing a relatively large impactor diameter in this study is

to induce greater membrane deformation in the samples to better simulate impacts in automotive applications. Eight socket cap bolts held the steel plates together and were tightened to 40 Nm with a torque wrench to provide a consistent clamping force. A hemispherical impactor head with a diameter of 30 mm and mass of 0.272 kg was used to impact the lightweight materials considered in this study at the energy levels given in Table 3.

Energy levels were chosen based on achieving a full spectrum of damage progression in each material leading up to full penetration. Each material was impacted at least once at the energy level given in Table 3 and in some cases multiple times to establish confidence in the reliability of test results. This is similar to the impact energy methods used by Parnanen *et al.* and Mugica *et al.* [26, 27]. Due to limitations in space for mounting the lead masses, producing impact energies in a 5-155 joule range required altering the impactor height and ultimately the impactor velocities. Thus, the potential for strain rate dependent mechanical responses in the materials may be present at impactor velocities exceeding 3 m/s [27].

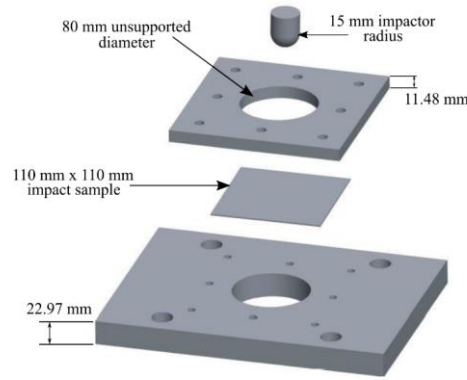


Figure 2: Clamping fixture and impactor for LVI test.

Table 3: Test matrix for LVI tests

	Carbon/Epoxy	Carbon/Nylon	CARAL 5 2/1-0.3	GLARE 5 2/1-0.3	Al 2024-T3
Impact Energies (J)	5	5	5	2	2
	15	15	12.55	5	5
	25	25	15	15	15
	35	35	25	25	25
	45	45	35	35	35
	55	55	45	55	55
			55	65	65
				75	75
				85	85
				95	95
				105	105
					125
					135
					145
					155

II. RESULTS AND DISCUSSION STYLING

One of the objectives of this research is to correlate low-velocity impact tests performed with the impact requirements for potential automotive applications. A biaxial bending/tension impact test method was chosen to rate the impact resistance of lightweight materials in this study since it creates stress states that are of high design priority [24]. The impact characterization of six different automotive applications will be used as a guide towards determining what failure limits with corresponding controlling variables are of interest for the impact testing method chosen. These applications consist of the fender, door outer panel, front fascia, hood outer panel, engine side panel, and the grill.

A. Failure Limits

According to Roche and Kakarala [24], failure limits are defined from force-deflection curves developed at the low-velocity impact energy near or at puncture as depicted in Figure 3. Total and break responses are well defined from the force-deflection curve, however, incipency and yield are less obvious. Incipency, defined as early property degradation, is difficult to differentiate from early vibrations as the impactor strikes the plate. Since incipency is of little importance for automotive application, it is not considered here. The yield failure limit is defined as the proportional limit and is shown to be indistinguishable from the break response for the majority of the materials tested. This can be attributed to the small thickness chosen for the specimens where membrane deformation and fast fracture are promoted versus delayed damage regions found in impact testing thick panels [18]. CARAL is the only material that has a definitive yield response where break exceeds the proportional limit. Roche and Kakarala [24] associate the yield failure limit with the formation of visible damage in the impact specimen; however, this does not coincide with the proportional limit in the force-deflection plot. Visible damage was present in each material at impact energy levels well below that to cause the proportional limit or break responses. The first occurrence of visible damage on the non-impact side is shown in Figure 4 for the carbon composite panels in the form of fibre breakage and matrix cracking. For comparison, the impact energy required to obtain the visible damage was 15 joules while the proportional limits were 35 joules and 24.6 joules for the carbon/nylon and carbon/epoxy, respectively.

In the FML and monolithic aluminum 2024-T3, damage progresses from a small plasticity zone at the site of impact, to a plastic deformation of the entire 80 mm diameter unsupported region. The final damage state before a full puncture takes place is when cracking of the aluminum layer on the non-impact side for FML samples or complete through thickness cracking in the monolithic aluminum 2024-T3 occurs. Cracking takes place near the break point in both FML and aluminum 2024-T3 samples. Examples of

plastic deformation occurring at lower impact energies than the proportional limit are given in Figure 5.

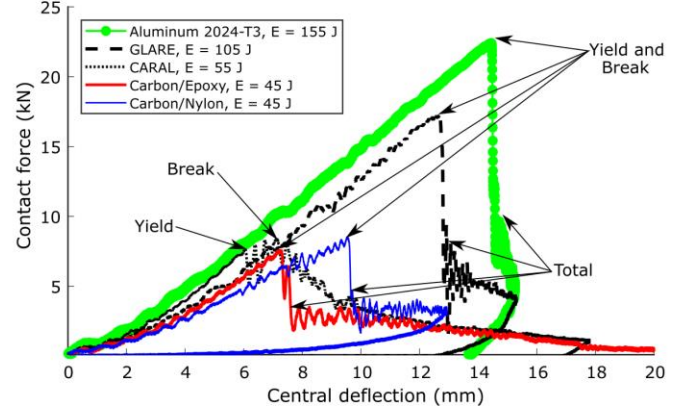
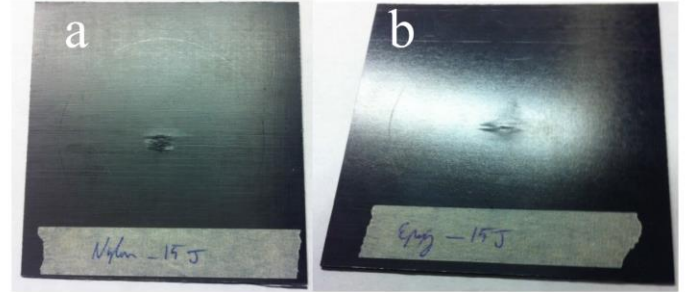
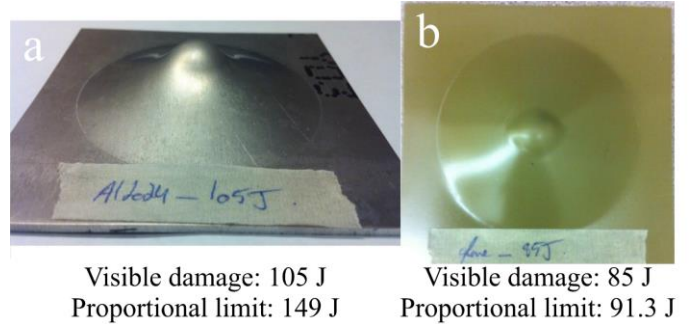


Figure 3: Force-deflection plots from LVI tests highlighting failure limits.



Proportional limit: 35 J Proportional limit: 24.6 J

Figure 4: Visible damage states on the non-impact side that occur before the proportional limit is reached for a) carbon/nylon and b) carbon/epoxy.



Visible damage: 105 J Visible damage: 85 J
Proportional limit: 149 J Proportional limit: 91.3 J

Figure 5: Plastic deformation at impact site and in the unsupported region of a) aluminum 2024-T3 panel and b) GLARE panel.

It is realized that the start of visible damage cannot be determined from a force-deflection curve and this failure limit must be assessed based on the inspection of the impacted samples at each energy level. This failure limit is in fact a visible damage instead of yield. A visible damage indicates that structural integrity is compromised.

B. Deflection

The maximum panel deflection (under the load) and permanent panel deflection (when load is released), are presented in Figure 6, at lower impact energy levels for the materials tested. The dent size (permanent deflection) is shown using the solid color while the maximum recorded deflection is shown as the total bar height. A dent is visible to the human eye when it exceeds 1 mm. At an impact energy of 5 joules, visible damage can be seen on the impact side of the GLARE, CARAL, and aluminum 2024-T3 specimens. GLARE and aluminum 2024-T3 are shown to have similar deflections and permanent dent depths at every energy level. Carbon composite panels are better at resisting denting as they can reach an impact energy of 25 joules while maintaining a barely visible damage. Carbon/nylon has the best dent resistance at low impact energies but also the highest deflections at the maximum load. Of the materials tested, carbon/nylon is the most efficient at storing and transferring the energy from and back to the impactor without incurring visible damage on the impact side. Similarly, at low impact energies carbon/epoxy is also resistant to visible damage on the impact side, however, as the impact energy approaches 35 joules, a greater permanent deflection was observed. In the aluminum 2024-T3 and FML panels, the impact energy is dissipated as plastic deformation at the site of impact.

Although carbon/epoxy and carbon/nylon have barely visible dent depths at an impact energy of 25 joules, from Figure 7 it is shown that at 15 joules, visible damage (cracks) exists in both materials on the non-impact side based on a visual observation. In fact, first signs of visible damage on the non-impact side were found to be more severe than that on the impacted side. This is due to the high tensile bending stresses present on the non-impact side where the largest out-of-plane deformations are present. For carbon/epoxy and carbon/nylon panels, the first visible

damage is represented by matrix cracking and fibre breakage in the layer farthest from the impact point. For FMLs and aluminum 2024-T3 the first sign of visible damage on the non-impacted side is in the form of plastic deformation followed by the development of a crack. Plasticity is difficult to quantify in this test; thus, the first signs of cracking in the aluminum will be used for comparison purposes. Measured impact parameters at the first sign of visual damage on the non-impacted side are given in Figure 7.

When examining visible damage on the non-impact side, aluminum 2024-T3 and GLARE perform better than the other materials. Considering the material properties in Table 1, it is clear that strain to failure plays an important role in delaying visible damage on the highly tension strained non-impact side. Aluminum 2024-T3, with the highest strain to failure of 18%, requires the largest deflection to crack. In GLARE, although it shows delayed crack occurrence, comparing to other three materials, the sandwiched layers of S2 glass with an inferior strain to failure of 4.5% cause cracking in the outer aluminum layer at an impact energy 65 joules. Panels containing carbon fibres, with the lowest failure strain of approximately 1.6%, are the first to show cracking/fibre breakage. The addition of thin aluminum 2024-T3 layers did not help to increase the visible damage limit of carbon/epoxy in CARAL.

The MVF also had an effect on the impact resistance of the laminates. As the MVF increased, so did the impact energy at the first sign of visible damage. The carbon/nylon and carbon/epoxy laminates contain no metal and thus have the lowest visible damage energy limits. The tested CARAL and GLARE laminates have a MVF of approximately 50% and have a higher visible damage energy limit than the Nylon and epoxy reinforced carbon fibre laminates. Typically, the visible damage failure limit peaks around a MVF of 60% [36].

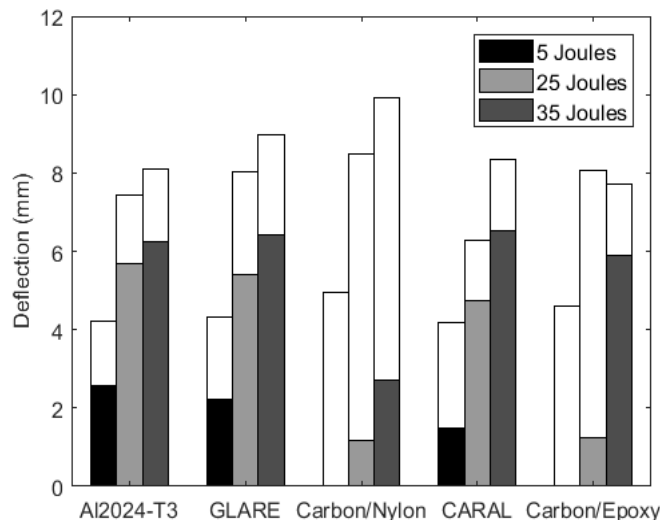


Figure 6: Maximum and permanent deflections under low energy impacts.

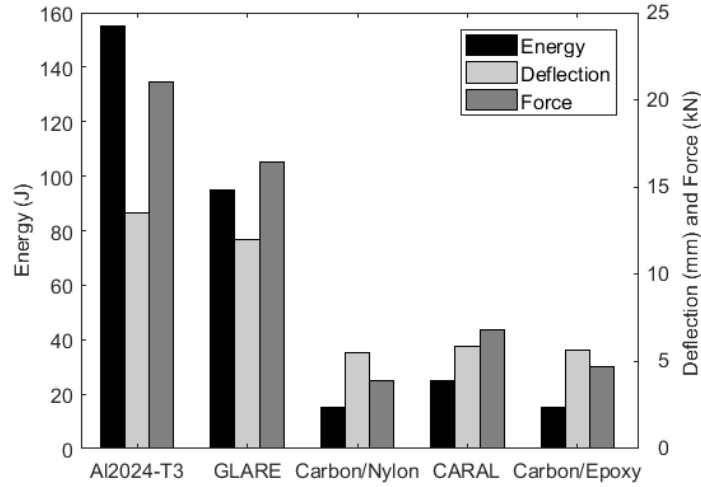


Figure 7: Impact parameters at the first sign of visible damage on the non-impact side.

C. Breaking Failure Limit

The break failure limit represents the start of severe damage during impact, at which point the material is no longer able to sustain impact loads. There is great interest in this failure limit for automotive applications that are required to absorb significant deflection at high impact energies before losing their load carrying capacity. Energy, deflection, and load at the break point in each material are compared in Figure 8. It is noticed that the measured impact parameters at the break point are similar in magnitude to those gathered at the first sign of visible damage on the non-impact side in Figure 7, for each panel type. This is expected as visible damage on the non-impact side represented fibre breakage and/or metal cracking after which greater loads cannot be sustained at much higher impact energies. It is realized that visible damage on the non-impact side signifies the start of final failure of the panels.

Aluminum 2024-T3 is observed to have superior impact resistance in terms of breaking failure limit over all other materials tested, followed by GLARE. Higher break failure limit is achieved by carbon/nylon over the CARAL fibre metal laminate. It is realized that the outer aluminum panels in CARAL have offered little improvement over a carbon/epoxy panel of the same thickness at the break point. The difference in maximum deflection between aluminum 2024-T3 and GLARE is less than 2 mm at the break point.

D. Total Failure Limit

Lightweight materials with high energy absorption capabilities are critical for impact damage prone structures. The energy absorption limits produced from clamped boundary conditions and biaxial membrane deformation caused by impact in this research can be used as a preliminary evaluation step in choosing materials for crash structures. The energy absorption and puncture characteristics of the lightweight materials in this study are

presented in an energy profile diagram in Figure 9. Here the complete puncture is defined for each material as energy absorption that equals the impact energy and when the impactor does not rebound [8]. Aluminum 2024-T3 requires the greatest energy to puncture, at a value of 155 joules, 50 joules more than the next best material, GLARE. Carbon/epoxy represents the lowest end of the energy absorption spectrum with a value of 45 joules, while CARAL and carbon/nylon both puncture at 55 joules.

Carbon/epoxy has superior stiffness to monolithic aluminum 2024-T3 (as shown in Figure 10), yet it punctures at a lower impact energy (Figure 9). This is a result of differences in strain energy densities or area under the stress-strain curves. Monolithic aluminum 2024-T3 has large plastic deformation region and high failure strain; it provides the best strain energy density among all materials tested. This correlates well with its superior energy absorption and puncture limits in the LVI impact tests. GLARE and CARAL's inner composite cores are shown to gain strain energy density with the addition of aluminum layers as they develop a plastic deformation region and higher failure strains.

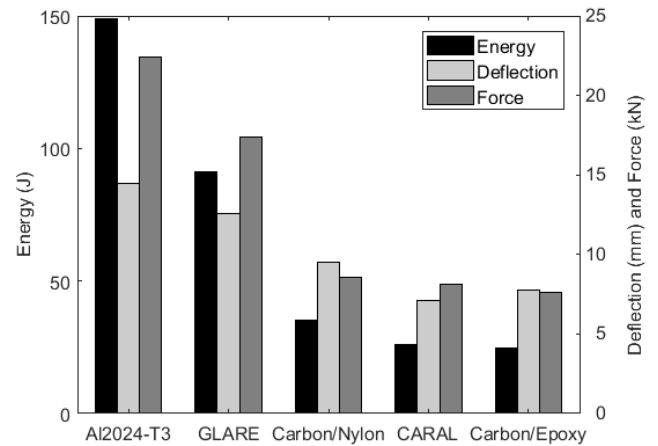


Figure 8: Impact parameters at the break point.

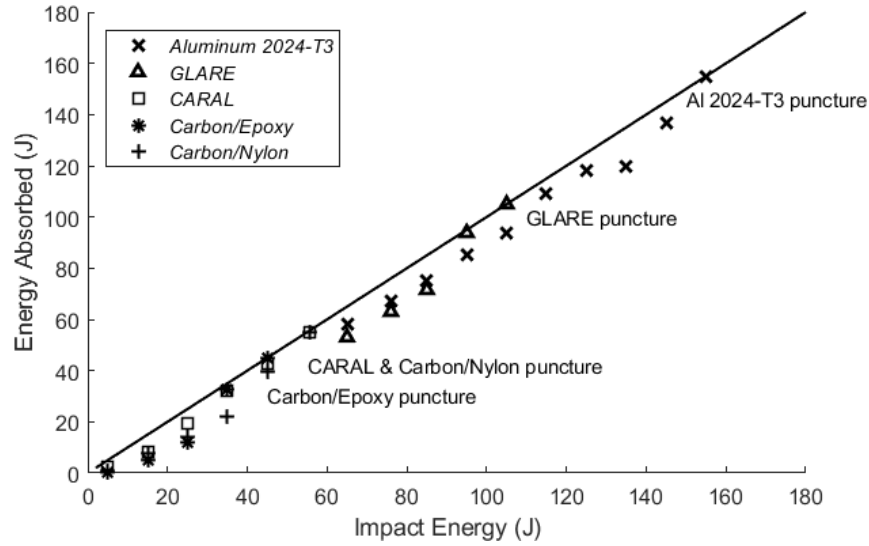


Figure 9: Energy profile diagram emphasizing impact energies to cause puncture.

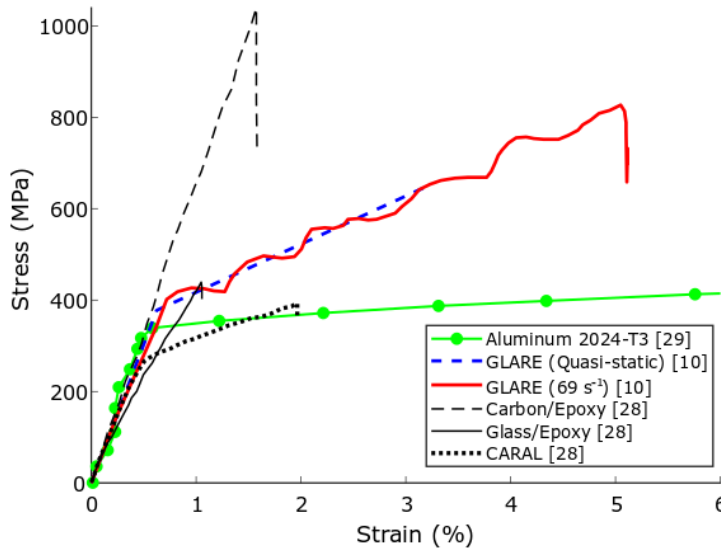


Figure 10: Stress-strain curves of the various laminates and their constituents [10, 28, 29]. The curve for aluminum has been reduced to 6% strain.

GLARE's energy absorption and puncture limit dominance over the other laminate is related to the strain rate dependent mechanical properties of glass-fibres that can increase the strain energy density at higher strain rates (Figure 10). Carbon/nylon is estimated to have a similar stress-strain curve to that of CARAL since they puncture at the same impact energy. Improved toughness with a thermoplastic matrix is beneficial for increasing strain energy density, similar to that achieved by adding the aluminum layers to the thermoset matrix in a CARAL panel.

An approximate panel thickness of 1.2 mm was selected in this study to simulate automotive body panels. It has been

shown that under constant thickness, aluminum 2024-T3 provides the highest energy absorption before puncture; however, it is also the heaviest panel among the five evaluated. Specific energy absorptions of each material are thus contrasted in Figure 11 where it is evident that aluminum 2024-T3 is the most efficient energy absorbing material based on energy absorption per mass. Energy absorption has been normalized based on density since it accounts for the differences in mass and slight variations in panel thicknesses.

To match the energy absorption of aluminum, the thickness of all other lightweight materials would have to be

increased. It is interesting to realize that the specific energy absorption of FMLs has been found to increase with thickness [23]. Also, the layered configuration can be varied. For instance, Moriniere *et al.* [30] conducted LVI impact test with 1.3 mm thick aluminum 2024-T3 and GLARE 5 2/1-0.4 panels and concluded that GLARE had a specific energy absorption 72% greater than the aluminum panels. Reasons for discrepancies among Moriniere [30] and the results obtained in this study could be due to the increased thickness of aluminum layers used in GLARE. The differences in boundary conditions and shape of impactor head employed may also play a role.

This work used circular unsupported areas whereas Moriniere opted for rectangular unsupported areas. The clamping fixture used in this study has an unsupported region smaller than what may be found in some automotive applications such as an outer door panel or vehicle fender, potentially providing a conservative estimate for the energy absorption. With a larger amount of material supported at the boundary, the samples tested by Moriniere *et al.* [30] can more effectively distribute the load. Studies using different sized laminates have shown that the maximum impact energy increases with larger laminates [31]. Increasing the in-plane dimensions of the laminate increases the maximum impact energy of the sample, whereas increasing laminate thickness increases both the maximum impact energy and maximum contact force [32]. The boundary conditions also play a role as square fixtures introduce stress concentrations in the corners of the specimen while circular samples are free from these fixture-induced effects [2].

The hemispherical impactor head was selected instead of a conical head to distribute the load at the point of impact and avoid causing point damage due to the sharp tip. Studies have shown that altering the shape of the impactor head causes a change in the initiation and propagation characteristics in the sample [33, 34]. In addition, even if the impact energy is consistent between tests, differences in impactor mass may alter results. Impactors with a larger mass promote energy absorption in the panel by introducing larger bending strains. On the contrary, an impactor with a lower mass results in a decreased penetration threshold energy, as the impact force is subjected to a smaller area [35]. Therefore, when considering materials for impact resistance the dependence of impact response on clamping boundary conditions, impactor geometry, and other impact test variables must be accurately assessed. Future studies should strive to simulate the boundary conditions expected for the panel when it is in service.

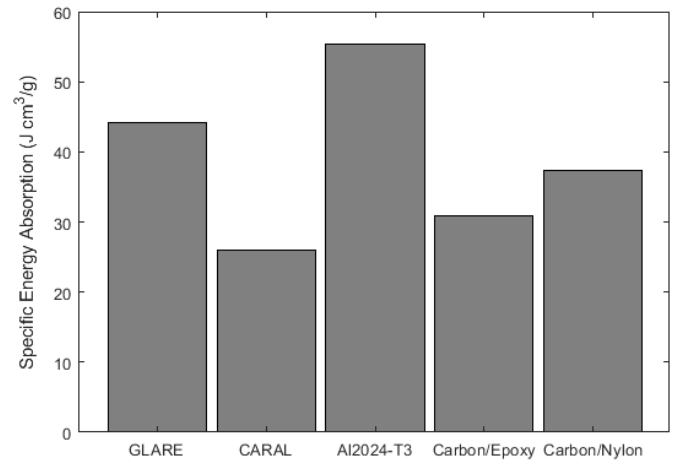


Figure 11: Specific energy absorption calculated at puncture.

III. CONCLUSION

This research assessed the low-velocity impact performance of carbon/epoxy, carbon/nylon, CARAL and GLARE lightweight materials against a baseline aluminum 2024-T3 alloy. The following conclusions are made based on the results obtained. Aluminum 2024-T3 was found to perform better in impact than the other four lightweight materials considered in this study. More specifically, aluminum 2024-T3 was superior in its ability to sustain load before puncturing and had the ability to absorb more energy per unit mass. Strain rate strengthening mechanisms of the glass fibres in the GLARE FML set it apart from the other lightweight materials impacted and gave it deflection capabilities similar to aluminum 2024-T3. It was also determined that the impact characteristics of each material matched their respective stress-strain responses especially when comparing strain energy density to energy absorption achieved. Differences between previous work may be due to the sample size, boundary conditions and impactor head used. When applied to transport vehicles, the use of FML materials in applications including door and hood panels is advantageous when compared with aluminium since larger FML panels are able to better distribute impact loads. For smaller vehicle sections such as fenders and grills, the panel size and mounting must be carefully considered before committing to using FML materials instead of conventional materials.

ACKNOWLEDGMENT

Funding for this work was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

1. C. Rans, "The role of rivet installation on the fatigue performance of riveted lap joints," Doctoral Thesis, Mechanical and Aerospace Engineering, Carleton University, Ottawa, Canada, 2007.
2. J. Laliberte, "Investigation of low-velocity impact damage in fibre metal laminates," Doctoral Thesis, Mechanical and Aerospace Engineering, Carleton University, Ottawa, 2002.
3. C. Pitzer and J. M. Yang, "Hybrid metal laminate manufacturing planning- evaluation and assessment," Department of Materials Science and Engineering, University of California, Los Angeles, CA, Tech. Rep., 2008.
4. F. Moriniere, "Low-velocity impact on fibre-metal laminates," Doctoral Thesis, Faculty Mechanical, Maritime and Materials, Delft University of Technology, Delft, NL, 2014.
5. A. Vlot and J. W. Gunnink, *Fibre Metal Laminates: An Introduction*, Norwell, MA, Kluwer Academic Publishers, 2001.
6. R. C. Alderliesten, "Fatigue crack propagation and delamination growth in Glare," Doctoral Thesis, Faculty Mechanical, Maritime and Materials, Delft University of Technology, Delft, NL, 2005.
7. C. C. Chamis, M. P. Hanson, and T. T. Serafini, "Impact resistance of unidirectional fiber composites," in *Composite Materials: Testing and Design Conference*, Anaheim, CA, April 1971.
8. Y. Liu and B. Liaw, "Effects of constituents and lay-up configuration on drop-weight tests of fiber-metal laminates," *Appl Comp Mater*, 17, 43-62, 2010.
9. A. Vlot, "Impact tests on aluminium 2024-T3, Aramid and glass reinforced laminates and thermoplastic composites," Delft University of Technology, Report LR-534.
10. A. Vlot, "Impact loading on fiber-metal laminates," *Int J Impact Eng*, 18, 291-307, 1996.
11. M. Sadighi, R. C. Alderliesten and R. Benedictus, "Impact resistance of fiber-metal laminates: A review," *Int J Impact Eng*, 49, 77-90, 2012.
12. P. J. Shadbolt, R. S. J. Corran and C. Ruiz, "A comparison of plate perforation models in the sub-ordnance impact velocity range," *Int J Impact Eng*, 1, 23-49, 1983.
13. F. J. Bradshaw, "Impact Resistance of Carbon Reinforced Plastics," MOD, Rep. RAE TR 72240, 1972.
14. J. G. Avery, "Design Manual for Impact Damage Tolerant Aircraft Structures," NATO, Rep. AGARD-AG-238, 1981.
15. P. W. Manders and W. C. Harris, "A parametric study of composite performance in compression-after-impact testing," *SAMPE J*, 22, 47-51, 1986.
16. A. Rotem, "Residual flexural strength of FRP composite specimens subjected to transverse impact loading," *SAMPE J*, 24, 19-25, 1988.
17. S. Agrawal, K. K. Singh and P. K. Sarkar, "Impact damage on fibre-reinforced polymer matrix composite - A review," *J Comp Mater*, 48, 317-332, 2013.
18. W. J. Cantwell and J. Morton, "The Impact Resistance of Composite Materials-A Review," *Comp*, 22, 347-362, 1991.
19. S. Abrate, "Impact on Composite Structures," New York, NY, Cambridge University Press, 1998.
20. G. B. Chai and P. Manikandan, "Low velocity impact response of fibre-metal laminates – A review," *Comp Struct*, 107, 363-381, 2014.
21. M. O. W. Richardson and M. J. Wisheart, "Review of low-velocity impact properties of composite materials," *Comp Part A*, 96, 1123-1131, 1996.
22. B. D. Agarwal, L. J. Broutman and K. Chandrashekhara, "Analysis and Performance of Fiber Composites," 3rd ed. New Jersey, NY, Wiley and Sons, 2006.
23. J. Zhou, Z. W. Guan, and W. J. Cantwell, "The influence of strain-rate on the perforation resistance of fiber metal laminates," *Comp Struct*, 125, 247-255, 2015.
24. J.L. Roche and S. N. Kakarala, "Methodology for Selecting Impact Tests of Composite Materials in Automotive Applications," in *Instrumented Impact Testing of Plastics and Composite Materials Symposium ASTM STP 936*, Houston, TX, Mar. 1985.
25. RTP Imagineering Plastics. (2005). *TP 287 H Impact-Modified Nylon 6/6 (PA) Carbon Fiber* [Online]. Available: <http://web.rtpcompany.com/info/data/0200H/RTP287H.htm>.
26. T. Pärnänen, R. Alderliesten, C. Rans, T. Brander, and O. Saarela, "Applicability of AZ31B-H24 magnesium in Fibre Metal Laminates – An experimental impact research," *Comp Part A*, 43, 1578-1586, 2012.
27. J.I. Múgica, L. Aretxabaleta, I. Ulacia, and J. Aurekoetxea, "Impact characterization of thermoformable fibre metal laminates of 2024-T3 aluminium and AZ31B-H24 magnesium based on self-reinforced polypropylene," *Comp Part A*, 61, 67-75, 2014.
28. E. C. Botelho, R. A. Silva, L. C. Pardini, and M. C. Rezende, "A review on the development and properties of continuous fiber/epoxy/aluminum hybrid composites for aircraft structures," *Mater Res*, 9, 247-256, 2006.
29. Military Handbook: Metallic Materials and Elements for Aerospace Vehicle Structures, Department of Defense and Federal Aviation Administration, 1998.

30. F. Moriniere, "Energy distribution in GLARE and 2024-T3 aluminum during low-velocity impact," in 28th International Congress of the Aeronautical Sciences Energy, Brisbane, AU, 2012.
31. D. Liu, B. B. Raju, and X. Dang, "Size effects on impact response of composite laminates," *Int J Impact Eng*, 21(10), 837-854, 1998.
32. W. J. Cantwell and J. Morton, "Geometrical effects in the low velocity impact response of CFRP," *Compos. Struct.*, vol. 12, pp. 39-59, 1989.
33. P. Robinson and G. A. O. Davies, "Impactor mass and specimen geometry effects in low velocity impact of laminated composites," *Int J Impact Eng*, 12(2), 189-207, 1992.
34. F. Elaldi, B. Baykan, and C. Akto, "Experimental Analysis for the Effect of Impactor Geometry on Carbon Reinforced Composite Materials," *Polymers & Polymer Composites*, 25(9), 2017.
35. S. Abrate, *Impact Engineering of Composite Structures*, Springer-Verlag Wien, New York, 2011.
36. J. Zhu, "Improving the impact resistance of CFRP laminates by embedding metal constituents," Master's Thesis, Faculty of Aerospace Engineering, Delft University of Technology, Delft, NL, 2018.

Experimental characterization of anisotropic mechanical properties of extruded AA6061-T6

M. Bondy, W. Altenhof*, J. Magliaro**

Mechanical, Automotive, and Materials Engineering
University of Windsor
Windsor, ON, Canada
*altenh1@uwindsor.ca
**presenting author

M. R. Jensen
Certasim LLC
Saratoga, CA, USA
morten@certasim.com

ABSTRACT— An aluminum extrusion of a practical geometry for automotive crashworthiness applications was acquired for both mechanical testing of the full extrusion and material characterization specimen extraction. The purpose of this endeavor was to collect data on the anisotropic mechanical properties resulting from the extrusion process. In terms of uniaxial tension data, there was a 3% difference in yield stress between 0° and 90° with 2% coefficient of variation. However, plastic strain ratios were distinct in the 0°, 45°, and 90° directions (average plastic strain ratios: 0.51, 0.23, and 1.43 for 0°, 45°, and 90°, respectively). VDA plate bending tests were distinct in terms of force-deflection responses after the onset of failure (peak load) for all directions considered (0°, 15°, 30°, 45°, and 90°). Force-deflection responses after the onset of significant plastic deformation were distinct for 0°, 15°, and 30° specimens but essentially identical for 45° and 90°. Repeatable force-deflection and failure mechanisms were observed for three point bending of an aluminum extrusion of a practical geometry for automotive crashworthiness applications. Three dimensional deflection data was acquired by digital image correlation to allow rigorous finite element model validation.

Keywords: *aluminum; extrusion; anisotropy; failure; fracture*

I. INTRODUCTION

Between 1985 and 1995 the mass of aluminum in an average automobile in Western Europe, Japan, and North America increased approximately 30% [1] with the expectation that this trend would accelerate over the next 10 years. In the same period of time it has been observed that vehicle mass has increased approximately 17% [2]. A general trend has been observed in the automotive industry where lightweighting materials become viable and are implemented when costs drop below approximately \$4.50 (USD) per kilogram mass reduction (\$2.00 USD per pound) [3]. Aluminum has found applications in closures (i.e. hoods, trunk lids) at a cost of \$4.00 per kilogram and, more recently, in vehicle chassis and body components (e.g. the aluminum bodied 2015 Ford F-150, aluminum chassis of the 2016 Jaguar XE, and the aluminum chassis and body of the Tesla Model S).

As novel materials are employed to reduce vehicle mass, structural requirements for crashworthiness, durability, and noise/vibration/harshness must be maintained or improved upon. Motor vehicle collisions in Canada resulted in over 2000 fatalities, 160,000 ER visits, and 170,000 injuries in 2010 at a cost of \$2.2 Billion [21]. For less ductile alloys with sufficient strength for structural applications, accurate predictions of material failure may be critical for numerical modelling in the area of vehicle crashworthiness. Furthermore, it has been shown that for many materials, the accuracy of a finite element model may be strongly influenced by the manufacturing process in the form of significant anisotropy in mechanical characteristics. Therefore, this study was developed with the following objectives:

- To fundamentally mechanically characterize a commercially common aluminum extrusion at large deformations applying standardized methodologies.
- Complete fundamental material characterization on standardized specimens extracted from this extrusion at different orientations with respect to the extrusion direction.
- Employ digital image correlation (DIC) to acquire high resolution displacement fields from extrusions exposed to three point bending for purposes including, but not limited to, comprehensive finite element model validation.

II. LITERATURE REVIEW

A number of publications exist documenting experimental characterization and modelling of the failure and anisotropy of aluminum alloys. Luo et al. [4] acquired surface strain fields and load-deflection data for notched tensile specimens, tensile specimens with a central hole, and butterfly shear specimens for 6260-T6 sheets (2 mm thickness). Novel yield [5] and fracture models were proposed and implemented in ABAQUS with the ability to accurately capture the onset of fracture for all specimen configurations. Wadley et al. [6] developed a finite element model of an extruded 6061-T6 sandwich panel subjected to blast loading. Tensile tests were completed for three orientations (0°, 45°, & 90° to extrusion direction).

Significant variation between tests with the same specimen orientation was observed. The Johnson and Cook [7] isotropic constitutive equation was used for the relationship between effective plastic strain and effective stress. Failure was captured with the isotropic Cockroft-Latham failure model [8].

Beese et al. were the first to publish a detailed investigation of the anisotropy of AA6061-T6 sheets [9]. A number of specimen configurations, including but not limited to, tensile tests and butterfly shear specimens, were employed to obtain failure strain as a function of stress triaxiality in the range of -0.2 to 0.7. Fracture strains at the specimen surface were measured with DIC software. Average fracture strains were determined from specimen thickness reduction. Local fracture strains were estimated through an inverse method with finite element models.

The influence of stress triaxiality on failure for AA 6061-T6 has been investigated at the length scale of the grain structure in two publications by Ghahremaninezhad and Ravi-Chandar [10, 11]. Uniaxial tension, notched tension, and Arcan specimen pure shear/superposed tension/compression experiments were completed with commercial and custom developed DIC software. The latter software acquired strain from grain deformation. For stress triaxialities in the ranges of -0.1 to 0 and 0.4 to 1, lower bounds on strain to failure were identified. The observed strain to failure is significantly larger, almost an order of magnitude, with respect to the strain to failure predicted by the Johnson-Cook damage model parameters identified by Lesuer et al. [12] for cross rolled AA6061-T6 plate.

The mechanisms for damage associated with ductile fracture are void growth, nucleation, and coalescence [13, 14]. However, in a precipitate hardened alloy like 6061-T6, damage evolution is influenced by particle/inclusion size in addition to grain size. The earliest damage models were porous plasticity models accounting for the effect of void growth [15] and extended to include nucleation and coalescence [16, 17]. Phenomenological damage models are commonly found in commercial finite element models and range from continuum damage mechanics (CDM) models which consider statistical thermodynamics to models employing a damage indicator/parameter, D , which is a function of stress and/or strain states. Two such common models of the latter category are the Johnson-Cook [7] and Cockroft-Latham [8] models.

The field of characterization of failure of anisotropic metals is relatively novel with minimal standardization. The German Automotive Industry Association (Verband der Automobilindustrie or VDA) has developed a standard test for assessing formability: *VDA 238-100: Plate bending test for metallic materials* [18]. This standardized test cannot replace the specimen configurations referenced previously while maintaining equal fidelity in terms of strain to failure as a function of stress triaxiality. However, in many industrial applications material characterization may have practical limitations in terms of time and financial resources that do not permit the dependence of failure on both stress triaxiality and Lode parameter to be explicitly identified experimentally. Larour et al. [19] completed a sensitivity analysis of the VDA 238-100 test noting a number of factors which influence the

bending angle (e.g. elastic deformation of the apparatus, specimen curvature) and recommending this angle be measured optically since the equation in [18], to analytically compute the angle, neglects the punch thickness which can result in up to 10% error.

III. METHODOLOGY

A. Uniaxial tensile tests

Uniaxial tensile tests were completed for 9 specimens in each of the three orientations with respect to the extrusion direction (0° , 45° , and 90° where 0° denotes a specimen where the load is applied in the extrusion direction) for a total of 27 specimens. All specimens were extracted by wire electrical discharge machining (EDM) from one 6.1 m (20 ft) length of 101.6 mm (4 inch) by 101.6 mm (3.175 mm wall thickness) AA6061-T6 square tube extrusion with each set of specimens (3 @ 0° , 3 @ 45° , and 3 @ 90°) extracted from a different region of the extrusion length (each end and midspan). ASTM standard B557 (*Standard Test Methods for Tension Testing Wrought and Cast Aluminum and Magnesium Alloy Products*) was followed for all uniaxial tensile tests. A subsize specimen geometry was selected given the 101.6 mm width of the extrusion. Tests were completed at a constant crosshead speed of 5 mm/min on a 50 kN MTS Criterion electromechanical load frame. Lankford coefficients (as functions of effective plastic strain) and Poisson's ratio were acquired from post-processing the displacement field acquired with Correlated Solutions VIC-2D DIC software and a 1.3 MP Allied Vision Manta monochrome camera with a Sill Optics S5LPJ5160 large working distance, high magnification lens at a frame rate of 17 fps. Poisson's ratio was calculated consistent with the procedure in ASTM standard D638. Lankford coefficients were computed consistent with ASTM E517.

B. VDA 238-100 Plate Bending Test for Metallic Materials

The number/orientation of specimens for the VDA238-100 plate bending test was consistent with the uniaxial tensile test methodology. Standard specimens (60 mm by 60 mm) were extracted by wire EDM. A fixture consistent with the requirements of the VDA 238-100 standard was constructed (Figure 1). A custom procedure for the MTS load frame was developed consistent with the VDA standard such that the crosshead speed was 10 mm/min until the 100 N preload specified in the standard was observed. When this preload was achieved the load frame automatically increased the crosshead speed to 20 mm/min. Load-displacement data for both phases were acquired but for post-processing the punch displacement was set to zero at the 100 N preload. MTS Advantage video extensometer software, with identical camera hardware to that employed for the tensile tests, was used to acquire 12 points on the edge of the specimen to track the bending angle. For each set of 6 points, separated by the axis of the punch, a line was fit using least squares regression. These two lines were used to compute the bending angle as a function of punch displacement.



Figure 1. VDA 238 plate bending test.

C. Three Point Bending of an Aluminum Extrusion

Three point bending of 609.6 mm (24 inch) lengths of the 101.6 mm (4 inch) by 101.6 mm by 3.175 mm (1/8 inch) wall thickness extrusion (identical to that from which tensile and plate bending specimens were extracted) was completed on a 150 kN MTS Criterion electromechanical load frame with an MTS 642.25 three point bending fixture. The roller diameter of the three point bending apparatus was 50.8 mm (2 inch) with a spacing of 457.2 mm (18 inch). The extrusion had sharp corners; the radius was approximately 0.4 mm (measured with minimal precision using available radius gauges). The loading rate was 20 mm/min with a maximum crosshead displacement of 80 mm. The 3D displacement field on the surface of the specimen was acquired with two 5 MP (2096 x 2048 pixels²) Point Grey Research Grasshopper cameras, Schneider-Kreuznach 30 mm lenses, and Correlated Solutions VIC-3D software. Images were captured with VIC-Snap from Correlated Solutions at a frame rate of 1 fps. Synchronization of the load-deflection data and displacement field from DIC was accomplished by generating a TTL signal at one of the user defined outputs of the MTS load frame to initiate and terminate image acquisition by the DIC system.

D. Axial Compression of an Aluminum Extrusion

Axial compression of 300 mm lengths of aluminum extrusion was completed on a Tinius Olsen load frame. Load was measured with two 220 kN load cells in parallel. Displacement was measured with an Acuity 300 mm laser displacement transducer. Both transducers were connected to a National Instruments CompactDAQ chassis with analog input modules and custom software developed in LabView. The data acquisition rate was 3000 samples per second and the average crosshead displacement speed was approximately 16 mm/min.

IV. RESULTS AND DISCUSSION

A. Uniaxial tensile tests

Engineering stress-strain responses for 0° tensile specimen orientation are provided in Figure 2. Mean elastic modulus, Poisson's ratio, yield stress, tensile strength, and strain at failure are also provided in Table 1. The values in bracket are coefficient of variation expressed as a percentage. As can be observed through analysis of the data in the table, significant anisotropy was not observed in terms of the uniaxial stress-

strain data (3.4% increase in yield stress from 90° to 0° with a coefficient of variation of approximately 2%) However, very distinct responses between directions were observed in terms of plastic strain ratios (also known as R-values or Lankford coefficients). Lankford coefficients plotted as a function of effective plastic strain are given in Figure 3 for the 0°, 45°, and 90° directions. Significant noise in the Lankford coefficient data necessitated post-processing consisting of fitting a 6th degree polynomial to the averaged Lankford coefficient versus effective plastic strain curve.

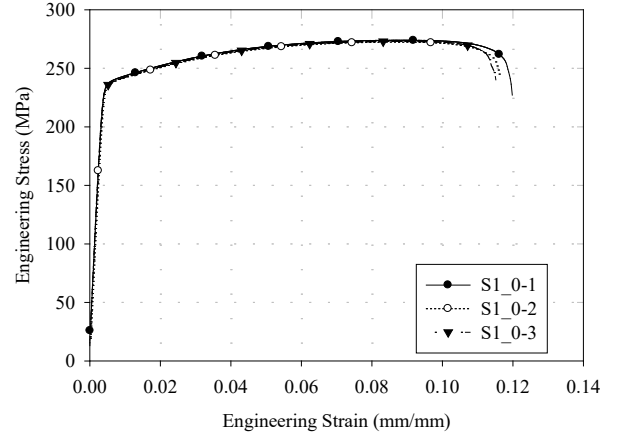


Figure 2. 0 degree uniaxial tensile stress-strain data.

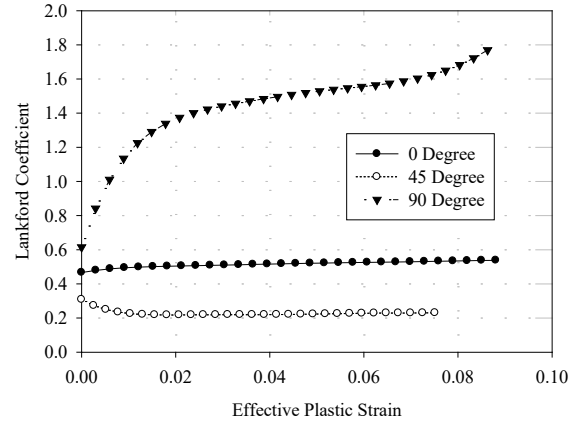


Figure 3. Lankford coefficients

TABLE 1. TENSILE DATA SUMMARY

	Modulus [GPa]	Yield Stress [MPa]	Ultimate Strength [MPa]	Strain at Fracture	Poisson's Ratio
0°	63.41 (1.93%)	247.3 (2.45%)	280.5 (1.81%)	0.118 (4.69%)	0.334 (10.2%)
45°	68.12 (1.03%)	246.2 (1.37%)	274.5 (0.85%)	0.098 (6.09%)	0.317 (9.19%)
90°	65.00 (1.57%)	239.1 (1.83%)	273.3 (1.60%)	0.113 (4.41%)	0.355 (6.81%)

B. VDA 238-100 Plate Bending Test for Metallic Materials

Force-deflection responses for the VDA plate bending tests are presented for the 0° direction in Figure 4 to show the level of consistency between tests of the same direction. To facilitate comparison with other mechanical data it is noted that the 0° specimen configuration was orientated consistent with the VDA standard such that the normal stress (bending) was in the 90° direction. Distinct responses were observed as a function of direction, particularly in terms of force-deflection responses post peak load (onset and propagation of failure). However, the force-deflection responses are not identical between the onset of significant plastic deformation (approximate deflection of 0.75 mm) and peak load. Considering micrographs presented in [20], there may be a skin-core grain structure resulting in this observation of anisotropy under bending but not under uniaxial tension, both in terms of plastic deformation and failure.

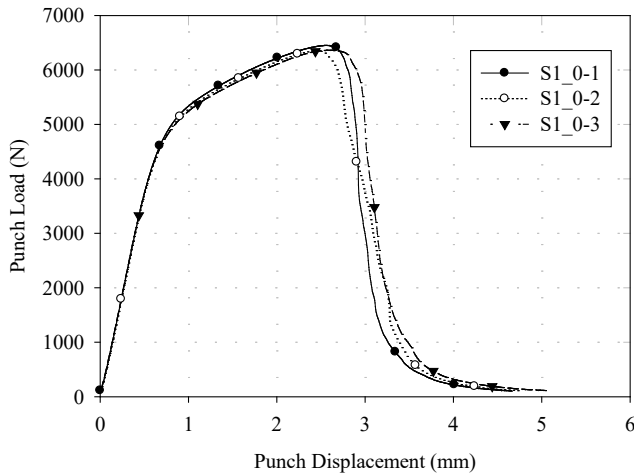


Figure 4. 0 degree VDA force-deflection responses.

Based upon the difference in pre-peak force behavior between 0° and both 45° and 90° (which exhibit similar force-deflection behavior prior to the onset of failure), additional directions were considered as shown in 5. Where the direction was the same, mechanical responses were essentially identical to specimens characterized in the first round of testing. The mechanical behavior in the additional 15° and 30° directions were consistent with expectations: the force deflection responses progressively shift towards the behavior at 45°.

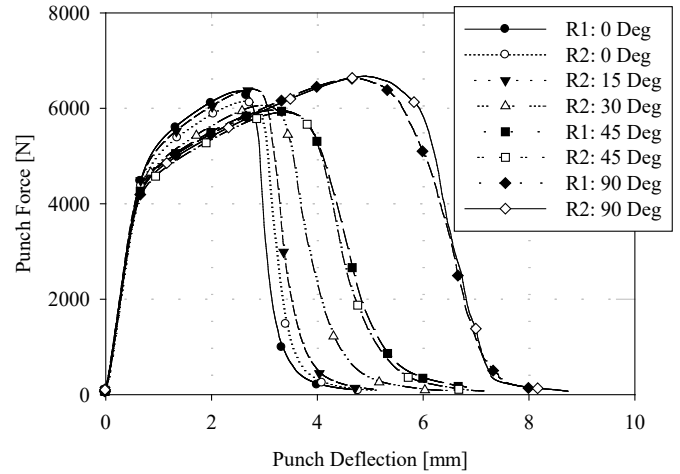


Figure 5. VDA force deflection responses at 0°, 15°, 30°, 45°, and 90°.

C. Three Point Bending of an Aluminum Extrusion

Force-deflection responses for three point bending are shown in Figure 66. All specimens were very consistent in terms of both force-deflection response and failure location/propagation. Fracture locations are shown in Figure 7. Significant bending/curvature at failure locations is noted suggesting that the failure mechanisms of the VDA238 tests may be very relevant in tuning a damage model for this application.

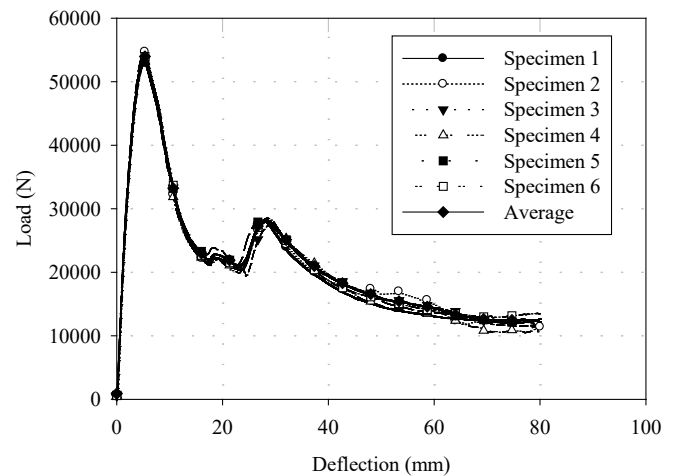


Figure 6. Force-deflection responses of extrusion three point bending tests.

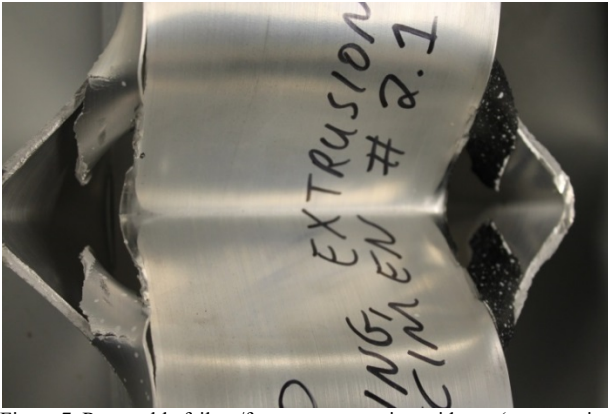


Figure 7. Repeatable failure/fracture at extrusion midspan (compressive surface - in contact with midspan roller).

Deflection measurements (X, Y, and Z, respectively) from digital image correlation for three point bending of an extrusion are presented in Figure 8, Figure 9, and Figure 10 for a crosshead displacement of 10 mm. Deflections at a crosshead displacement of 20 mm are shown in Figure 11, Figure 12, and Figure 13 and for 30 mm in Figure 14, Figure 15, and Figure 16. Projection error was less than 0.1 pixels for all data sets presented here.

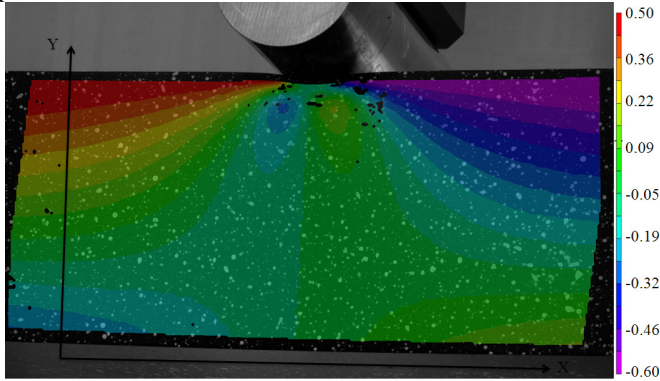


Figure 8. X deflection [mm] at a crosshead displacement of 10 mm.

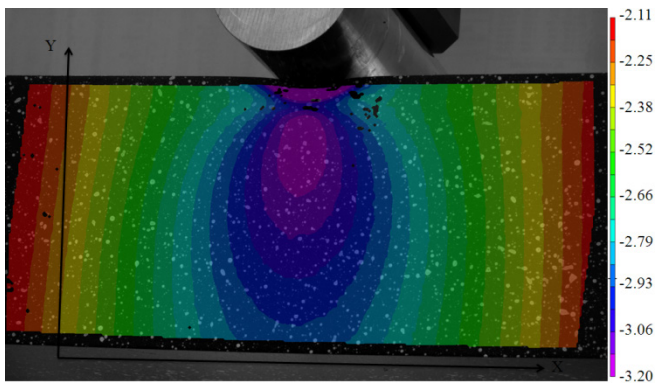


Figure 9. Y deflection [mm] at a crosshead displacement of 10 mm.

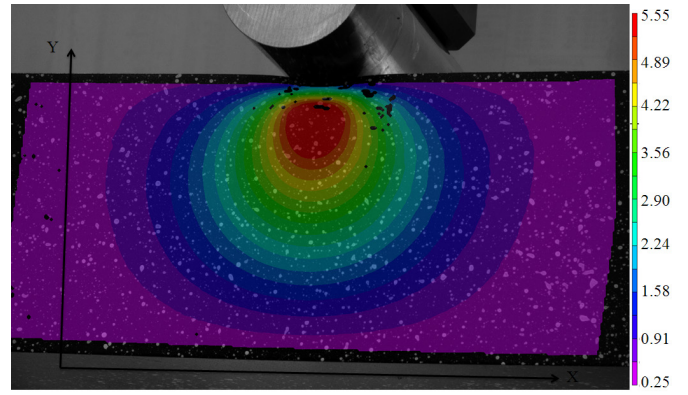


Figure 10. Z deflection [mm] at a crosshead displacement of 10 mm.

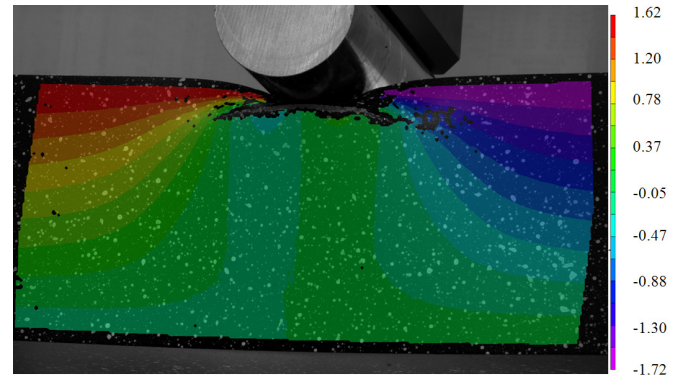


Figure 11. X deflection [mm] at a crosshead displacement of 20 mm.

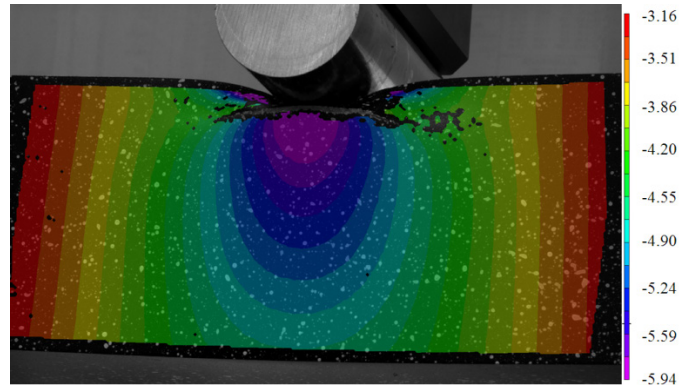


Figure 12. Y deflection [mm] at a crosshead displacement of 20 mm.

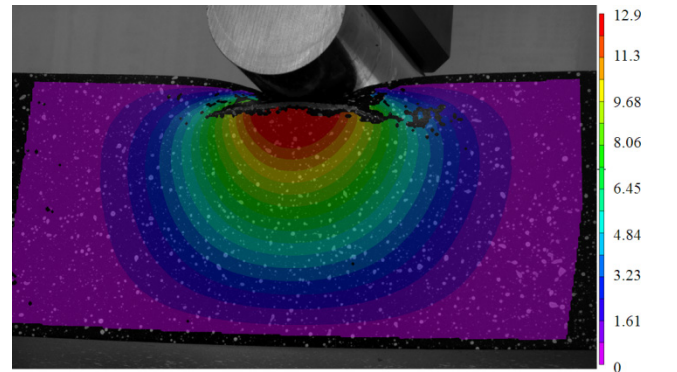


Figure 13. Z deflection [mm] at a crosshead displacement of 20 mm.

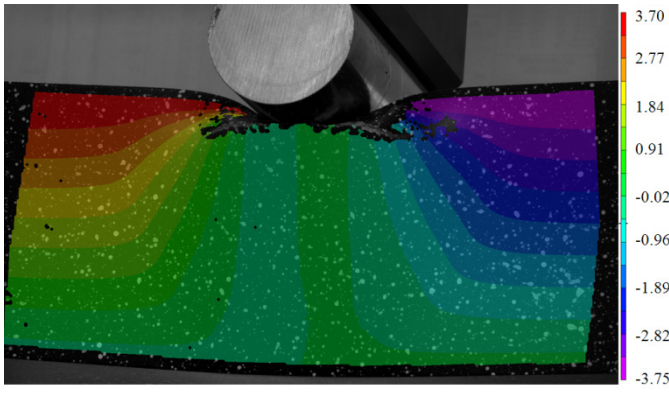


Figure 14. X deflection [mm] at a crosshead displacement of 30 mm.

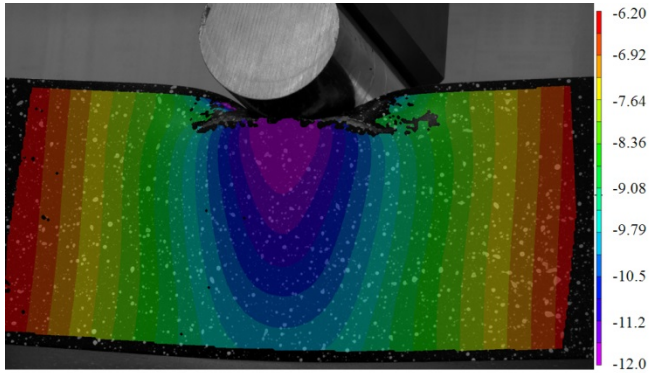


Figure 15. Y deflection [mm] at a crosshead displacement of 30 mm.

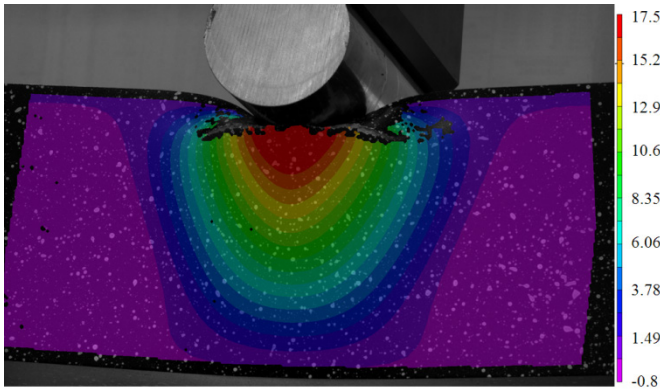


Figure 16. Z deflection [mm] at a crosshead displacement of 30 mm.

D. Axial Compression of Aluminum Extrusions

Load deflection responses for axial compression of 300 mm lengths of aluminum extrusion are given in Figure 17. The first three of nine specimen responses are included as well as an average of all nine responses. All nine specimens post-test are shown in Figure 18. Most specimens buckled near their ends but a small number buckled at approximately midspan. This did not noticeably affect the load-deflection responses but presents challenges for use of the deflection fields acquired through DIC for finite element validation.

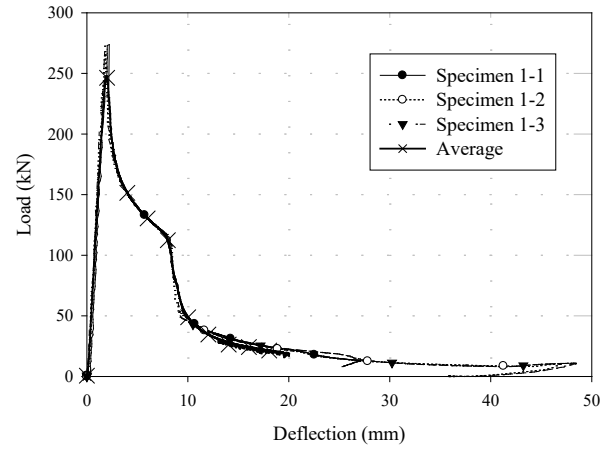


Figure 17. Force-deflection responses of extrusion axial crush tests.

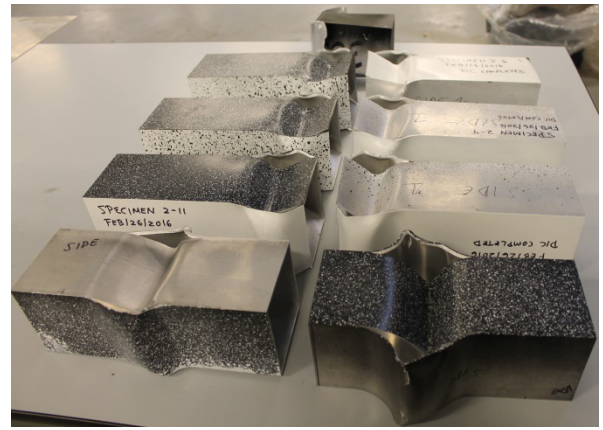


Figure 18. Axial compression specimens post-test

V. CONCLUSIONS & FUTURE WORK

Uniaxial tensile tests did not exhibit anisotropy in terms of stress (3% difference in yield stress with 2% coefficient of variation) but plastic strain ratios were distinct in the 0°, 45°, and 90° directions. Average plastic strain ratios were 0.51, 0.23, and 1.43 for 0°, 45°, and 90°, respectively. VDA plate bending tests were distinct in terms of force-deflection responses after the onset of failure (peak load) for all directions considered (0°, 15°, 30°, 45°, and 90°). Force-deflection responses after the onset of significant plastic deformation were distinct for 0°, 15°, and 30° specimens but essentially identical for 45° and 90°. Repeatable force-deflection and failure mechanisms were observed for three point bending of an aluminum extrusion of a practical geometry for automotive crashworthiness applications. Three dimensional deflection data was acquired by digital image correlation to allow rigorous (future) finite element model validation. Axial compression of 300 mm lengths of aluminum extrusion was also completed. While the force-deflection responses were consistent, the location at which buckling occurred varied between specimens.

This project was completed in partnership with Certasim, a software distributor/support service provider for the finite element modelling software IMPETUS. One outcome was the

enhancement of a Cockcroft-Latham damage model to improve the fidelity of the anisotropy capabilities of the IMPETUS implementation of this damage model. The previous version of this implementation included 2 parameters (0° and 90°). The revised implementation includes a third parameter for 45° . Future and completed but unpublished work includes: finite element modelling of the three point bending test documented here; mechanical testing of this extrusion under axial compressive loading (axial crush); and finite element modelling of axial crush.

REFERENCES

- [1] O. Hopperstad, B. Leira, S. Remseth, and E. Trømborg, "Reliability-based analysis of a stretch-bending process for aluminium extrusions." *Computers & Structures*, 1999. 71(1): p. 63-75.
- [2] L. Chauvet, "Trelleborg Weight Reduction Strategy-Utilization of Plastic for Structural Components for Power Train Mounting System," 2012, SAE Technical Paper # 2012-28-0017.
- [3] A. I. Taub and A. A. Luo, "Advanced lightweight materials and manufacturing processes for automotive applications." *MRS Bulletin*, 2015. 40(12): p. 1045-1054.
- [4] M. Luo, M. Dunand, and D. Mohr, "Experiments and modeling of anisotropic aluminum extrusions under multi-axial loading—Part II: Ductile fracture." *Int. J. Plast.*, 2012. 32: p. 36-58.
- [5] M. Dunand, A. P. Maertens, M. Luo and D. Mohr, "Experiments and modeling of anisotropic aluminum extrusions under multi-axial loading—Part I: Plasticity." *Int. J. Plast.*, 2012. 36: p. 34-49.
- [6] H. N. Wadley, T. Børvik, L. Olovsson, J. J. Wetzel, K. P. Dharmasena, O. S. Hopperstad, V. Deshpande, and J. W. Hutchinson, "Deformation and fracture of impulsively loaded sandwich panels." *J. Mech. Phys. Solids*, 2013. 61(2): p. 674-699.
- [7] G. R. Johnson and W. H. Cook, "Fracture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures." *Eng. Fract. Mech.*, 1985. 21(1): p. 31-48.
- [8] M. Cockcroft and D. Latham, "Ductility and the workability of metals." *J. Inst. Met.*, 1968. 96(1): p. 33-39.
- [9] A. M. Beese, M. Luo, Y. Li, Y. Bai, and T. Wierzbicki, "Partially coupled anisotropic fracture model for aluminum sheets." *Eng. Fract. Mech.*, 2010. 77(7): p. 1128-1152.
- [10] A. Ghahremaninezhad and K. Ravi-Chandar, "Ductile failure behavior of polycrystalline Al 6061-T6." *Int. J. Fract. Mech.*, 2012. 174(2): p. 177-202.
- [11] A. Ghahremaninezhad and K. Ravi-Chandar, "Ductile failure behavior of polycrystalline Al 6061-T6 under shear dominant loading." *Int. J. Fract. Mech.*, 2013. 180(1): p. 23-39.
- [12] D. R. Lesuer, G. Kay, and M. LeBlanc, "Modeling large strain, high rate deformation in metals." *Engineering research, development and technology*, 1999.
- [13] F. A. McClintock, "A criterion for ductile fracture by the growth of holes." *J. Appl. Mech.*, 1968. 35(2): p. 363-371.
- [14] J. R. Rice and D. M. Tracey, "On the ductile enlargement of voids in triaxial stress fields." *J. Mech. Phys. Solids*, 1969. 17(3): p. 201-217.
- [15] A. L. Gurson, "Continuum theory of ductile rupture by void nucleation and growth: Part I—Yield criteria and flow rules for porous ductile media." *J. Eng. Mat. Tech.*, 1977. 99(1): p. 2-15.
- [16] C. Chu and A. Needleman, "Void nucleation effects in biaxially stretched sheets." *J. Eng. Mat. Tech.*, 1980. 102(3): p. 249-256.
- [17] V. Tvergaard and A. Needleman, "Analysis of the cup-cone fracture in a round tensile bar." *Acta Metall.*, 1984. 32(1): p. 157-169.
- [18] Plate bending test for metallic materials, in VDA 238-1002010, VERBAND DER AUTOMOBILINDUSTRIE E.V.: Berlin. p. 1-13.
- [19] P. Larour, B. Hackl, and F. Leomann, "Sensitivity Analysis on the Calculated Bending Angle in the Instrumented Bending Test." in IDDRG 2013. Zurich, Switzerland.
- [20] I. Westermann, K. E. Snilsberg, Z. Sharifi, O. S. Hopperstad, K. Marthinsen, B. Holmedal, "Three-point bending of heat-treatable aluminum alloys: influence of microstructure and texture on bendability and fracture behavior." *Metall. Mater. Trans. A*. 2011. 42(11): p. 3386-98.
- [21] Parachute. (2015). *The Cost of Injury in Canada*. Parachute: Toronto, ON

MANUFACTURING PROCESSES INTERACTION OVER THE FINAL RESIDUAL STRESS STATE OF CRANKSHAFTS

Luiz G. Aun Fonseca*

Instituto Tecnológico de Aeronáutica, Brazil
University of Waterloo, Canada
luizaun@ita.br / luiz.aun.fonseca@uwaterloo.ca

Alfredo R. de Faria

Divisão de Engenharia Mecânica
Instituto Tecnológico de Aeronáutica - ITA
São José dos Campos, Brazil

Mário H. F. Batalha

Laboratório de Estruturas Leves - LEL
Instituto de Pesquisas Tecnológicas - IPT
São José dos Campos, Brazil

Hamid Jahed

Mechanical and Mechatronics Engineering
University of Waterloo - UW
Waterloo, Canada

Abstract— Different residual stress magnitude and profile are generated after every crankshaft manufacturing stage. The processes are often investigated separately, regardless of their possible interaction. The acknowledgement of such interactions can be useful when investigating crankshaft fatigue behavior. This study contributes to the matter by evaluating the resultant residual stress generated at the fillet region after crankshaft's machining and deep rolling processes. Prior to it, a procedure was developed to guarantee the position of the samples in the X-ray diffractometer after surface electropolishing and depth measurement. Repeatability and reproducibility analyses attest the measurements' accuracy. An in-depth profile assessment was made at the crankpin top and bottom regions in two stress directions. The influence of the machining process over the residual stress left after deep rolling can be inferred. Moreover, unexpected residual stress results, and consequently fatigue behavior, can be better understood with more data regarding the whole manufacturing chain.

Keywords: *crankshaft, residual stress, machining, deep rolling, interaction, X-ray diffractometry*

I. INTRODUCTION

A crankshaft can follow several of manufacturing routes. Forged or casted, heat-treated or deep rolled, machine finished or ground, the selection of these variations depends on general design requirements. Each one will imprint a different residual stress (RS) state in the component, which can be negligible or not. There are a few studies related to crankshaft manufacturing processes, especially on residual stress generation at fatigue critical areas such as the bearings' fillets. However, they usually treat the targeted process separately and do not consider prior influence of the rest of the manufacturing chain [1].

This research, therefore, is focused on the possible interactions that may arise related to residual stress generation inside a specific crankshaft manufacturing route. Here it will be

treated the case of a spheroidal cast iron crankshaft that underwent machining, deep rolling (DR) and grinding processes. The hardening treatment is the main concern as it drives the introduction of compressive residual stresses for component fatigue enhancement.

The manufacturing chain can be separated in stages, as described by [2]. The first covers all steps before a rough casted crankshaft is obtained. The temperature gradient during cooling can leave a residual stress fingerprint in the component. According to [3], temperature-controlled stress relief process is optional but mostly indicated after the component is knocked-out. The first stage is considered not to leave any relevant residual stresses into the component, which will be accordingly shown further on for this investigation.

Next phase comprises bearing turning or milling, oil hole drilling, filleting and bearing grinding, among general machining processes performed at the crankshaft counterweights and extremities [2]. Filleting consists of any surface treatment or the combination of them made particularly at the fillets, which can be either thermo-chemical or mechanical. The final stage comprehends balancing, which is done by removing material from the counterweights, quality inspection and aftermarket tracking markup.

Crankshaft machining is considered non-trivial because of its complex geometry. Turn-broaching was implemented for this study, which is able to machine non-concentric features concomitantly [4]. The machine in Figure 1(a) [5] is equipped with broaching spindles packed with a large number of tools around its periphery. The spindles rotate around their own axes at high speeds while it transversely translates to follow crankpin eccentricity. The lathe headstock fixes and rotates the crankshaft at low speeds to allow full access of the turn-broaching tooling to the component.

This manufacturing process potentially leaves a trace of residual stresses in the component. A few parallels can be drawn based on related literature. Reference [6] gather

investigations on general machining processes and their impact over the surface residual stress state. Since turn-broaching can be correlated to milling, outcome similarities are expected. The residual stress state is shifted from tension to compression depending on the machining parameters, tool and workpiece material. Compressive RS are expected because of considerable plastic deformation produced by milling [6]. However, high temperatures favor less compressive or even tensile stresses at the surface. Up cut milling generates cold plastic deformation and introduces less heat to the workpiece, favoring the appearance of compressive stresses. Down cut milling is the opposite, since the cut starts with the maximum chip thickness, transducing more heat to the surface.

Crankshaft turn-broaching is similar to down cut milling. For this reason and for the lack of references on the subject, it is not trivial to predict the magnitude of tractive stresses induced by the process near the surface. Since machining is not the last stage of the crankshaft manufacturing chain, the following processes will be disturbed by its outcomes, as shown by [7] for gear production. For the present study, only deep rolling will affect fillet region residual stresses after machining. If superficial compressive residual stresses arise at the fillet from turn-broaching, it will be theoretically beneficial for the component fatigue performance. In case tensile RS are generated, then the effect will be inversed. Either way, the lack of information on the matter constitutes an opening to be addressed.

The next process in the investigated crankshaft manufacturing chain is deep rolling. It is commonly the

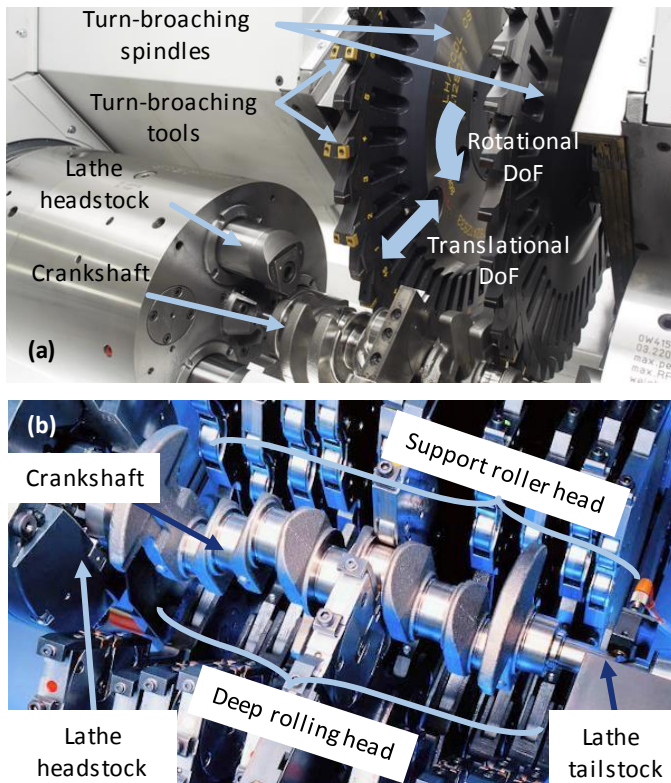


Figure 1 –Crankshaft (a) turn-broaching (b) deep rolling machines [5], [10].

only mechanism employed to improve fatigue performance for cast iron components [8]. Compressive residual stresses are introduced by pure mechanically induced plastic deformation [9]. It consists of a forming process that generates plastification and compressive residual stresses at the bearing fillets. The component is fixed at one end by a lathe headstock and supported at the other extremity by a tailstock, as shown in Figure 1(b) [10]. The rollers are housed in cages at the DR head. To counterbalance the forces applied at the rollers, there are supports in contact with the bearings' surfaces. The crankshaft is driven at speeds around 60 rpm while hydraulic pistons close both the deep rolling head and support against the main journal and crankpin fillets.

In the manufacturing chain targeted in the present study there are no other processes that significantly modify the residual stress state of the fillet region. Grinding is performed only at the pins and main bearings surfaces, but not at the fillets. Additionally, balancing commonly removes material from the counterweights, also far from the fillets. Thus, bearing in mind that casting does not induce any relevant residual stress state, it is the interaction between machining and deep rolling that must be studied in order to understand the RS generation in the fillet zone throughout the crankshaft manufacturing chain. Focus will be given to DR, which predominantly shapes the RS profile along depth. The understanding of residual stress origins and measurement principles is thus necessary.

X-ray diffraction (XRD) is a common and useful method employed to assess the residual stress field generated in mechanical components from different sources [11], [12]. This nondestructive technique works with an embedded statistical approach for the crystalline lattice evaluation [13]. One advantage of XRD, especially portable diffractometers, is its application in components of complex geometry [14]. Here the emphasis is towards the assessment of residual stress profile along depth of automotive components.

In this specific study, a crankpin fillet section was chosen, as it is an area that have an inherent uncertainty about its residual stress distribution, according to [15]. This is mainly because of the component's geometry, which may hamper measurements, aside from the operating loads that contributes to the difficulty in accurately predicting its life. Crankshafts support cyclic loads, which makes it susceptible to damage generated by fatigue [16].

Although residual stress evaluation has an intrinsic variability related to specimen material and manufacturing, most studies do not evaluate or substantiate the measurement procedure as one possible source of uncertainty. This lack of information can lead to challenges on validating results from diverse sources, which compromise confidence and, possibly, industrial application. A few researches have attended this issue. Reference [17] have proposed a method to assess the residual stress heterogeneity state of automotive gears, approaching the deviation brought by sample repositioning. Their experimental methodology included a mathematical correlation of micro residual stress intensity and macro RS deviation, which were also dependent on repeatable results. The Margin of Random Error (MRE) was calculated and the

scenario with sample repositioning considered for conservative results [17].

In the light of the above, this study aims at filling the existing gap of residual stress interaction between the machining and deep rolling processes. Additionally, it contributes to understand the impact of possible sources of measurements scattering with focusing in specimen position and operator interference. This investigation is inserted in a broader research topic that encompasses deep rolling modeling for RS prediction and correlation with crankshaft fatigue performance [18]. Thus, the determination of accurate residual stress values is vital for its success.

A. OBJECTIVE AND APPROACH

The objective of this study is the verification of residual stresses generated by crankshaft machining process and its translation to the final RS state after deep rolling. It is also intended to develop an experimental apparatus for sample positioning in residual stress assessment made on crankshafts. The contribution lies in the establishment of a procedure that enables accurate X-ray diffraction measurements for future finite element model validation.

First, a brief explanation is given on the deep rolling model development and on the necessity of its validation. Description on the crankshaft specimen manufacturing route is also given. Next, the templates' and sample holders' design is covered, with the corresponding measurement requirements as input. Finally, a repeatability and reproducibility procedure was conducted to validate the developed experimental apparatus.

II. EXPERIMENTAL APPARATUS AND PROCEDURE DEVELOPMENT

A. MODEL STATUS AND CRANKSHAFT SPECIMENS

As mentioned before, a finite element model of the deep rolling process is under development at this time. Although several advancements were already made and verified [18], the continuation of the model validation is still necessary. The simulation outputs related to energy balance showed themselves to be in accordance with expected process dynamics. Also, geometrical measurements made at the fillet matched the model's nodal displacements. Figure 2 brings a summary on the model development so far. The residual stress assessment began to be covered in [18], however the authors stated clearly that a thorough validation must be comprehended in the research scope.

Crankshafts target of this study were casted, with a final microstructure presenting graphite nodules. After casting, some samples had its bearings and fillet radii machined and were separated. Other specimen followed the rest of the manufacturing chain, passing through deep rolling and bearing grinding. Regarding the regions of interest, residual stresses were evaluated at the crankpin top and bottom areas. That is related to the fatigue requirements for each region. As the bottom of the crankpin suffers tensile stresses due to bending loads when the piston is on top dead center and combustion occurs, that is considered the critical fatigue site. Thus, the rolling process favors the introduction of higher compressive

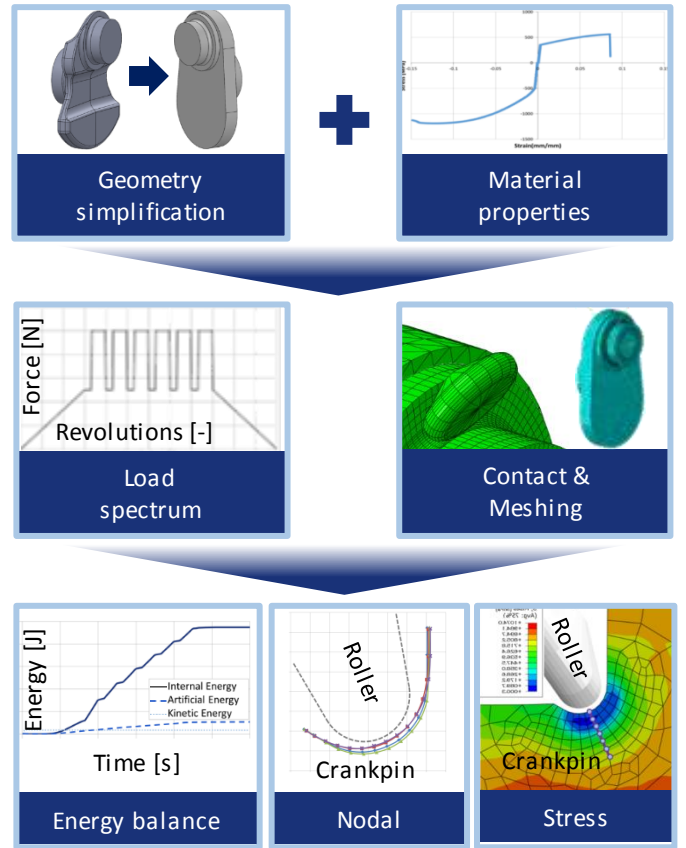


Figure 2 - Crankshaft deep rolling modeling workflow [18].

residual stresses in this region. This is done through the load spectrum shown in Figure 2, which applies twice the load at the crankpin bottom if compared to the top half. Machining theoretically introduces the same amount of RS through the bearing fillet periphery, which should be verified. Therefore, both top and bottom areas will be evaluated in order to understand how the processes differently imprint residual stresses near and at the surface.

B. RESIDUAL STRESS MEASUREMENT PROCEDURE

As the literature survey indicated, crankshaft residual stress measurement along depth is not trivial. Researches were not clear on which conditions and procedures their RS measurements were taken [19], [20]. This research intends to create a detailed methodology of crankshaft fillet in-depth RS assessment. To accomplish that, it relied on partners for expert assistance in the field.

In a first moment, a residual stresses measurement procedure was developed together with *Stresstech GmbH*. The objective was not to produce massive RS data, but to elaborate a solid methodology for the challenging task in hand. Only the crankpin fillet top region was assessed at this time. All research requirements were discussed, and a feasible procedure was drawn. The equipment used for residual stress measurements was the *Xstress 3000 G2R* diffractometer. Its main functionalities include portability and ability to perform ϕ -rotation and measure RS in two directions without sample repositioning. A cylindrical coordinate system is defined at the bearing fillet, with radial (r), hoop (θ) and axial (z) axes. The

verification of the in-plane stresses, or θ and z directions in this case, will be compared with the simulation's outputs in following investigations. Figure 3(a) illustrates a transverse cut of the crankpin to better illustrate the assessed directions. The X-ray collimator is aligned with the radial direction, so the plane stresses measured are σ_θ and σ_z .

Twelve points were measured along fillet depth, spaced progressively at 0, 50, 100, 250, 500, 750, 1000, 1250, 1500, 2000, 2500, 3000 μm . These points were approximately aligned to the roller angle or radial direction, for it is the region that experiences deep rolling direct force application. Electropolishing was used to remove layers of the material. The space between the evaluated points gradually increases in order to capture any steep characteristic of the stress profile near the surface. The surface was electropolished deeply enough to observe residual stress transition between compressive and tractive states. The equipment used for the task was the *Struers Movipol 5* with an acid solution AC2. The electrolyte is a mix of perchloric and citric acids, ethanol, propanol, sodium thiocyanate anhydrous, hydroxyquinoline and distilled water. Grinding was necessary near the etched region to avoid blocking of the X-ray beam before it could reach the fillet surface, as it is shown in Figure 3(b).

As the material matrix is predominantly $\alpha\text{-Fe}$, a Cr K- α X-ray tube anode was used to measure this phase at the microstructure lattice plane $\langle 211 \rangle$. The collimator utilized had a 0.8 mm diameter. A powder sample with virtually no stress was previously measured for equipment calibration. In this process, the angular position of the two detectors was set. The scanning range of the diffracted ray 2θ was set to $146^\circ - 164^\circ$. Tilt in the χ angle was the method chosen. A total of nine χ angles varying from -39° to $+39^\circ$ were set in both hoop and axial directions. Equipment for determining X-ray elastic constants was not available for this research, so these properties were taken from [21]. Besides the impact in stress accuracy by up to 10% [22], it does not impair the conclusions drawn here.

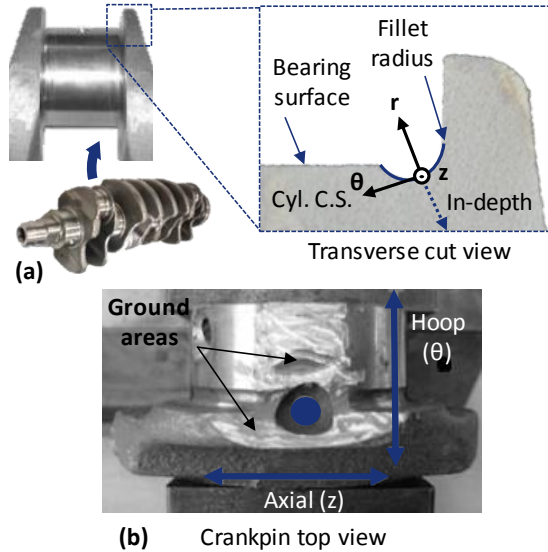


Figure 3 - Top region (a) cylindrical coordinate system and (b) etched configuration.

The interest relies on the tendency shown by numerical and experimental results rather than absolute value accuracy.

The first residual stress measurements were performed by *Stresstech GmbH* in their facilities in Germany. They were important to observe the RS tendency and initiate the model validation phase [18]. The elaborated procedure served as a baseline for the measurements to be performed with the *Xstress 3000 G2R* diffractometer available in the *Lightweight Structures Laboratory* at the *Institute for Technological Research (LEL-IPT)*.

Care was taken to accurately position the crankshaft sample for the residual stress measurements. The diffractometer table is equipped with built-in linear guide and indexers. It allowed table translation and indexing for the creation of three basic stations: RS measurement, depth determination and electropolishing. Figure 4 brings an illustration of the stations set up. Depth was determined using a dial gauge that resulted in an accuracy of approximately 10 μm for each measured point.

C. DIFFRACTOMETER'S TEMPLATE DESIGN

Fixture templates were designed to position and fix the crankshaft sample over the diffractometer table. Figure 5 shows a schematic of these indexers. They were developed to allow repositioning of the crankpin depending on whether the top or bottom regions need to be evaluated. A base fixed to the machine's table had sidewalls to prevent sample movement in the linear guide axis, as the table has its own fine adjustment and fixture features. The sidewalls enable movement only in the transverse direction and screws and nuts lock the sample holder in place. An assistance template was developed to help with sample alignment, which is ultimately visual. Once positioned over the crankpin, the apparatus marks the center in both axes, so the sample can be aligned with the machine's collimator and locked down in place. The sample holder was designed to support the crankpin inclined and aligned with the

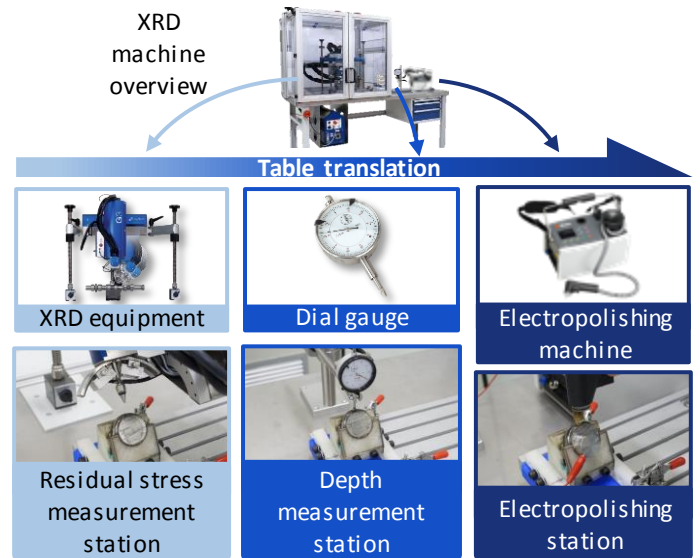


Figure 4 - Residual stress measurement, sample electropolishing and depth measuring stations in the diffractometer equipment.

same angle as the roller during the deep rolling process. As the template parts have quite complex geometries, all pieces were produced in Polylactic Acid (PLA) through 3D printing.

Another point of concern was related to crankpin geometry interference with X-ray emission and diffraction. As the fillet has a natural undercut radius, the counterweight and bearing region could block the X-ray beam when the diffractometer head is tilted. The same challenge was faced by [20] on their crankshaft fillet radius residual stress investigation. That demanded a similar solution, based on an analysis in order to determine up to which degree the collimator can be angled and if any adaptations are required. As it can be observed in Figure 6, the collimator cannot be tilted up to the machine's maximum angle of 45° without X-ray blocking when stress measurements are being made in the hoop direction. The bearing surface and counterweight wall must be slightly ground, as shown in Figure 6(a), to permit free path to the X-ray beam up to the measurement region. Additionally, in order to maximize the collimator tilting, asymmetric angles ranging from -40° up to 35° were determined, illustrated in Figure 6(b). This action, in combination with the necessary grinding, guarantees that the radiation signal is hitting and being scattered at the correct area. Measurements in the axial direction do not face such challenges, as the collimator can freely tilt without any blockage.

Interesting observation that could be drawn from the collimator interference study regards the mesh of the target region. The projection of the X-ray beam over the measuring surface indicates the area that will be primarily assessed. The overlapping of this image by the fillet-structured mesh gives an indication of the elements and nodes that should be closely evaluated. Evidently, the system positioning has intrinsic errors and a good practice implemented was to analyze the adjacent region also. Figure 6(c) shows the X-ray beam over the fillet surface and highlights the node rows to be correlated with future model validation and crankshaft fatigue behavior.

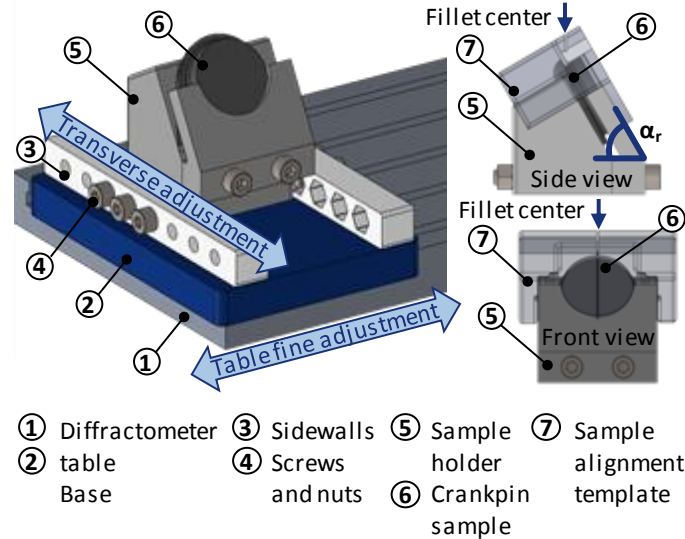


Figure 5 - Templates designed to position the crankpin specimen in the diffractometer.

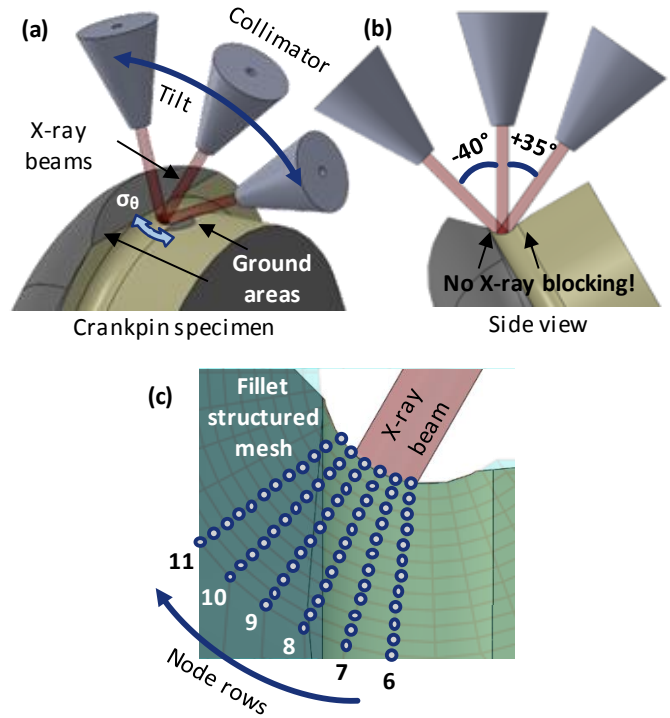


Figure 6 - Collimator interference verification: (a) general view of ground areas; (b) asymmetric tilting; (c) visual identification of mesh region of interest with X-ray beam.

To verify the template's necessity, a repeatability and reproducibility analysis was conducted. For the repeatability test, the same operator made six measurements with and without the 3D printed indexers, removing the crankshaft sample from the diffractometer table every time. The operator used a manual protractor and modeling mass to position the specimen at a determined angle without the templates. For the reproducibility evaluation, measurements performed by two operators were confronted. They followed the same procedure described above. Results were statistically verified through the analysis of average value, standard deviation σ , p-value and coefficient of determination R^2 [23].

III. RESULTS AND DISCUSSION

At a first moment, it was necessary to attest that the XRD templates indeed benefits the residual stress outputs' accuracy. Figure 7 summarizes the results obtained from measurements in both stress directions evaluated at the crankpin top region. The two operators trained to implement the procedure performed them with and without the designed indexers. As the results with the aiding apparatus fell together inside a quite small deviation, it was chosen to display the results here as a single group. However, without the templates the outcomes for both operators diverged. For hoop residual stress, although the mean values' magnitudes were similar to measurement with template, the standard deviations reveal a lack of repeatability. The analysis of p-value and R^2 attest that all three groups are indeed statistically equal. The analysis of the axial residual stresses reveals another condition. The groups can be

considered statistically different at a confidence level of 95%, as the calculated p-value falls below 0.05. Sources for the relatively high standard deviation and lack of reliability for the measurements without the templates are possibly related to protractor miss positioning and modeling mass compliance. These facts strengthen the conclusion of the template's requirement for this case.

The residual stress measurement procedure validation enabled its usage for crankshafts that passed only through turn-broaching. Figure 8 summarizes the results found so far. Both top and bottom regions were assessed up to one millimeter in depth at this point. Figure 8(a) shows that the hoop residual stresses are tractive at the surface, both near 450 MPa. For the top region, the value rapidly drops to a value close to zero before 0.1 mm and oscillates to negative stress in order to compensate surface tensile and account for the RS equilibrium principle [24]. The bottom curve also starts at an elevated value but reduces quite slowly and only reaches null stress close to 0.8 mm. Axial residual stresses depicted in Figure 8(b) are similar in both top and bottom regions. They also start with a tractive value near 300 MPa at the surface and quickly drop

and alternates between positive and negative stress up to 1.0 inward.

As exposed in the literature survey, the topic for RS state left after crankshaft fillet machining was little explored, so there are few parallels for comparison. Some aspects deserve notice here. Surface residual stresses in the hoop direction appears to be consistently higher than in axial. Stress drops to null values in-depth, which contributes to acknowledging that the casting process is irrelevant in terms of residual stress generation. One point of concern is related to the bottom region hoop RS. It is particularly different from the other measured data, as its drop is not abrupt, but slow and steady. This reveals that the turn-broaching process can introduce a disturbance factor in the manufacturing chain. Further measurements should be conducted to verify rather if this was an isolated condition or an intrinsic process characteristic.

Surface and in-depth residual stresses measured at the crankpin top and bottom region after the specimens passed through deep rolling are depicted in Figure 9. Both directions have compressive values at the surface and show a tendency to transit to tractive states at three millimeters in depth, which is expected according to the RS equilibrium principle.

A closer look at Figure 9(a) discloses particularities worth of comment. For both regions analyzed there is a sudden decrease in RS value just below surface at 50 μm , which is quickly recovered afterwards. One hypothesis raised for this

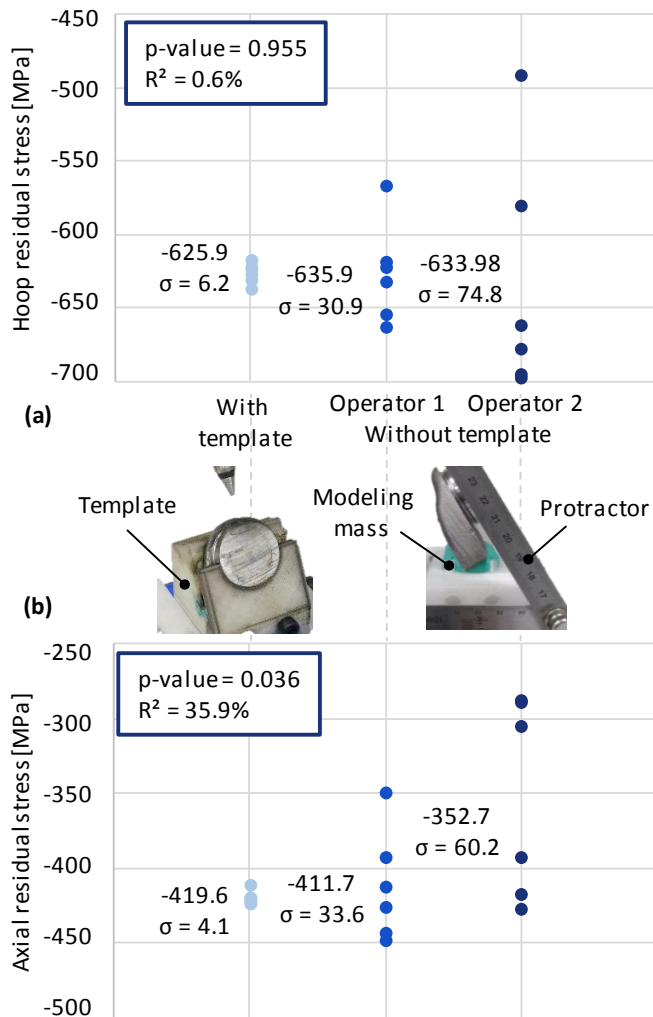


Figure 7 - Repeatability and reproducibility analysis for (a) hoop and (b) axial residual stress directions.

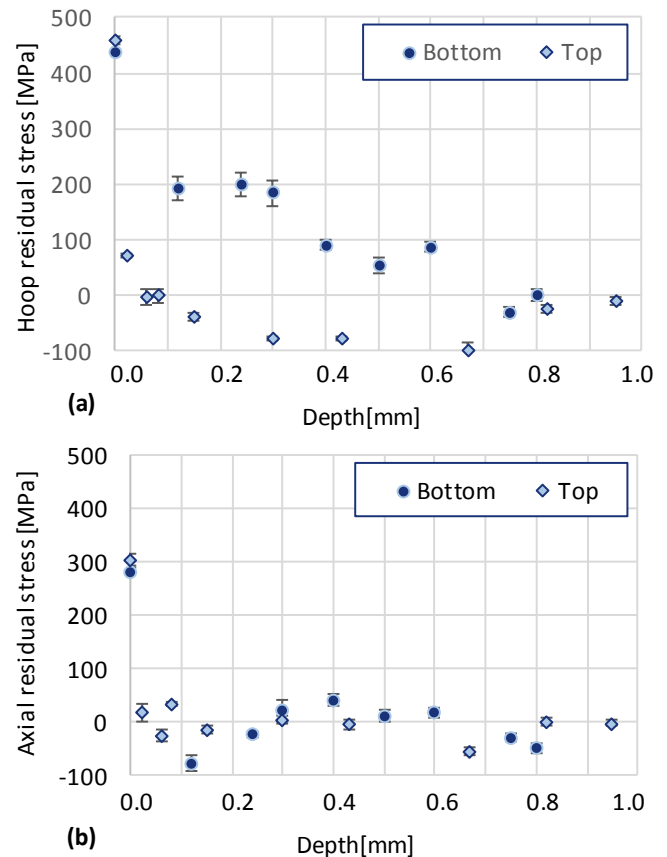


Figure 8 – (a) Hoop and (b) axial RS measurements after turn-broaching.

behavior is the influence of the residual stress trace left by the machining process. Moreover, a comparison between the two regions responses reveal another point of concern. The magnitude of the bottom stress should, theoretically, be higher than at the top. This is because of the fatigue requirements for the regions, which is imposed by the load spectrum applied to the specimen, as depicted in Figure 2. Nonetheless, the opposite is observed inside the one millimeter range. This can be also explained by the influence of turn-broaching RS over the next process in the chain. As Figure 8(a) showed, the hoop stress at the bottom region was particularly high up to a depth of approximately 0.5 mm. This coincides with the range of Figure 9(a) bottom curve that has less residual stress magnitude compared to the top region. Another cause for this observation would naturally be a deviation produced at the deep rolling process itself. However, there is no positive evidence that supports this hypothesis at this time.

IV. CONCLUSIONS

Referring back to the objective of this investigation, the hypothesis for residual stress disturbance generated at the machining stage possibly affecting deep rolling outcomes was raised. The effect was more pronounced for bottom hoop RS,

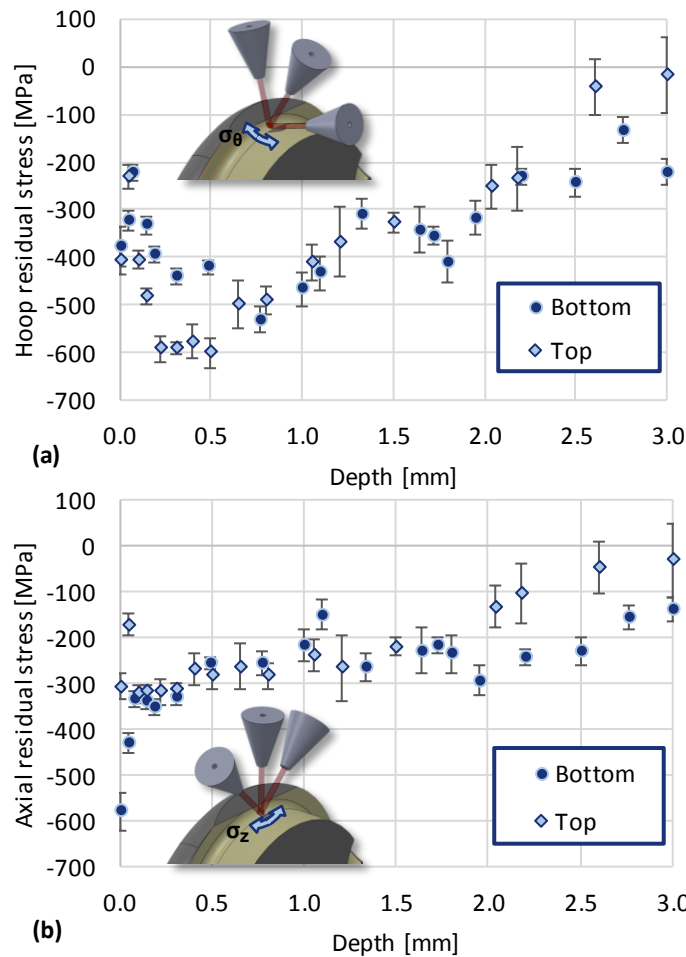


Figure 9 – (a) Hoop and (b) axial RS measurements after deep rolling.

which is coincidentally the critical region for fatigue failure. This indicates that attention should be given towards the evaluation of every process that can potentially change the workpiece overall residual stress state. The following remarks can be highlighted at this point:

- The repeatability and reproducibility analysis revealed the need for the use of positioning templates in this case. The sample had to be often manipulated for electropolishing, so the development of indexers became imperative;
- The residual stress measurements with the specimens that passed only through turn-broaching indicated tractive RS at surface, with a natural transition to compressive, which was well expected according to the equilibrium principle;
- The stress results for the bottom region in the hoop direction raised a point of concern, as the machining process seemed to have imprinted an undesired RS state in the samples;
- Additionally, the first measurements indicated that the casting process does not leave any relevant amount of residual stresses;
- For the analysis of the deep rolled samples, a compressive RS field was observed as anticipated;
- Again, for the bottom region in the hoop direction, the residual stresses appeared to have been diminished in terms of absolute value. The hypothesis raised was that the stress state was directly influenced by the previous manufacturing stage.

The results obtained in this study have contributed to shed light over the crankshaft manufacturing chain residual stress generation. The next steps include the incorporation of such results in the deep rolling finite element model validation. This should be done with proper caution related to the possible influence that the intermediate processes may have over the final residual stress state.

ACKNOWLEDGMENT

The authors would like to thank the company *Stresstech Group* for the residual stress procedure development and measurements. This study was partially funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under grants 147267/2015-3 and 306193/2017-5, and also by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) under grant 88881.187305/2018-01 for the Programa Institucional de Bolsas de Doutorado Sanduíche no Exterior - PDSE.

REFERENCES

- [1] DELGADO, P; CUESTA, I.I.; ALEGRE, J.M.; DÍAZ, A. State of the art of Deep Rolling. *Precision Engineering*, vol. 46, p. 1-10, 2016.
- [2] NALLICHERI, N.; CLARK, J.; FIELD, F. Material Alternatives for the Automotive Crankshaft; A competitive assessment based on manufacturing economics. *SAE Technical Paper n. 910139*, 12p., 1991.
- [3] DAVIS, J. R. *ASM Specialty handbook cast irons*. ASM International, 494 p., 1996.

- [4] SCHROETER, R. B.; BASTOS, C. M.; CRICHIGNO FILHO, J. M. Simulation of the main cutting force in crankshaft turn broaching. *International Journal of Machine Tools & Manufacturing*, vol 47, p. 1884-1892.
- [5] HELLER MACHINE TOOLS L.P. Available at: <www.heller-us.com>. Accessed on 09 Sep. 2018.
- [6] BRINKSMEIER, E.; CAMMETT, J.T.; KÖNIG, W.; LESKOVAR, P.; PETERS, J.; TÖNSHOFF, H.K. Residual stresses – measurement and causes in machining processes. *Annals of the CIRP*. Vol. 31/2. 1982.
- [7] REGO, R.; LÖPENHAUS, C.; GOMES, J.; KLOCKE, F. Residual stress interaction on gear manufacturing. *Journal of Materials Processing Technology*, v. 252, p. 249-258, 2018.
- [8] CEVIK, C.; HOCHBEIN, H.; REBBERT, M. Potentials of crankshaft fillet rolling process. *SAE Int. J. Engines*, v. 5, n. 2, p. 622-632, 2012.
- [9] CEVIK, C. An Efficient Methodology for Borderline Design of Crankshafts. 2011. 140p. Dissertation (Doctor of Engineering Sciences) – Faculty of Mechanical Engineering, RWTH Aachen University, Germany.
- [10] NILES-SIMMONS-HEGENSCHEIDT (NSH). Available at: <www.hegenscheidt-mfd.com/en/automotive-technology>. Accessed on: 30 Aug. 2017.
- [11] R. HUSSON, J. DANTAN, C. BAUDOUIN, S. SILVANI, T. SCHEER and R. BIGOT, Evaluation of process causes and influences of residual stress on gear distortion, *CIRP Annals – Manufacturing Technology*, vol. 61, 2012, pp. 551-554.
- [12] J. LAN, S. FENG and L. HUA, The residual stress of the cold rolled bearing race, *Procedia Engineering*, vol. 207, 2017, pp. 1254-1259.
- [13] WITHERS, P. J.; PREUSS, M.; STEUWER, A.; PANG, J. W. L. Methods for obtaining the strain-free lattice parameter when using diffraction to determine residual stress. *J. Appl. Cryst.* (2007). V. 40, p. 891-904.
- [14] FITZPATRICK, M.E.; FRY, A.T.; HOLDWAY, P.; KANDIL, F.A.; SHACKLETON, J.; SUOMINEN, L. Measurement Good Practice Guide: Determination of Residual Stresses by X-ray Diffraction. National Physical Laboratory, United Kingdom, Issue 2, No 52, 2005.
- [15] P. SPITERI and S. HO, Y. LEE, Assessment of bending fatigue limit for crankshaft sections with inclusion of residual stresses. *International Journal of Fatigue*, vol. 29, 2007, pp. 318-329.
- [16] W. Y. CHIEN, J. PAN, D. CLOSE and S. HO, Fatigue analysis of crankshaft sections under bending with consideration of residual stresses. *International Journal of Fatigue*, vol. 29, 2005, pp. 1-19.
- [17] F. KLOCKE, J. GOMES, C. LÖPENHAUS and R. R. REGO, Assessing the heterogeneity of residual stress for complementing the fatigue performance comprehension. *The Journal of Strain Analysis for Engineering Design*, vol. 51, 2016, pp. 347-357.
- [18] FONSECA, L.; FARIA, A. A deep rolling finite element analysis procedure for automotive crankshafts. *Journal of Strain Analysis in Engineering Design*. Vol. 53(3) p. 178-188, 2018.
- [19] MICHAUD H.; SPRAUEL, J.M.; GALZY, F. The Residual Stresses Generated by Deep Rolling and their Stability in Fatigue & Application to Deep-rolled Crankshafts. *Materials Science Forum*, vol 524-525, p. 45-49, 2006.
- [20] GALZY, F., MICHAUD, H., SPRAUEL, J.M., Approach of Residual Stress Generated by Deep Rolling Application to the Reinforcement of the Fatigue Resistance of Crankshafts, *Material Science Forum*, vol.: 490-491 p. 384-389, 2005.
- [21] CULLITY, B.D.; STOCK, S.R. Elements of X-Ray Diffraction. Essex, England: Pearson, 2014. 649 p.
- [22] PREVÉY, P.S. A method of determining the elastic properties of alloys in selected crystallographic directions for X-ray diffraction residual stress measurement. *Advances in X-Ray Analysis*, vol. 20, 345-354, 1977.
- [23] MONTGOMERY, D. C. Design and analysis of experiments. New York: John Wiley, 752 p., 2001.
- [24] HAUKE, V. Structural and Residual Stress Analysis by Nondestructive Methods. Elsevier, Amsterdam, 1997.

Failure Characterization of Heavily Cross-linked Epoxy Part I: Testing

A. Y. Elruby

Ph.D. Candidate

Mechanical Engineering, Faculty of Engineering

Memorial University of Newfoundland

St. John's, NL, Canada

E-mail: aleruby@mun.ca

Sam Nakhla

Assistant Professor

Mechanical Engineering, Faculty of Engineering

Memorial University of Newfoundland

St. John's, NL, Canada

E-mail: satnakhla@mun.ca

Abstract— This is part I of two companion papers on failure characterization of heavily cross-linked epoxy. Fiber reinforced polymers (FRPs) are widely used in many engineering fields such as marine, automotive and aerospace industries. The matrix material dominates two failure modes of fiber reinforced polymers (FRPs). Epoxy resins represent more than 90% of matrix materials in FRPs. The objective of the current work is to enhance failure predictions in FRP relying on the precise characterization of failure and multiscale finite element analysis. In part I, the failure of heavily cross-linked epoxy is thoroughly inspected under different types of loading, namely uniaxial tension and three-point bending. In part I, detailed testing procedures for conducted standard testing are provided with insight on fractured surfaces. Digital image correlation (DIC) was used for accurate strain measurements in all conducted tests. Also, optical microscopy was employed to conducted fractographic analysis on fractured surfaces.

Keywords—Heavily cross-linked epoxy; Fractographic analysis; Characterization

I. INTRODUCTION

Epoxy resins represent the majority of matrix and materials in fiber reinforced polymers (FRPs). Also, epoxy resins are widely used over other resins as lamination adhesive for low-shrinkage characteristics [1]. In many industrial applications such as automotive, marine and aerospace FRPs have been replacing conventional materials [2]. FRPs are preferred for their enhanced mechanical properties such as stiffness-to-weight ratio, also outstanding capability in resisting harsh environmental conditions [3]. However, manufacturing defects such as inclusions in FRPs fabrication processes are inescapable. FRPs are characterized by complex failure mechanisms owed to heterogeneity nature of fabricated composites as well as manufacturing imperfections. FRPs mechanical behavior and failure mechanisms are strongly related to these imperfections. Failure mechanisms are somehow problematic to deal with, particularly in early design stages due to two main reasons, anisotropic nature and material imperfections [4].

Two out of four typical failure modes of FRPs are dominated by matrix material [5]. Moreover, ply debonding in laminated composites are dominated by lamination resin. Voids in matrix materials have proven to deteriorate mechanical behavior [6–8]. The polymerization process of resin curing usually results in undesired inclusions such as micro-voids [6,9,10]. Some methods were developed to reduce void formation such as vacuum bagging and autoclaving, yet it remains unavoidable. Several studies have shown that microvoids could possibly lead to micro-cavitation [11–13]. In turn, possible nucleation and micro-cracks will initiate and under different loading conditions may propagate to macro-cracks [14]. Therefore, efforts have been made to account for the effect of voids on composite failure [9,15–17]. Few studies were mainly focused on resin materials failure. For example, the work by Asp et al. [13,14] provided an insight into the epoxy failure under composite like stress state.

Ply-delamination, matrix cracking and fiber debonding are typical failure modes which are dependent on resin materials [18]. Precise identification of failure mechanisms in composite materials would be of paramount usage for design stages. As a result, there exist several studies in the literature on failure modes of composites. For example, Lachaud et al. [19] modeled matrix cracking and ply delamination in laminated composites using cohesive models and continuous damage models. They showed in their work that matrix cracking and delamination were principally initiating failure in laminated composites. Also, Pollayi and Yu [20] studied failure onset in composite rotor blades which can be found in a helicopter rotor or wind turbine blades. They showed that matrix micro-cracking maybe considered as the first damage event in FRPs. Riaño et al. [21], utilized the representative volume element (RVE) method to predict the mechanical behavior of unidirectional glass fiber composites. They also modeled damage phenomena such as matrix cracking and fiber/matrix debonding at a mesoscale to enhance macro-mechanical numerical predictions. Their study yielded approximately 25% error when compared to testing.

As shown in the literature, failure predictions in FRPs are usually experiencing a relatively high bound. Also, neat resin materials failure has been partially studied. Therefore, a closer insight into the failure of neat epoxy resin similar to the one

used in FRPs and composite layups is extensively characterized and examined. Neat epoxy samples were prepared from the same patch to be tested under both uniaxial and three-point bending loads. Efficient identification method of material yielding is proposed. Also, fractographic analysis for fractured specimens was conducted and results are provided. In part II, modeling and simulation of heavily cross-linked epoxy are provided utilizing multiscale finite element analysis.

II. SPECIMENS PREPARATION AND MECHANICAL TESTING

A. specimens preparation

Commonly used in aerospace FRPs neat epoxy was cast and cured by the manufacturer (i.e., Polynt Composites Canada, Inc.). The provided casted epoxy plate had a thickness of 10mm which was milled down to a thickness of 9mm to ensure mirror-like finish surfaces from both sides. The epoxy mixture had a resin to hardener weight ratio of 6:1. All specimens were prepared from the same neat epoxy plate. Three dog-bone specimens (i.e., type III in the testing standard) were machined to be tested under uniaxial tension according to ASTM D638 recommendations. Another set of three prismatic specimens were fabricated to be tested under three-point loading according to ASTM D790 recommendations. Both types of specimens ahead of testing are shown in Figure 1.



Figure 1. Machined specimens before testing.

B. Standard testing procedures

Both uniaxial and three-point loading tests were conducted on an electric Instron load frame with a 10 KN capacity. The

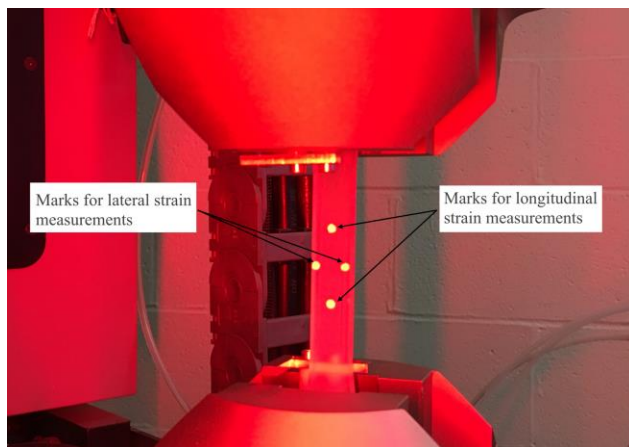


Figure 2. Dog-bone specimen marking and test setup.

load frame is equipped with a high precision non-contact

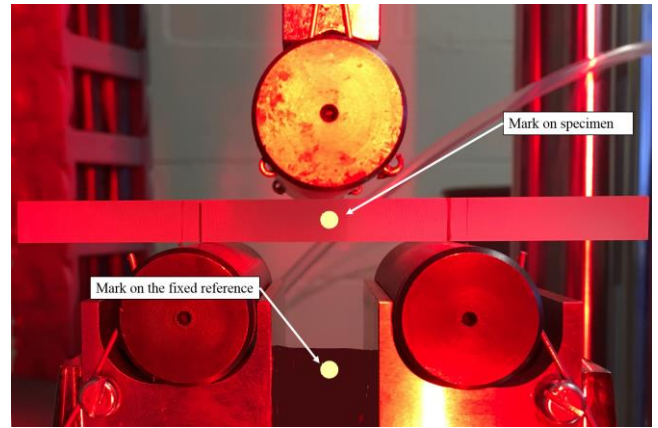


Figure 3. Prismatic specimen marking and test-setup.

digital image correlation (DIC) system for strain measurements. The DIC system has a resolution of $0.5\mu\text{m} \pm 1\%$. Each dog-bone specimen was precisely marked with two circles along the longitudinal and lateral specimens axes as shown in Figure 2. Longitudinal markings were placed as per the recommended gage length of the testing standard. DIC system used these marking to measure longitudinal strains. The lateral markings were used to evaluate the Poisson's effect. The lower grip was fixed while a displacement control was applied on the upper grip at a fixed rate of 1mm/min. All specimens were loaded until fracture. The uniaxial stress-strain results were used to evaluate the modulus of elasticity and Poisson's ratio.

Similarly, three-point loading specimen's marking and setup are presented in Figure 3. One mark was placed at mid-span of the specimen while the other was placed on a fixed frame of reference. Standard metric rollers were used as recommended by the testing standard. The relative deflection was used to generate load-deflection curves. The lower rollers were fixed while a displacement control was applied on the upper roller at a fixed rate of 1mm/min. All specimens were loaded until final failure.

C. Efficient identification of yielding

A glass-like behavior characterizes heavily cross-linked epoxy due to low strain-to-failure capacity. Identifying the onset of yielding for such brittle material is quite difficult. In the current work, an efficient method for yield identification from testing is proposed. Neat epoxy has a near transparent color. This color change into white color upon plastically deforming. A close definition to the yield point on a stress-strain curve would be the point at which non-linear elastic behavior ends and the plastic deformation is begun. Upon which the strains are no longer purely elastic and residual inelastic strains will not be recovered. Those unrecovered strain components cause what is known as "stress-whitening." This phenomenon is commonly observed in almost all thermosets. Testing procedures were video recorded synchronically with DIC strain measurements. Recorded videos were analyzed at different time frames to determine the first event of stress-whitening. The identified time frame was utilized to determine the corresponding stress level at the exact time.

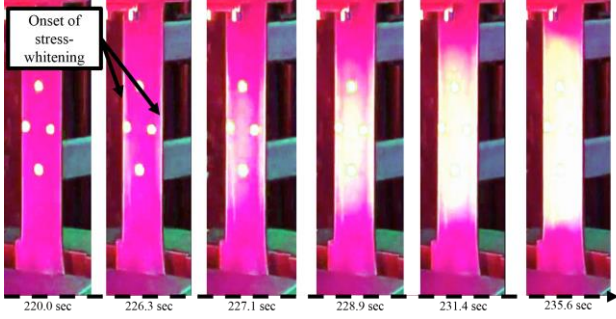


Figure 4. Images from synchronized video recording showing onset of plastic deformation.

This stress level surely corresponds to the onset of yielding. Therefore, yield stress was efficiently determined from testing rather than relying on a method of evaluation (e.g. 0.2% offset). For illustration purposes, different images at different time frames showing the gradual plastic deformation are portrayed in Figure 4. Images at different time frames were filtered to demonstrate the stress-whitening effect better.

III. TESTING RESULTS

In FRPs the resin material (i.e., epoxy for current study) are subject to different types of loading. Neat epoxy was tested under two types of loading, namely uniaxial tension and three-point bending loads. Both testing results were used to assess the behavior of neat epoxy as a commonly used resin material with the objective of attaining a grasp understanding of failure mechanism. This can possibly lead to a significant enhancement in failure predictions of resin materials in FRPs.

Uniaxial tensile testing results are documented in terms of stress-strain curves as shown in Figure 5. Three-point bending test results are presented by load-deflection curves as presented in Figure 7. An insight on failure mechanisms was enabled through optical microscopic fractographic analysis.

A. Uniaxial tension testing results

Following the testing standard, the minimum number of specimens were used to characterize neat epoxy resin. Mechanical properties were characterized using the mean values of all tested specimens. As can be observed from stress-

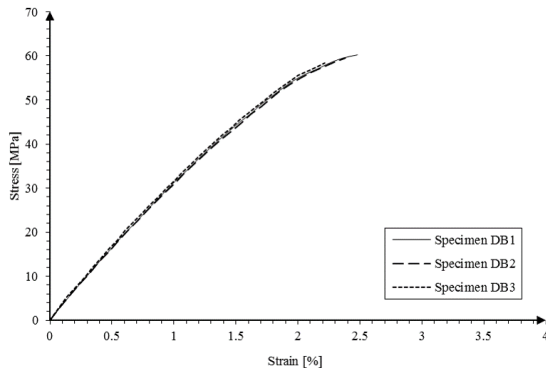


Figure 5. Stress-strain results from uniaxial testing.

strain curves, Figure 5, the material behavior is almost the same with a slight change towards failure limits. Stress-strain results yielded a modulus of elasticity and Poisson's ratio of 3.301GPa and 0.36, respectively. The mean value of fracture energies was evaluated as 2.054N.m. Specimens DB1 and DB3, respectively recorded the maximum and the minimum failure stresses and strains. Table I documents the specific failure limits and fracture energy of each specimen.

TABLE I. UNIAXIAL TENSION TESTING FAILURE LIMITS

Specimen	Failure limits		
	Fracture energy [N.m]	Failure stress [MPa]	Failure strain [%]
DB1	2.114	60.223	2.480
DB2	2.108	59.520	2.383
DB3	1.941	58.284	2.217

As can be observed from stress-strain curves, all specimens held a linear-elastic relation just before undergoing in minor plastic deformation. The hardening behavior was almost inexistent. Finally, specimens experienced sudden failure at deficient strains, i.e. less than 2.5%.

Initial inspection of fractured specimens indicated that neat epoxy fails in a brittle manner. This can be observed by a fracture surface that is normal to the load application direction. As shown in Figure 6, all three specimens had fractured surfaces normal to the load application direction signifying that final fracture was dominated by brittleness even if specimens exhibited slight plastic deformation ahead of failure. Also, stress-whitening caused by plastic deformation can be observed. The whitening is uniform throughout specimens' narrow section with the tendency of vanishing towards the rounded edges. Evidence indicates that standard specimens' preparation was followed and as a result, only the narrow section was affected by plasticity. Closer inspection to fractured surfaces is provided in the fractographic analysis section.



Figure 6. Dog-bone specimens after failure.

B. Three-point bending testing results

The three-point bending test was used to assess the behavior of neat epoxy under the combined state of stresses. Moreover, the mid-span of a prism under three-point loading is experiencing both maximum normal and shear stresses.

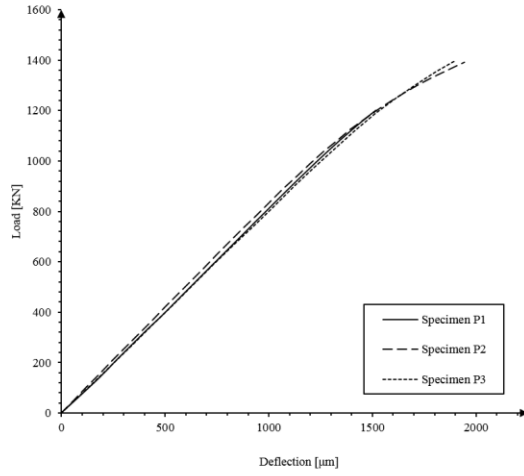


Figure 7. Load-deflection curves from three-point bending tests.

These conditions may represent severe loading conditions which may help to better investigate failure of heavily cross-linked epoxy. The load-deflection results are shown in Figure 7 where deflections were measured in μm . As can be observed the three specimens have a very similar behavior in the linear region. Regarding failure limits the minimum and the maximum deflections we recorded by specimens P1 and P3, respectively. The mean value of fracture energy was evaluated as 1.8N.m . Failure limits details are documented in Table II. Regarding fracture energy, results are in very good agreement with characterization results. Failure loads and deflections results were moderately varied. Also, a minimal plastic plateau can be observed in the behavior of specimen P2 and P3. This signifies that specimen P1 was the one with least plastic deformation before exhibiting final failure.

TABLE II. THREE-POINT BENDING TESTING FAILURE LIMITS

Specimen	Failure limits		
	Fracture energy [N.m]	Failure load [N]	Failure δ [μm]
P1	1.686	1200	1516
P2	1.871	1392	1944
P3	1.911	1511	2190

Initial inspection of fractured prisms showed that failure initiated in the tension side. Fractured specimens are presented in Figure 8. Failed surfaces were not perfectly flat, but in general, they were approximately normal to the maximum

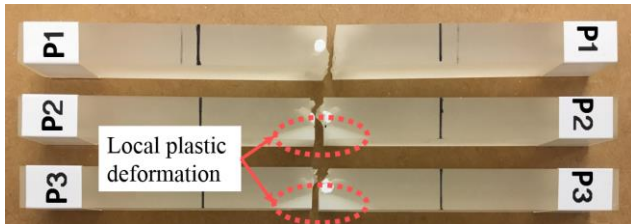


Figure 8. Failed prismatic specimens showing local plastic deformation.

principal stress direction. Noteworthy to mention, two specimens showed higher intensity of stress-whitening, namely P2 and P3. In fact, these specimens recorded relatively higher failure limits than specimen P1. This can be signified by undergoing slightly higher plastic deformations. Also, it can be observed that these plastic deformations were localized. Mid-span of specimens exhibited this local deformation, i.e. tension side. It seems that the final failure of specimens was a combination of two types of failure. Approximately 40% of the depth plastically deformed while the rest of the depth exhibited almost pure brittle failure. The fractographic analysis is provided in the following section to understand the failure mechanism better.

IV. FRACTOGRAPHY

In the current section, a closer insight on the failed surfaces from two types of conducted testing is provided. First, are the dog-bone specimens from uniaxial testing. Specimen DB2 is used as an illustrative example on the epoxy behavior under uniaxial tension. The magnified cross-sectional image of DB2 fractured surface is presented in Figure 9. As can be observed, cracks seem to have an origin or an onset location. This location is highlighted by a red dashed line on figure. In fact this agreeing with the literature findings micro-cavitation may lead to final failure. Also, the cracking crazes originating from this cavitation signifies that brittleness dominated the failure



Figure 9. Fractured surface from specimen DB2 showing crazes of cracking.

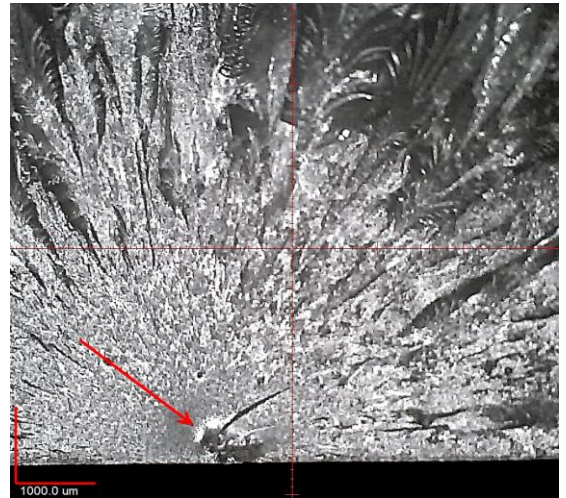


Figure 10. Zoomed in image in the crack onset vicinity for DB2.

mechanism while necking behavior did not exist. A zoomed in image is provided for the vicinity of the crack onset location in Figure 10. The onset of crazing is much clearer in the zoomed in image. A minor transition region between nucleation and crazes of cracking can be observed. Undoubtedly, brittleness dominated the fracture mechanism of epoxy under uniaxial loading scenario with very slight plastic deformation ahead of fracture.

On the other hand, the three-point loading scenario was a little bit different. Specimen P1 and P3 are taken as the extreme cases for illustration purposes. Figure 11 presents a schematic center line for the zoomed microscopic images. First and second images presents fractured surfaces of specimens P1 and P3, respectively. As can be observed, both specimens showed crack initiation in the lower segment (i.e. tension side) of the prism. The exact segment that is undergoing maximum normal and shear stresses. While specimen P1 exhibited slight local plastic deformation, specimen P3 showed a plastic deformation region that is approximately more than 40% of the prism depth. Also, crazes of brittle cracking seems to be radially originating from towards the center of lower edge in specimen P1. On the other hand, specimen P3 seems to have locus region for the initiation of cracks. Cracking direction seems to be inclined on the horizontal axis of the cross-section. Yet, final failure is still dominated by brittleness. Noteworthy to mention that micro-cavitation average diameter was in the order of 70 μ m. From conducted fractographic analysis it seems that while plastic deformations minimally/locally exists, brittle behavior is the one dominating final failure. This was observed in combined stress state that of a prism under three-point loading and was obvious in case of uniaxial loading.

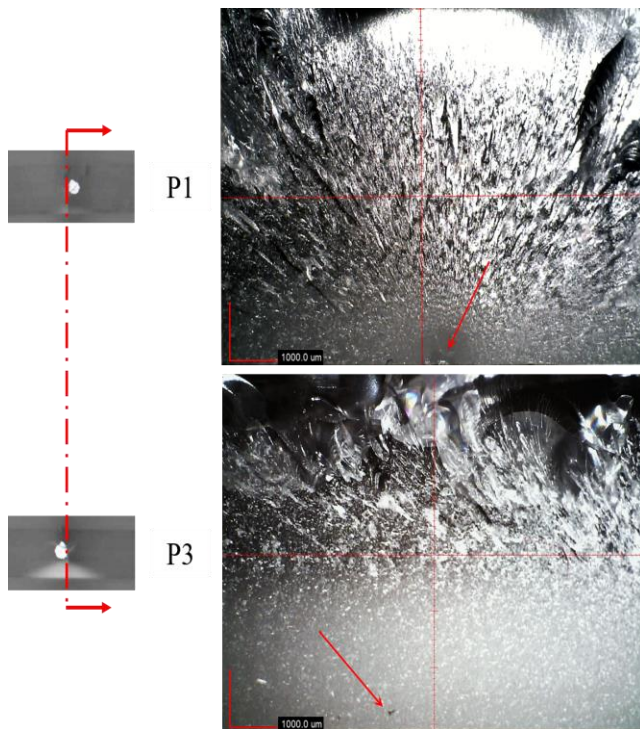


Figure 11. Zoomed in images for fractured surfaces of specimens P1 and P3.

V. CONCLUSIONS

In part I of this work, neat heavily cross-linked epoxy was prepared and test under two type of loading, uniaxial tension and three-point loads. The objective of part I was to characterize the neat epoxy resin plate by testing using a high precision non-contact strain measurement system utilizing DIC. Also, the fractographic analysis was provided for all failed specimens. Fractured surfaces were examined using optical microscopic analysis. A new method was proposed for efficient yield stress identification. The method utilizes synchronized video recording with strain measurements that are later used for image analysis to identify the onset of plastic deformation using stress-whitening phenomenon. Agreeing with literature findings neat epoxy resin exhibited brittle final failure under both uniaxial as well as three-point loads. As elaborated in this part, minimal/local plastic deformations existed yet, none of the dog-bone specimens showed a necking type behavior. It can be concluded that this work provided a better understanding of failure mechanisms of neat epoxy resins used FRPs and in layup laminations. In part II, multiscale modeling of the same tested material is provided.

ACKNOWLEDGMENT

Authors would like to acknowledge Polynt Composites Canada, Inc. for providing the neat epoxy resin slab.

REFERENCES

- [1] T.K.U. Uygunoglu, I.K.U. Gunes, W. of N.T. Brostov, physical and Mechanical properties of polymer composite with high content, *Mater. Res.* 18 (2015) 1188–1196. doi:10.1590/1516-1439.009815.
- [2] P.K. Mallick, *Fiber-Reinforced Composites: Materials, Manufacturing, and Design*, 2007.
- [3] S. Budhe, M.D. Banea, S. de Barros, L.F.M. da Silva, An updated review of adhesively bonded joints in composite materials, *Int. J. Adhes. Adhes.* 72 (2017) 30–42. doi:10.1016/j.ijadhadh.2016.10.010.
- [4] D. Pulungan, G. Lubineau, A. Yudhanto, R. Yaldiz, W. Schijve, Identifying design parameters controlling damage behaviors of continuous fiber-reinforced thermoplastic composites using micromechanics as a virtual testing tool, *Int. J. Solids Struct.* 117 (2017). doi:10.1016/j.ijsolstr.2017.03.026.
- [5] T. Yokozeki, Y. Iwahori, S. Ishiwata, Matrix cracking behaviors in carbon fiber/epoxy laminates filled with cup-stacked carbon nanotubes (CSCNTs), *Compos. Part A Appl. Sci. Manuf.* 38 (2007) 917–924. doi:10.1016/j.compositesa.2006.07.005.
- [6] L. Di Landro, A. Montalto, P. Bettini, S. Guerra, F. Montagnoli, M. Rigamonti, Detection of voids in carbon/epoxy laminates and their influence on mechanical properties, *Polym. Polym. Compos.* (2017).
- [7] Z. Wang, L. Ma, L. Wu, H. Yu, Numerical simulation of crack growth in brittle matrix of particle reinforced composites using the xfem technique, *Acta Mech. Solida Sin.* 25 (2012) 9–21. doi:10.1016/S0894-9166(12)60002-0.
- [8] N.H. Kim, H.S. Kim, Micro-void toughening of thermosets and its mechanism, *J. Appl. Polym. Sci.* 98 (2005) 1290–1295. doi:10.1002/app.22262.
- [9] O.F.J.T. Bressers, Craze initiation in glassy polymer systems, (n.d.).
- [10] B. Joseph, F.W. Hanratty, J.L. Kardos, Model-Based Control of Voids and Product Thickness during Autoclave Curing of Carbon/Epoxy Composite Laminates, *J. Compos. Mater.* 29 (1995) 1000–1024. doi:10.1177/002199839502900801.
- [11] R. Talreja, Assessment of the fundamentals of failure theories for

- composite materials, *Compos. Sci. Technol.* 105 (2014) 190–201. doi:10.1016/j.compscitech.2014.10.014.
- [12] J. Chevalier, X.P. Morelle, C. Bailly, P.P. Camanho, T. Pardoen, F. Lani, Micro-mechanics based pressure dependent failure model for highly cross-linked epoxy resins, *Eng. Fract. Mech.* 158 (2016) 1–12. doi:10.1016/j.engfracmech.2016.02.039.
- [13] L.E. Asp, L.A. Berglund, P. Gudmundson, Effects of a composite-like stress state on the fracture of epoxies, *Compos. Sci. Technol.* 53 (1995) 27–37. doi:10.1016/0266-3538(94)00075-1.
- [14] L.E. Asp, L.A. Berglund, R. Talreja, A criterion for crack initiation in glassy polymers subjected to a composite-like stress state, *Compos. Sci. Technol.* 56 (1996) 1291–1301. doi:10.1016/S0266-3538(96)00090-5.
- [15] V. Tvergaard, Influence of voids on shear band instabilities under plane strain conditions, *Int. J. Fract.* 17 (1981) 389–407. doi:10.1007/BF00036191.
- [16] Z. Yosibash, A. Bussiba, I. Gilad, Failure criteria for brittle elastic materials, *Int. J. Fract.* (2004) 307–333. doi:10.1007/978-1-4614-1508-4.
- [17] K. Matouš, M.G.D. Geers, V.G. Kouznetsova, A. Gillman, A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials, *J. Comput. Phys.* 330 (2017) 192–220. doi:10.1016/j.jcp.2016.10.070.
- [18] D.J. Mortell, D.A. Tanner, C.T. McCarthy, In-situ SEM study of transverse cracking and delamination in laminated composite materials, *Compos. Sci. Technol.* 105 (2014). doi:10.1016/j.compscitech.2014.10.012.
- [19] F. Lachaud, C. Espinosa, L. Michel, P. Rahme, R. Piquet, Modelling Strategies for Simulating Delamination and Matrix Cracking in Composite Laminates, *Appl. Compos. Mater.* 22 (2015) 377–403. doi:10.1007/s10443-014-9413-4.
- [20] H. Pollayi, W. Yu, Modeling matrix cracking in composite rotor blades within VABS framework, *Compos. Struct.* 110 (2014) 62–76. doi:10.1016/j.compstruct.2013.11.012.
- [21] L. Riaño, L. Belec, Y. Joliff, Validation of a Representative Volume Element for unidirectional fiber-reinforced composites: Case of a monotonic traction in its cross section, *Compos. Struct.* 154 (2016) 11–16. doi:10.1016/j.compstruct.2016.07.020.

Performance of Metals in Space: A molecular dynamics simulation of atomic oxygen impacts on silver

Liam S. Morrissey, Stephen M. Handrigan, Sam Nakhla¹

Department of Mechanical Engineering
Memorial University of Newfoundland
St. John's, Canada

Ali Rahnamoun

Centre for Sustainable Nanotechnology
Johns Hopkins University
Baltimore, United States of America

¹ Department of Mechanical Engineering/ Department of Emergency Medicine

I. Abstract— Atomic oxygen (AO) impacts on spacecraft materials can be a source of significant erosion and material performance degradation in the low Earth orbit (LEO). Moreover, testing in the LEO is difficult and cost prohibitive. As a result the purpose of this study was to use the ReaxFF potential to simulate the AO degradation of metals commonly used on space crafts. These simulations studied both the erosion and temperature evolution of the slab as a function of the number of impacting oxygen atoms. Overall, after impact with 100 oxygen atoms the erosion coefficient of silver closely matched previously reported results for erosion in LEO. These results clearly demonstrate the potential of ReaxFF as a cost effective method to simulate test conditions in the LEO.

Keywords- *molecular dynamics, atomic oxygen, erosion*

II. INTRODUCTION

A. Background

Due to the constant uncertainty in the budgets for the world's major space agencies there is a need for more cost effective methods of testing materials in the space environment. One such option is using simulations to test performance, degradation and remaining life of various materials that are typically used in space. However, the degradation of a material's performance often begins far before any macroscale visible flaws are introduced. Therefore, simulations must be able to capture the atomic scale interactions that preclude performance degradation and, ultimately, failure. Atomic oxygen (AO) impacts on spacecraft materials represent one of the biggest threats to material performance in low Earth orbit (LEO). Within Earth's atmosphere oxygen typically exists as the O₂ molecule, however, the high radiation levels in space separate these molecules and produce significant amounts of AO. Collisions with this AO can reach energies as high as 4.5 eV due to the high relative speeds between orbiting space crafts and the atmospheric oxygen [1, 2]. These high energy collisions can, depending on the material, result in material loss through erosion and a rapid decline in instrumentation performance

over longer durations. Overall, significant in-situ measurements have been conducted to measure the effect of AO on a wide range of materials [3, 4]. Results show that while AO is reactive with many polymers, its effect on metals varies significantly. For example, silver and gold are two of the most commonly found metals on space crafts. Silver is typically used for electrical components while gold is often used as a coating due to its ability to absorb visible light and not reflect. Tests conducted in LEO have shown that silver is highly reactive with AO whereas little to no erosion was observed for gold [3, 4]. Therefore, due to the large differences in performance that can exist amongst various materials, they must each be tested in LEO for their performance prior to application. However, due to the high cost of space launches these tests are not always feasible. Moreover, due to the atomistic scale of the collisions, these studies are only able to observe the eroding effects of AO and do not actually observe the effect of each individual impact. A way to more accurately simulate the effect of impact of AO on various materials could enhance both the accuracy of remaining life predictions and alleviate the costly need to test new materials in LEO. It is important to note that as opposed to simulations or in situ testing, AO testing facilities are another viable alternative. However, they too are costly, require equipment that is often not readily available, and cannot always replicate the harsh and complex space environment.

One, more recent, alternative is molecular dynamics (MD) simulations. MD has been used since the early 1960s to simulate and, depending on the force field used, model the bond breaking and formation during chemical reactions. Therefore, MD provides an exciting option to actually model the atomic interactions during AO impact, as opposed to simply viewing the macroscale surface after the collisions have occurred. As an example, studies by Rahnamoun and van Duin and Zeng et al. [5, 6] have used MD to model the effect of AO on several commonly used polymers in the space environment. First, Rahnamoun and van Duin used the ReaxFF force field to study the effect of AO impacts on

Kapton, POSS polyamide, and amorphous silica [5]. In this study AO was propelled towards a small substrate of each material and the subsequent mass loss, temperature growth, and erosion coefficients were calculated and compared to experimental values. Results showed good agreement with experimental data and concluded that the temperature evolution on the surface was critical to predicting whether erosion and material loss would occur [5]. Similar to this study, Zeng et al. also used the ReaxFF forcefield to study the breakdown and disintegration of two commonly used spacecraft polymers, PVDF and FP-POSS [6]. Again, small simulated substrates were impacted with AO until material loss occurred. Results showed that PVDF continuously eroded due to AO while FP-POSS did not erode until after a specific number of AO impacts [6]. Together, these two studies demonstrate the potential for MD simulations to predict the effect of AO on spacecraft polymers.

However, while research has been conducted on the MD simulation of the AO erosion of polymers, the effect of AO on spacecraft metals has yet to be simulated. Further, the crystalline and ordered bonding of metals as compared to polymers means that accuracy for polymers cannot be used to assume accuracy for metals as well. Moreover, because molecular dynamics is force field dependent, the force fields used to simulate metals will be drastically different than the polymer based force fields that consist mainly of carbon, hydrogen and oxygen. Force fields are also sensitive to the partial charges of the atoms in the simulation model. However, for metals this charge is not constant and is affected by factors like environmental condition and oxidation state of the metal. In addition, the outermost orbital for metals (d- or f-) can cause more complicated chemical bonding characteristics. All of these factors increase the challenge of simulating systems that include metals.

B. Purpose of Study

Overall, many spacecraft metals are assumed to be non-reactive with AO. However, this conclusion is only applicable to the specific energy found in LEO collisions. There may exist an energy barrier, that, with a harsh enough environment, fast enough collisions, or significant exposure, could be overcome. For example, as space missions continue to move further into the solar system there becomes an increased risk of significantly higher speed collisions with both interstellar dust and heavy atoms [7]. As a result, these AO ‘immune’ metals must be tested again to determine their performance with even higher energy collisions. Testing these situations in-situ is extremely difficult and MD simulations represent a much more viable alternative. However, before this can be done, the accuracy of using molecular dynamics for high-speed collisions with metals must first be verified by comparing with readily available experimental data from LEO.

As a result, the purpose of this paper was to investigate whether MD simulations can be used to model the effect of repeated AO impact on silver in the space environment. Silver

was chosen as it is one of the most commonly used metals in space and is known to be highly reactive to AO. Simulation results will then be compared to previously reported test data in the LEO to determine the accuracy of this method.

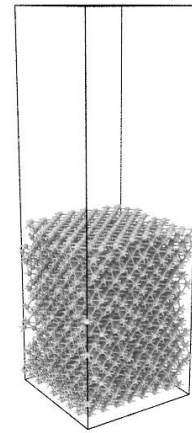


Fig. 1 – FCC silver substrate equilibrated to 200K prior to impacts with atomic oxygen

III. PROCEDURES

A. Molecular Dynamics

MD simulates the movement of atoms and molecules within a body. A typical simulation begins with known initial positions and velocities of all atoms within the system. The atomic accelerations are then calculated using an interatomic force field that defines the forces between interacting particles. With the atomic acceleration known, Newton’s equations of motion are then used to predict subsequent positions and velocities. Therefore, the interatomic force field is a critical component to any MD simulation. However, what makes the AO simulation unique is that it requires a force field that can describe the bond breaking and formation during repeated impacts. As a result, several traditional force fields are not applicable as they do not consider the breakage of bonds or reactions. In contrast, the reactive force field (ReaxFF) is a bond order based force field method that allows for the dynamic simulation of bond breaking and reformation in a body. In the ReaxFF force field, the energy of the system is calculated as a combination of the partial energy contributions from the bond, over- and under-coordination, lone pair, valence, torsion, van der Waals, and Coulomb energy [8] (Equation 1):

$$E_{system} = E_{bond} + E_{over} + E_{under} + E_{lp} + E_{val} + E_{tor} + E_{vdWaals} + E_{Coul} \quad (1)$$

Bond order is then used to determine the interactions between all atoms in the system. The bond order accounts for contributions from sigma, pi, and double pi-bonds as a

continuous function of the distance between atoms via equation 2:

$$BO_{ij}^I = BO_{ij}^S + BO_{ij}^T + BO_{ij}^{\pi} \quad (2)$$

B. Substrate Preparation

First, a 32x32x40 Å slab of pure FCC silver was developed using molecular dynamics (Figure 1). After preparing the slab it was then placed in a simulation box that was 100 Å high with fixed boundary conditions along its height (z direction) and periodic boundary conditions in the x and y. Therefore, the slab had an open surface in the z direction. Next, to obtain a realistic temperature, a NVT thermostat was used to equilibrate the slab in the periodic box to 200K. This thermostat kept the number of atoms, and volume of the system constant while changing the pressure to reach the desired temperature.

C. Oxygen Impact

After the system was stabilized to 200K it was then ready for impact with AO. Following previous work, at every 200 fs (400 simulation steps), one AO was deposited into the system with a speed of -7.4 km/s (4.5 eV) in the z-direction at a randomly generated x- and y-position [5, 6]. During deposition and subsequent impact, the thermostat was removed and an NVE (constant number of atoms; N, Volume; V, and Energy; E) simulation was run to fix the number and volume of the system while allowing the temperate to increase. In total, 100 oxygen atoms were emitted during the first 40000 time steps for a total simulation time of 20 ps. Table 1 summarizes the parameters used for the molecular dynamics model.

Table 1: LAMMPS input parameters for simulation of atomic oxygen impacts on a steel substrate

	Silver
Lattice Spacing (Å)	4.08 Å
Timestep (fs)	0.5
Density (g/cm ³)	10.5
Potential	Lloyd 2016
AO Impacts	100
AO Speed (km/s)	7.4
Atomic Mass (amu)	107.9

Erosion rate of the substrate was tracked using a cut-off distance from the original surface that was defined based on a multiple of the maximum bonding distance as per a similar study by Yuan et al. [9]. However, while the maximum bond distance of Ag-Ag bonding is approximately 3Å, it is important to also consider the potential for AO adsorption and bonding onto the surface of the substrate due to high

temperatures [10]. In this case, typical Ag-O bond length is approximately 2Å [10]. Therefore, the maximum distance from the substrate was calculated as twice the sum of the Ag-Ag and Ag-O bond length (i.e., 10Å). This value is similar to the 9Å cut-off used by Rahmani et al. to define erosion during AO impact [11]. Any substrate atoms further than 10Å from the original surface were considered eroded from the surface.

IV. RESULTS AND DISCUSSION

A. Erosion due to Atomic Oxygen

Referring to Figure 2, the equilibrated substrate begins as well ordered FCC silver. However, upon impact damage begins to propagate throughout the depth of the slab. As impact continues atoms begin to move beyond the cut-off distance and are hence no longer bonded to the slab and considered as eroded. This cut-off was used to determine the number of eroded silver atoms as a function of the number of AO impacts. The silver substrate had only one atom eroded from the substrate beyond the cut-off distance after the first 57 AO impacts. However, after reaching this point, the last 43 AO impacts caused a loss of a further 64 silver atoms from the substrate. Silver atoms were eroded as both metallic species and oxides. Therefore, there appears to be a critical energy value that, when reached, leads to a significant increase in erosion rate. In total, the impact of 100 oxygen atom lead to the loss of 66 silver atoms from the substrate and a sputtering rate of 0.65 Ag atoms/Oxygen atom or 11.6×10^{-23} grams of Ag/Oxygen atom.

Previous studies on AO erosion of silver agree that silver has a high rate of erosion when compared with other metals. Research has demonstrated that the source of these different erosion rates is likely in the different free energy of oxide formation values for the various oxides. Free energy of oxide formation can be used to predict the energy required to break oxide bonds. Therefore, a lower free energy of oxide formation means less energy is needed to break oxide bonds, resulting in a higher erosion rate due to the cyclic process of oxide break-up and reformation. As an example, silver has a free energy of formation of 0.32 eV, whereas metals like iron and nickel are 2.79 and 2.53, respectively [12]. Moreover, silver oxides also have a high tendency for spalling from the silver substrate, meaning the oxide is not stable or protective from further AO erosion [12]. Following this logic, at the end of the simulation, several of the silver atoms that have eroded from the surface appear to be bonded with oxygen atoms, likely demonstrating the initial spalling of the oxide later. As a result, while silver is a crucial metal for several spacecraft mechanisms, constant monitoring, short flight durations, or coatings must be used.

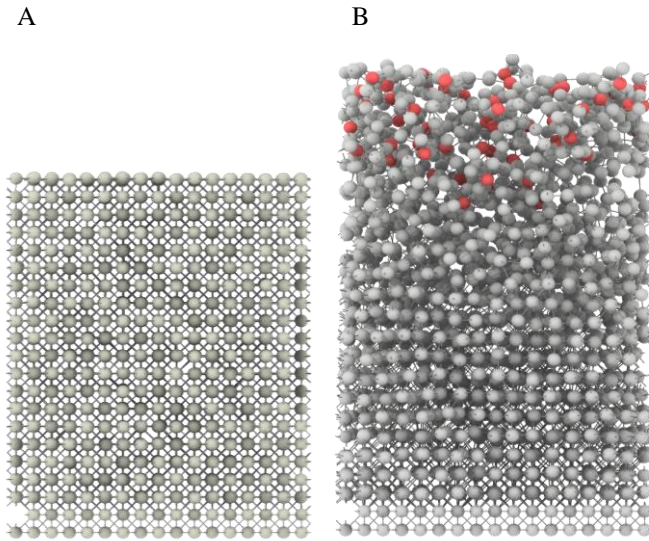


Fig. 2 – FCC silver substrate and lattice before (A) and after (B) atomic oxygen impact. Grey spheres are silver and red spheres are oxygen.

B. Comparison with Test Data

Next, to comment on the accuracy of the MD simulation method simulation results were compared with erosion yield observations from flight tests of silver in LEO. Referring to Banks et al. [3], the erosion yield for silver was measured to be $10.5 \times 10^{-24} \text{ cm}^3/\text{oxygen atom}$. To allow for an accurate comparison, this value was converted to a mass loss per oxygen atom by multiplying by the density of silver ($11 \times 10^{-23} \text{ cm}^3/\text{atom}$). Therefore, the silver erosion rate predicted by the MD model showed strong agreement with experimental data and had a relative error of only 5% ($11.6 \times 10^{-23} \text{ g/oxygen atom}$ and $11 \times 10^{-23} \text{ cm}^3/\text{oxygen atom}$, respectively). As a result, not only do the MD simulations capture the expected trend of erosion with AO impact, but the predicted erosion rate agrees well with test data.

C. Temperature of Substrate

While test data simply tracks the reduced weight as a function of atomic impact, MD simulations allow for a deeper investigation into the atomistic processes at play that may precipitate erosion. Therefore, the final step of this study was to study the temperature evolution of the substrate during impact. First, prior to impact the silver substrate was equilibrated to 200K to mimic space conditions. During impact the thermostats were removed and the substrate was allowed to heat up. Referring to Figure 3A, the silver substrate temperature increased almost linearly with AO impact to a final temperature of 1700K. Moreover, there was a large temperature spike for silver at approximately 60 AO impacts. When investigating further, it can be seen that this spike

corresponds to the beginning of rapid substrate erosion (Figure 3B). Therefore, this indicates that an energy barrier must first be reached prior to erosion.

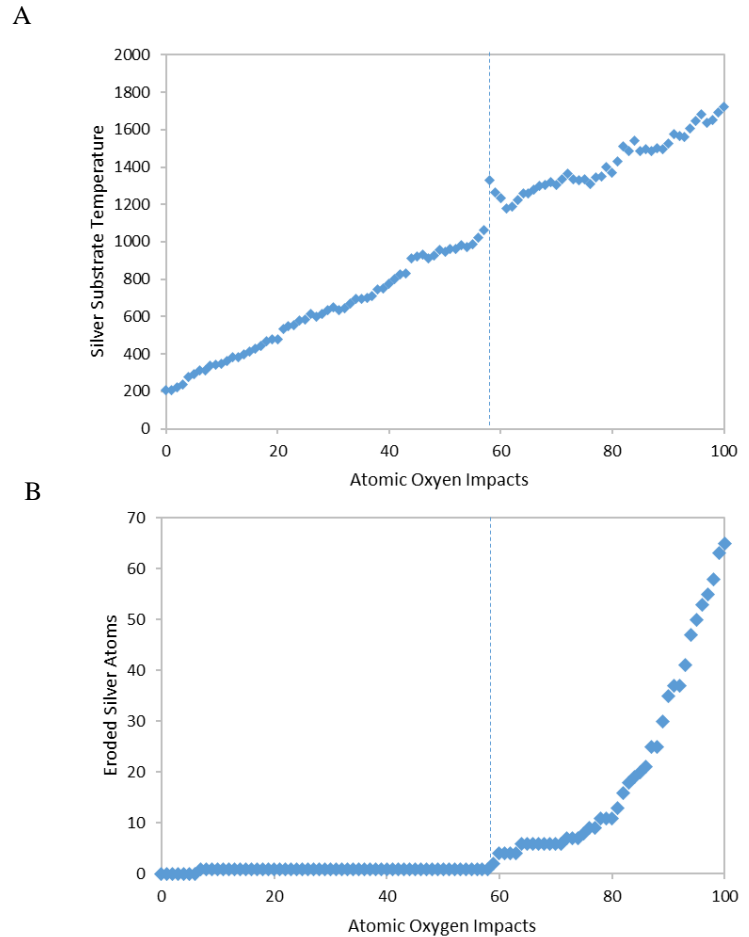


Fig. 3 – Silver substrate temperature as a function of atomic oxygen impacts (A) and number of eroded silver atoms as a function of atomic oxygen impacts (B). The dashed line indicates a spike in both temperature and number of atomic oxygen impacts

It is important to also consider the temperature evolution of the nanoscale substrate and how this corresponds to bulk temperatures and melting points. First, upon initial inspection it appears that the substrate temperature exceeds the melting point of silver, approximately 1200K. However, melting points predicted from atomistic simulations exceed the actual macroscale value [13]. This can be seen in Figure 2B where after impact there is still bonding in the substrate. Moreover, previous MD simulations of AO erosion also exceeded the macroscale melting points of the substrate without actually observing melting [5, 6]. To verify this another simulation was run where the substrate was raised to the maximum temperature observed and equilibrated for 15000 time steps. Referring to Figure 4, the substrate remains almost completely FCC bonded at this temperature, with only the fixed bottom

layer and a small exposed top layer being non-FCC. This finding was also confirmed through a fragment analysis to detect any non bonded atoms or molecules. Again, before and after equilibration at 1700K the substrate contained 2688 bonded silver atoms with no fragments. Therefore, while the extreme temperature evolution caused by the AO impacts is certainly important in making the substrate susceptible to damage and erosion, it does not induce a melted substrate.

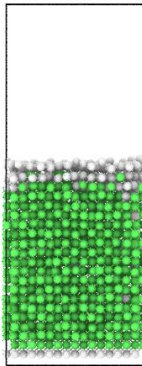


Fig. 4 – Silver substrate bonding after equilibration to 1700K. Green atoms are FCC bonded, and top and bottom layers are non FCC.

I. CONCLUSIONS

In summary, this study used molecular dynamics to simulate the effect of AO impacts on silver in space. While research has shown that these methods are accurate for polymers, due to the different crystal structure of metals and the necessity for changing the force field, it is important to also test the accuracy with spacecraft metals. Results demonstrated that impacting a silver substrate with 100 AO atoms resulted in a breakup of the metal and an erosion coefficient that matched data from LEO. These findings demonstrate that the performance of metals in LEO can be simulated using molecular dynamics as opposed to traditional high-cost space environment tests. Therefore, as new metals and materials are developed, performance in the LEO environment can easily be checked prior to further space testing. Moreover, as missions move to deeper space where in-situ testing is extremely difficult and where even higher energy impacts occur, the potential for molecular dynamics to simulate these complex environments cannot be ignored.

REFERENCES

1. Banks, B.A., S.K. Rutledge, P.E. Paulsen, and T.J. Steuber, *Simulation of the low earth orbital atomic oxygen interaction with materials by means of an oxygen ion beam*. 1989.
2. Zhang, J. and T.K. Minton, *Production of volatile CO and CO₂ from oxidized polyethylene and graphite surfaces by hyperthermal atom-surface collisions*. High Performance Polymers, 2001. **13**(3): p. S467-S482.
3. Banks, B.A., S.K. Rutledge, J.A. Brady, and J.E. Merrow, *Atomic oxygen effects on materials*. 1989.
4. Samwel, S.W., *Low earth orbital atomic oxygen erosion effect on spacecraft materials*. Space Research Journal, 2014. **7**(1): p. 1-13.
5. Rahnamoun, A. and A.C.T. van Duin, *Reactive Molecular Dynamics Simulation on the Disintegration of Kapton, POSS Polyimide, Amorphous Silica, and Teflon during Atomic Oxygen Impact Using the Reaxff Reactive Force-Field Method*. The Journal of Physical Chemistry A, 2014. **118**(15): p. 2780-2787.
6. Zeng, F., C. Peng, Y. Liu, and J. Qu, *Reactive Molecular Dynamics Simulations on the Disintegration of PVDF, FP-POSS, and Their Composite during Atomic Oxygen Impact*. The Journal of Physical Chemistry A, 2015. **119**(30): p. 8359-8368.
7. Srama, R., T.J. Ahrens, N. Altobelli, S. Auer, J.G. Bradley, M. Burton, V.V. Dikarev, T. Economou, H. Fechtig, M. Görlich, M. Grande, A. Graps, E. Grün, O. Havnes, S. Helfert, M. Horanyi, E. Igenbergs, E.K. Jessberger, T.V. Johnson, S. Kempf, A.V. Krivov, H. Krüger, A. Mocker-Ahlreep, G. Moragas-Klostermeyer, P. Lamy, M. Landgraf, D. Linkert, G. Linkert, F. Lura, J.A.M. McDonnell, D. Möhlmann, G.E. Morfill, M. Müller, M. Roy, G. Schäfer, G. Schlotzhauer, G.H. Schwehm, F. Spahn, M. Stübig, J. Svestka, V. Tschernjawski, A.J. Tuzzolino, R. Wäsch, and H.A. Zook, *The Cassini Cosmic Dust Analyzer*, in *The Cassini-Huygens Mission: Orbiter In Situ Investigations Volume 2*, C.T. Russell, Editor. 2004, Springer Netherlands: Dordrecht. p. 465-518.
8. van Duin, A.C.T., S. Dasgupta, F. Lorant, and W.A. Goddard, *ReaxFF: A Reactive Force Field for Hydrocarbons*. The Journal of Physical Chemistry A, 2001. **105**(41): p. 9396-9409.
9. Yuan, B., F. Zeng, C. Peng, and Y. Wang, *A reactive molecular dynamics simulation study to the disintegration of PVDF and its composite under the impact of a single silicon-oxygen cluster*. Computational Materials Science, 2018. **154**: p. 14-24.
10. Trushin, E.V. and I.L. Zilberberg, *Anion-radical oxygen centers in small (AgO)_n clusters: Density functional theory predictions*. Chemical Physics Letters, 2013. **560**: p. 37-41.

11. Rahmani, F., S. Nouranian, X. Li, and A. Al-Ostaz, *Reactive molecular simulation of the damage mitigation efficacy of POSS-, graphene-, and carbon nanotube-loaded polyimide coatings exposed to atomic oxygen bombardment*. ACS applied materials & interfaces, 2017. **9**(14): p. 12802-12811.
12. De Rooy, A. *The degradation of metal surfaces by atomic oxygen*. in *Proceedings of the 3ed European Symposium on Spacecraft Materials in a Space Environment*, ESA SP-232. 1985.
13. Dozhdikov, V. S., A. Yu Basharin, and P. R. Levashov. *Two-phase simulation of the crystalline silicon melting line at pressures from -1 to 3 GPa*. The Journal of chemical physics 137.5 (2012): 054502.

Molecular Dynamics Investigation on the Effect of Size and Temperature on Young's Modulus of Iron

Stephen M. Handrigan*, Liam S. Morrissey, and Sam Nakhla¹

¹. Faculty of Engineering and Applied Science (Mechanical) and Faculty of Medicine (Emergency Medicine)
 Faculty of Engineering and Applied Science (Mechanical)
 Memorial University of Newfoundland
 St. John's, Canada
 Corresponding author: smh113@mun.ca

Abstract—Molecular dynamics is utilized to investigate the effect of temperature on Young's modulus of single-crystal pure iron using two (2) many-body force fields – Tersoff and ReaxFF. First, an optimized simulation size is developed. Next, the effect of aspect ratio is investigated. Afterwards, a study on the effect of temperature on Young's modulus is performed. Results are compared to empirical equation from literature for single-crystal pure iron. Tersoff fails to capture trend while ReaxFF captures trend within average error of 11%.

Keywords- *molecular dynamics, Young's modulus, force field, temperature*

I. INTRODUCTION

Molecular dynamics (MD) is a useful tool to researchers for studying various interactions of atoms and molecules. More recently, it has become a useful tool for predicting mechanical properties of materials and has been used to study how atomistic defects [1]-[4], strain rate [5], and various other variables influence this prediction. Unfortunately, the popularization of MD to predict mechanical properties could lead to incorrect usage and selection of force fields, resulting in incorrect and erroneous predictions. Several papers utilizing MD to simulate, and compare to, macroscale experiments on the nanoscale or sub-nanoscale perform simulations at unrealizable temperatures – i.e., at 0K. Since it is well-known that temperature affects the Young's modulus, any MD results of mechanical properties performed at 0K and compared to room temperature are either unreliable or factually improbable. One justification for performing simulations at 0K could be that the force field of choice is temperature independent or was developed for 0K. As such, authors may believe that MD simulations can be performed at any temperature and the results are independent from temperature. The work of Morrissey et al. [6] demonstrated how significant the selection of force field influences the prediction in Young's modulus of pure iron. In addition, the work of Rajabpour et al. [7] also demonstrated how material properties of iron and iron-carbon are influenced by force field type. As such, it is critical to understand if MD simulations can accurately capture the effect of temperature on mechanical properties and clearly demonstrate the effect force field selection has on the results.

Another crucial aspect of a MD study is the simulation size – a representation of the number of atoms studied. As the simulation size increases, the accuracy in predictions become more accurate as values are averaged over more atoms. Unfortunately, computational demand drastically increases as a result. Hence, a balance between accuracy and computational efficiency must be determined. With many-body force fields, such as Tersoff [8] and Reactive Force Field (ReaxFF) [9], the degree to which computational demand increases is significant even compared to other many-body force fields, such as Embedded Atom Method (EAM) [10] and Modified Embedded Atom Method (MEAM) [11]. This is likely due to the inclusion of bond-order, but also ReaxFF's inclusion of breaking and creating bonds. As such, a study into which many-body force field provides the most accurate prediction in Young's modulus of iron with the most efficient simulation size must be completed. It was demonstrated in [6] that the Aryanpour et al. [12] ReaxFF parameter set was the most accurate for predicting elastic modulus, with the Henriksson et al. [13] Tersoff parameter set being a close second.

Once a sufficient simulation size is determined, many mechanical properties can be studied to determine the accuracy of the selected force field. Of key importance is the influence of temperature on the mechanical properties of iron. Expanding upon the work of Morrissey et al. [6], an investigation into the effect of temperature on Young's modulus of pure iron was performed. It is known that as temperature increases, the Young's modulus of a material decreases. For pure iron specifically, Fields [14] showed Young's modulus (GigaPascals) was a function of temperature in Equation (1).

$$E = 196 \times (1 - 2.1 \times 10^{-4}(T - 300) - 1.7 \times 10^{-7}(T - 300)^2 - 3.2 \times 10^{-10}(T - 300)^3) \quad (1)$$

Several studies have investigated the effect of size or temperature on the predictions of mechanical properties. Fang et al. [15] studied the effect temperature has on predicting the elastic modulus and hardness of pure copper undergoing diamond-tipped nano-indentation for temperatures ranging from 150K to 600K. Their results demonstrated that as temperature increased, both Young's modulus and hardness decreased. The effect of temperature on elastic modulus is not compared to experimental results or an empirical equation for the change in modulus with respect to temperature. However,

Fang et al. [15] do use MD, coupled with the Morse potential, to predict elastic moduli and hardness values which are consistent with that of pure copper. Similarly, the work of Chang [4] investigated the mechanical properties of pure copper with vacancies and various temperatures undergoing static- and cyclic-loading uniaxial tension. Chang [4] also demonstrated that as temperature increased, the elastic modulus decreases. Using a Lennard-Jones potential, Chang's [4] reported elastic modulus at 300K is approximately equal to reported values for the elastic modulus of copper. However, Chang [4] does not compare the effect of temperature on elastic modulus predicted in MD to macro-scale experimental results or an empirical equation in literature. The work of Zhang and Shen [16] demonstrated the effect temperature has on the elastic properties of single-walled carbon nanotubes undergoing axial compression and torsion. Using the reactive empirical bond order (REBO) potential, it is demonstrated that elastic modulus decreases as temperature is increased. Lastly, the use of MD for studying nanowires has become more prolific in recent years. Two studies, Koh et al. [17] and Wu [18], study platinum and copper nanowires, respectively, while observing the effects of both size and temperature on the predicted elastic modulus using different forms of potentials. Koh et al. [17] used pair functional and Wu [18] used an EAM. Both studies demonstrated that size has an effect on the predicted elastic modulus and that temperature also effects the elastic modulus. However, a similar trend is observed that results are not compared to experimental data or an empirical equation.

While it has been demonstrated in MD for only a handful of materials that force field, size, and temperature have an effect on the prediction of Young's modulus, no single study investigates the effect of all three on the prediction of Young's modulus. In addition, many studies lack comparisons to macro-scale experimental results, and as such, it must be determined if MD is capable of accurately predicting the Young's modulus as temperature increases for pure iron in comparison to data found in literature (Equation 1 [14]). As such, the current study investigated the effect simulation size and temperature have on Young's modulus for two force fields – Tersoff and ReaxFF.

II. MOLECULAR DYNAMICS

A. Many-body Potentials

The simulation of atom and molecular movement is the basis of Molecular Dynamics. Using Newton's equations of motion, with known position and velocity of the atoms, future positions and velocities are predicted utilizing each atom's acceleration. The force field provides information regarding how the atoms and molecules will react and as such, allows for the determination of atom acceleration from the interaction of particles. As such, the force field is a crucial component of any MD simulation. Using the updated acceleration, new positions and velocities are determined repeatedly until the simulation is complete. Initially, the force fields that were first developed were two-body styled – Leonard-Jones and Coloumbic, for Van der Waals and charged particles, respectively [19]. These force fields, while useful for non-bonded interactions, are unable to incorporate the interactions which occur in large, bonded

systems. As such, many-body force fields were developed to incorporate the effects of clusters of atoms. In the current study, two types of many-body force fields, Tersoff [8], and ReaxFF [9], are investigated.

B. Tersoff

The Tersoff potential is bond-order dependent. Bond order for each atom depends upon the neighbouring atoms and is highly influential to the physics of the system. Tersoff stated the most important factor to be determined is the coordination number, where the coordination number is said to be the number of neighbouring atoms close enough to form bonds [8]. As such, an atom with many neighbours is said to form weaker bonds than atoms with few neighbours [8]. A detailed explanation of the Tersoff force field can be found in [8].

C. Reactive Force Field (ReaxFF)

Originally proposed by van Duin et al. [9], the ReaxFF force field is a bond-order dependent force field, similar to the Tersoff force field. The inclusion of bond order allows for contributions due to various types of covalent bonding (sigma, pi, and double-pi) over the distance between atoms. In addition to bond order inclusion, the added benefit of the ReaxFF force field is the ability to simulate dynamic formation and breaking of bonds within the system. The energy defined in the ReaxFF force field is shown in Equation (2), while bond order is shown in Equation (3).

$$E_{system} = E_{bond} + E_{over} + E_{under} + E_{val} + E_{pen} + E_{tors} + E_{conj} + E_{vdWaals} + E_{Coulomb} \quad (2)$$

where E_{bond} is the bond-order dependent energy, E_{over} is the energy penalization term for over-coordinated atoms, E_{under} is the under-coordinated atoms energy, E_{val} is the energy from the valence angles, E_{pen} is another energy penalty to provide stability, E_{tor} is the energy from the torsion angles, E_{conj} is the energy of the conjugation effects, $E_{vdWaals}$ is the van der Waals energy term, and $E_{Coulomb}$ is the Coulomb energy term [9]

$$BO_{ij}^i = BO_{ij}^\sigma + BO_{ij}^\pi + BO_{ij}^{\pi\pi} \quad (3)$$

where bond order is determined from the interatomic distance between a pair of atoms for the various types of covalent bonding (sigma, pi, and double-pi). More detail on the ReaxFF force field can be found in [9].

III. METHODOLOGY

Building upon the work of Morrissey et al. [6], the current study investigated the effect of simulation size on the prediction of Young's modulus for pure iron to determine an ideal simulation size which balances accuracy and computational efficiency. With a simulation size determined, the effect of temperature on Young's modulus was investigated and compared to an empirical equation describing this effect for pure iron (Equation 1). To perform this work, a body-centered cubic (BCC) body of pure iron atoms was placed in tension at room temperature to replicate a macroscale uniaxial tension test. This was repeated for simulation sizes ranging from 5x5x5 lattice units to 30x30x30 lattice units, where one

lattice unit is 2.856Å for pure iron. From this, the stress-strain curves were generated and analyzed to determine the mechanical properties. After selecting the most ideal simulation size, a study into the aspect ratio was performed to observe differences in shape on the prediction of Young's modulus. Next, the study on the effect of temperature was performed for pure iron. The Young's modulus predicted in MD was then compared to the empirical equation from Fields [14] (Equation 1). Molecular dynamics simulations were carried out using Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS), as developed by Sandia National Laboratories [20].

First, a 5x5x5 lattice units BCC pure iron structure, with $\langle 100 \rangle$ orientations in the x-, y-, and z-directions, was created. The geometry and volume are representative of the gauge section in a typical tensile specimen [6]. This block of iron atoms contained 250 atoms and periodic boundary conditions were applied in all directions. Next, structures of 10x10x10, 15x15x15, 20x20x20, 25x25x25, and 30x30x30 were created, with 2000, 6750, 16000, 31250, 54000 atoms, respectively.

A molecular dynamics script was developed for use with LAMMPS which replicated a uniaxial tension test. For the initial size determination, a temperature of 300K was selected to replicate realistic testing. The time step and equilibration time were derived from balancing computational efficiency, minimal energy fluctuations, and ensuring energy conservation. Table 1 provides a summary of the selected values for simulation components.

As stated prior, two parameter sets were chosen to compare results. Both parameter sets studied were obtained from the National Institute of Standards and Technology (NIST) interatomic potentials repository [21].

Each system was equilibrated for 50 picoseconds at 300K using an isothermal-isobaric ensemble (NPT). The systems were allowed to adjust the lattice under natural expansion and contraction throughout the equilibration process. Each force field reached equilibrium well within this time period, with pressure and temperature damped using coefficients equal to 1000 and 100 times the time step, respectively.

Once the system was equilibrated, it was then deformed in the x-direction, using NPT. Deformation occurred at 300K with zero pressure on transverse sides to allow for natural Poisson contractions [5]. To obtain an accurate Young's modulus, the linear portion of the stress-strain curve must be studied such that Hooke's law can be utilized. As such, strains of 3% were utilized to allow for linear approximation of the stress-strain curve. In addition, it has been shown that the selected strain rate of 10ns-1 is well within the range at which minimal variation occurs for low strains of 3% [5].

This process was repeated for each simulation size to determine the ideal size. Next, using the ideal simulation size, the effect of temperature was investigated. Temperatures ranging from 100K to 600K were studied with the remaining simulation parameters remaining the same. The results were then compared to an empirical equation (Equation 1) found in literature [14].

Table 1 – LAMMPS Input Parameters

Parameter	Tersoff	ReaxFF
Temperature (K)	300	300
Time step (fs)	1	0.25
Equilibration (fs)	50,000	50,000
Strain Rate (1/fs)	0.00001	0.00001
Strain (%)	2	2

IV. RESULTS

A. Size Effect on Prediction of Young's Modulus

Prior to observing the effect of temperature on Young's modulus, an adequate simulation size was required to ensure accurate predictions. First, multiple cubic structures, with side lengths ranging from 5 lattice units to 30 lattice units, were tested to determine the effect that simulation size has on predicting Young's modulus. This was completed using both a Tersoff parameter set [13] and a ReaxFF parameter set [12]. For predicting an accurate modulus, the first 3% strain is observed in the stress-strain curve. Figure 1 shows the results from the Tersoff parameter set [13] while Figure 2 shows the results from the ReaxFF parameter set [12].

When observing the slope of the stress-strain curves in Figure 1, it was noted that the data points have more variance in stress at extremely small simulation sizes (5x5x5), thus negatively affecting the prediction in Young's modulus. As the size was increased, the variance in stress at each strain becomes smaller, allowing for a more accurate prediction as expected. Similar trends are observed in Figure 2; however, ReaxFF predictions contained less variance for all sizes, even for the smallest simulation size. A comparison of the predicted Young's moduli and simulation size for both parameter sets was performed to obtain a balance between computational efficiency and accuracy of prediction (Table 2).

Table 2 - Prediction in Young's Modulus from MD

Simulation Size (lattice units)	Tersoff Parameter Set [4]		ReaxFF Parameter Set [5]	
	Young's Modulus (GPa)	Error from Nominal (%) [6]	Young's Modulus (GPa)	Error from Nominal (%) [6]
5x5x5	239.12	22.0	178.26	-9.05
10x10x10	226.25	15.4	175.4	-10.5
15x15x15	223.61	14.1	174.0	-11.2
20x20x20	225.84	15.2	174.1	-11.2
25x25x25	225.22	14.9	174.3	-11.1
30x30x30	225.18	14.9	174.3	-11.1

It was determined that a 15x15x15 lattice units cube simulation size was the most ideal (containing 6750 atoms), allowing for a converged prediction in Young's modulus from stress-strain while also allowing for computational efficiency. With the convergence of the Young's modulus (Table 2), it is demonstrated that sizes above 15x15x15 lattice units yield no marginal increase in the accuracy of prediction. This size allows for a sufficient number of atoms to enhance the

prediction of Young's modulus. Computing over more atoms in larger simulation sizes did not further enhance predictions as the inaccuracy in predicting the true value for Young's modulus is due to 1) the parameterization of the force field utilized, and 2) the simulations were single-crystal and not polycrystalline.

It is also noted in Table 2 that the ReaxFF parameter set [12] was capable of predicting an accurate Young's modulus with less error than the Tersoff parameter set [13]. Overall, the ReaxFF parameter set predicts the Young's modulus within -10.7% error, with a minimum and maximum error of -9.05% and -11.2%, respectively (Table 2). Overall, the Tersoff parameter set predicts the Young's modulus within 16.1% error; however, there is a larger variance than ReaxFF, with a minimum error of 14.1% and a maximum of 22% (Table 2). While the ReaxFF under predicts Young's modulus at all sizes tested, Tersoff over predicts Young's modulus. Regardless of the accuracy in predicting Young's Modulus, the current study only required a converged modulus to understand the effect of temperature, since the reduction in modulus due to temperature is investigated. As such, the 15x15x15 lattice units cube is selected for the current study.

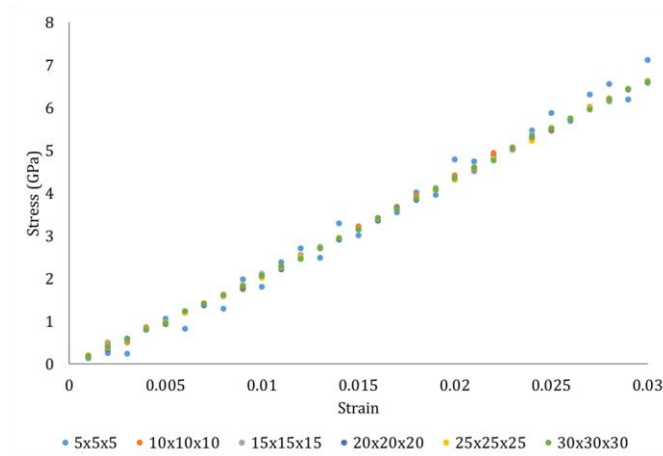


Figure 1 - Stress-strain results for Tersoff parameter set [13] versus Simulation Size

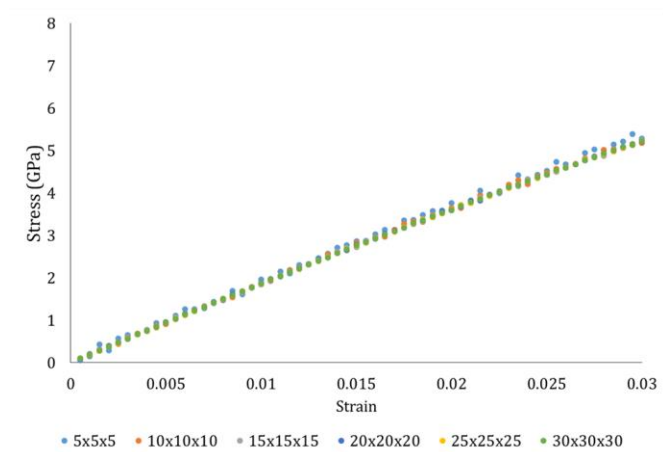


Figure 2 - Stress-strain results for ReaxFF parameter set [12] versus Simulation Size

B. Effect of Aspect Ratio

After considering the effect of force field selection on cubic samples, i.e. a 1:1:1 aspect ratio, the next step was to determine if non-cubic samples with various aspect ratios still followed the same trend. As above, samples with various aspect ratios were equilibrated and then placed in tension. Five different aspect ratios with a similar number of atoms were tested, ranging from a 22x22x22 lattice unit cube, to a 200x10x5 lattice unit slender beam (**Error! Reference source not found.**). Referring to **Error! Reference source not found.**, for the Tersoff parameter set [13] there was no observable difference in the Young's modulus for all aspect ratios tested. Similarly, for the ReaxFF parameter set [12] all aspect ratios followed an identical Young's modulus (**Error! Reference source not found.**). These results are expected as the boundary conditions for all cases were periodic, meaning that, regardless of aspect ratio, it was still simulating a bulk sample. Therefore, for both the ReaxFF and Tersoff parameter sets the aspect ratio did not change the loading behavior, as such, a simple 15x15x15 lattice unit cubic structure was utilized for studying the effect due to temperature.

Table 3 - Comparison of Aspect Ratios

Case	Dimension (lattice units)			Number of atoms
	x	y	z	
1	100	10	10	20000
2	25	20	20	20000
3	50	20	10	20000
4	200	10	5	20000
5	22	22	22	21296

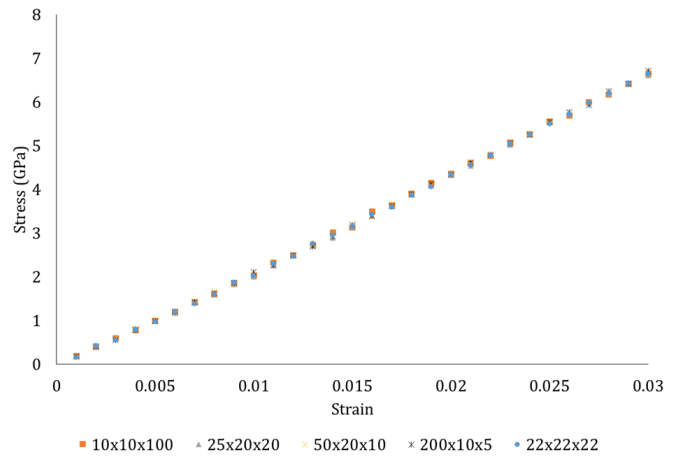


Figure 3 – Effect of Aspect Ratio for Tersoff parameter set [13]

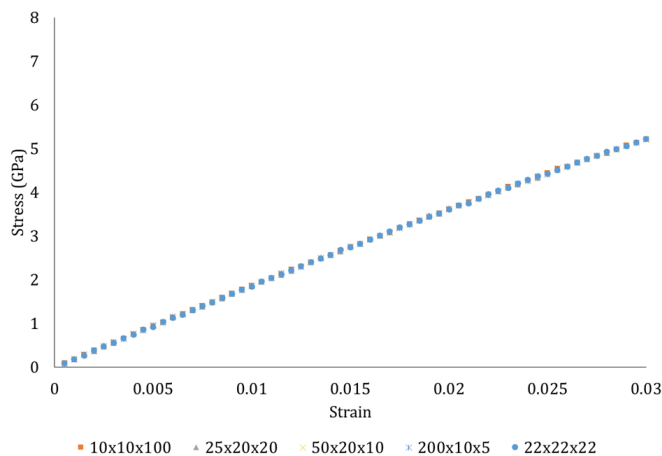


Figure 4 - Effect of Aspect Ratio for ReaxFF parameter set [12]

C. Temperature Effect on Young's Modulus

With a sufficient simulation size determined, the effect of temperature on Young's modulus was investigated. It is known that as temperature increases, the Young's modulus decreases, as shown in Equation 1; however, it is not known if MD is capable of accurately predicting the reduction in Young's modulus at various temperatures. As such, temperatures ranging from 0K to 600K, in 100K increments, were studied. A comparison of the reduced Young's modulus versus temperature for Fields, ReaxFF, and Tersoff is shown in Figure 5. The Fields' [14] equation showed a highly linear relationship for the reduction in Young's modulus as temperature increased. The Tersoff parameter set [13] showed a highly non-linear prediction in Young's modulus, with a reduction of 12.1% and 43.5% in Young's modulus for 300K and 600K, respectively. Compared to the reduction of 3.6% and 12% predicted by Fields' equation at 300K and 600K, respectively.

Conversely, the ReaxFF parameter set [12] was highly linear and approximately parallel to Fields' [14] equation except ReaxFF predicted approximately 2.5% more reduction than Fields at every temperature (as evident in Table 4).

Table 3 - Temperature effect on Young's modulus

Temperature (K)	Reduction in Young's modulus		
	Fields' Equation [14]	Tersoff Parameter Set [13]	ReaxFF Parameter Set [12]
100	0%	0%	0%
200	1.7%	3.4%	3.7%
300	3.6%	12.1%	6.3%
400	5.9%	20.4%	9.1%
500	8.6%	32.6%	11.5%
600	12.0%	43.5%	13.7%

It appears that the Tersoff parameter set [13] is not well suited for studying mechanical properties undergoing varying temperatures due to the highly non-linear reduction in Young's modulus with temperature. Overall, the Tersoff parameter set over predicts the reduction of Young's modulus when compared to Fields' equation [14] (Table 4).

In contrast, it appears that the ReaxFF parameter set [12] is acceptable for studying mechanical properties undergoing varying temperatures due to the highly linear reduction in Young's modulus with temperature. Overall, the ReaxFF parameter set slightly over predicts the reduction of Young's modulus when compared to Fields' equation [14] (Table 4).

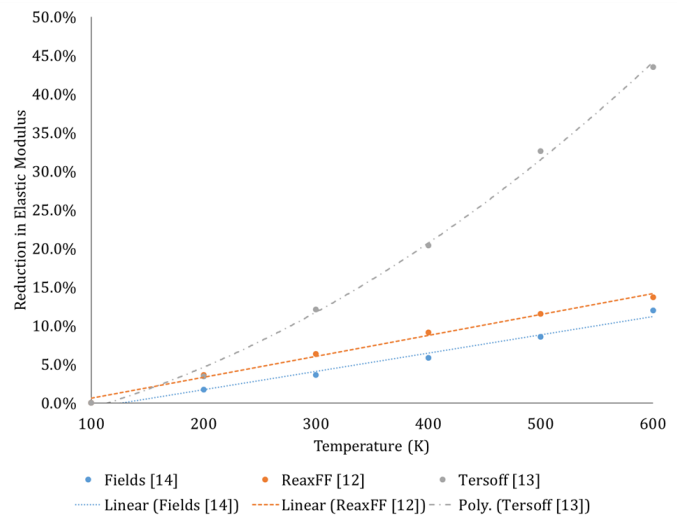


Figure 5 - Reduction in Modulus versus Temperature for Fields [14], ReaxFF [12], and Tersoff [13]

V. CONCLUSION

The current work presents a study on size effect and temperature effect in molecular dynamics simulation for the purpose of demonstrating the impact of force field selection on predicting the Young's modulus of pure iron. Molecular dynamics simulations were performed using two many-body force fields to determine the ideal simulation size for accurately predicting the Young's modulus of perfect crystalline iron. Next, molecular dynamics simulations were performed to study the effect temperature had on the prediction of the Young's modulus. Comparisons were made to an empirical equation from literature for pure single-crystal iron. Predicted values of reduction in Young's modulus for ReaxFF showed good agreement to the empirical equation reported in literature, while Tersoff captured the trend, but had much more error in prediction.

ACKNOWLEDGMENT

The authors are grateful for the financial support from Suncor Energy Inc.

REFERENCES

- [1] J. Zhu, M. He and F. Qiu, "Effect of Vacancy Defects on the Young's Modulus and Fracture Strength of Graphene: A Molecular Dynamics Study," *Chin. J. Chem.*, vol. 30, (7), pp. 1399-1404, 2012. Available: <https://doi.org/10.1002/cjoc.201200505>. DOI: 10.1002/cjoc.201200505.
- [2] X. Sun et al, "Effects of vacancy defect on the tensile behavior of graphene," *Theoretical and Applied Mechanics Letters*, vol. 4, (5), pp. 051002, 2014. . DOI: <https://doi.org/10.1063/2.1405102>.
- [3] V. Vijayaraghavan and L. Zhang, "Effective Mechanical Properties and Thickness Determination of Boron Nitride Nanosheets Using Molecular Dynamics Simulation," *Nanomaterials*, vol. 8, (7), pp. 546, 2018.

- Available: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC6071296/>. DOI: 10.3390/nano8070546.
- [4] W. Chang, "Molecular-dynamics study of mechanical properties of nanoscale copper with vacancies under static and cyclic loading," *Microelectronic Engineering*, vol. 65, (1), pp. 239-246, 2003. . DOI: [https://doi.org/10.1016/S0167-9317\(02\)00887-0](https://doi.org/10.1016/S0167-9317(02)00887-0).
 - [5] B. D. Jensen, K. E. Wise and G. M. Odegard, "The effect of time step, thermostat, and strain rate on ReaxFF simulations of mechanical failure in diamond, graphene, and carbon nanotube," *J. Comput. Chem.*, vol. 36, (21), pp. 1587-1596, 2015. Available: <https://doi.org/10.1002/jcc.23970>. DOI: 10.1002/jcc.23970.
 - [6] L. S. Morrissey et al, "Atomistic uniaxial tension tests: investigating various many-body potentials for their ability to produce accurate stress strain curves using molecular dynamics simulations," *Molecular Simulation*, pp. 1-8, 2018. Available: <https://doi.org/10.1080/08927022.2018.1557333>. DOI: 10.1080/08927022.2018.1557333.
 - [7] A. Rajabpour, L. Seidabadi and M. Soltanpour, "Calculating the Bulk Modulus of Iron and Steel Using Equilibrium Molecular Dynamics Simulation," *Procedia Materials Science*, vol. 11, pp. 391-396, 2015. . DOI: <https://doi.org/10.1016/j.mspro.2015.11.005>.
 - [8] J. Tersoff, "New empirical approach for the structure and energy of covalent systems," *Phys. Rev. B*, vol. 37, (12), pp. 6991-7000, 1988. Available: <https://link.aps.org/doi/10.1103/PhysRevB.37.6991>. DOI: 10.1103/PhysRevB.37.6991.
 - [9] A. C. T. van Duin et al, "ReaxFF: A Reactive Force Field for Hydrocarbons," *J Phys Chem A*, vol. 105, (41), pp. 9396-9409, 2001. Available: <https://doi.org/10.1021/jp004368u>. DOI: 10.1021/jp004368u.
 - [10] M. S. Daw and M. I. Baskes, "Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals," *Phys. Rev. B*, vol. 29, (12), pp. 6443-6453, 1984. Available: <https://link.aps.org/doi/10.1103/PhysRevB.29.6443>. DOI: 10.1103/PhysRevB.29.6443.
 - [11] B. J. Lee et al, "Second nearest-neighbor modified embedded atom method potentials for bcc transition metals," *Phys. Rev. B*, vol. 64, (18), pp. 184102, 2001. Available: <https://link.aps.org/doi/10.1103/PhysRevB.64.184102>. DOI: 10.1103/PhysRevB.64.184102.
 - [12] M. Aryanpour, A. C. T. van Duin and J. D. Kubicki, "Development of a Reactive Force Field for Iron-Oxyhydroxide Systems," *J Phys Chem A*, vol. 114, (21), pp. 6298-6307, 2010. Available: <https://doi.org/10.1021/jp101332k>. DOI: 10.1021/jp101332k.
 - [13] K. O. E. Henriksson, C. Björkas and K. Nordlund, "Atomistic simulations of stainless steels: a many-body potential for the Fe-Cr-C system," *Journal of Physics: Condensed Matter*, vol. 25, (44), pp. 445401, 2013. Available: <http://stacks.iop.org/0953-8984/25/i=44/a=445401>.
 - [14] R. J. Fields, T. Weerasooriya and M. F. Ashby, "Fracture-mechanisms in pure iron, two austenitic steels, and one ferritic steel," *Metallurgical Transactions A*, vol. 11, (2), pp. 333-347, 1980. Available: <https://doi.org/10.1007/BF02660638>. DOI: 10.1007/BF02660638.
 - [15] T. Fang, C. Weng and J. Chang, "Molecular dynamics analysis of temperature effects on nanoindentation measurement," *Materials Science and Engineering: A*, vol. 357, (1), pp. 7-12, 2003. . DOI: [https://doi.org/10.1016/S0921-5093\(03\)00219-3](https://doi.org/10.1016/S0921-5093(03)00219-3).
 - [16] C. Zhang and H. Shen, "Temperature-dependent elastic properties of single-walled carbon nanotubes: Prediction from molecular dynamics simulation," *Appl. Phys. Lett.*, vol. 89, (8), pp. 081904, 2006. Available: <https://doi.org/10.1063/1.2336622>. DOI: 10.1063/1.2336622.
 - [17] S. J. A. Koh et al, "Molecular dynamics simulation of a solid platinum nanowire under uniaxial tensile strain: Temperature and strain-rate effects," *Phys. Rev. B*, vol. 72, (8), pp. 085414, 2005. Available: <https://link.aps.org/doi/10.1103/PhysRevB.72.085414>. DOI: 10.1103/PhysRevB.72.085414.
 - [18] H. A. Wu, "Molecular dynamics study on mechanics of metal nanowire," *Mechanics Research Communications*, vol. 33, (1), pp. 9-16, 2006. . DOI: <https://doi.org/10.1016/j.mechrescom.2005.05.012>.
 - [19] S. J. Plimpton and A. P. Thompson, "Computational aspects of many-body potentials," *MRS Bull.*, vol. 37, (5), pp. 513-521, 2012. . DOI: 10.1557/mrs.2012.96.
 - [20] S. J. Plimpton, "Fast Parallel Algorithms for Short-Range Molecular Dynamics," *Journal of Computational Physics*, vol. 117, (1), pp. 1-19, 1995. . DOI: <https://doi.org/10.1006/jcph.1995.1039>.
 - [21] C. A. Becker et al, "Considerations for choosing and using force fields and interatomic potentials in materials science and engineering," *Current Opinion in Solid State and Materials Science*, vol. 17, (6), pp. 277-283, 2013. . DOI: <https://doi.org/10.1016/j.cossms.2013.10.001>.

On the Influence of the Environmental Factors on Corrosion Behavior of 13Cr Stainless Steel Using Box-Behnken Design

Mostafa Kazemipour, Salar Salahi, Ali Nasiri
Faculty of Engineering and Applied Science
Memorial University of Newfoundland,
St. John's, NL, A1B 3X5, Canada
mkazemipour@mun.ca

Abstract

Recognizing the dominating environmental factor(s) in controlling the corrosion behaviour of materials is extremely crucial for the service life prediction of industrial alloys, such as stainless steels, in harsh environments. Among various grades of corrosion resistant alloys, 13Cr martensitic stainless steel has increasingly attracted attention from various industrial sectors due to its suitable mechanical and corrosion properties. However, the sensitivity of the alloy to localized corrosion particularly in harsh offshore and downhole environments limits the service life of the alloy. The localized corrosion of the alloy is affected by both the material's intrinsic properties and the peripheral conditions. Focusing on the latter, in this study, to evaluate the effects of environmental factors on the corrosion rate and pitting resistance of 13Cr stainless steel, the single and interactive effects of three main environmental parameters specific to a downhole service, including the environment's pH (4-7), chloride concentration (1000-22000 mg/L), and temperature (22-80 °C), were studied by conducting cyclic potentiodynamic polarization (CPP) testing. Using a response surface methodology, the experiments were performed according to the Box-Behnken design, and was further optimized using Analysis of Variance (ANOVA). To correlate the above-mentioned environmental factors to the corrosion response of 13Cr stainless steel, a quadratic model for predicting the pitting potential of the alloy was developed. The results revealed the destructive effects of decreasing pH and increasing the chloride concentration in the environment on the pitting corrosion resistance of the alloy. Although the increase of the temperature was found to initially deteriorate the alloy's pitting resistivity up to the mid-temperature of 51 °C, at temperatures higher than 51 °C, an improved pitting resistance was predicted and measured, associated with the increased content of the corrosion products on the surface at higher temperatures, mitigating the diffusion of aggressive ions through the passive layer.

Keywords: 13Cr Stainless steel, Corrosion properties, Box-Behnken design, Environmental factors

1. Introduction

Corrosion is a significant challenge to most of the industrial sectors and in particular oil and gas related industries, causing serious economic and environmental issues, primarily due to its basic roll in the failure of the collection and transportation infrastructures [1,2]. Having a highly complex nature, many different factors, including temperature, pH, partial pressure of the existing gases, flow velocity, etc. can affect corrosion [1,2].

Corrosion resistance alloys (CRAs) are usually the most commonly used materials for harsh environments and severe operational conditions [3]. Stainless steels provide the largest family of CRAs containing usually more than 12 wt.% chromium dissolved in the structure of the steel in a form of a solid solution. This element forms a passive thin layer of its oxide on the steel surface, leading to its corrosion resistivity and protect the steel against uniform corrosion [4]. However, in the environments containing aggressive ions, such as chloride, the local breakdown of the passive layer at the defect sites by the ions, makes them prone and susceptible to a localized type corrosion attack called pitting corrosion, which is usually difficult to be mitigated especially in off-shore environments [5,6]. Such localized corroded sites can subsequently be accompanied by applied stresses and act as stress concentrators, leading to a crack initiation named stress corrosion cracking (SCC), which is the most frequent reason for the sudden failure of casings, tubing, and pipes in field applications [5,7].

In comparison with the other types of CRAs, such as dual phase duplex stainless steels, 13Cr (AISI 420) martensitic stainless steel is widely used as the casings and tubing material for harsh oil and gas downhole environment at elevated temperatures up to 150°C due to its good corrosion and mechanical properties with lower cost [3].

It is well established that both the material's inherent properties, such as hardness and tensile strength of the stainless steels, and the environmental factors, including H₂S/CO₂ partial pressures, pH, the total pressure of the environment, chloride concentration, and temperature, act as

dominating factors in controlling the corrosion of a structure over time [8]. Among these factors, three of them, *i.e.* chloride concentration, temperature, and pH, have the most profound influence on the material's failure in harsh downhole environments [9]. Therefore, a deep understanding of the impacts of these factors is a key step in service life prediction and failure analysis of the oil country tubular goods (OCTG), such as 13Cr stainless steel casting and tubing [8,9].

The corrosion resistance of 13Cr stainless steel is drastically affected by any small variation of the environmental conditions [3]. Determining the existing correlations between the corrosion response and the altering factors is fundamentally required to be able to develop an ultimate comprehensive corrosion prediction model. Modeling and optimization approach is a low-cost method to evaluate the effects of the environmental factors on the corrosion and degradation behaviour of the material systematically, and could potentially lead to a methodology to assess the material's remaining life and strength in a harsh environment after many years of service [10].

Prediction of a stainless steel's failure as a result of electrochemical reactions is practically difficult and highly complex, and the developed empirical models are mostly on the basis of statistical analysis of real data [11]. For instance, in order to model the pitting corrosion, maximum pit depth has been modeled as a function of environmental parameters [11]. Furthermore, providing the input data to develop an empirical model capable of predicting the pitting corrosion from real structures in service, such as operating casings and tubing, is extremely difficult [12]. In a study performed by Galvele [13], a logarithmic relationship between the pitting potential and the solution chloride concentration was predicted. This dependency prediction was then evaluated and confirmed in a range of temperature, from ambient to higher temperatures [5]. Contrarily, some investigations reported a linear correlation between pitting potential and chloride concentration [14].

Sakamoto *et al.* [15] proposed equations to predict the effect of temperature and pH on the corrosion rate of 13Cr stainless steel under CO₂ environment. Their reported results indicated that the corrosion rate of the stainless steel first increases with an increase in the environment temperature, and then decreases after reaching 200 °C [15]. The proposed equation showed a logarithmic relation between the corrosion rate and pH. In a recent study, Ossai *et al.* [16] developed a multivariate regression model using the maximum pit depth factor and by considering the pipeline operational parameters, including temperature, CO₂ partial pressure, pH, fluid flow rate, chloride ion, and sulfate ion.

Although, the corrosion and concerning prediction models have been extensively studied, most of them are adjusted for a special case study and not a comprehensive model. In spite of

so many offered models, there is no standard approach for prediction of pitting corrosion in downhole environments [12].

Thus, along with previous attempts to predict corrosion properties of stainless steels, in this study as the first step for the final goal of establishing a comprehensive prediction model, a mathematical model was developed using a design of experiment (DOE) methodology. For this purpose, three environmental parameters, *i.e.* temperature, chloride concentration, and pH, were considered as the variable of the design. The measured pitting potentials from cyclic potentiodynamic polarization (CPP) scans at different levels of the parameters were decided as the response.

2. Materials and Methods

2.1. Specimen preparation

For the experimental work, small cubic samples were sectioned from a 13Cr-L80 stainless steel tube, provided by Suncor Energy. The measured chemical composition of the tube is presented in Table 1. The samples were hot mounted into an epoxy resin, confining the exposed area of the samples to only inner surface area of the tube. Following the hot mounting, the side of each mounted sample was drilled, and a stainless steel screw was connected to the samples' side to provide a conductive path to facilitate the passage of current during corrosion testing. Then, the samples were ground and polished to a mirror-like surface finish following standard sample preparation procedures for stainless steels, followed by an ultrasonic cleaning in ethanol.

2.2. Electrochemical Testing

The electrochemical experiments were conducted at different temperatures (22-80 °C), pH (4-7), and chloride concentrations (1000-22000 mg/Lit), covering a range for each environmental factor that simulates the real working conditions of the tube in downhole service (see Table 2). This set of the parameters were measured and calculated by Suncor Energy. The electrolyte solutions were prepared by dissolving different amount of high purity NaCl crystals into 500 ml of distilled water. A temperature-controlled water bath was used to control the testing solution's temperature, and the pH was adjusted by adding hydrochloric acid or sodium hydroxide to the solution. All the experiments were conducted in a standard three-electrode corrosion cell, containing the specimen as the working electrode, saturated Ag/AgCl as the reference electrode, and a graphite rod as the auxiliary electrode. To mitigate the side effect from the sample-mounting resin interface, all edges of the sample were covered by highly resistant non-conductive polyester resin. Prior to each electrochemical test, the open circuit potential (OCP) was monitored for 3600 s to ensure the samples reach the stability. CPP experiments were performed at starting potential of -0.2 V versus OCP up to the vertex current of 0.1 mA at a scanning rate of 1 mV/s.

Table 1. The measured chemical composition of 13Cr stainless steel used in this study (wt.%)

Material	Cr	Mn	P	Si	S	C	Ni	Cu	Fe
13Cr (AISI 420)	13.4	0.85	0.003	0.8	0.004	0.19	0.42	0.24	Bal.

2.3. Design of experiment method

To be able to model the corrosion response of the 13Cr stainless steel as a function of different environmental factors, *i.e.* temperature, pH, and chloride concentration, and determine the interactions and the single effect of the factors on the corrosion behaviour of the alloy, a response surface methodology (RSM) using Box-Behnken experimental design was adopted. As a basic statistical method for the experimental analysis and modeling, RSM was used to find a relationship between the 3 input factors and the response variables. By performing multiple regression analysis on the obtained data, a second order mathematical model capable of predicting the response variable can be developed. The model follows the general form of:

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=2}^k \beta_{ij} x_i x_j + \varepsilon \quad (1)$$

where y represents the model response, ε is the error term, β_i , β_{ii} , β_{ij} are the linear coefficients, quadratic coefficients, and the interaction coefficients, respectively, which should be determined by the least square method, x_i and x_j are independent input variables.

The statistical data analysis was performed by Design-Expert version 11, which uses analyses of variance (ANOVA) to determine the significance of the factors. The parameters of the model with p -value < 0.05 (confidence level 95%) were considered to be significant.

Table 2. Levels of variables chosen for Box-Behnken design

Factor	pH	Chloride concentration (mg/L)	Temperature (°C)
Low level	4	1000	22
High level	7	22000	80

3. Results and discussion

The Box-Behnken design and the response of each experiment, defined as the pitting potential and measured by the CPP testing, are summarized in Table 3. The CPP graphs of some the tests are depicted in Fig. 1. The best corrosion response, indicated by the highest pitting potential, was obtained for the environmental conditions of the temperature of 22 °C, solution pH of 5.5, and chloride concentration of 1000 mg/L. The worst corrosion behaviour accompanied by the lowest pitting potential was observed at 51 °C, 4, 22000 mg/L for the temperature, pH, and chloride concentration, respectively. To determine the right combination of the factors that lead to the best and the worst corrosion response, a careful optimization of the results should be performed. Therefore, RSM method was employed to analyze the obtained results using the Design-Expert software, aiming to develop a mathematical model capable of defining the relationship between the variables. Based on the experimental results, the best-matched model was recognized to be a quadratic (second order) model, which was used for the subsequent analyses. The analysis of variance (ANOVA) was used to examine the obtained regression quadratic model, and to determine the significance of the model main factors and their interactions. These results are summarized in Table 4. The statistical analyses were done at a confidence level of 95% (level of significance of 5%). For determining the significance of each factor in the model, p -value was used, and for the examination of the statistical significance of the regression model, F -value was applied.

Table 3. Box-Behnken experimental design layout and the corresponding experimental results

Run	A: Chloride Concentration (mg/L)	B: Temperature (°C)	C: pH	Pitting Potential ($V_{Ag/AgCl}$)
1	22000	51	7	0.029
2	11500	51	5.5	0.098
3	1000	51	7	0.17
4	1000	80	5.5	0.15
5	1000	51	4	0.164
6	22000	80	5.5	0.038
7	1000	22	5.5	0.235
8	11500	80	4	0.058
9	11500	51	5.5	0.083
10	11500	22	4	0.12
11	11500	51	5.5	0.079
12	22000	51	4	-0.07
13	11500	51	5.5	0.072
14	22000	22	5.5	0.044
15	11500	22	7	0.17
16	11500	51	5.5	0.094
17	11500	80	7	0.14

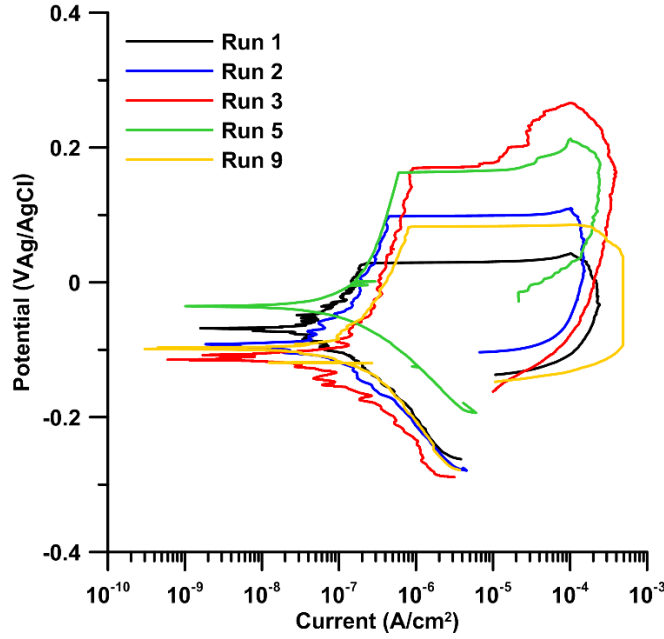


Fig. 1. Cyclic polarization graphs of Runs 1, 2, 3, 5, and 9 experiments.

If the p-values for the model, factors, and interactions are less than 0.05, then they are statistically significant. Lack of fit parameter shows the significance of the replicate error in comparison to the model dependent error. In other words, it indicates the model adequacy. If this parameter is significant, it implies that the predicted model might not consider some contributions between the independent variables and the response.

The R^2 from the ANOVA analysis shows how the fitted model is close to the actual data. It is a measure of the quality of the model. Usually models with an R^2 more than 80% are considered good fitting models capable of providing an acceptable relationship between the independent variables and the response. Also, R^2_{adj} corresponds to the adjusted value of R^2 , which decreases by increasing the number of terms in the model, if the added terms do not add value to the model. It should be also noted that R^2_{pred} is a measure of the variations in the data predicted by the model. Generally, the closer all these R-square values to each other and 100% are, the more accurate the predicted model is. Finally, in order to measure the signal to the noise ratio, which is a comparison between the range of the predicted values at a design point and the average prediction model, the adequate precision value was used. This number should be greater than 4 for a desirable model.

Herein, for the developed RSM model, the F-value of 72.08 and the corresponding P-value of less than 0.0001 indicate that the model is significant. There is only a 0.01% chance that such large F-value has occurred due to the noise. All the model terms, including the main factors of A, B, C, and the interactions of AB, AC, and B^2 are significant because their P-values are less than 0.05. Also, the F-value for lack of fit of the

model is 1.99, which is insignificant. There is a 26.37% chance that a Lack of Fit F-value this large could occur due to the noise.

The R^2 value for the model was calculated to be 0.9774, indicating that it fits the experimental and predicted values competently. The other adequacy measures for the model, namely $R^2_{adj}=0.9638$ and $R^2_{pred}=0.9046$, both are reasonably high and in agreement with each other because their difference is less than 0.2. The adequate precision value of 35.11, which is significantly greater than 4, confirmed an adequate signal, indicating that the model can be used for the design ranges. The final model co-relating the pitting potential (E_{pit}) and the actual factors (variables), *i.e.* chloride concentration (A), temperature (B), and pH value (C), was obtained as follows:

$$E_{pit} = 0.357844 - 0.000019A - 0.006324B + 0.002774C + 6.48604 \times 10^{-8}AB + 1.47619 \times 10^{-6}AC + 0.000047B^2 \quad (2)$$

To examine the assumptions of ANOVA and validate the model assumptions in regression, the residuals plots should be analyzed. Figure 2(a) represents the normal probability plot of residuals. It is evident that the residuals fall on a straight line suggesting that the distribution of the errors is relatively normal and the model well fitted with the observed values. The results of residuals versus the predicted response for this experiment are shown in Figure 2(b). According to this figure, it can be seen that the data have a random scatter and the assumption of constant variance is approved. Figure 2(c) compares residuals and experimental run order, which investigates the assumption of the independence of the observations. It also checks for the variables that may have influenced the response in the experiment.

Table 4. ANOVA and regression analyses for the quadratic model

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	0.0790	6	0.0132	72.08	< 0.0001	significant
A-Chloride	0.0575	1	0.0575	314.59	< 0.0001	
B-Temp.	0.0042	1	0.0042	22.92	0.0007	
C-pH	0.0070	1	0.0070	38.44	0.0001	
AB	0.0016	1	0.0016	8.54	0.0152	
AC	0.0022	1	0.0022	11.84	0.0063	
B²	0.0066	1	0.0066	36.15	0.0001	
Residual	0.0018	10	0.0002			
Lack of Fit	0.0014	6	0.0002	1.99	0.2637	not significant
Cor Total	0.0808	16				
R²	0.9774					
Adjusted R²	0.9638					
Predicted R²	0.9046					
Adeq. Precision	35.11					

This plot also reveals a random scatter without any particular trend for the observations. In a further analysis, each of the obtained values for the pitting potential was compared with the value predicted by the model in Figure 2(d). It can be seen that all the points lie fairly close to a line with an angle of 45°, with respect to the plot axis, implying a good consistency between the actual and the predicted values.

From the performed ANOVA analysis presented in Table 4, it can be noted that the model contains two important interactions, including AB and AC. The plots for these interaction effects are shown in Figure 3. The interaction between temperature and chloride concentration (AC) illustrated at Figure 3(a) indicates that at a high or a low level of pH of the environment, the pitting potential decreases with the increase in chloride concentration. However, the sensitivity of the pitting potential to chloride concentration at higher values of pH is lower compared to lower pH values. The same behavior can be seen for the other interaction, AB. Based on this interaction, at different level of temperature, by an increase in the value of chloride concentration, pitting potential decreases, implying that the corrosion resistance is degraded. However, at higher level of temperature, the changes in chloride concentration does not affect the pitting potential significantly. Plausibly, at higher temperatures the uniform corrosion dominates over the pitting corrosion, and acts as the controlling corrosion mechanism.

The surface plots of the response (the pitting potential) with respect to the independent variables of temperature and chloride concentration at different levels of pH are shown in Figure 4.

Figure 4 indicates that the general shape of the response surface is relatively identical for all pH values representing a gradual increase in the pitting potential by reducing the chloride concentration or increasing the pH value. The impact of pH on the pitting potential was found to be more noticeable at the end of the surface, where the chloride concentration reaches to 22000 mg/L. Furthermore, from the surface plots in Figure 4, it can be inferred that in all the pH values, the minimum pitting potential takes place around the medium temperature of 51 °C.

To define the exact condition in which the best and the worst corrosion behaviors occur, corresponding to the highest and the lowest pitting potentials, respectively, optimization was performed in the model using two different criteria of the maximum and minimum pitting potentials. The results for the optimization are shown in Table 5. As expected, the maximum pitting potential occurs at lower bound of two parameters of chloride concentration and temperature, and the higher limit of pH. However, the minimum pitting potential should be expected at the highest concentration of chloride, the lowest value of pH, and the medium temperature.

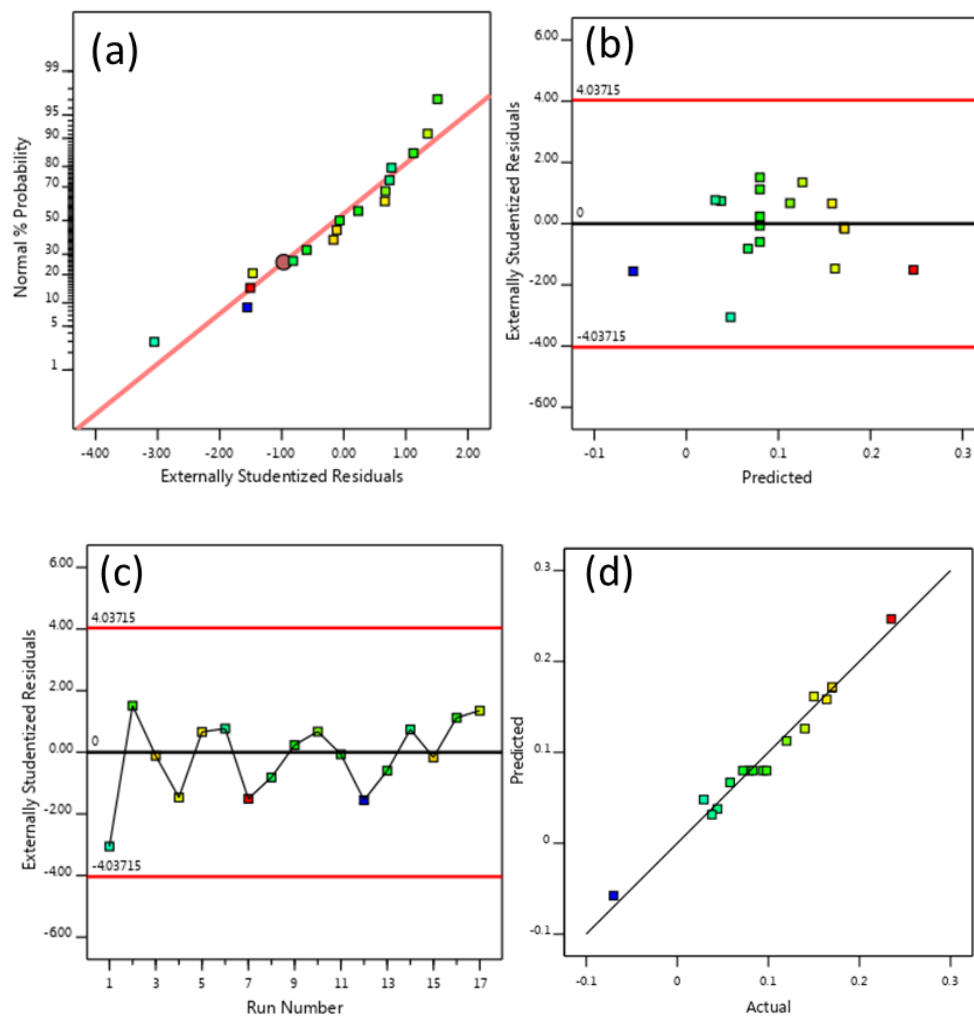


Fig. 2. (a) Normal probability plot of residuals, (b) plot of residuals versus the predicted values, (c) plot of residuals versus run, and (d) plot of the predicted values by the model versus the actual values

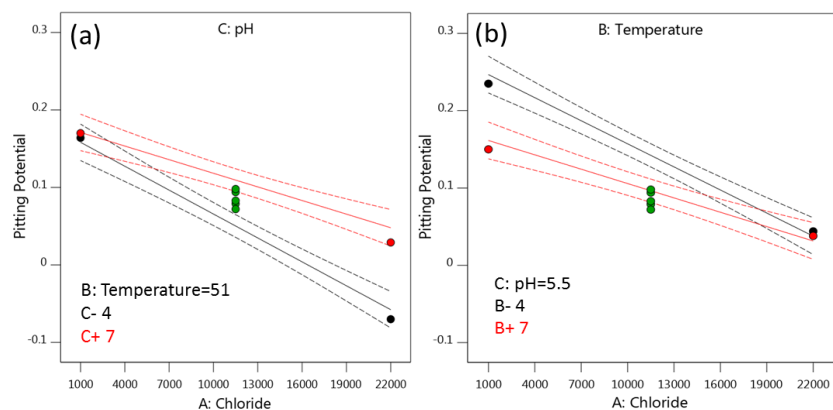


Fig. 3. Interaction effects between (a) chloride concentration (A) and pH (C), (b) chloride concentration (A) and temperature (B).

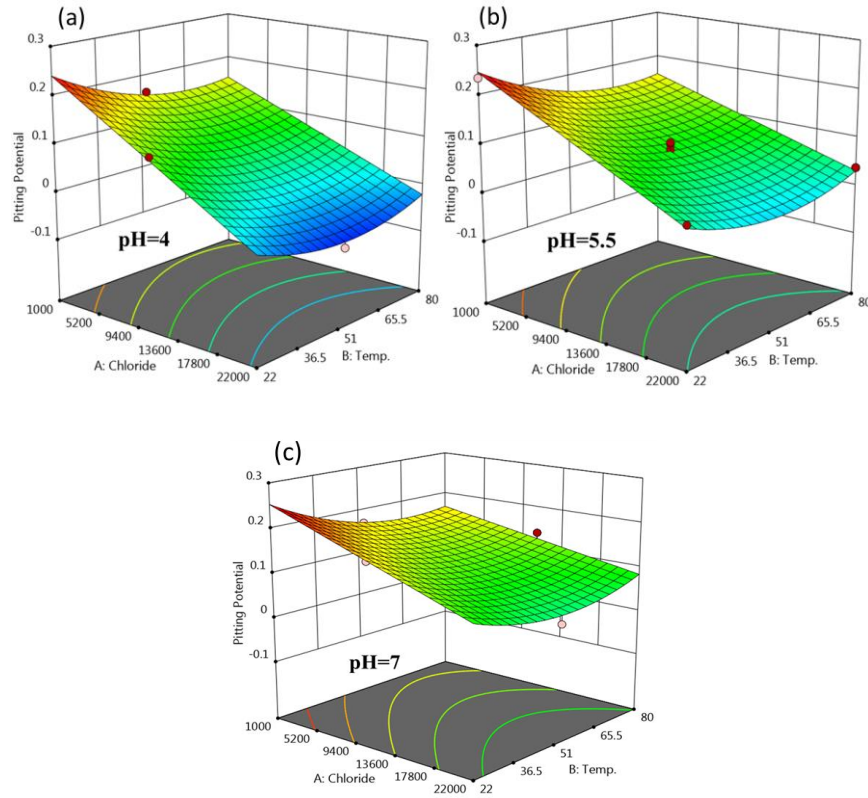


Fig. 4. Surface plots for the pitting potential with respect to the chloride concentration and temperature at (a) pH = 4, (b) pH = 5.5, and (c) pH = 7.

Table 5. The model optimization

Optimization	Temperature	Chloride Concentration	pH	Pitting Potential ($V_{Ag/AgCl}$)
Max. Pitting Potential	23.7	2240	7.0	0.237
Min. Pitting Potential	52.1	22000	4.0	-0.058

It is well known that the pH of the environment has usually a profound effect on the stability of the corrosion products on a steel surface [17]. The accumulation of corrosion products on the surface can suppress the diffusion of aggressive ions through the passive layer, leading to an improved corrosion response for the alloy [17]. By decreasing the pH of the environment, the stability of the corrosion products will be lower, and therefore the pitting potential decreases [18]. In addition, at a higher level of chloride concentrations due to the increased content of the aggressive ions in the electrolyte, the metal dissolution would accelerate [19].

Temperature is a crucial environmental factor, affecting phase content, acid gas solubility, and corrosion product formation [20]. Generally, at higher temperatures, an increase in the corrosion rate is anticipated, associated with the accelerated rates of the electrochemical and chemical reactions in the environment [12]. Therefore, it is expected that by increase in the temperature, the pitting potential decreases. However, the rate of corrosion products precipitation also increases at higher temperatures, leading to a lower corrosion rate [12]. If the temperature is sufficiently high, a protective layer of corrosion

products can form, contributing to lowering the corrosion rate of the steel [20].

It should be also mentioned that with increasing the temperature, the microstructure of the surface products and scales could potentially modify, leading to a reduced susceptibility of the surface to the pitting corrosion, and further provoking the tendency for the uniform corrosion [21]. Therefore, it can be concluded that at temperatures near 80 °C, the excessive formation of the corrosion products can plausibly increase the pitting potential. Nevertheless, the overall effect of the temperature on the protective film formation is highly complex and very much dependant on the nature of the layer. For instance, in the case of physically adsorbed layers, any increase in the environment's temperature would deteriorate the protectiveness of the layer, because higher temperatures would accelerate the layer's desorption. On the other hand, for layers that possess a chemisorbed nature on the surface of the steels, the increase of the temperature can potentially strengthen the chemical bonds in the layer and enhance its electrochemical stability up to a certain temperature, where above that the thermal dissolution of the layer is triggered [12].

4. Conclusion

In this study, the synergic effects of the environmental factors specific to the downhole environments, including pH (4-7), temperature (22-80 °C), and chloride concentration (1000-22000 mg/L), on corrosion behaviour of 13Cr stainless steel tubing were investigated. To design the experimentation plan, a response surface methodology using the Box-Behnken design method was adopted. The pitting potential obtained from the cyclic potentiodynamic polarization testing was defined as the response of the designed experiments. The main conclusions from this study can be summarized as follows:

- A quadratic mathematical model was developed to predict the pitting potential of 13Cr stainless steel as a function of pH, temperature, and chloride concentration ranges, specific to the downhole environments.
- The results of ANOVA confirmed that the developed quadratic model for the 13Cr stainless steel pitting potential has a fairly well fit with the experimental data within 95% confidence level and the error parameter value of above 0.97.
- The model optimization results confirmed that the maximum pitting potential, corresponding to the best corrosion resistance of the alloy, occurs at the lowest temperature of 22 °C, highest pH of 7, and the lowest chloride concentration of 1000 mg/L.
- The medium temperature of 51 °C, the highest chloride concentration of 22000 mg/L, and the lowest pH value of 4 resulted in the lowest pitting potential, correlated to the worst corrosion resistance of the 13Cr stainless steel.

5. Acknowledgment

The authors would like to thank the support of Suncor Energy and the Memorial University of Newfoundland for sponsoring this work.

6. References

- [1] B. Luo, J. Zhou, P. Bai, S. Zheng, T. An, X. Wen, Comparative study on the corrosion behavior of X52, 3Cr, and 13Cr steel in an O₂ – H₂O – CO₂ system: products, reaction kinetics, and pitting sensitivity, 24 (2017) 646–656. doi:10.1007/s12613-017-1447-9.
- [2] L. Calabrese, M. Galeano, E. Proverbio, D. Di, F. Cappuccini, A. Donato, Monitoring of 13 % Cr martensitic stainless steel corrosion in chloride solution in presence of thiosulphate by acoustic emission technique, 111 (2016) 151–161.
- [3] H. Zhang, Y.L. Zhao, Z.D. Jiang, Effects of temperature on the corrosion behavior of 13Cr martensitic stainless steel during exposure to CO₂ and Cl⁻ environment, 59 (2005) 3370–3374. doi:10.1016/j.matlet.2005.06.002.
- [4] A. Dalmau, C. Richard, A.I. Mu, Tribology International Degradation mechanisms in martensitic stainless steels: Wear, corrosion and tribocorrosion appraisal, 121 (2018) 167–179.
- [5] J. Jun, Localized Corrosion of Super 13Cr Stainless Steel in 1D Pit for H₂S-free and Sour Brines at Elevated Temperature Dissertation, (2016).
- [6] S. Marcelin, N. Pébère, S. Régnier, Electrochimica Acta Electrochemical characterisation of a martensitic stainless steel in a neutral chloride solution, 87 (2013) 32–40.
- [7] M. Cristina, S. Nakamatsu, A. Laura, P. Rueda, Resistance to Pitting Corrosion in Steels Based on the Fe-Cr-Ni-C System, 20 (2017) 115–119.
- [8] B. Mishra, Corrosion Characterization of Advanced Steels for Use in the Oil & Gas Industry Sweet Corrosion of Drill Pipe Steels, 2 (2013) 221–229. doi:10.5923/j.ijmee.20130202.14.
- [9] Y. Prawoto, K. Ibrahim, W.B. Wan Nik, Effect of pH and Chloride Concentration on the Corrosion of Duplex Stainless Steel, Arab. J. Sci. Eng. 34 (n.d.) 115–127.
- [10] F. Gapsari, R. Soenoko, A. Suprpto, W. Suprpto, Minimization of Corrosion Rate Using Response Surface Methodology, Eng. Rev. XX (2018) 115–119.
- [11] J.C. Velazquez, F. Caleyó, A. Valor, J. m. Hallen, Predictive Model for Pitting Corrosion in Buried Oil and Gas Pipelines, Corrosion. (2009) 332.
- [12] S. Papavinasam, A. Doiron, R.W. Reivie, V. Sizov, Model predicts internal pitting corrosion of oil, gas pipelines, Oil Gas J. (2007) 68.
- [13] J.R. Galvele, Transport Processes and the Mechanism of Pitting of Metals, (1976) 464–474.
- [14] C. Mohammad, M. Islam, Effects of Concentration of Sodium Chloride Solution on the Pitting Corrosion Behavior of AISI-304L, 17 (2011) 477–483. doi:10.2298/CICEQ110406032A.
- [15] S. Sakamoto, K. Maruyama, Corrosion Property of API and Modified 13Cr Steels in Oil and Gas Environment, in: NACE Int. Annu. Conf. Expo., 1996.
- [16] C.I. Ossai, B. Boswell, I.J. Davies, Predictive Modelling of Internal Pitting Corrosion of Aged Non-Piggable Pipelines, 162 (2015). doi:10.1149/2.0701506jes.
- [17] W. Yan, P. Zhu, J. Deng, Corrosion Behaviors of SMSS 13Cr and DSS 22Cr in H₂S / CO₂ - Oil - Water Environment, 11 (2016) 9542–9558. doi:10.20964/2016.11.13.
- [18] S. Nesic, Key issues related to modelling of internal corrosion of oil and gas pipelines – A review, 49 (2007) 4308–4338. doi:10.1016/j.corsci.2007.06.006.
- [19] J.O. Park, S. Matsch, H. Bo, Effects of Temperature and Chloride Concentration on Pit Initiation and

Early Pit Growth of Stainless Steel, (2002) 34–39.
doi:10.1149/1.1430415.

- [20] S. Richter, M. Achour, K. Addis, M. Singer, S. Nesic, M. Technology, Development and Application of a Downhole Corrosion Prediction Model, in: Int. Annu. Conf. Expo., 2016; pp. 1–15.
- [21] G.X. Zhao, M. Zheng, X.H. Lv, X.H. Dong, H.L. Li, Effect of Temperature on Anodic Behavior of 13Cr Martensitic Steel in CO₂ Environment, 11 (2005).

Study on the Wear behaviour of Aluminium Metal Matrix Composites Reinforced with Hexagonal Boron Nitride and Cubic Boron Nitride

E. Naveen^{1*}, G.B.Rohith¹, R.Sudharshan¹,
Department of Mechanical Engineering,
Sri Sairam Engineering College,
Chennai, Tamil Nadu, India

N.Ramanan²,
R&D Division,
Sharda Motors Industries Ltd,
Chengalpet, Tamil Nadu, India

Jagathish Umapathy³,
Department of Industrial Engineering,
University of Windsor,
Ontario, Canada

Corresponding Author's Mail: irumbalayar@gmail.com

Abstract—Aluminium Alloy 6061 was mixed with different %wt of Hexagonal Boron Nitride (hBN) and Cubic Boron Nitride (CBN) through Stir Casting process. The Cast Aluminium Metal Matrix Composites were subsequently machined to get the required test specimens. The experiments were conducted on a pin-on-disc Tribometer, conforming to ASTM G99 standards, on a rotating EN32 disc. Using Minitab 16 software, the Dry Sliding wear experiments were planned using L27 Orthogonal Array. The % hBN Addition, % CBN Addition, Load and Sliding Speed were taken as input parameters and Wear Rate was the output parameter. It was observed that the actual density of the Cast specimens were around more than 90% of their theoretical density. It was found that the addition of hBN and CBN significantly improves the wear resistance of the Aluminium Metal Matrix Composites. Empirical correlations for Wear Rate with respect to input parameters had been developed using Regression Analysis. The morphology of the worn out surfaces were analyzed using Scanning Electron Microscope (SEM).

Keywords- Aluminium Metal Matrix Composites and Wear

I. INTRODUCTION

Current requirement of automotive industries in to improve the fuel efficiency with improved properties like friction and wear resistance that led to the development of low cost and lightweight materials such as cylinders block, liners, piston, cam shafts, lifters, brake drum materials, etc. Selection of the right material for an engineering application depends on properties such as strength, density, light weight, melting point, conductivity, cost, etc. Density of Aluminium being one-third of the steel makes it still a better choice for high strength and low weight requirements. Though Aluminium and its alloys represent better mechanical properties, it shows degradation of Wear resistance, which limits its application in engineering. In order to overcome these limitations, addition of hard material (reinforcement) into soft (matrix) material improves the mechanical and Tribological properties, termed

as Metal Matrix Composites (MMC). MMCs are known for its combined properties of toughness and ductility of metal alloy with strength and stiffness of reinforcement. Tailoring the mechanical properties of MMC for specific application is identified by the alloy matrix, type and amount of reinforcement, and fabrication technique. The light weight matrix materials used are Aluminium, Magnesium and Titanium. Among them Aluminium and Magnesium draws the attentions as preferred matrix material due to its several attractive properties.

Aluminium Metal Matrix Composites (AMMCs) have reported better strength at high temperatures, low coefficient of friction, thermal expansion, good wear resistance, stiffness. The hardness of the AMC increase with decrease in density for the increasing wt% of reinforcements. AMMCs has found its application in aircraft, aerospace, automobiles and various other fields. The commercial application being the diesel engine piston for Toyota engineering as it offers better wear and strength at high temperature compared with Cast Iron.

II. LITERATURE SURVEY

Self-lubricating materials form a broad class of compounds by incorporating of one or more lubricants for decreasing friction and wear in industrial applications involving severe sliding contacts. There are some components that are difficult to lubricate via external lubrication, Therefore some Tribo Resistant Material must be developed in such cases, such that those materials should have Self-Lubricating capacity. When the sliding of Aluminium on steel takes place without any external lubrication, the adhesion of Aluminium to the steel takes place, thus creating on interface of low shear strength. There are chances of formation of wear debris as a result, soft aluminium surface gets ploughed by the asperities of steel and flaking of particles flakes from transfer film. The ceramic reinforcement addition reduces the friction and wear of matrix of aluminum. In response to these challenges, the successful

incorporation of solid lubricants like Molybdenum Disulfide (MoS_2), Tungsten Disulfide (WS_2), Graphite, or Boron Nitride (BN) are introduced into both matrix material for the production of self-lubricating materials.

The special bonding behaviors of Boron and Nitrogen makes the Boron Nitride exist in many different structures. The well-defined crystallographic structures are Hexagonal Boron Nitride (hBN), Rhombohedral Boron Nitride (RBN), Wurtzitic Boron Nitride (wBN), and Cubic Boron Nitride (CBN)

Hexagonal boron nitride (hBN) has unique characteristics which makes it a suitable alternative to inorganic solid lubricants, such as Molybdenum Disulfide and Graphite. hBN is also called White Graphite. It has a lamellar structure in which van der Waals forces exist between sheets of covalently bonded boron and nitrogen atoms. The lubricating performance of hexagonal boron nitride arises from the easy shearing along the basal plane. The crystalline structure of hBN shears easily along the basal plane that makes it a lubricant. hBN is used in many forms such as solid lubricant/ release/ powders. hBN is widely used in the Aluminium casting industry due to their excellent release and lubrication properties. Other properties of boron nitride are its high-temperature resistance, thermal shock resistance, high-thermal conductivity, its chemical inertness, non-toxicity and environmental safety. On the other hand Cubic Boron Nitride (CBN) is a very hard phase. It is also called Polycrystalline Diamond (PCD) as it offers high resistance to oxidation and its lesser reactivity with Ferrous materials. Also it is used as a Cutting Tool material and finds application in Wear resistant coatings. The below literature gives a brief idea about the Wear Experiments done with hBN and CBN.

Vara Prasad Kaviti et al [1] investigated the effect of Nano-Boron Nitride additions (0, 0.5, 1.5 and 2.5 wt% of Boron Nitride) on the wear behaviour of Magnesium Metal Matrix Composites that were prepared through Powder Metallurgy Technique and Hot Extrusion technique. Three Loading Conditions (5N, 7N and 10N) and Three Sliding Velocities (0.6 m/s, 0.9 m/s and 1.2 m/s) were taken as Input parameters. Wear Rate for the Mg-BN Nano-Composites were calculated for certain Sliding Distance. The Wear test was carried at Room Temperature conditions. It was inferred that the increase in sliding speed and load proportionally increases the wear rate and loss of volume. For Mg-0.5 wt% BN Nano-Composites the raise of rate of wear and loss of volume were very less compared with other compositions of Mg-BN Nano-Composites. The increase in sliding distance resulted in low Coefficient of Friction (μ). Higher Coefficient of Friction was observed at a load of 7 N compared with other working loads.

Çelik et al [2] investigated the effects of nano hexagonal boron nitride(hBN) particles on the friction and wear properties of AISI 4140 steel material when the hBN Nano particles were added to engine oil (SAE10 W) from 0 to 10%

by volume, for enhancing lubrication. Four different lubricant samples were prepared. Wear tests were conducted using ball-on-disc geometry. The experiments proved that the addition of nano hexagonal boron nitride particles to the engine oil decreased the Wear Rate upto 65% with 14.4% improvement in the Coefficient of Friction without affecting the Viscosity of Engine Oil. The asperities were completely covered by hBN Nanoparticles such that the wear-track widths and lowest wear rate were reported along with Lower value of surface roughness. This is because the Direct contact between the mating surfaces.

Bak et al [3] investigated the Wear behaviour of Hybrid Aluminium Metal Matrix Composites reinforced with Aluminium Borate whisker ($\text{Al}_{18}\text{B}_4\text{O}_{33}$, ABO), hexagonal boron nitride (BN) and carbon nanotubes (CNTs) that were fabricated through Squeeze Infiltration Technique with Al-Sn Alloy as Matrix. The wear properties of the hybrid MMCs were evaluated using a ball-on-disk tester. Four Samples with different Proportions of reinforcements were subjected to Wear Tests (Al-Sn+20% ABO, Al-Sn+20% ABO+5% hBN, Al-Sn+20% ABO+5% CNT and Al-Sn+20% ABO+5% hBN +5% CNT). The uniform distribution of the reinforcements were observed in the Microstructure. The wear resistance of the Al-Sn alloy improved with the addition of 20% ABO. The wear properties of the Al-Sn+20% ABO+5% hBN and Al-Sn+20% ABO+5% CNT further improved. The friction coefficient and wear rate of Al-Sn+20% ABO+5% hBN +5% CNT sample decreased by $1/4^{\text{th}}$ and $1/20^{\text{th}}$, respectively when compared to that of the Al-Sn+20% ABO. This Sample was found to be best among the Four Compositions as the worn out surfaces were found to be smoothest without any grooves and delamination pits.

Chen et al [4] studied the Unlubricated tribological behaviour of Silicon Nitride reinforced with various % of Hexagonal Boron Nitride (0%, 5%, 10%, 20% and 30%) with two test modes(Upper Disc on Bottom pin and Upper Pin on Bottom Disc)in Room Temperature by using a pin-on-disc tribometer. In Upper Disc on Bottom pin mode, the addition of 20% hBN to Si_3N_4 resulted in a significant decrease of the Coefficient of Friction, from 0.54 to 0.19 for Si_3N_4 -20% hBN against Si_3N_4 because of the formation of tribochemical film consisting of SiO_2 and H_3BO_3 on the worn out surfaces due to the embedment of wear debris into the spalling pits on the wear surfaces of Si_3N_4 -hBN specimen. For Upper Pin on Bottom Disc mode ,Abrasive wear mode was dominant and no tribochemical products were formed. A slight decrease of the friction coefficient from 0.85 to 0.56 for Si_3N_4 -30% hBN, against Si_3N_4 was observed due to the layered structure of hBN.

Haito Chi et al [5] studied the wear behaviour of Al2024-15% TiB_2 -15% hBN and Al2024-30% TiB_2 composites fabricated by pressure infiltration technique. With the addition of hBN particles, the TiB_2 particles were uniformly dispersed in the Aluminium matrix. The results of dry sliding tests showed

that the addition of hBN improved the tribological performance significantly especially at low sliding speed and low loading conditions, which was attributed to the role of hBN on hindering the formation of compacted layer. Three types of Wear Regimes were observed in the worn out surfaces of the Composites.

Harichandran and Selvakumar [6] prepared Aluminium Metal Matrix composites with different B_4C content (2%, 4% and 6%) with 2% h-BN nanoparticles in all the three compositions. These were compared with their Unreinforced counterpart. The Composites were prepared by Ultrasonic assisted casting. The role of the B_4C and hBN nanoparticles on the mechanical and tribological properties of the aluminium composites were evaluated. The Sonication lead to uniform distribution, grain refinement and low porosity in all the specimens. The wear properties of the hybrid Nano Composites, containing Al-4% B_4C -2 % hBN Nano particles exhibited the superior wear resistance property among the other two specimens. The wear rate increases with increasing applied load and it was minimum for hybrid Nano Composites as compared to unreinforced aluminium matrix due to the presence of h-BN in addition to hard B_4C particles which acts as solid lubricant that decreased the wear rate of the composites by forming a thin protective lubricant layer between pin and disc during sliding. This phenomenon attributed to the strengthening of matrix due to hard particle dispersion, which resulted in increase of dislocation density with increase in reinforcement content.

Ghalame et al [7] investigated the wear loss behaviour of Si_3N_4 -hBN composite against Alumina and evaluated the effect of hBN addition in Si_3N_4 to minimize the wear loss. DoE-Taguchi technique was used to plan and analyze the experimental results. L25 Orthogonal Array of 5 Level, 2 factor was chosen to perform the experiments. The Load and % hBN addition were taken as Input Parameters and the Wear Loss was taken as response. Analysis of experimental results shows the Wear Loss was minimum at 15 N load and 8 % of hBN addition. It is evident that percentage of hBN had a significant effect on wear volume loss, and combined load and percentage of hBN have a major role in wear of composite.

Wu Pan et al [8] investigated the Wear behaviour of Four different Sintered Boron carbide/Boron nitride (B_4C /hBN) ceramic composites against Gray Cast Iron disc on a Pin-on-Disc Tribometer. 4% of Al_2O_3 and 6% Y_2O_3 are constant in all four samples. The hBN was added from 0-30% with an increment of 10%. Under 5 N Loading, the Coefficient of Friction of B_4C -0% hBN and B_4C -10% hBN reached up to 0.5. However, as hBN contents reach to 20 and 30 %, the Coefficient of Friction was reduced to 0.2. Adhesive wear with severe plastic deformation of Cast Iron Disc was observed for B_4C -0% hBN composites. On the contrary, a continuous surface film is promoted on the worn-surface B_4C -30% hBN composites. With an increase in the test load, the wear mechanism of B_4C -0% hBN and B_4C -10% hBN is still adhesive wear with similar Coefficient of Friction compared to

5 N load. However, as test load increases, the film formed on B_4C -30% hBN worn surface is damaged, resulting in the rise of the friction coefficient. The friction coefficient increases to 0.6 at the load of 10 N, indicating that the addition of hBN has no improvement on lubrication in this load regime

Tyagi et al [9] studied the wear behaviour of Nickel-based self-lubricating composites with the addition of Silver and hBN by testing them up to 600°C using a unidirectional ring-on-disc configuration. Silver-containing samples performed well in terms of friction and wear. The samples containing a mixture of 12% Ag and 4 %hBN had lesser Wear resistance when compared to the ones with different %Ag additions at Higher Temperatures. The diffusion of Ag and its transfer into the Steel Disc provides low shear strength at the interface. This study shows the limitations of hBN as a High Temperature solid lubricant.

Wozniak et al [10] studied the Wear behaviour of AA6061-base composites reinforced with Graphene, Hexagonal Boron Nitride (hBN) and Molybdenum Disulphide (MoS_2). The composites were prepared by powder metallurgy route using the Spark Plasma Sintering for consolidating the powders. The increase in the reinforcement content lead to the decrease of Wear Rate and Coefficient of Friction of the Composites. The addition of solid lubricants reduced plastic deformation of the surface caused by temperature softening. The dominating mechanisms in the composites were Abrasive type. The intensity of the Plastic deformation was found to decrease after the formation of the solid film. Delamination and grooves were observed for all composites containing hBN. Deeper and Wider Grooves were observed for the composites with a higher content of hBN.

Chen et al [11] studied the wear behaviour of 663-Tin Bronze powder reinforced with Boron Nitride (BN) from 0-6 wt% in increment of 1% .The Coefficient of Friction and Wear Loss both reduced after BN was introduced in the Cu alloy matrix due to BN forming a lubricating film on the friction surface. With a BN content of 2 wt%, the sleeve showed the best combination of properties with a hardness of 55 HV, crushing strength of 302 MPa, density of 7.069 g/cm³, a relatively small friction coefficient and lowest wear loss. The new copper-based/BN self-lubricating composite material has good prospects for application in the manufacturing of frictional sleeves of high-speed sewing machinery.

Kimura et al [12] studied the lubricating properties of hBN when it was added to lubricating oil in sliding line contact experiments of Bearing Steel vs. itself and Bearing Steel vs. Cast Iron. In the first case the addition of hBN resulted in the reduction of wear at higher concentrations of hBN. The Coefficient of Friction increased slightly with the addition of hBN. The second case also the reduction of Wear was proportional to increase in hBN content. However, the Coefficient of Friction in Second case was reduced by the

addition of hBN. The presence of Boron Oxide ultimately lead to the decrease in the Wear rate in both the cases.

Podgornik et al [13] investigated the lubricating properties of the hBN and the effect of the particle size (0.5 μm , 5 μm and 30 μm) and concentration (5%, 10% and 20%) when added to a Lithium based Grease. These were compared with the Graphite containing Grease (5%, 10% and 20% with 5 μm Graphite size). Standard Tribological tests were performed through reciprocating sliding motion using a ball-on-flat contact configuration at an average sliding speed of 0.05 m/s, a maximum Hertzian contact pressure of 2N and a total sliding distance of 30 m. AA2014 disc was loaded against an oscillating steel ball with a diameter of 10 mm and an average surface roughness of 0.05 μm . The hBN particles were. By increasing the hBN particle size and concentration the friction and wear between AA2014- Mild Steel were reduced when the concentration of the grease does not exceed 10 %. The hBN powder of 30 μm size showed approximately 25 % lesser wear when compared to pure grease and graphite-containing grease. The results also show that hBN is a solid lubricant capable of replacing graphite during forming operations of Aluminium especially during Extrusion.

Carlton Reeves and Pradeep Menezes [14] conducted the pin-on-disc experiment by making Oxygen Free Copper pin sliding against Al2024 Aluminium Alloy disc to evaluate the friction and wear properties of the Avocado and Canola oil mixed with hBN particles of 5% addition with at different sizes (5 μm , 1.5 μm , 0.5 μm and 70 nm). Multi-phase lubricant, hBN particles with varying sizes (70 nm+1.5 μm , 70 nm+0.5 μm and 1.5 μm +0.5 μm) were incorporated as additives in the avocado oil. Tests were performed at 10N Loading, 36 mm/s Sliding Speed and 2100 m Sliding Distance in Room temperature conditions. The Nano-sized hBN particles offered the best tribological performance when compared to its micron- and submicron counterparts. The avocado oil with hBN particulate additives had significantly lower Coefficient of Friction values with less surface damage and reduced wear when using nanometer- and submicron-sized hBN when compared to Canola Oil added with same hBN concentration. The 1.5 μm +0.5 μm hBN combination showed best wear resistance when added to avocado oil.

Elkady et al [15] studied the wear behaviour of Nano Copper powder reinforced with 0-10% hBN with increment of 2.5% hBN addition that were prepared through Powder Metallurgy route. The Wear tests of the sintered specimens were carried out using a pin-on-ring Tribometer under dry condition. The specimens were loaded against the ring under normal loads of 10, 15, and 30 N at sliding speed of 0.2 m/s. The addition of hBN has a negative effect on the Wear Rate. But had a optimistic effect on lowering the Co-Efficient of Friction.

Rajesh Purohit et al [16] investigated the Wear behaviour of Copper Metal Matrix reinforced with 5% Al_2O_3 and CBN with different % by wt (0%,1%,2% and 3%) by powder metallurgy

route. It was found that the Composite with 3% CBN had better resistance to wear. However the increase in Load and Sliding Velocity increases the Wear Rate. Also the 3% CBN reinforced composite showed lesser Coefficient of Friction.

From the Literature Survey it was found that the Wear behaviour of Aluminium reinforced with Hexagonal Boron Nitride and Cubic Boron Nitride were not reported anywhere. A new attempt was made to study the wear behaviour of such composite with different %wt of hBN and CBN addition.

III. SPECIMEN PREPARATION

The Aluminium reinforced with hBN and CBN were fabricated through Stir Casting technique as it was found to be more economical when compared with other methods of fabrication. In this process, parent matrix alloy AA6061 was first superheated to 820°C above its melting temperature 750°C and then temperature was lowered below the liquids temperature to keep the matrix alloy in the semisolid state. The wettability of the reinforcement with the Aluminium matrix was improved by incorporating magnesium into the melt as the hBN particles would not get wet easily by the Liquid Aluminium as the contact angle is around 160°. Similarly the Graphite Crucible was also preheated upto 200°C . Extreme care was taken while adding Magnesium because of its explosive nature. It was wrapped in an Aluminium foil and then added to the molten melt. Cubic Boron Nitride reinforcements which were preheated at 700°C, was introduced into the Molten Aluminium in their respective weight fraction and mixed. The Stirring was carried at a stirring speed of 250 rpm under protected argon gas atmosphere. The Hexagonal Boron Nitride particles were then introduced into the melt. Then with a rigorous stirring at 400 rpm was continued for five minutes which is shown in Figure 3.1. After that the Composite Slurry was poured into a Cast Iron Die for solidification.



Fig 3.1: Rigorous Stirring of Composite Slurry

The Actual Density was measured using Archimedes principle using Mitutoyo™ make electronic weighing machine which has an option to measure the Actual Density. The Theoretical Density of the Composites were calculated by the Rule of Mixtures. The Actual Density of the Composite will

always be lesser than its Theoretical Counterpart because of the formations of Porosity in the Cast samples.

The Hardness of the Composites was measured by Micro Vickers Hardness tester with a pyramidal diamond indenter with the square base having an angle of 136° between the opposite diagonals according to the ASTM E 384 standard. Hardness was measured at Five Locations. The values of Density, Porosity and Hardness of the Composites are Tabulated in Table 3.1.

Table 3.1: Density, Porosity and Hardness of Al-hBN-CBN Composites

% hBN	% CBN	Theoretical Density (g/cc)	Actual Density (g/cc)	% Porosity	Micro Vicker's Hardness
0.5	1.0	2.4671	2.3798	3.54	125
0.5	2.0	2.4724	2.3950	3.13	132
0.5	3.0	2.4778	2.4294	1.95	140
1.0	1.0	2.4635	2.3273	5.53	123
1.0	2.0	2.4689	2.3609	4.37	128
1.0	3.0	2.4743	2.4035	2.86	135
1.5	1.0	2.4600	2.3402	4.87	121
1.5	2.0	2.4654	2.3413	5.03	123
1.5	3.0	2.4707	2.3475	4.99	133

IV. WEAR TESTING

In this work, a dry sliding wear behavior of Aluminium-Hexagonal Boron Nitride- Cubic Boron Nitride were studied using a pin on-disc apparatus (DUCOM™ make). Figure 4.1 shows the arrangement of pin-on-disc apparatus. The disc material was made of EN-32 steel with a hardness of 65 HRC. The pin specimen is pressed against disc at a specified load usually by means of an arm and attached weights. The dry sliding wear tests were carried out at room temperature (30°C ± 3°C, RH 55 % ± 5%) under dry sliding condition in accordance with the ASTM G 99 standard. Cylindrical pins of 10 mm diameter and 25 mm long were machined and polished by metallographic method. Immediately prior to testing, were cleaned and dried using acetone to remove all dirt and foreign matter from the specimens. Initial and final weight of the specimen was measured using a Mitutoyo™ make electronic weighing machine with an accuracy of 0.0001 grams. Wear measurement is carried out to determine the amount of materials removed (or worn away) after a wear test, (and in reality after a part in service for a period of time). The material worn away can be expressed either as weight (mass) loss, volume loss, or linear dimension change depending on the purpose of the test, the type of wear, the geometry and size of the test specimens, and sometimes on the availability of a measurement facility. The specimen was fixed in the arm and the disc was made to rotate below the pin. Mass of the specimen before and after the test is measured using digital balance and mass loss is found out for each set of input

parameters. Wear track images and SEM images were taken to analyze the maximum wear.



Fig 4.1: Pin-on-Disc Apparatus

The Experiments were conducted according to the Taguchi's L27 Orthogonal Array- 3 level and 4 Factors. The Input Parameters, their Levels and Responses are given in Table 4.1.

Table 4.1 Input parameters and Responses for Wear Test

Input Parameter	Level 1	Level 2	Level 3
% hBN Addition	0.5	1.0	1.5
% CBN Addition	1.0	2.0	3.0
Load in Newton	10	30	50
Sliding Speed (SS) in m/s	1	2	3
Sliding Distance in m	2000		
Response (Output)	Wear Rate (mm ³ /m) and Co-Efficient of Friction (CoF)		

The Wear Rate of the Composites is calculated by the formula:

$$\text{Wear Rate} = \text{Mass Loss} / (\text{Density} * \text{Sliding Distance})$$

Mass Loss is expressed in Grams. The Actual Density of the Composite is taken into account. The Sliding Distance is fixed and is given by the Formula:

$$\text{Sliding Distance} = 3.14 * d * n * t$$

d= Track Diameter in mm

n= Disc Speed in rpm

t= Time of Exposure of the Pin with Steel Disc in Minutes

The Co-Efficient of Friction was displayed in the Computer Screen which was connected to the Load cell of the Tribometer. A Graphical representation of Coefficient of Friction Vs Time was displayed. The Average of the Coefficient of Friction was taken.

V. RESULTS AND DISCUSSIONS

Table 5.1 gives the Experimental details which were planned according to Taguchi's L27 Orthogonal Array. The result of the Wear Test is given in Table 5.2. Minitab 16 software was used for further analysis. The Wear Rate and Coefficient of Friction has to be minimized

Table 5.1 Experiment Plan according to L27

Expt.No	% hBN	% CBN	Load (N)	Sliding Speed (m/s)
1	0.5	1	10	1
2	0.5	1	30	2
3	0.5	1	50	3
4	0.5	2	10	2
5	0.5	2	30	3
6	0.5	2	50	1
7	0.5	3	10	3
8	0.5	3	30	1
9	0.5	3	50	2
10	1	1	10	2
11	1	1	30	3
12	1	1	50	1
13	1	2	10	3
14	1	2	30	1
15	1	2	50	2
16	1	3	10	1
17	1	3	30	2
18	1	3	50	3
19	1.5	1	10	3
20	1.5	1	30	1
21	1.5	1	50	2
22	1.5	2	10	1
23	1.5	2	30	2
24	1.5	2	50	3
25	1.5	3	10	2
26	1.5	3	30	3
27	1.5	3	50	1

Table 5.2 Results of the Wear Test

Expt.No	Wear Rate * 10 ⁻³ (mm ³ /m)	Coefficient of Friction (CoF-μ)
1	0.4515	0.3217
2	0.4966	0.3410
3	0.5546	0.3680
4	0.3936	0.3815
5	0.4369	0.4298
6	0.4416	0.4702
7	0.3397	0.5021
8	0.3445	0.6211
9	0.3691	0.6770
10	0.3437	0.2584
11	0.3846	0.2851
12	0.3798	0.2816
13	0.2965	0.3220
14	0.3161	0.3645
15	0.3264	0.3961
16	0.2699	0.4571
17	0.2727	0.4869
18	0.2833	0.5840
19	0.3212	0.2475
20	0.3223	0.2068
21	0.3268	0.2266
22	0.2773	0.2695
23	0.2842	0.2770
24	0.2979	0.3098
25	0.2070	0.3767
26	0.2059	0.4281
27	0.2354	0.4653

The Main Effect plots of S/N Ratios for Wear Rate of the Composites is given in Figure 5.1

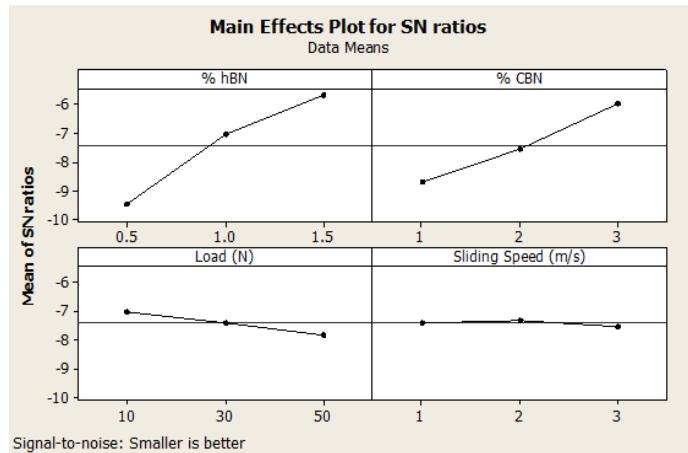


Fig 5.1: Main Effect Plots for S/N Ratios for Wear Rate

The regression equation for the Wear Rate of Al-hBN-CBN Metal Matrix Composites is given by:

$$\text{Wear Rate} = [0.64631 - (0.378127 * \% \text{ hBN}) - (0.0451929 * \% \text{ CBN}) + (0.00144872 * \text{Load}) - (0.00146963 * \text{SS}) + (0.124477 * \% \text{ hBN} * \% \text{ hBN}) - (0.0018006 * \% \text{ CBN} * \% \text{ CBN}) - (1.68476e-006 * \text{Load} * \text{Load}) + (0.00661419 * \text{SS} * \text{SS}) + (0.0172865 * \% \text{ hBN} * \% \text{ CBN}) - (0.000837525 * \% \text{ hBN} * \text{Load}) - (0.0151358 * \% \text{ hBN} * \text{SS}) - (0.000211329 * \% \text{ CBN} * \text{Load}) - (0.00854244 * \% \text{ CBN} * \text{SS}) + (0.000393297 * \text{Load} * \text{SS})] * 10^{-3}$$

$$S = 0.0109444 \quad R\text{-Sq} = 99.21\% \quad R\text{-Sq}(\text{adj}) = 98.29\%$$

Table 5.3: ANOVA table for % contribution of inputs for Wear Rate

Source	DF	Seq SS	F	P	%Contribution
% hBN	1	0.10125	82.13400	0.00000	55.749
% CBN	1	0.06167	4.69300	0.05113	33.957
Load	1	0.00550	2.67300	0.12802	3.028
SS	1	0.00038	0.00500	0.94500	0.206
% hBN* %hBN	1	0.00581	48.50900	0.00002	3.199
%CBN* %CBN	1	0.00002	0.16200	0.69404	0.010
Load*Load	1	0.00000	0.02300	0.88262	0.002
SS*SS	1	0.00026	2.19100	0.16456	0.144
%hBN* %CBN	1	0.00135	7.01600	0.02122	0.744
%hBN*Load	1	0.00133	6.58800	0.02469	0.730
%hBN*SS	1	0.00089	5.37900	0.03882	0.492
%CBN*Load	1	0.00020	1.67800	0.21958	0.111
%CBN*SS	1	0.00082	6.85400	0.02247	0.452
Load*SS	1	0.00070	5.81100	0.03287	0.383
Error	12	0.00144			0.791
Total	26	0.18161			100.00

The Main Effect plots of S/N Ratios for Coefficient of Friction of the Composites is given in Figure 5.2

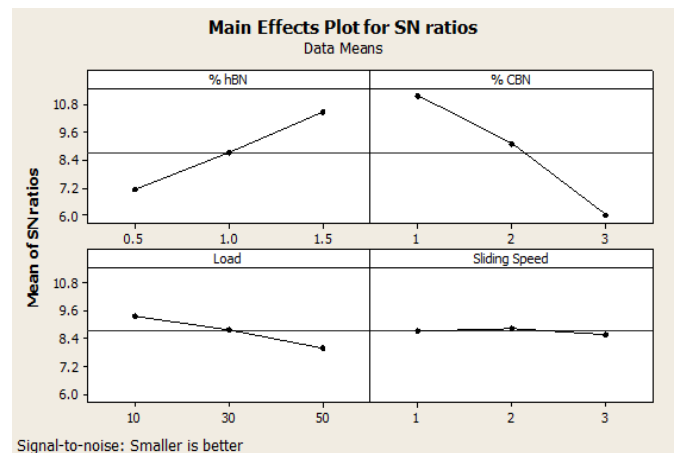


Fig 5.2: Main Effect Plots for S/N Ratios for CoF

The regression equation for the Coefficient of Friction of Al-hBN-CBN Metal Matrix Composites is given by:

$$\text{CoF} = 0.388797 - (0.0800551 * \% \text{ hBN}) - (0.0347609 * \% \text{ CBN}) - (0.0002578 * \text{Load}) - (0.0286133 * \text{SS}) + (0.0107511 * \% \text{ hBN} * \% \text{ hBN}) + (0.0385744 * \% \text{ CBN} * \% \text{ CBN}) + (4.81944e-006 * \text{Load} * \text{Load}) + (0.00510111 * \text{SS} * \text{SS}) - (0.0330071 * \% \text{ hBN} * \% \text{ CBN}) - (0.0015292 * \% \text{ hBN} * \text{Load}) + (0.0127138 * \% \text{ hBN} * \text{SS}) + (0.00134496 * \% \text{ CBN} * \text{Load}) - (0.006172 * \% \text{ CBN} * \text{SS}) + (0.000295711 * \text{Load} * \text{SS})$$

$$S = 0.0108702 \quad R\text{-Sq} = 99.63\% \quad R\text{-Sq}(\text{adj}) = 99.19\%$$

Table 5.4: ANOVA table for % contribution of inputs for Coefficient of Friction

Source	DF	Seq SS	F	P	%Contribution
% hBN	1	0.09463	3.73200	0.07733	24.814
% CBN	1	0.23613	2.81500	0.11925	61.915
Load	1	0.02290	0.08600	0.77459	6.005
SS	1	0.00002	1.90700	0.19247	0.005
% hBN* %hBN	1	0.00004	0.36700	0.55602	0.011
%CBN* %CBN	1	0.00893	75.55800	0.00000	2.341
Load*Load	1	0.00002	0.18900	0.67170	0.006
SS*SS	1	0.00016	1.32100	0.27275	0.041
%hBN* %CBN	1	0.00271	25.93200	0.00027	0.710
%hBN*Load	1	0.00340	22.26400	0.00050	0.892
%hBN*SS	1	0.00205	3.84700	0.07344	0.539
%CBN*Load	1	0.00814	68.89000	0.00000	2.134
%CBN*SS	1	0.00043	3.62700	0.08110	0.112
Load*SS	1	0.00039	3.33000	0.09300	0.103
Error	12	0.00142			0.372
Total	26	0.38137			100.000

Table 5.3 shows the Analysis of Variance (ANOVA) table for % contribution of inputs for Wear Rate. From the Figure 5.1, it is inferred that the Wear Rate is found to be minimum at 1.5% hBN addition, 3% CBN addition, 10N Load and 1 m/s Sliding Speed. From Table 5.3, it can be inferred that the %hBN addition contributes around 55.749% followed by %CBN addition 33.957%. The %hBN, Square of %hBN, product of %hBN & Load, %hBN & %CBN, %hBN & SS, %CBN & Load and %CBN & SS were found to be significant factor influencing the Wear Rate of AA6061-hBN-CBN Composites.

Table 5.4 shows the Analysis of Variance (ANOVA) table for % contribution of inputs for Coefficient of Friction. From the Figure 5.2, it is inferred that the Coefficient of Friction is found to be minimum at 1.5% hBN addition, 1% CBN addition, 10N Load and 1 m/s Sliding Speed. The %CBN addition contributes around 61.915% followed by %hBN addition 24.814%. The Square of %CBN, product of %hBN & Load, %hBN & %CBN and %CBN & Load were found to be significant factor influencing the Coefficient of Friction of AA6061-hBN-CBN Composites.

Effect of Reinforcement addition: The lowering in wear rate of composites can be raised due addition of Hexagonal Boron Nitride (Self-Lubricating) and Cubic Boron Nitride in a soft and ductile matrix which thereby increases the hardness and hence reduces the wear rate of the matrix alloy. The thermal mismatch (difference in coefficient of thermal expansion) between the matrix alloy and the reinforcements results in an increase in the dislocation density. The resistance of the materials to plastic deformation and reduction the friction during sliding motion is due to increase in dislocation density thereby preventing the propagation of Crack during wear. The hardness of the Composites is found to be increasing with % CBN addition. But as the %hBN addition is increased, the Hardness of the Composite is found to decrease. The strengthening of the metallic matrix due to the solubility of hBN solid lubricant into the metallic matrix, in particular through diffusion, thereby reducing the degree of continuity of the metallic matrix. The stacking of the friction products, which come from the shearing of hBN, resulted in a formation of Surface film. The hard loose asperities leads to the formation of tribolayer film around them by reducing the impact of micro cutting at higher loading conditions. Stacked friction products were pulled over leading to Oxidation reaction such that a Smooth Film was formed thereby resulting in decrease in Coefficient of Friction. But on the contrary the addition of CBN lead to the increase of Coefficient of Friction. This is because it is a harder than hBN such that it protects the Matrix from Thermal softening while it is present in the interface region ultimately reducing the Wear Rate. As it contacts the Steel Disc, the Coefficient of Friction between them gets increased.

Effect of Applied Load: It may be observed that the as the application load increases, the rate of wear also increases that can be explained by basis of Archard's theory. The amount of sliding wear is proportional to the applied load and the sliding distance and inversely proportional to the hardness of the surface being worn away. The penetration of hard asperities from the counter surface into the softer pin surface increases with the plastic deformation by fracturing the softer surface. At higher loads, the thermal softening and severe plastic deformation prevails as shown in Figure 5.3. The transfer material from the pin to disc results in the Debris that get welded to the Steel Disc as shown in Figure 5.4 called Galling Effect.

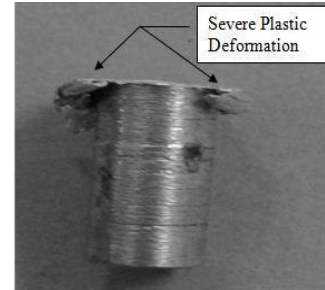


Figure 5.3: Severe Plastic Deformation

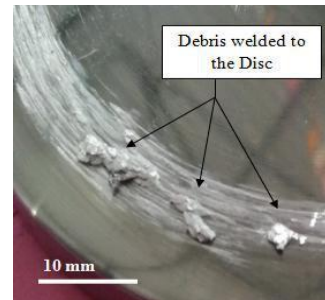


Figure 5.4: Galling Effect

As the load increases, the movement of dislocations past the reinforcements is by Orowan over looping mechanisms. Thus, the matrix strengthening by CBN enhances the load bearing ability of the Aluminium MMC. The pile-up dislocations at the grain boundary and the hBN & CBN particles in the sub surface raise the stress concentration above the fracture stress of the Aluminium matrix initiating a delamination crack at the particle/matrix interface by a shear deformation at the sub surface

Effect of Sliding Speed: It is proved from the results that the sliding speed influences Wear Rate with the combination of %hBN addition, %CBN addition and Applied load where the frictional heat generated between pin and the steel disc is influenced. The degree of softening and the frictional heat are directly proportional to each other, where the penetration of the reinforcements from the pin to the counter plate takes place. But the increase in the sliding speed increases the wear rate. With increase in Sliding Speed, rapid removal of oxide scales takes place and exposes direct metal to metal contact for

fresh oxidation. The wear rate of the alloy increases with an increase in speed. As the Sliding Speed increases, temperature at the interface rises resulting in softening of the surface and sub surface of the wear surface. The absence of reinforcements in the Aluminium alloy accelerated the softening effects with the increasing sliding speed.

SEM Analysis Wear surface morphology of Aluminium Metal Matrix Composites:

After the wear test, the worn-out surfaces of some specimens were observed under Scanning Electron Microscope (SEM). Figure 5.5 a, b and c shows the surface morphology of the hybrid Composite of AA6061-1.5% hBN-3% CBN tested at low load i.e. at 10 N. Figure 5.5a clearly exhibits the formation of which permanent and deep and permanently grooves leading to increase of wear loss. Figure 5.5b shows the worn out surfaces of the exhibited finer grooves and slight plastic deformation near its edge appearing to be smooth because of the presence of the hBN reinforcement. Also the width between the grooves is also increased. When the applied load is smaller, fracture of the reinforcements occurs along the crack lines. It is observed that cavities are formed in the composite matrix and have aligned parallel to the direction of sliding. Particles chopped off during sliding were observed in Figure 5.5c. Also the CBN is clearly visible in this figure.

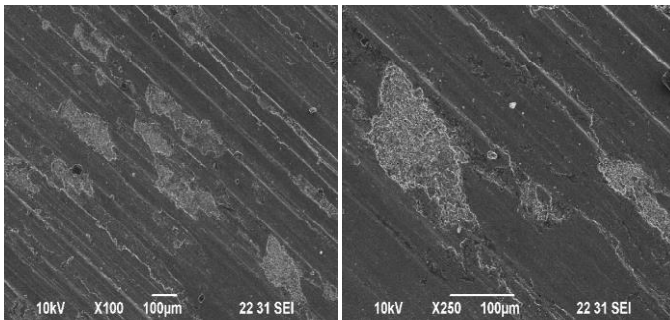


Fig5.5 a: Formation of Deep grooves

Fig5.5 b: SEM Image worn out Surface near the edge

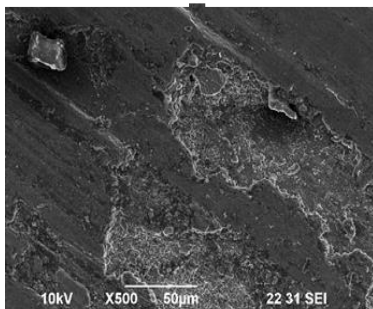


Fig5.5 c: Particle chopping during sliding

Figure 5.6 a, b and c shows the SEM images for the worn out surfaces of AA6061-1.5% hBN-3% CBN 50N Load. With increase in load to 50 N worn surface shows a different appearance. In Figure 5.6a the amount of cavitations is more than that of the previous case. The ploughing action of hard asperities on the counter disc is observed. It can be visualized that when the sample is rubbed, against steel disc, at higher sliding speeds, the hard particles might have chipped off such that the matrix material smeared along the direction of the sliding. Amount of cavitations also have increased. Some cavities appear to be formed around the hard particles, (i.e. CBN particulates). Figure 5.6b shows the presence of Matrix breakage where large plastics strains causing subsurface crack propagation and subsurface delamination. From the Figure 5.6 c shows the expulsion of reinforcements from the aluminium matrix with increasing the load that caused embrittlement of hard particles during sliding.

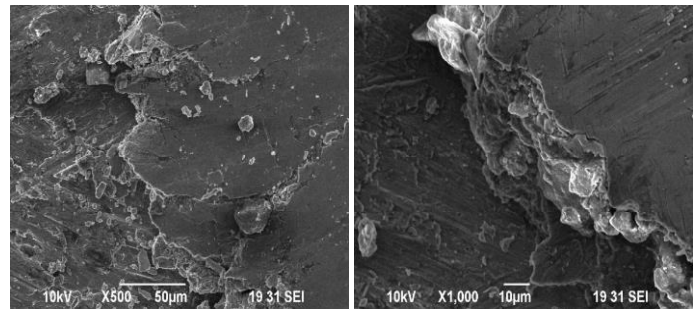


Fig5.6 a: Ploughing & Smearing

Fig5.6 b: Crack propagation at subsurface

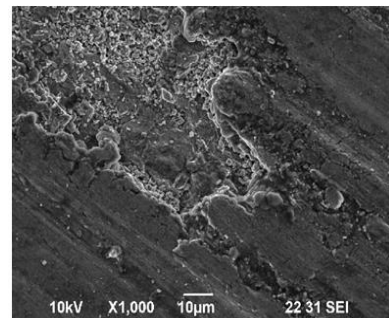


Fig5.6 c: Expulsion of Reinforcements

VI. CONCLUSION

The wear resistance of the Aluminium AA6061 is increased by addition of Hexagonal Boron Nitride and Cubic Boron Nitride. Among these two, the contribution of Hexagonal Boron Nitride is more than its Cubic structured counterpart. Irrespective of the composition as the Load increases Wear resistance decreases and Mass Loss increases. The hBN reinforcements removed forms a thin tribo film. It prevents the direct contact between the metal and the sliding surface, also prevents the breaking of CBN particles. Also these

reinforcements act as a barrier preventing the movement of dislocation that appears as small patches. The highest wt% of hBN and CBN resulted in the decreased material removal during wear. It can be noted that the interesting part of this wear study is deformation resistance offered by CBN, which fracture into small pieces producing wear debris particle. This debris prevents the disk from penetrating into the metal protecting the soft aluminium matrix thus increasing the wear resistance. Confirmation Experiments were carried with the Empirical Equation obtained. The deviation from the Actual Readings were less than 5%.

REFERENCES

- [1] **R.Vara Prasad Kaviti, D.Jeyasimman, GururajParande , Manoj Gupta , R. Narayanasamy:** “Investigation on dry sliding wear behavior of Mg/BN nanocomposites”, *Journal of Magnesium and Alloys* 6 (2018) 263–276
- [2] **O. N. Çelik ,N. Ay & Y. Göncü:** “Effect of Nano Hexagonal Boron Nitride Lubricant Additives on the Friction and Wear Properties of AISI 4140 Steel”, *Particulate Science and Technology, Volume 31, 2013 - Issue 5*
- [3] **Ji HyeonBak, Dae Hyun Cho, Sunmi Shin, Jin Young Park and Ik Min Park:** “Wear properties of hybrid ABO+BN+CNT/Al-Sn alloy matrix composites for engine bearing materials”*Met. Mater. Int.* (2018) 24: 205. (<https://doi.org/10.1007/s12540-017-7236-1>)
- [4] **Wei Chen, Yimin Gao, Yong Wang and Huangqiang Li:** “Tribological Behavior of Si₃N₄-hBN Ceramic Materials without Lubrication under Different Test Modes”, *Tribology Transactions Volume 53, 2010 - Issue 6 Pages: 787-798*
- [5] **Haitao Chi, Longtao Jiang, Guoqin Chen, PengchaoKang,Xiu Lin and Gaohui Wu:** “Dry sliding friction and wear behavior of (TiB₂ + h-BN)/2024Al composites”, *Materials & Design*Volume 87, 15 December 2015, Pages 960-968
- [6] **R. Harichandran, and N. Selvakumar:** “Effect of h-BN solid nano lubricant on the dry sliding wear behaviour of Al-B₄C nanocomposites”, *Archives of Materials Science and Engineering* 77/1 (2016) 5-11.
- [7] **Ghalme.S, Mankar.A. and Bhalerao.Y.J:** “Optimization of wear loss in silicon nitride (Si₃N₄)–hexagonal boron nitride (hBN) composite using DoE–Taguchi method”, *SpringerPlus* (2016) 5: 1671. (<https://doi.org/10.1186/s40064-016-3379-7>)
- [8] **Wu Pan, Yimin Gao, Xiuqing Li, Shanshan Wu, Liancheng Song and ZhichaoZhong:** “Tribological Behavior of B₄C/hBN Ceramic Composites Sliding Against Gray Cast Irons Without Lubrication” *Tribology Letters* (2015) 60: 10. <https://doi.org/10.1007/s11249-015-0579-1>
- [9] **Rajnesh Tyagi, Dang Sheng Xiong, Jian Liang Li, Jihuai Dai:** “High-temperature friction and wear of Ag/h-BN-containing Ni-based composites against steel”, *Tribology Letters* 2010;40:181–186.
- [10] **Wozniak J, Kostecki M, Cygan T, Buczek M, Olszyna A:** “Self-lubricating aluminium matrix composites reinforced with 2D crystals”, *Composites Part B: Engineering, Volume 111, 15 February 2017, Pages 1-9*
- [11] **Suiyuan Chen, Yuning Bi, Honlon Zhang, Jing Liang, Dan Wellburn, Liu Chang-sheng:** “Effect of BN fraction on the mechanical and tribological properties of Cu alloy/BN self-lubricating sleeves”, *Journal of Composite Materials : Vol 49, Issue 30, 2015*
- [12] **Yoshitsugu Kimura , Toshiaki Wakabayashi, Kazumi Okada, Tetsuya Wada, Hiroshi Nishikawa:** “Boron nitride as a lubricant additive”, *Wear* 232 (1999) 199–206
- [13] **B. Podgornik, T. Kosec, A. Kocijan, Č. Donik :** “Tribological behaviour and lubrication performance of hexagonal boron nitride (h-BN) as a replacement for graphite in aluminium forming”, *Tribology International, Volume 81, January 2015, Pages 267-275*
- [14] **Carlton J. Reeves and Pradeep L. Menezes:** “Evaluation of boron nitride particles on the tribological performance of avocado and canola oil for energy conservation and sustainability”, *International Journal of Advanced Manufacturing Technology* (2017) 89: 3475. (<https://doi.org/10.1007/s00170-016-9354>)
- [15] **Omayma A. M. Elkady, Ahmed Abu-Oqail, Emad M.M. Ewais, M. El-Sheikh:** “Physico-mechanical and tribological properties of Cu/h-BN nanocomposites synthesized by PM route”, *Journal of Alloys and Compounds, Volume 625, 15 March 2015, Pages 309-317* (<https://doi.org/10.1016/j.jallcom.2014.10.171>)
- [16] **Rajesh Purohit, Navin Kumar Solanki, Gaurav Bajpayee, R.S.Rana, G.Hemanth Kumar, Raman Nateria:** “Development of Cu-Al₂O₃-CBN Hybrid Composite through Powder Metallurgy Process and Analysis of Mechanical Properties” *Materials Today Proceedings, Volume 4, Issue 2, Part A, 2017, Pages 3270-3279*

Effect of Plastic Deformation on the Corrosion Behavior of 13Cr Stainless Steel

Salar Salahi, Mostafa Kazemipour, Ali Nasiri

*Faculty of Engineering and Applied Science,
Memorial University of Newfoundland,
St. John's, NL, A1B 3X5, Canada*

Abstract— Although several investigations have been carried out to understand the effects of environmental factors and the inherent properties of Oil Country Tubular Goods (OCTG) materials on their corrosion behavior in downhole environments, the correlation between the manufacturing process induced plastic deformation in the material and its corrosion behavior has never been comprehensively studied. As one of the most commonly used materials for casing and tubing in downhole environment, 13Cr martensitic stainless steel was selected for this study and comparisons were made between a texture free heat-treated sample and preferentially-textured plastically deformed samples in terms of microstructure, crystallographic orientation, and corrosion performance. Cyclic potentiodynamic polarization tests were performed on all samples in aerated 3.5 wt.% NaCl electrolyte at room temperature to characterize the corrosion resistance. Scanning electron microscopy (SEM) and electron backscatter diffraction (EBSD) were utilized to analyze the microstructure and crystallographic orientation of all samples. The galvanic coupling between the existing carbide precipitates in the alloy and their surrounding matrix was found to be the dominant factor in controlling the electrochemical behavior of the alloy in both non-deformed and plastically deformed conditions. Cold deformation of 13Cr stainless steel was found to improve the pitting resistance of the alloy. As the deformation level increases, the breakage of coarse brittle carbide precipitates into smaller idiomorphic particles improves distribution uniformity of small carbide particles in the matrix, leading to a reduced galvanic potential difference between the precipitates and the matrix.

Keywords: plastic deformation, crystallographic orientation, electrochemical testing, corrosion, 13Cr stainless steel

I. INTRODUCTION

Due to the increased development of deep hot wells and their extremely harsh and corrosive environments applications of corrosion resistant martensitic stainless steels (MSS) have been increased drastically considering their economic benefit with a price nearly half of its duplex stainless steels

counterparts [1]. The usage of MSS tubing has not been limited to only sweet applications, but also has been expanded to numerous sour service applications. The increasing demand in oil and gas industries has promoted the exploitation of materials for higher temperature applications at various chloride concentrations, and even in sour environments [2-4].

The fabrication process of martensitic stainless steels typically involves warm or cold deformation of the annealed material followed by applying a post-deformation heat treatment cycle, commonly hardening and tempering, leading to formation of different types of precipitates embedded in a ferritic/martensitic matrix. While annealing leads to formation of localized spherical carbides in the microstructure, hardening and subsequent tempering form undissolved carbides along with fine precipitated carbides [5-7]. It is well reported that the martensitic transformation associated with the hardening of MSS can be accompanied with formation of undissolved carbides, unwanted retained austenite and in some scale with the ferrite phase [2]. Although the carbide dissolution can be accelerated by increasing austenitizing temperature, the accompanied austenite grain coarsening during austenitization that can potentially lead to stability of some volume fraction of austenite at room temperature (the retained austenite) can deteriorate the mechanical properties of the alloy. The attained martensitic phase from the hardening cycle possess high hardness and strength with extremely low ductility. Therefore, to improve the toughness and ductility and to relieve the residual stresses induced by the martensitic transformation, the tempering process is applied [8]. Each of the individual constituents in the microstructure of MSS, *i.e.* carbides, retained austenite, ferritic and martensitic phases can significantly affect both mechanical and corrosion properties of the alloy [5, 7].

Extensive studies have been done to characterize the mechanical properties of MSS and to discuss the effect of tempering conditions on the microstructure and corrosion properties. Lu *et al.* [9] studied the effect of tempering temperature on the corrosion properties of 13Cr alloy and have reported a higher corrosion resistance for samples tempered at 300 °C relative to the samples tempered at 500 °C due to the sensitization of the matrix adjacent to the chromium carbide

precipitates. On the other hand, few studies have been done to develop a correlation between the microstructure and corrosion properties of MSS. Anatha *et al.* [2] studied the effect of various types of precipitates in the microstructure on the localized corrosion for the AISI 420 martensitic stainless steel and reported that regions adjacent to the carbides are less noble due to formation of chromium depleted zones at the periphery of the chromium carbides.

It is known that the applied plastic deformation during cold working can modify the microstructure and grain orientation drastically by turning a texture free microstructure into a strongly textured one after deformation. Consequently, the electrochemical stability of the passive layer on the stainless steel can potentially change. In a recent study on the API X46 steel, it has been reported that the preferred grain oriented with plane {111} parallel to the normal direction ({111}//ND) increased the corrosion resistance [10, 11]. Previous studies revealed that the effect of texture on the corrosion properties is significantly dependent on the slip system. For instance, the crystallographic textures associated with {110}//ND and {332}//ND reduce the sensitivity to corrosion resistance in ferritic pearlite steels, while {001}//ND and {112}//ND textures have the opposite effect [12]. In case of austenitic stainless steels, the corrosion behavior after deformation has been discussed based on the developed modifications in the structure of the passive film, *i.e.* its thickness and composition, while contradictory results have been reported in case of pitting corrosion. Favorable passivation was reported for the rolled 304 stainless steel due to formation of a thicker passive film, whereas for a high nitrogen 30Cr15Mo1N stainless steel, a passive layer with a less protective nature has been reported to form after deformation [13, 14].

To the best of our knowledge, no studies has been done to characterize the effect of plastic deformation on the corrosion properties of MSS. Characterizing the impact of plastic deformation associated with the manufacturing process on the corrosion property of the alloy is extremely vital to be able to oversee the components' integrity and remaining life after a long service life particularly in harsh environments. In particular, the effect of plastic deformation on both the grains orientation and distribution of precipitates and the resultant corrosion behavior are required to be investigated. It is well established that the galvanic potential difference between the micro-constituents and their surrounding matrix can provoke micro-galvanic coupling, deteriorating the corrosion performance[15, 16]. In this study, the impacts of deformation on both the microstructure and corrosion behavior of 13Cr stainless steel samples were investigated and compared with the deformation-free samples.

II. EXPERIMENTAL PROCEDURE

A. Sample preparation

A 13Cr martensitic stainless steel with the measured composition summarized in Table 1 was used in this study. The material of study was sectioned from a long 13Cr tube that was

austenitized at 1020 °C for 30 min in a vacuum furnace, followed by rapid cooling. The material was then hot deformed and tempered at 250 °C, followed by cooling in air.

To induce plastic deformation in the samples and create a microstructure containing a stronger texture, the samples were plastically deformed under uniaxial tension up to elongations of 5% and 19% (approximately fracture elongation). The sample preparation procedures followed by the corrosion testing are illustrated schematically in Figure 1.

TABLE I. THE MEASURED CHEMICAL COMPOSITION (WT %) OF 13CR STAINLESS STEEL USED IN THIS STUDY

C	Si	Mn	Cr	V
0.4	0.96	0.47	13.8	0.3

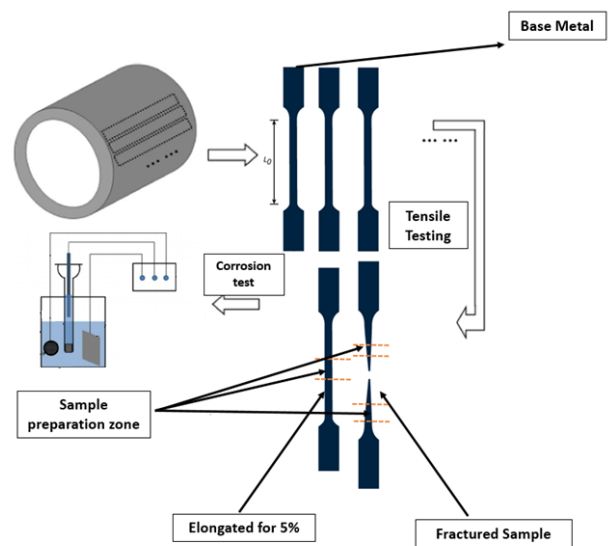


Figure 1. Schematic of the sample preparation procedure herein for the subsequent corrosion testing.

B. Electrochemical analysis

The exposed surface of the samples was ground using 500 grit sandpaper, ultrasonically cleaned in acetone, followed by polishing. All electrochemical measurements in this study were performed in aerated 3.5 wt.% NaCl solution. The solution's temperature was maintained at 25±0.5 °C using a temperature-controlled water bath. A standard three-electrode cell set-up containing a graphite rod as the counter electrode, a saturated Ag/AgCl electrode as the reference electrode, and the 13Cr stainless steel specimen as the working electrode. The open circuit potential (OCP) was monitored for one hour prior to each test. Cyclic potentiodynamic polarization (CPP) tests were performed at a scanning rate of 0.5 mV/s, commencing at -0.2 VAg/AgCl vs OCP. The CPP experiments for each condition were conducted at least 3 times to ensure the reproducibility and reliability of the results.

C. Microstructural characterization

To reveal the microstructural characteristics of the samples, samples were cut and mounted in an epoxy resin from both the non-deformed sample and the gauge length of the deformed samples. The mounted samples were prepared following standard grinding and polishing procedures for the MSS materials. Vilella's reagent was used for etching of the polished surfaces. Microstructural details of the samples were revealed using a FEI MLA 650F scanning electron microscope (SEM) equipped with a high throughput Bruker energy dispersive X-ray (EDX) analytical system and an HKL EBSD system. EBSD analysis was also performed on polished samples to acquire crystallographic orientation micrographs (inverse pole figures (IPF)), pole figures (PF), and grain boundary maps. Channel 5 software was used to post process and analyse the EBSD data. XRD analysis was also performed at 40 kV and 35 mA with an aperture of 2 mm over a diffraction angle range of 40° - 90° with a step size of 0.01° . Peak positions of index data were compared with those obtained from standard powder diffraction file (PDF) cards.

III. RESULTS

A. Mechanical properties characterization

Figure 2 shows the engineering stress-elongation diagram along with the hardness variation of the alloy at different levels of plastic strain. Accordingly, the base metal revealed the yield strength (YS) of 365 ± 6 MPa and ultimate tensile strength (UTS) of 769 ± 4 MPa. Homogenous plastic deformation has occurred around $1 \pm 0.15\%$ elongation and inhomogeneous plastic deformation around $10 \pm 0.20\%$ elongation. As depicted in Figure 2, the plastically deformed samples for the purpose of electrochemical testing were prepared from 5% elongated and the fractured samples possessing the largest possible plastic strain.

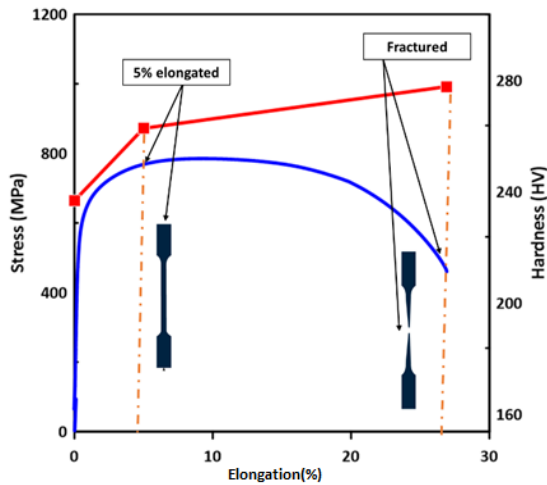


Figure 2. The stress-strain curve and the variation of hardness values at different strain levels, indicating the degree of plastic deformation in the deformed samples prior to the corrosion testing.

The variation of Vickers hardness data indicates that by increasing the deformation level from 5% to 19%, the hardness increases from 245 ± 1 HV to 280 ± 1 HV, attributed to the increased level of strain hardening in the alloy.

B. Microstructure Characterization

Figure 3 shows the XRD results for the base metal. Three peaks corresponding to the martensite phase (M) were observed for the base metal in angles ranging from 40° - 90° , introducing martensite as the dominant phase in the microstructure. The highest intensity peak of $(110)_{\text{martensitic}}$ was observed in the 44. degrees. Lower intensity peaks of $(200)_{\text{martensitic}}$ and $(211)_{\text{martensitic}}$ were also detected. Previous studies have also detected peaks corresponding to the formation of retained austenite and undissolved carbides in the XRD spectra [2]. However, none of those phases were detected on the XRD spectra herein, plausibly due to their drastically lower volume fraction as compared to the martensite phase.

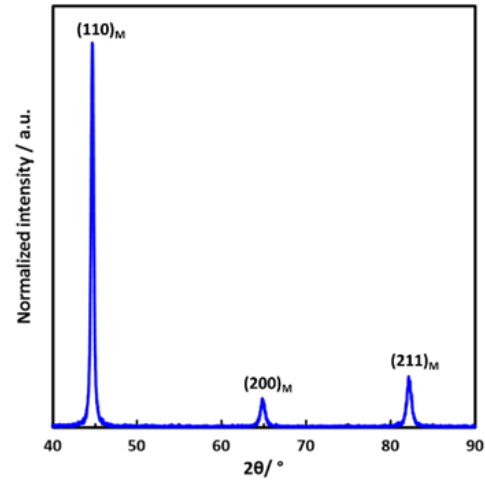


Figure 3. XRD results showing the presence of martensitic phases

Figure 4 illustrates the SEM micrographs taken from the prepared samples at different deformation levels. As shown in Figure 4a, the base metal sample exhibited a high number of martensitic lathes originated from the hierarchic tempered martensite. A high volume fraction of precipitates was also observed in the structure of the base metal. Precipitates were mainly coarse carbides distributed homogeneously in the Fe matrix and within the grain boundaries along with fine dissolved carbides with the diameter of nearly 200 nm propagated through grain boundaries possibly generated during the tempering process. The geometry of coarse undissolved precipitates is nearly round-shaped, while the fine precipitates have capsule-like shapes [2]. The EDX elemental mapping of the microstructure revealed higher concentration of chromium and carbon for the precipitates, where Fe was depleted, indicating the chromium carbide origin of the generated precipitates. Analogous to the non-deformed sample, the 5% elongated sample revealed a high-volume

fraction of undissolved carbides along with dissolved carbides in the matrix.

The fractured samples (Figures 4c and 4d) showed the preferential distribution of both dissolved and undissolved carbides along the loading direction. Breakage of some undissolved precipitates during the high level of deformation were evident at higher magnification in the fractured sample (shown in Figure 4d), while the preferential orientation of the dissolved carbides along the loading direction and expanded interspaces between the precipitates were observed.

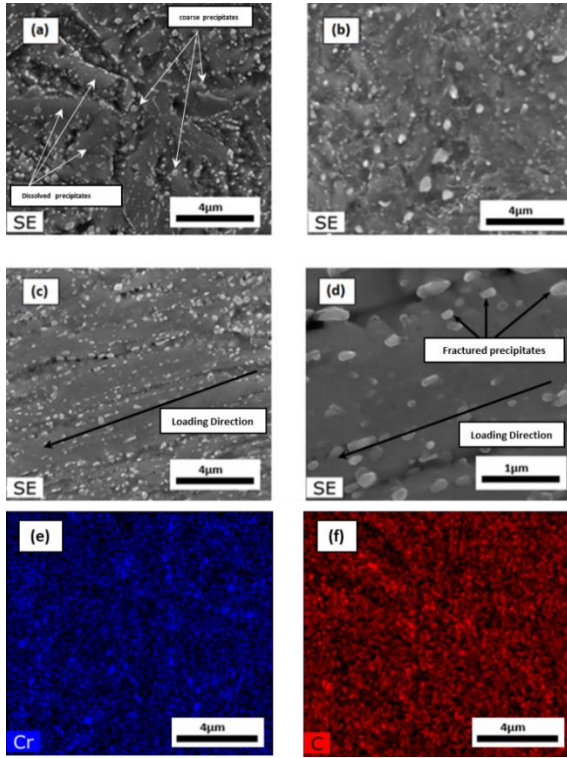


Figure 4. SEM micrographs of the (a) base metal, (b) 5% elongated sample, (c) and (d) fractured sample in different magnifications.

Figure 5 shows the inverse pole figures (IPFs) from the microstructure of the base metal, the 5% elongated sample, and the fractured sample, revealing an extremely fine grain structure (average grain size $< 1 \mu\text{m}$) in all samples. A slight grain refinement was detected in the samples after deformation. The average grain size in the non-deformed base metal was $0.79 \mu\text{m}$, and it decreases to $0.76 \mu\text{m}$ and $0.64 \mu\text{m}$ for the 5% deformed and fractured sample, respectively. Ueji *et al.*[17] discussed the grain refinement of low carbon steel during cold deformation by rolling and related the formation of ultra-fine-grained structure to characteristics of the martensite starting structure. Bracke *et al.*[18] investigated the effect of cold deformation on the texture and microstructure of the austenitic Fe–Mn–C alloy and reported that the mechanism behind this phenomenon was based on the nucleation and growth of the recrystallized state in an energetically relatively homogeneous microstructure.

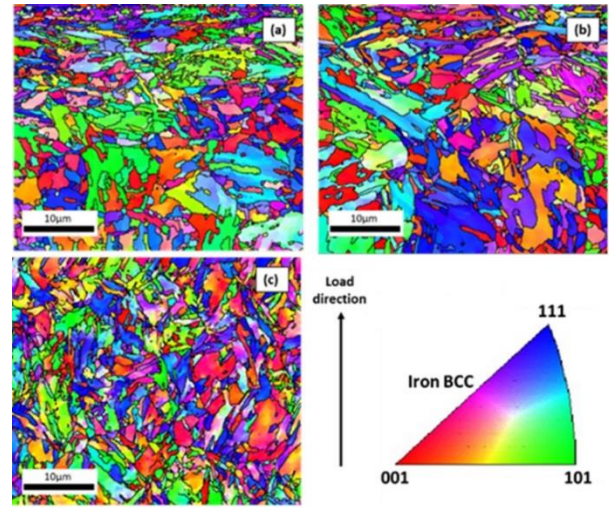


Figure 5. EBSD inverse pole figure maps of the (a) base metal, (b) 5% elongated sample, and (c) fractured sample.

Figure 6 shows the recrystallization fraction for samples during the deformation process. It is seen that recrystallization fraction and the size of recrystallized grains decreases significantly by increasing the deformation level. While the base metal has a high fraction of recrystallized grains generated plausibly during the tempering process, the size and fraction of recrystallized grains decrease drastically in the fractured sample. The highest volume fraction of deformed grains was detected in the fractured sample while the deformation level is nearly the same in the 5% deformed and the base metal microstructure. It can be concluded that the severe deformation is needed to demonstrate the deformed grains in the obtained texture results. The deformation of the grains and their orientation along the applied load can be seen in the IPF of the fractured sample shown in Figure 5c.

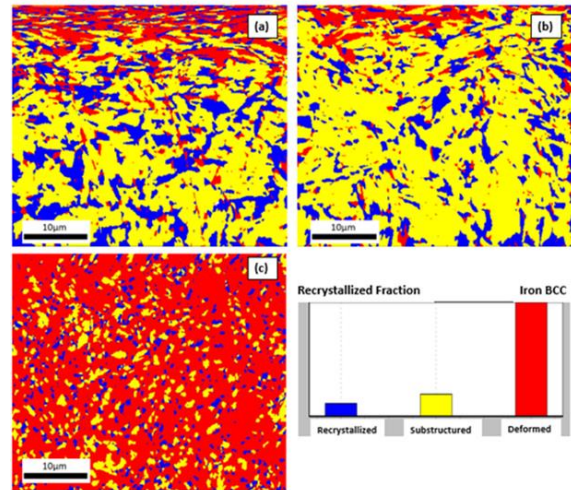


Figure 6. Recrystallization factor maps for the (a) base metal, (b) 5% elongated sample, and (c) fractured sample.

Figure 7 exhibits the distribution of high-angle grain boundaries (HAGBs), medium-angle grain boundaries (MAGBs), and low-angle grain boundaries (LAGBs) in all samples. While the fraction of LAGBs are nearly the same for the base metal and the 5% elongated sample, high fraction of LAGBs were observed in the fractured samples. Fraction of HAGBs remains relatively constant in all samples with various deformation levels.

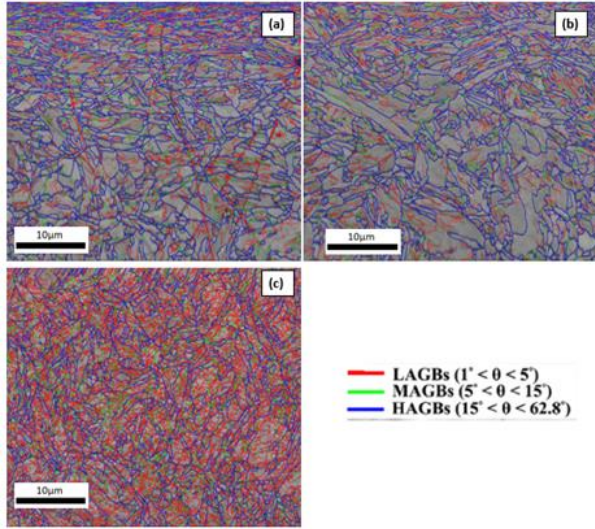


Figure 7. EBSD grain boundary maps showing distribution of HAGBs, MAGBs, and LAGBs in the (a) base metal, (b) 5% elongated sample, and (c) fractured sample.

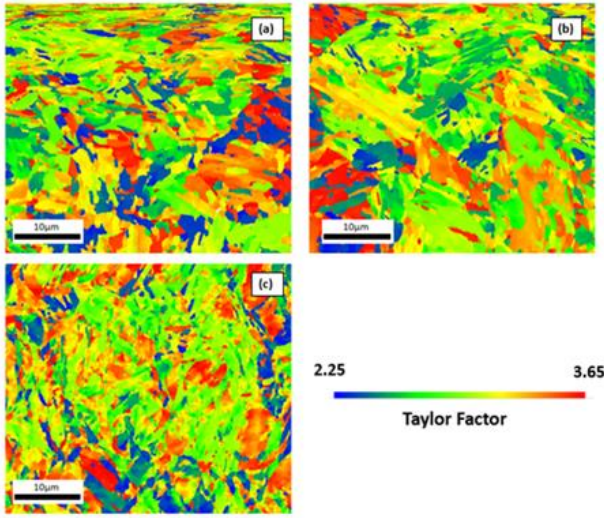


Figure 8. Taylor factor maps of the (a) base metal, (b) 5% elongated sample, and (c) fractured sample.

Taylor factor displays a correlation between the yield strength and misorientation of the grains and is useful for characterizing the level of plastic deformation in the samples [19, 20]. The corresponding Taylor factor distribution maps for the samples in this study are shown in Figure 8. It is clearly visible that the non-deformed base metal and 5% elongated

samples contained grains with higher Taylor factors, whereas the volume fraction of grains with high Taylor factor decreases in the fractured sample. Grains with higher Taylor factor are known as hard grains, characterized by having lower tendency to deform, while grains with lower Taylor factor tend to deform easier and are denoted as soft grains [19]. As shown in Figure 8c, a more uniform distribution of Taylor factor in the grains of fractured sample along with a higher density of soft grains were detected in the fractured sample.

The pole figures of the samples obtained from the EBSD data are shown in Figure 9. The active slip planes in the BCC materials are generally $\{110\}$ and $\{111\}$ families of planes while the $\{100\}$ planes are responsible for phase transformation textures resulted from processes, such as heat treatment [20]. The $\{110\}$ and $\{111\}$ planes for the base metal show random textures with an intermediate intensity along the generated poles. The fractured sample depicted a strong texture at the $\{110\}$ plane with one pole concentrated along the direction normal to the loading direction. An analogous texture distribution is reported in BCC materials during the rolling process [21]. The 5% elongated sample revealed an intensity aligned with the direction normal to the loading direction in $\{111\}$ planes.

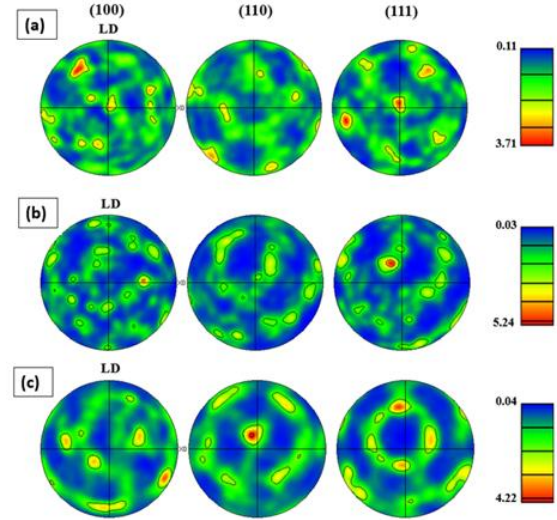


Figure 9. Pole figures (PFs) for the (a) base metal, (b) 5% elongated sample, and (c) fractured sample.

C. Corrosion Results

The open-circuit potential (OCP) trends for all samples are displayed in Figure 10. A significant fluctuation of OCP was recorded for the base metal and 5% elongated samples from the onset through the whole measurement. Anantha *et al.* [2] showed that fluctuation of the OCP represents a continuous occurrence of metastable localized (pitting) corrosion during the OCP monitoring. The intensity of fluctuations remains constant over the entire measurement period in the base metal, while for the 5% elongated sample, it diminishes after 2000 s. Contrarily, the fractured sample revealed significantly more

stable OCP over the monitoring time relative to the base metal and 5% elongated sample, indicating the highest stability of the passive film on the fractured sample. The measured OCP value for the fractured sample was $-0.13 \text{ V}_{\text{Ag/AgCl}}$, while it decreases to $-0.15 \text{ V}_{\text{Ag/AgCl}}$ and $-0.17 \text{ V}_{\text{Ag/AgCl}}$ for the 5% deformed and the non-deformed base metal samples, respectively, confirming the improved stability of the passive film by increasing the degree of plastic deformation in the sample.

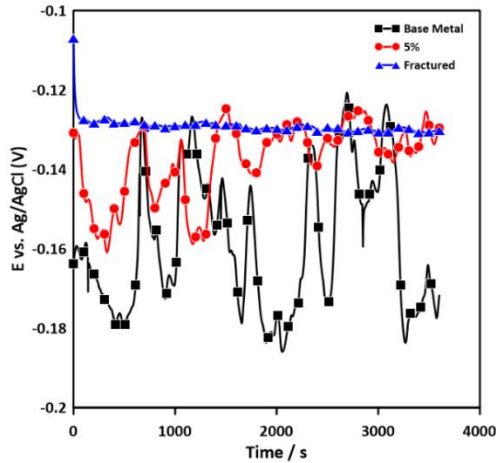


Figure 10. The open-circuit potential (OCP) values over time for all samples.

To study the electrochemical behaviour of all samples, cyclic polarization polarization (CPP) testing in aerated 3.5 wt.% NaCl solution at 25 °C was conducted and the results are shown in Figure 11. As expected, all samples revealed a clear passive region with a similar corrosion ($\sim -175 \text{ mV}_{\text{Ag/AgCl}}$). The CPP graphs of both deformed samples revealed approximately the same corrosion current density, whereas the non-deformed base metal was characterized by having a higher corrosion current density. The anodic branch of the CPP graphs indicated a high frequency of fluctuations originating from the metastable corrosion activities between the corrosion potential and the pitting potential. A pitting potential of $28.2 \text{ mV}_{\text{Ag/AgCl}}$ was recorded for the base metal. The pitting potential was increased to $134.3 \text{ mV}_{\text{Ag/AgCl}}$ for the 5% elongated sample and to $252.3 \text{ mV}_{\text{Ag/AgCl}}$ for the fractured sample. The detected higher pitting potential value for the fractured sample is in good agreement with the OCP results, exhibiting higher stability of the passive film on the severely deformed sample, leading to its higher corrosion resistance against pitting corrosion.

Figure 12 shows the SEM micrograph of samples after CPP testing and their corresponding elemental maps for chromium concentration. The micrograph of the corroded base metal depicts randomly distributed corrosion pits, while the 5% elongated sample revealed a lower number of pits per unit area. The intensity of localized pitting on the surface of the fractured sample after the CPP testing was found to be much lower than that of the other two samples, and stable pits were barely detected on its surface (see Figure 12e). The chromium concentration maps for the base metal and 5% deformed

sample (Figures 12b and 12d) revealed a high concentration of large chromium rich zones (possibly carbides) adjacent to the pitted spots, implying that corrosion pit initiation starts along the interface of the matrix and the carbides.

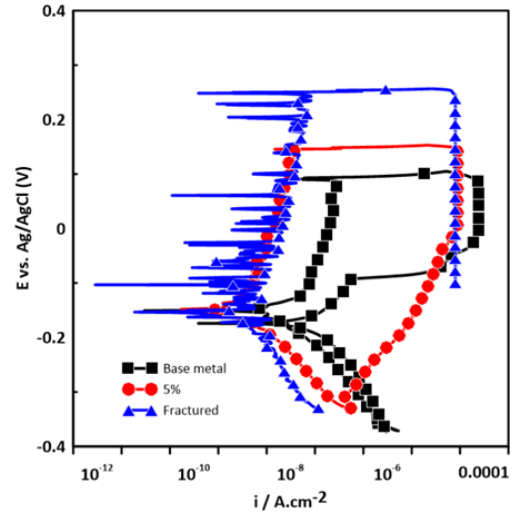


Figure 11. cyclic potentiodynamic polarization curves for 13Cr alloy at solution of 3.5% NaI and room temperature

IV. DISCUSSION

A. Effect of grain orientation and passive layer formation on the corrosion behavior

The microstructural analysis of the samples confirmed a slight grain refinement of the alloy through the plastic deformation of the samples. It is well known that finer grain structures provide larger grain boundary areas, and since most of the precipitates form along the grain boundaries, it can consequently lead to an increased galvanic coupling over the entire microstructure exposed to the corrosive electrolyte. On the other hand, the evolution of the corrosion potential (OCP results) is directly related to the integrity and density of the passive layer [22, 23]. On this context, previous studies have shown that deformation process can deteriorate the integrity and the thickness of the passive film, leading to the reduced corrosion potential of the surface. However, such detrimental effect of the applied plastic deformation on the electrochemical stability of the surface contradicts with the results presented herein. Hence, the passive film thickness reduction cannot act as the dominant factor in controlling the corrosion properties of the deformed 13Cr stainless steel samples.

Although the general impact of the grain boundaries on the plastic deformation of materials is well understood [24, 25], the effect of grain boundaries on the electrochemical stability of the surface is highly complex. In this study, severely deformed sample showed a higher volume fraction of LAGBs relative to that of the base metal and the 5% deformed sample. It has been reported that a high volume fraction of LAGBs can lead to corrosion properties enhancement [26]. Similarly, improvement in the corrosion resistance of three ferritic–martensitic (F/M) steels, HT9, T91, and NF616, and two

binary model alloys Fe-15%Cr and Fe-18%Cr was observed in grain-refined samples because of the enhanced diffusion of chromium on the surface, through a high density of grain boundaries [26].

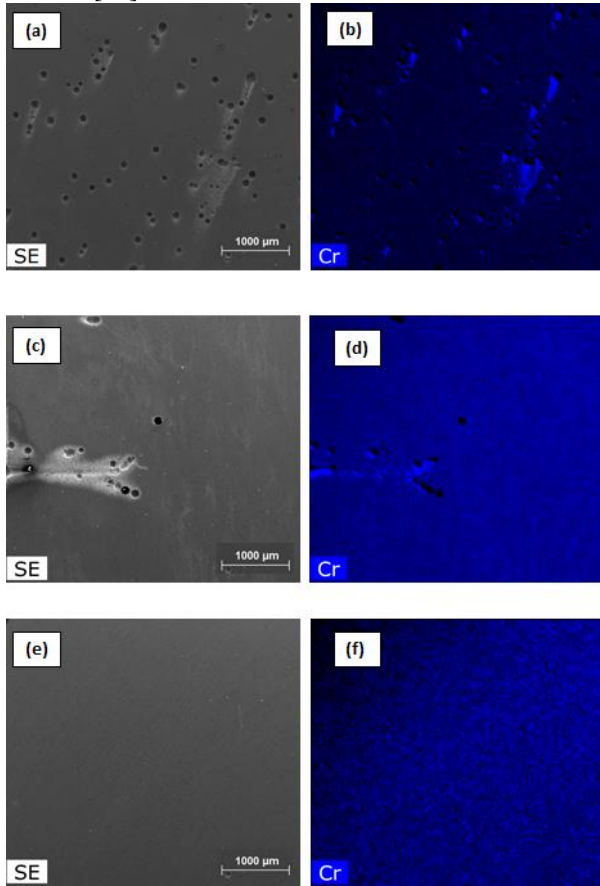


Figure 12. SEM micrographs from the samples after cyclic polarization tests and their corresponding Cr concentration maps: (a,b) base metal, (c,d) the 5% elongated sample, and (e,f) the fracture sample.

B. Effect of precipitation on the corrosion behavior

Considering previous studies on the effect of precipitates, particularly carbides, on the corrosion behavior of stainless steels, it is well proved that in most cases, the matrix phase is electrochemically less noble than the precipitates. Furthermore, it is known that the interface between the matrix and the precipitates is more susceptible to corrosion [27, 28]. Significantly high potential difference between the matrix and constituents can accelerate the micro-galvanic coupling in which matrix acts as local anode part and the carbides as the local cathode [27]. The microstructure of the non-deformed base metal in this study showed random distribution of larger carbides. These large precipitates were also visible on the 5% deformed sample. Existence of larger carbides in the structure is accompanied by a more expanded sensitized regions around these carbides, where chromium depletion was detected (shown in Figures 12b and 12d) [29, 30]. The sensitized regions are characterized by a high anodic activity relative to

the carbide precipitates, leading to the lower corrosion resistance of the base metal and 5% elongated samples [30].

On the other hand, the fractured sample possesses high fraction of fine precipitates generated through the breakage of coarser undissolved brittle carbides during plastic deformation of the sample. Figures 4c and 4d indicate that a preferential orientation was imposed for the precipitates aligned with the loading direction and larger interspacing between the precipitates has formed. Breakage of a coarse carbide precipitate surrounded by an expanded sensitized region into smaller precipitates also splits the large sensitized region into smaller sensitized zones, each adjacent to the finer precipitates. This could potentially increase the number of locations susceptible to meta-stable pitting on the surface; however, due to the confined size of these sensitized regions, stable growth of a pit can be significantly hindered. The potential difference between the precipitates and matrix can stimulate micro-galvanic coupling and act as the main driving force for the corrosion [2]. However, galvanic coupling is not only affected by the magnitude of potential difference, but also the distance between the anodic and cathodic sites can affect the severity of the corrosion drastically. Hence, as the deformation level increases, the magnitude of galvanic coupling decreases due to the decreased size of the precipitates and the increase in the interspacing of carbides [31].

V. CONCLUSIONS

In this study, the effect of plastic deformation on the corrosion behavior of 13Cr stainless steel is investigated. Samples with deformation levels of 5% and fractured samples were prepared using uniaxial tensile testing, and their microstructure and corrosion properties were investigated in detail and compared with the non-deformed base metal. The fractured sample revealed a high fraction of deformed LAGBs with a slight grain refinement, while the 5% elongated and the base metal samples showed lower densities of LAGBs. A more stable and higher corrosion potential value was measured for the fractured sample, indicating a better stability of the passive film on its surface, whereas the base metal and 5% elongated samples showed lower corrosion potential values over time with a higher fluctuation of the OCP values. From the CPP test results, a higher pitting potential value was detected for the fractured, consistent with the trends of OCPs. The dominant factor in controlling the corrosion behavior of 13Cr MSS in both non-deformed and deformed conditions was found to be formation of the micro-galvanic coupling between the existing precipitates and the surrounding matrix, significantly affected by the distribution and size of the existing precipitates in the microstructure of the alloy. As the deformation level increases, the magnitude of galvanic coupling decreases due to the decreased size of the precipitates and the increased interspacing of carbides. The corrosion of the non-deformed base metal and 5% deformed sample was found to be

controlled by sensitization in chromium-depleted zones adjacent to the larger size chromium carbides.

ACKNOWLEDGMENT

The authors would like to thank the support of Suncor Energy and the Memorial University of Newfoundland for sponsoring this work.

REFERENCES

- [1] M.S. Cayard, R.D. Kane, Serviceability of 13 Cr tubulars in oil and gas production environments, CORROSION-NATIONAL ASSOCIATION OF CORROSION ENGINEERS ANNUAL CONFERENCE-, NACE, 1998.
- [2] K.H. Anantha, C. Örnek, S. Ejnermark, A. Medvedeva, J. Sjöström, J. Pan, Correlative microstructure analysis and in situ corrosion study of AISI 420 martensitic stainless steel for plastic molding applications, Journal of The Electrochemical Society 164(4) (2017) C85-C93.
- [3] S. Zhu, A. Fu, J. Miao, Z. Yin, G. Zhou, J. Wei, Corrosion of N80 carbon steel in oil field formation water containing CO₂ in the absence and presence of acetic acid, Corrosion Science 53(10) (2011) 3156-3165.
- [4] X. Zhao, Y. Han, Z. Bai, B. Wei, The experiment research of corrosion behaviour about Ni-based alloys in simulant solution containing H₂S/CO₂, Electrochimica Acta 56(22) (2011) 7725-7731.
- [5] A.N. Isfahany, H. Saghaian, G. Borhani, The effect of heat treatment on mechanical properties and corrosion behavior of AISI420 martensitic stainless steel, Journal of Alloys and Compounds 509(9) (2011) 3931-3936.
- [6] A. Rajasekhar, G.M. Reddy, T. Mohandas, V. Murti, Influence of austenitizing temperature on microstructure and mechanical properties of AISI 431 martensitic stainless steel electron beam welds, Materials & Design 30(5) (2009) 1612-1624.
- [7] C.G. De Andrés, L. Alvarez, V. López, J.A. Jiménez, Effects of carbide-forming elements on the response to thermal treatment of the X45Cr13 martensitic stainless steel, Journal of materials science 33(16) (1998) 4095-4100.
- [8] L.D. Barlow, M. Du Toit, Effect of austenitizing heat treatment on the microstructure and hardness of martensitic stainless steel AISI 420, Journal of materials engineering and performance 21(7) (2012) 1327-1336.
- [9] S.-Y. Lu, K.-F. Yao, Y.-B. Chen, M.-H. Wang, X. Liu, X. Ge, The effect of tempering temperature on the microstructure and electrochemical properties of a 13 wt.% Cr-type martensitic stainless steel, Electrochimica Acta 165 (2015) 45-55.
- [10] V. Venegas, F. Caleyo, T. Baudin, J.H. Espina-Hernandez, J. Hallen, On the role of crystallographic texture in mitigating hydrogen-induced cracking in pipeline steels, Corrosion Science 53(12) (2011) 4204-4212.
- [11] V. Venegas, F. Caleyo, T. Baudin, J. Hallen, R. Penelle, Role of microtexture in the interaction and coalescence of hydrogen-induced cracks, Corrosion Science 51(5) (2009) 1140-1145.
- [12] J. Verdeja, J. Asensio, J. Pero-Sanz, Texture, formability, lamellar tearing and HIC susceptibility of ferritic and low-carbon HSLA steels, Materials characterization 50(1) (2003) 81-86.
- [13] X. Feng, X. Lu, Y. Zuo, N. Zhuang, D. Chen, The effect of deformation on metastable pitting of 304 stainless steel in chloride contaminated concrete pore solution, Corrosion Science 103 (2016) 223-229.
- [14] S. Phadnis, A. Satpati, K. Muthe, J. Vyas, R. Sundaresan, Comparison of rolled and heat treated SS304 in chloride solution using electrochemical and XPS techniques, Corrosion science 45(11) (2003) 2467-2483.
- [15] V. Guillaumin, P. Schmutz, G. Frankel, Characterization of corrosion interfaces by the scanning Kelvin probe force microscopy technique, Journal of the electrochemical society 148(5) (2001) B163-B173.
- [16] D.B. Blücher, J.-E. Svensson, L.-G. Johansson, M. Rohwerder, M. Stratmann, Scanning Kelvin Probe Force Microscopy A Useful Tool for Studying Atmospheric Corrosion of MgAl Alloys In Situ, Journal of the electrochemical society 151(12) (2004) B621-B626.
- [17] R. Ueji, N. Tsuji, Y. Minamino, Y. Koizumi, Ultragrain refinement of plain low carbon steel by cold-rolling and annealing of martensite, Acta Materialia 50(16) (2002) 4177-4189.
- [18] L. Bracke, K. Verbeken, L. Kestens, J. Penning, Microstructure and texture evolution during cold rolling and annealing of a high Mn TWIP steel, Acta Materialia 57(5) (2009) 1512-1524.
- [19] J. Rosenberg, H. Piehler, Calculation of the Taylor factor and lattice rotations for bcc metals deforming by pencil glide, Metallurgical Transactions 2(1) (1971) 257-259.
- [20] V. Randle, O. Engler, Introduction to texture analysis: macrotexture, microtexture and orientation mapping, CRC press 2014.
- [21] S.H. Lee, D.N. Lee, Analysis of deformation textures of asymmetrically rolled steel sheets, International journal of mechanical sciences 43(9) (2001) 1997-2015.
- [22] D. Song, A. Ma, W. Sun, J. Jiang, J. Jiang, D. Yang, G. Guo, Improved corrosion resistance in simulated concrete pore solution of surface nanocrystallized rebar fabricated by wire-brushing, Corrosion Science 82 (2014) 437-441.
- [23] Y. Fu, X. Wu, E. Han, W. Ke, K. Yang, Z. Jiang, Influence of cold work on pitting corrosion behavior of a high nitrogen stainless steel, Journal of The Electrochemical Society 155(8) (2008) C455-C463.
- [24] X. Sauvage, G. Wilde, S. Divinski, Z. Horita, R. Valiev, Grain boundaries in ultrafine grained materials processed by severe plastic deformation and related phenomena, Materials Science and Engineering: A 540 (2012) 1-12.
- [25] A. Cao, Y. Wei, E. Ma, Grain boundary effects on plastic deformation and fracture mechanisms in Cu nanowires: Molecular dynamics simulations, Physical Review B 77(19) (2008) 195429.
- [26] X. Ren, K. Sridharan, T. Allen, Effect of grain refinement on corrosion of ferritic-martensitic steels in supercritical water environment, Materials and Corrosion 61(9) (2010) 748-755.
- [27] C. Örnek, D.L. Engelberg, Correlative EBSD and SKPFM characterisation of microstructure development to assist determination of corrosion propensity in grade 2205 duplex stainless steel, Journal of materials science 51(4) (2016) 1931-1948.
- [28] C. Örnek, D.L. Engelberg, An experimental investigation into strain and stress partitioning of duplex stainless steel using digital image correlation, X-ray diffraction and scanning Kelvin probe force microscopy, The Journal of Strain Analysis for Engineering Design 51(3) (2016) 207-219.
- [29] O. Yevtushenko, D. Bettge, S. Bohraus, R. Bäßler, A. Pfennig, A. Kranzmann, Corrosion behavior of steels for CO₂ injection, Process Safety and Environmental Protection 92(1) (2014) 108-118.
- [30] S. Sharafat, N.M. Ghoniem, P.I. Cooke, R.C. Martin, F. Najmabadi, K.R. Schultz, C.P. Wong, T. Team, Materials selection criteria and performance analysis for the TITAN-II reversed-field-pinch fusion power core, Fusion engineering and design 23(2-3) (1993) 201-217.
- [31] G. Frankel, J. Scully, C. Jahnes, Repassivation of pits in aluminum thin films, Journal of The Electrochemical Society 143(6) (1996) 1834-1840.

Failure Characterization of Heavily Cross-linked Epoxy: Part II Multiscale Modeling

A. Y. Elruby

Ph.D. Candidate

Mechanical Engineering, Faculty of Engineering
Memorial University of Newfoundland
St. John's, NL, Canada
E-mail: aleruby@mun.ca

Sam Nakhla

Assistant Professor

Mechanical Engineering, Faculty of Engineering
Memorial University of Newfoundland
St. John's, NL, Canada
E-mail: satnakhla@mun.ca

Abstract— This is part II of two companion papers on failure characterization of heavily cross-linked epoxy. In part I, detailed testing procedures for both uniaxial tension and three-point bending tests were provided with insight on failure mechanisms. In the current part, multiscale finite element modeling utilizing the unit cell (UC) method is proposed to predict the mechanical behavior of epoxy. First, the UC method was used to investigate the effect of voids on the microstructural behavior of heavily cross-linked epoxy. Actual microscopic images showing voids were employed to construct the UC models. Convergence analysis regarding representative UC size and voids intensity was conducted. Effective mechanical behaviors obtained from micromechanical UC models proved prediction capabilities. Preliminary failure analysis in the microstructure was conducted using the framework of the extended finite element method (XFEM). Also, macro-mechanical specimen sized models are provided utilizing XFEM enrichments to predict final failure. Dilatational strain energy density (SED) criterion was implemented in a user-defined damage subroutine. For assessment and validation purposes, numerically obtained results are compared to those of testing from part I. Comparisons showed excellent agreement with testing results regarding failure limits as well as predicted fracture surfaces.

Keywords- *Unit cell method; multiscale modeling; XFEM; Failure*

I. INTRODUCTION

Fabrication of fiber reinforced polymers (FRPs) consistently results in manufacturing imperfections that are strongly associated with their mechanical behaviors [1]. These defects influence both micro and macro structural behaviors [2,3]. Void content resulting from the polymerization process in FRP composite is one of these manufacturing defects. The characteristic size of voids in thermosetting polymer are usually of microscopic dimensions [2]. The microscopic voids presence in FRPs can be minimized utilizing advanced manufacturing techniques such as vacuum bagging and autoclaving, yet they are unavoidable [3,4]. Voids have a

deteriorating effect on the mechanical behavior of epoxies especially regarding failure mechanism [5].

Ply-delamination and matrix cracking are typical failure modes which are associated with resin material failure [6]. Precise identification of failure mechanisms in composite materials would be of paramount usage for design stages. As a result, there exist several studies in the literature on failure modes of composites. For example, Lachaud et al. [7] modeled matrix cracking and ply delamination in laminated composites using cohesive models and continuous damage models. They showed in their work that matrix cracking and delamination were principally initiating failure in laminated composites. Also, Pollayi and Yu [8] studied failure onset in composite rotor blades which can be found in a helicopter rotor or wind turbine blades. They showed that matrix micro-cracking maybe considered as the first damage event in FRPs. Riaño et al. [9], utilized the representative volume element (RVE) method to predict the mechanical behavior of unidirectional glass fiber composites. They also modeled damage phenomena such as matrix cracking and fiber/matrix debonding at a mesoscale to enhance macro-mechanical numerical predictions. Their study yielded approximately 25% error when compared to testing.

Investigating microstructural behavior may be necessary to fully understand the mechanical behavior. Multiscale modeling approaches have been widely used to investigate the mechanical behavior of fiber reinforced polymers (FRPs) [10–12]. Modeling at different scales enable a broader understanding of material behavior. Two-scale modeling approaches such as micro-macro modeling have been proposed to systematically study various material behaviors [11,13–15]. For example, Mehdikhani et al. [16] utilized multiscale modeling at micro/mesoscale to study the effect of voids on matrix cracking in composite laminates. They showed that while matrix cracking occurs at the ply scale, yet it is strongly dependent on the ply microstructure. They concluded that the presence of voids influences both the onset and evolution of sporadic cracks. Also, they have concluded that further enhancement to the framework of XFEM is required for accurate crack predictions [16]. Also, Chevalier et al. [17] proposed a micro-mechanics based pressure dependent failure

model for highly cross-linked epoxies. In their work, a single ellipsoidal inclusion was implemented in a unit cell (UC) within the framework of finite element (FE) modeling. Their predictions showed an upper bound; in other words, failure was overestimated. They concluded that the exact nature of defects remains to be determined [17]. Asp et al. [18] proposed a crack initiation criterion for heavily cross-linked epoxy based on the dilatational strain energy density (SED). They showed that micro-cavitation dominate failure initiation in neat epoxy under a complex state of stress.

In Part II of these companion papers, the multiscale modeling approach is followed utilizing the framework of XFEM to investigate both micro and macro structural behavior of heavily cross-linked epoxy. Microstructural investigations were enabled employing the UC method within the framework of FE analysis to predict the effective elastic-plastic behavior of epoxy. The intensity of voids and distributions were studied to determine the representative size of the UC. Preliminary failure analysis using the UC model of the actual microstructure of voids is also provided. Maximum principal stress and fracture energy based damage mechanism were used in microstructural XFEM preliminary failure investigations. Micromechanical numerical results regarding effective stress-strain behavior were utilized as an input to macroscale specimen sized FE models. Also, the dilatational SED based failure initiation criterion from literature was implemented in a user-defined damage subroutine in Abaqus. Finally, comparisons with testing results from Part I are provided for validation purposes. Numerically obtained results are consistent with literature findings and testing results from Part I.

II. MICROSTRUCTURAL BEHAVIOR

A. Microstructural unit cell modeling

Microstructural behavior of heavily cross-linked epoxy is investigated using micromechanical modeling. For that purpose, the UC method was employed within the framework of finite element analyses to predict the mechanical behavior of epoxy. The representative size of a UC was determined to be 1000 μm which is sufficient for representing the microstructural features. Also, the voids intensity/distribution were varied to study their effect on mechanical behavior. Computed tomography (CT) scan of the epoxy resin showed a void content of approximately 0.5% by total volume. Also, voids were of regular spherical shapes with diameters varying from 5-80 μm . The total void content was introduced to UC models using a different distribution of holes in addition to actual distribution from microscopic imaging. The radius of each hole was evaluated to present the total volumetric fraction of voids. Figure 1 shows the uniformly distributed voids in different UC models. Each model has a different distribution of representative voids which was used to evaluate voids radii.

Necessary set of continuity boundary conditions were applied to enable predicting continuum mechanical behavior. This procedure was necessary and it is slightly different than homogenization techniques. This approach have the advantage of obtaining mechanical behavior in elastic-plastic regimes rather than discrete stiffness properties. A solid homogeneous

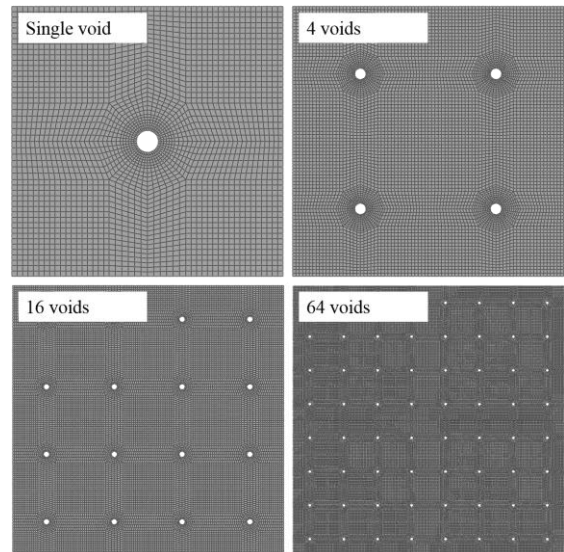


Figure 1. Uniformly distributed voids in different UC models.

section was used to apply the material definition to all UC models. The material behavior was defined using deformation plasticity theory utilizing the modulus of elasticity and Poisson's ratio for linear behavior while the yield stress together with the hardening exponent defines the plastic flow. Noteworthy to mention that efficient yield identification method presented in part I, enabled precise definition of utilized material model.

In the current study, the single void UC was found to be sufficient for the linear regime but not the plastic behavior. As the number of voids representing the total volumetric percentage increased, a convergence like behavior was observed. The convergence analysis in the current study showed that the 16 holes UC is satisfactory to determine the effective behavior in both elastic and plastic behaviors. To investigate the effect of actual voids on mechanical behavior, actual microscopic imaging was utilized to build a UC representing the actual microstructure regarding void sizes and locations. The procedures for generating the UC model are presented in Figure 2. First, the microscopic image was processed to a monochromatic image isolating the voids in the image as white circles. The isolated image had a size of 1000 μm and approximate total voids percentage of 0.53%. The processed image was converted to a drawing exchange format (DXF) which was finally imported to Abaqus to construct the UC of the actual microstructure. The convergent mesh of the actual microstructure had approximately 62000 triangular elements, namely CPS3 which is a plane stress 2D element. Triangular elements are best suited for this type of irregular geometries, i.e. holes with different radii. Conforming to the given complex geometry required denser mesh that uniformly distributed UCs.

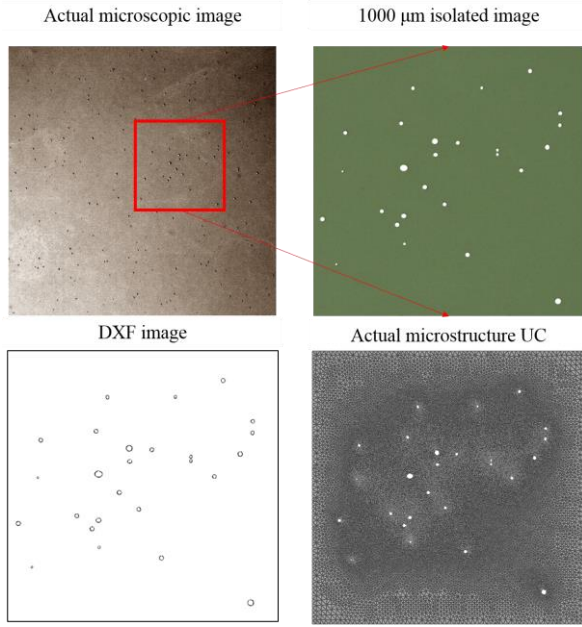


Figure 2. Actual microstructural UC model generation.

To delineate the effect of distributing voids in the microstructure, the full-field contour plots of von-Mises stress were used. Figure 3 shows a comparison von-Mises results obtained by the uniformly distributed UC vs the actual microstructural UC. As can be observed that the stress distribution in the uniform UC is symmetric with a tendency for concentration around the vicinity of holes. A zoomed in view is provided next to the figure for better illustrations. The symmetry is clear showing intensive stress concentrations around holes. These concentration regions are building up in a direction that is normal to the load application direction.

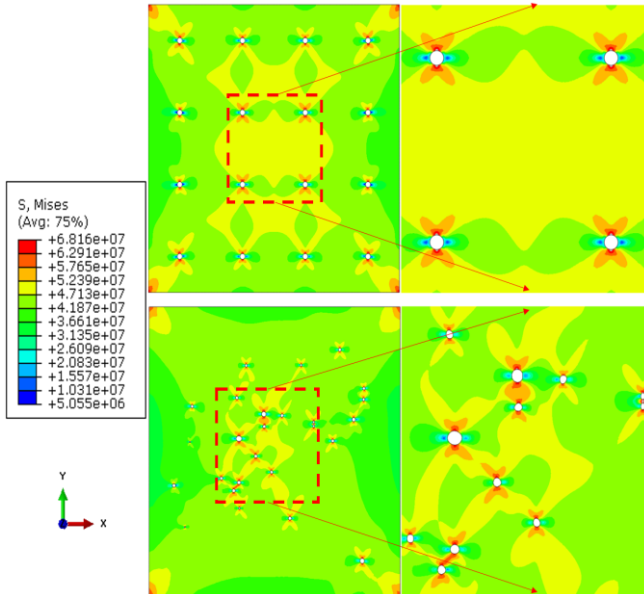


Figure 3. Von-Mises stress for uniform and actual voids distributions.

Inspecting the actual microstructural UC, it was observed that the tendency of stress concentration is more intense around holes with larger radii than with smaller ones. Also, the closer the holes the higher the tendency of building up stress was. Numerically obtained stress-strain behaviors are presented to delineate the effect of actual microstructure versus the uniform microstructure on mechanical behavior. As can be observed from Figure 4, both UC models (actual/uniform) distributions results are identical. Even though the sizes and distributions bias local stresses and plastic deformations, effective stress-strain behavior remains unaltered. Both numerically obtained stress-strain curves are compared to those of testing showing an excellent agreement. Therefore, the current investigation showed that actual/uniform distributions were sufficient to evaluate elastic-plastic behavior of the tested epoxy. It can be concluded that upon the determination of the representative UC size utilizing the proper characteristic length and the sufficient number of voids, the effective behavior can be precisely determined regardless of the distribution.

B. Microstructural failure initiation

XFEM framework was coupled with the actual UC model to investigate failure initiation at a microscale. The XFEM depends on enriching the conventional finite element domain with special shape functions accounting for crack initiation and evolution [19]. References [20–22] provide a detailed explanation to XFEM technique and its applications. Microstructural UC presenting actual voids distribution was utilized within the framework of XFEM to investigate failure initiation. Damage process is initialized depending on maximum principal stress while the evolution was enabled on energy basis. The traction separation law consists of two parts; initiation and evolution. The crack onset is defined as

$$f = \left\{ \frac{\langle \sigma_{max} \rangle}{\sigma_{max}^0} \right\} \quad (1)$$

where f represents the damage initiation condition, σ_{max} is the maximum principal stress at a material point and σ_{max}^0 is user-defined nominal value. The Macaulay brackets signify that a

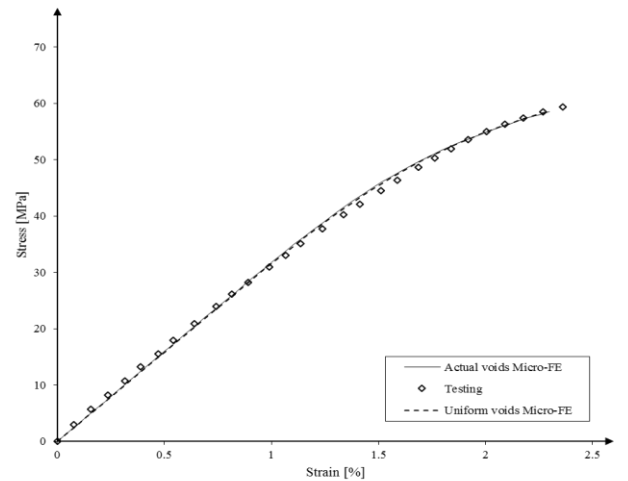


Figure 4. Numerically obtained stress-strain behaviors using UC models vs. testing.

pure compressive state does not initiate damage. A crack is introduced to the material domain once the condition is encountered. Crack propagation is controlled by fracture energy and a damage factor D . An element with no damage corresponds to a zero value while the entire damaged element corresponds to the unity value.

Figure 5 presents the preliminary micro-cracks in the UC model. The contour plot of von-Mises stress is provided along with a closer view on initiated cracks. The micro-cracks initiated at the largest hole/void signifying that failure onset at a micro-scale will possibly be in the vicinity of the larger voids. In fact, this observation was documented in Part I of this work. The micro-structural predictions in terms of crack initiation locations are in excellent agreement with experimental observations. Stress concentrations in the vicinity of holes triggered the onset of damage in microstructural UCs. There is a possibility that micro-cracks cause coalescence and activate damage at the macroscopic scale. To this end, the microstructural investigations were conducted and documented in details. Specimen sized model, i.e. macroscale utilizing the micromechanical numerical results are presented in the

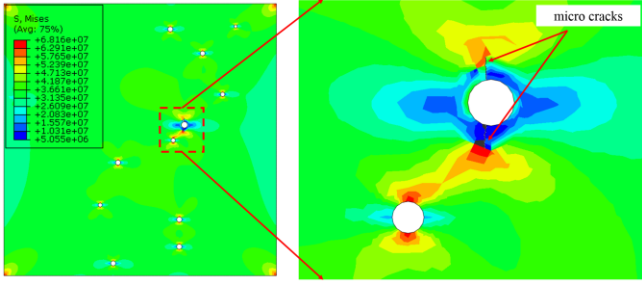


Figure 5. Von-Mises contour plot showing initiation of micro cracks.

following section.

III. SPECIMEN SIZED MODELING

To delineate multiscale modeling capabilities, specimen sized models simulating uniaxial and three-point bending tests are provided. The models were constructed in 3D space with the same geometry used for prepared specimens. Three-dimensional models showing their final convergent meshes are presented in Figure 6. The convergent mesh of the dog-bone specimen had approximately 47000 elements. Three-dimensional brick elements were used in meshing, namely C3D8R. The brick element is eight-noded and reduced integration option was used to minimize computational efforts. As for prismatic specimens, same element was used and the

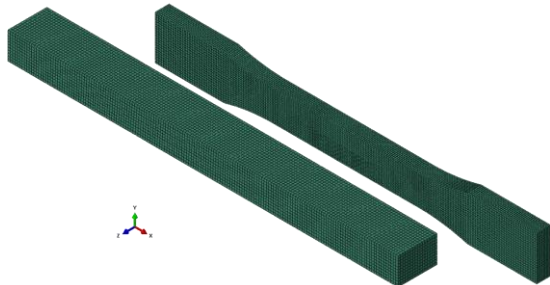


Figure 6. Three-dimensional finite element models.

convergent mesh had a size of approximately 29000 elements. ASTM standards testing procedures, documented in Part I, regarding loading conditions/rates were followed. Numerically obtained mechanical behaviors from microstructural UCs were used to define the macro scale material models.

The dog-bone specimen end tabs together with the cylinders used for three-point loading were constrained as rigid parts to mimic the testing conditions. The framework of XFEM is used to predict the actual macroscopic failure of both specimens. Critical dilatational strain energy density (SED) [18] was used as the crack initiation mechanism while the damage evolution was controlled using the fracture energy. The criterion was initially proposed by Asp et al. [18] for crack initiation in glassy polymers. The criterion was implemented in a user-defined subroutine to initiate damage using the framework of XFEM in Abaqus. The implemented critical dilatational SED criterion reads as

$$U_h = \frac{1 - 2\nu}{6E} (\sigma_1 + \sigma_2 + \sigma_3)^2 = U_h^{crit} \quad (2)$$

where U_h is the dilatational component of SED, σ_i represents the principal stresses along the 1-2-3 directions, respectively. Finally, U_h^{crit} represents the critical value of dilatational SED. Noteworthy to mention that the critical value can be readily determined from the uniaxial test of characterization. The failure criterion is checked upon each increment, if the condition is encountered, the damage is initiated in the corresponding enriched element. Damage will then evolve based on material's fracture energy which is known to be a material property.

Predicted crack results using the implemented criterion are presented in Figure 7. The criterion enabled predicting realistic fracture surfaces similar to those observed from testing. For example, the failure surface in the dog-bone specimen was normal to the load application direction signifying the brittle failure dominated the fracture mechanism. Also, for the prismatic beam under three-point loading, the fracture surface was perpendicular to the maximum principal stress direction. Also, the onset location of the crack was in the lower mid-segment of the prism, i.e. tension side. Predicted cracks demonstrate the predictive capability of utilizing the proposed methodology.

The residual plasticity contour plots for both models are compared to monochromatic scans of fractured specimens. The results showing actually fractured surfaces compared to prediction results are presented in Figure 8. As can be seen, the FE models using the proposed damage model enable precise predictions in terms of crack location as well as the predicted

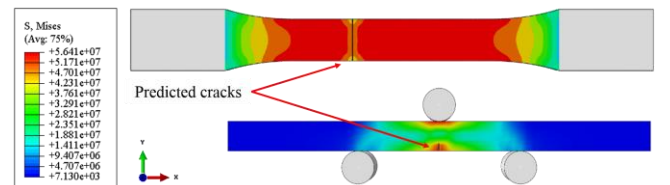


Figure 7. Predicted cracks using critical SED criterion defined by (2).

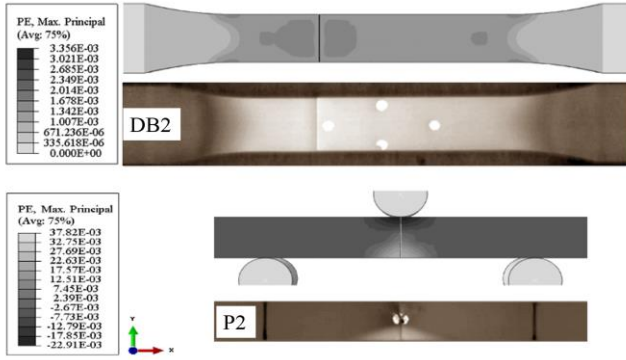


Figure 8. Residual plastic strains contour plots after failure.

fractured surfaces.

The prediction results are consistent with observations from testing. Also, monochromatic scans of failed specimens exposed stress-whitening caused by inelastic deformations. Comparing scans to contour plots of plastic strains shows excellent prediction capabilities for the proposed work. Finally, both FE numerical results are compared to those of testing. Uniaxial tension is compared regarding stress-strain results while three-point loading is compared in terms of load-displacement. Both comparisons are presented in Figure 9 and Figure 10, respectively.

The uniaxial tension numerical results are in excellent agreement with those of testing. The numerically obtained curve is closely following the testing curve in both elastic and plastic regimes. Moreover, the final failure is of lower bound, i.e. conservative in predicting failure limits. Both failure limits, stress/strain, were slightly underestimated signifying reliable and safe predictions. For the three-point bending model, it can be observed that numerically obtained load-displacement curve is successfully following the measured curve from testing. In the linear region, both curves are almost identical while in the plastic flow there exist a minimal deviation from testing results.

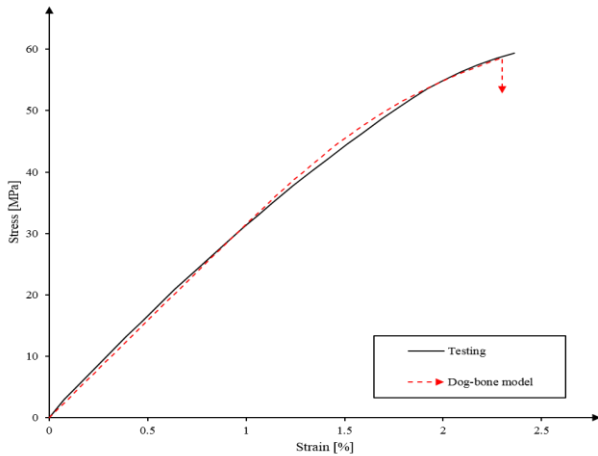


Figure 9. Uniaxial tension stress-strain curves: numerical vs. testing.

Noteworthy to mention that the failure limits regarding load

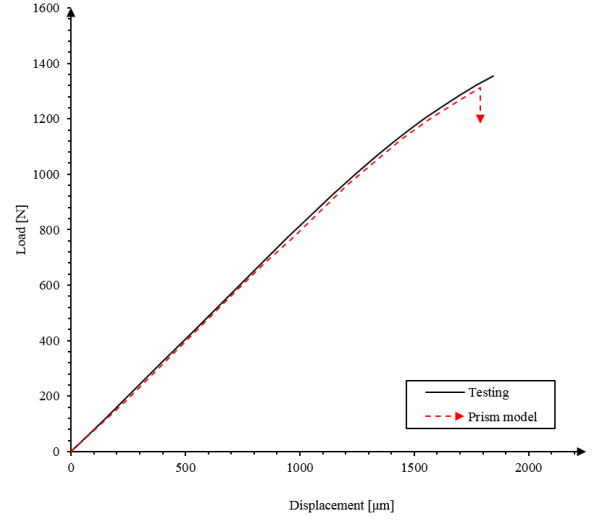


Figure 10. Uniaxial tension stress-strain curves: numerical vs. testing.

and displacements were of a lower bound compared to testing, i.e. conservative predictions.

Numerically obtained failure limits were quantitatively compared to those of testing and the comparison is documented in Table I. Failure stress and strain were efficiently predicted with errors of -1.20% and -2.54%, respectively. For three-point loading, the failure limits in terms of load and displacement are documented in Table II. The numerical predictions compared to testing recorded errors of -0.51% and -2.93% for failure load and displacement, respectively. In conclusion, the proposed modeling methodology has proven to be precise and consistent with testing results at different scales, i.e. micro and macro.

TABLE I. UNIAxIAL TENSION FAILURE LIMITS: NUMERICAL VS. TESTING.

Uniaxial	Failure limits	
	Failure stress [MPa]	Failure strain [%]
Numerical	58.63	2.30
Testing	59.34	2.36
Error	-1.20	-2.54

TABLE II. THREE-POINT BENDING FAILURE LIMITS: NUMERICAL VS. TESTING.

Uniaxial	Failure limits	
	Failure load [N]	Failure displacement [μm]
Numerical	1313.75	1790
Testing	1320.50	1844
Error	-0.51	-2.93

IV. CONCLUSIONS

Neat epoxy resin was prepared and tested in Part I of the current work. Also, an efficient yield identification method utilizing the stress-whitening phenomena was proposed and

validated. Fractographic analysis using optical microscopy enabled characterizing failure mechanisms under different types of loading. In part II, multiscale modeling using the framework of XFEM was provided. UC models were used to investigate preliminary failure on a microscale. Micromechanical modeling revealed that micro-cracks tend to initiate in the vicinity of larger voids. As for the macroscale modeling, specimen sized models were developed to simulate testing of uniaxial tension and three-point bending.

Moreover, the dilatational SED criterion was implemented in a user-defined damage subroutine within the framework of XFEM in Abaqus. Numerical results have proven accurate prediction capabilities. Failure surfaces were in excellent agreement with observations from testing. Besides, failure limits were of lower bound with a maximum error of 3% in predicting failure displacement, i.e. three-point bending.

Regarding the failure load, the maximum error recorded was -1.2%. Also, the residual plastic strains showed excellent agreement with monochromatic scans of failed specimens. It can be concluded that while epoxy resin failure is dominated by brittleness, yet there exist localized plastic deformations that should be taken into consideration. The damage initiation criterion proposed by Asp et al was implemented and proven excellent prediction capabilities.

ACKNOWLEDGMENT

Authors would like to acknowledge Polynt Composites Canada, Inc. for providing the neat epoxy resin slab.

REFERENCES

- [1] W. V. Liebig, C. Viets, K. Schulte, B. Fiedler, Influence of voids on the compressive failure behaviour of fibre-reinforced composites, *Compos. Sci. Technol.* 117 (2015) 225–233. doi:10.1016/j.compscitech.2015.06.020.
- [2] H. Zhu, B. Wu, D. Li, D. Zhang, Y. Chen, Influence of Voids on the Tensile Performance of Carbon/epoxy Fabric Laminates, *J. Mater. Sci. Technol.* 27 (2011) 69–73. doi:10.1016/S1005-0302(11)60028-5.
- [3] L. Di Landro, A. Montalto, P. Bettini, S. Guerra, F. Montagnoli, M. Rigamonti, Detection of voids in carbon/epoxy laminates and their influence on mechanical properties, *Polym. Polym. Compos.* (2017).
- [4] B. Joseph, F.W. Hanratty, J.L. Kardos, Model Based Control of Voids and Product Thickness during Autoclave Curing of Carbon/Epoxy Composite Laminates, *J. Compos. Mater.* 29 (1995) 1000–1024. doi:10.1177/002199839502900801.
- [5] M. Kalantari, C. Dong, I.J. Davies, Effect of matrix voids, fibre misalignment and thickness variation on multi-objective robust optimization of carbon/glass fibre-reinforced hybrid composites under flexural loading, *Compos. Part B Eng.* 123 (2017) 136–147. doi:10.1016/j.compositesb.2017.05.022.
- [6] D.J. Mortell, D.A. Tanner, C.T. McCarthy, In-situ SEM study of transverse cracking and delamination in laminated composite materials, *Compos. Sci. Technol.* 105 (2014). doi:10.1016/j.compscitech.2014.10.012.
- [7] F. Lachaud, C. Espinosa, L. Michel, P. Rahme, R. Piquet, Modelling Strategies for Simulating Delamination and Matrix Cracking in Composite Laminates, *Appl. Compos. Mater.* 22 (2015) 377–403. doi:10.1007/s10443-014-9413-4.
- [8] H. Pollayi, W. Yu, Modeling matrix cracking in composite rotor blades within VABS framework, *Compos. Struct.* 110 (2014) 62–76. doi:10.1016/j.compstruct.2013.11.012.
- [9] L. Riaño, L. Belec, Y. Joliff, Validation of a Representative Volume Element for unidirectional fiber-reinforced composites: Case of a monotonic traction in its cross section, *Compos. Struct.* 154 (2016) 11–16. doi:10.1016/j.compstruct.2016.07.020.
- [10] H. Talebi, M. Silani, S.P.A. Bordas, P. Kerfriden, T. Rabczuk, A computational library for multiscale modeling of material failure, *Comput. Mech.* 53 (2014) 1047–1071. doi:10.1007/s00466-013-0948-2.
- [11] J.M.W.Æ.S.A. Meguid, Coupling atomistics and continuum in solids : status , prospects , and challenges, (2009) 79–110. doi:10.1007/s10999-008-9087-x.
- [12] P.F. Liu, J.Y. Zheng, Recent developments on damage modeling and finite element analysis for composite laminates: A review, *Mater. Des.* 31 (2010) 3825–3834. doi:10.1016/j.matdes.2010.03.031.
- [13] E. Svenning, F. Larsson, M. Fagerström, Two-scale modeling of fracturing solids using a smeared macro-to-micro discontinuity transition, *Comput. Mech.* 60 (2017). doi:10.1007/s00466-017-1426-z.
- [14] Q. Puydt, S. Flouriot, S. Ringeval, F. De Geuser, R. Estevez, G. Parry, A. Deschamps, Relationship Between Microstructure, Strength, and Fracture in an Al-Zn-Mg Electron Beam Weld: Part II: Mechanical Characterization and Modeling, *Metall. Mater. Trans. A.* 45 (2014) 6141–6152. doi:10.1007/s11661-014-2567-8.
- [15] Z. Wang, T. Yu, T.Q. Bui, N.A. Trinh, N.T.H. Luong, N.D. Duc, D.H. Doan, Numerical modeling of 3-D inclusions and voids by a novel adaptive XFEM, *Adv. Eng. Softw.* 102 (2016) 105–122. doi:10.1016/j.advengsoft.2016.09.007.
- [16] M. Mehdikhani, N.A. Petrov, I. Straumit, A.R. Melro, S. V. Lomov, L. Gorbatikh, The effect of voids on matrix cracking in composite laminates as revealed by combined computations at the micro- and meso-scales, *Compos. Part A Appl. Sci. Manuf.* 117 (2018) 180–192. doi:10.1016/j.compositesa.2018.11.009.
- [17] J. Chevalier, X.P. Morelle, C. Bailly, P.P. Camanho, T. Pardoën, F. Lani, Micro-mechanics based pressure dependent failure model for highly cross-linked epoxy resins, *Eng. Fract. Mech.* 158 (2016) 1–12. doi:10.1016/j.engfracmech.2016.02.039.
- [18] L.E. Asp, L.A. Berglund, R. Talreja, A criterion for crack initiation in glassy polymers subjected to a composite-like stress state, *Compos. Sci. Technol.* 56 (1996) 1291–1301. doi:10.1016/S0266-3538(96)00090-5.
- [19] N. Moës, J. Dolbow, T. Belytschko, A finite element method for crack growth without remeshing, *Int. J. Numer. Methods Eng.* 46 (1999) 131–150. doi:10.1002/(SICI)1097-0207(19990910)46:1<131::AID-NME726>3.0.CO;2-J.
- [20] Y. Abdelaziz, A. Hamouine, A survey of the extended finite element, *Comput. Struct.* 86 (2008) 1141–1151. doi:10.1016/j.compstruc.2007.11.001.
- [21] A. Yazid, N. Abdelkader, H. Abdelmadjid, A state-of-the-art review of the X-FEM for computational fracture mechanics, *Appl. Math. Model.* 33 (2009) 4269–4282. doi:10.1016/j.apm.2009.02.010.
- [22] T.-P. Fries, T. Belytschko, The extended/generalized finite element method: An overview of the method and its applications, *Int. J. Numer. Methods Eng.* 84 (2010) 253–304. doi:10.1002/nme.

Investigative Experimental Study of the Wetting Behavior of Transparent Coatings of PMMA Blends with PVC/PS/PVDF/PDMS

Wing Yi Pao, Remon Pop-Iliev, Ghaus Rizvi
Faculty of Engineering and Applied Science
UOIT – University of Ontario Institute of Technology
Oshawa, Canada
wingyi.pao@uoit.net

Abstract— This paper focuses on studying the function of self-cleaning ability of polymer coatings. Self-cleaning is an important property for surfaces in different applications such as corrosion and chemical resistance, abrasion protection, and maintaining high transmittance. An intrinsically hydrophilic polymer coating made of poly (methyl methacrylate) (PMMA) coating has been converted into a hydrophobic coating. Conventionally, this can be achieved by introduction of secondary phase nanoparticles, however the production cost for these nano-materials are usually very high. The presented study aims to develop a self-cleaning coating at low cost by blending PMMA with another polymer, which has higher water contact angle (WCA) and lower surface tension. The polymers blended with PMMA included poly (vinyl chloride) (PVC), polystyrene (PS), poly (vinylidene fluoride) (PVDF), and poly (dimethyl siloxane) (PDMS). The coatings were fabricated by spray deposition with compressed air on a glass substrate. The angle achieved was 98° by blending PMMA with PDMS, as compared to 60° when coated only with PMMA at 5 wt% in N, N-Dimethylformamide (DMF). As a result, trends were observed where the effective WCA of blends lie between the WCAs of PMMA and the other blended polymers. It was also observed that the WCA of the blend increased in the order of lower surface tension of the second polymer in the blends.

Water contact angle; surface tension; hydrophobic; self-cleaning; polymer blends; spray deposition.

I. INTRODUCTION

This study aims to develop a transparent hydrophobic self-cleaning coating, and investigate the effect of blending PMMA with other polymers with higher WCAs and lower surface energies. The microstructure of specimen was evaluated to relate with hydrophobicity, static WCA was also measured to determine the self-cleaning abilities of the coatings. Spray coating is inexpensive compared to other available surface engineering techniques to produce hydrophobic property through structuring, including chemical vapor deposition, plasma etching, and templating. The main advantage of spray

deposition is the capability of coating on surfaces of various materials of all shapes and sizes. Coating is quite a common practice for protecting a surface, such as prevention of abrasion on vehicle body, corrosion on pipelines, and chemical spill on flooring. Some coatings are functionalized to perform different tasks. A few classes of high performance coatings can be differentiated: *structural* – by tailoring the surface structure combined with tuning the surface tension self-cleaning surfaces that allow dirt and unwanted substances to be washed away by liquids can be achieved [1]; *electrochemical* – an electroactive coating can act as solid electrolyte to reduce overheat issues in batteries [2]; *optical* – photochromic coating can change color on windows under different light intensities [3]; and *smart materials* – piezoelectric coatings where materials are coated on flexible or moving parts to exploit its sensitivity to displacement and strain that results in harvesting electrical current [4].

A coating is essentially a layer of material lying on top of a base substrate, deposition of materials can be in the form of particles, clusters, and fibers that form a film that adhere onto the substrate. Polymers have long been in use in the coating industry due to low production cost, excellent adhesion, high compatibility of blends, high tolerance for engineering properties by synthetic route, most of them also contain useful functional groups. In addition, polymers can relatively easily accommodate other classes of materials such as carbon nanotubes or ceramic nanoparticles for applications requiring resistance to heat [5].

The self-cleaning ability of surfaces can be classified by the wetting mechanisms, and there are two primary categories: hydrophobicity (repel-water) and hydrophilicity (like-water), and are determined by the water contact angle (WCA). Hydrophobic surfaces would have a WCA of $> 90^\circ$, whereas hydrophilic surfaces are $< 90^\circ$. These are often intrinsic properties that are related to the material's surface tension, where low energy materials would show hydrophobicity [6].

II. BACKGROUND

Extreme conditions of superhydrophobicity ($\text{WCA} > 150^\circ$) and superhydrophilicity ($\text{WCA} \sim 0^\circ$) can be obtained through introducing surface roughness and hierarchical structures as shown in Figure 1. Superhydrophobic coating allows dirt to be carried away by water droplets, whereas superhydrophilic coatings cause complete wetting with a thin layer of water covering the surface, and dirt is readily being washed away by incoming stream of water [1]. The related phenomenon is explained by the Cassie-Baxter model in Equation (1) [7].

There are two volume fractions f_1 (rough solid) and f_2 (air), where they sum up to 1. θ_1 and θ_2 are contact angles with their respective entity. θ_2 is close to 180° as it has complete contact with air in between the roughness, therefore according to this model, a hydrophobic surface will approach superhydrophobicity for observed contact angle θ , and vice versa for hydrophilicity. In general, the introduction of roughness would cause a hydrophobic surface to become more hydrophobic, and respectively, a hydrophilic surface to become more hydrophilic. This relationship is explained by the Wenzel model presented in Equation (2). The roughness factor $r = 1$ when the surface is completely flat, whereas increasing the surface roughness will result in $r > 1$, therefore the Wenzel angle θ_w will increase if Young's angle $\theta_Y > 90^\circ$, and decrease if $\theta_Y < 90^\circ$ [8].

$$\cos\theta = f_1 \cos\theta_1 + f_2 \cos\theta_2 \quad (1)$$

$$\cos\theta_w = r \cos\theta_Y \quad (2)$$

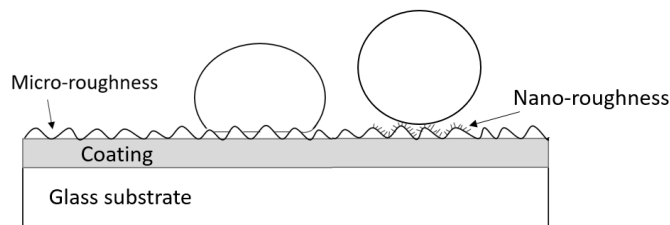


Figure 1. Wenzel model (left), and Cassie-Baxter model (right)

Poly (methyl methacrylate) (PMMA) has excellent transmittance and adhesion to substrate, it is also durable and chemical resistant, making it suitable for transparent coating application. It has been intensively studied for coating applications. [9] However, the nature of PMMA is hydrophilic, with a WCA of $\sim 70^\circ$. In order to convert PMMA to hydrophobic coating, the use of PMMA itself is often not sufficient.

X. Wang, et al., have mixed PMMA and PS at 30/70 (wt/wt) in THF to form 5wt% solution that was spin casted onto glass slides, and employed phase separation technique by exposing the coating in cyclohexane at elevated temperature, followed by immersing in water to remove the PS contents, resulted in superhydrophobicity with WCA of 153.2° . [10] H. Liu, et al., have used a one-step spraying process of 8 wt% of THF dissolved PDMS/PMMA at 50/50 (wt/wt) mixture, the WCA observed was 157.7° without pre- and post-treatments.

The coating was also able to transform into superhydrophilic state by hot press. [11] Other combinations have also been studied, R.P.S. Chakradhar, et al., have produced a stable superhydrophobic surface by spray coating PVDF/MWCNT nanocomposite, it was found that CNT promoted phase separation and had an effect of increasing hydrophobicity. [12] The use of ionic compounds was also found to be effective in assisting in obtaining superhydrophobicity. C. Peng, et al., have spray coated a wind turbine blade with mixture of PVDF/Sodium bicarbonate in DMF, the coating was observed to be superhydrophobic with WCA of 156° and had anti-icing property. [13]

III. EXPERIMENTAL

All experimental materials were used as received without pre- and post-treatments. The WCAs and calculated surface tensions from literature are listed in Table 1 [14, 15, 16]. PMMA average Mw ~ 350000 g/mol; PVC high molecular weight grade; PS average Mw ~ 350000 g/mol, Mn ~ 170000 g/mol; PVDF average Mw ~ 275000 g/mol, Mn ~ 107000 g/mol; PDMS vinyl terminated, average Mw ~ 25000 g/mol, viscosity 850 – 1150 cSt were acquired from Sigma Aldrich. N, N-Dimethylformamide (DMF) reagent A.C.S. grade was purchased from Produits Chimiques ACP Chemicals Inc. Chemical structures of the polymers used are shown in Figure 2.

Figure 3 shows the spray setup, the dual trigger spray gun was connected to the compressed air source through cotton, wire mesh filter, and desiccants. Air pressure was regulated by a gas regulator. 5 wt% solutions were prepared with 50/50 (wt/wt) polymer blends in 10 mL DMF. Solutions were stirred on magnetic stirrer at 200 RPM for 12 hours to obtain a homogeneous mixture. Heating to 40°C was required for PMMA/PVDF to aid swelling of PVDF. Solutions were deposited onto glass substrate at 20 psi, 15 cm from tip of nozzle (0.5 mm). It was allowed for 30 seconds of drying time in between every 10 sprays horizontally, this process was repeated for 20 times. The environment around the substrate was kept at 30°C .

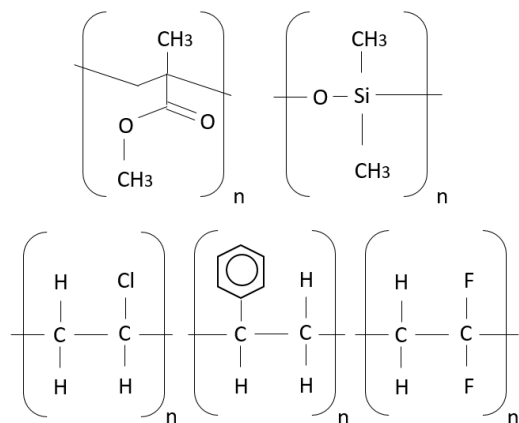


Figure 2. Polymer chemical structures of (Top-left) PMMA, (Top-right) PDMS, (Bottom-left) PVC, (Bottom-middle) PS, and (Bottom-right) PVDF.

TABLE I. WATER CONTACT ANGLES AND SURFACE TENSIONS OF POLYMERS USED

Polymer	WCA [14, 15]	Surface tension (mN/m) [16]
PMMA	68°	41.0
PVC	86°	40.0
PS	87°	39.7
PVDF	89°	29.2
PDMS	107°	22.8

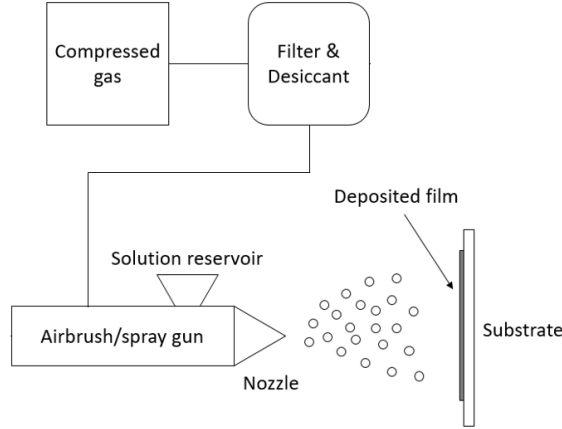


Figure 3. Spray deposition setup.

Digital microscope was used for the characterization of the experimentally obtained microstructures at 100x magnification and static WCA at 20x magnification. The microscope was tilted to 90° to obtain images that are parallel to the stage. The WCA measurements were performed using the built-in automatic calibrated features in the microscope system.

IV. RESULTS AND DISCUSSIONS

A. Water Contact Angles

Static WCAs were measured in this study to observe the effect of the blended polymers and their functional groups. Uncoated glass was observed to have a WCA of 36° as shown in Figure 4 (a), whereas pure PMMA coating in DMF solvent was measured to be 60°, similar to the published value in the literature. An improvement of the wetting behavior towards hydrophobicity was observed when PMMA was blended with another polymer of lower surface tension, the resulting measured WCAs are presented in Table 2.

In general, the measured WCAs for the experimental polymer blends were found to have a give-and-take relationship between PMMA and the second polymer, where all the WCAs achieved were larger than 60° but smaller than their individual WCA of the second polymer participating in the blend. Among all second choices of polymers, PDMS showed to be the most effective in converting hydrophilic

PMMA to hydrophobic with a WCA close to 100° as shown in Figure 4 (b).

B. Effect of Polymer Blends and End Functional Groups

PVC and PVDF are semi-crystalline polymers, therefore resulted in translucent coating, as crystals deflect light. On the other hand, PMMA, PS, and PDMS are amorphous polymers, which allowed better transmittance. Figure 5 shows the as sprayed samples on top of printed paper to demonstrate the degree of transparency, and Figure 6 shows the microstructures of the coatings fabricated with polymer blends. Although PS and PVC have characterized with a surface tension of a similar value, PS was not miscible with PMMA. Consequently, it created more phase separation into PMMA- and PS-rich regions. These regions were also observed to be larger than those of PMMA/PVC as shown in Figure 6 (top-left) and (top-right), therefore PS was more effective in increasing the WCA.

PVC and PVDF have similar structure, differing only by two components in the tertiary structure, all of H, Cl, and F molecules are small in size unlike the benzene ring in PS. Both PVC and PVDF mixed well with PMMA which was evident by the relatively homogenous microstructure produced as shown in Figure 6 (top left) and (bottom-left).

Since PVDF contains two fluorides, contributing to lower intermolecular forces than the single chloride found in PVC. This resulted in low surface tension for PVDF, which is one of the criterion to be water-repellent, such that PMMA/PVDF was found to be close to hydrophobic state with 20° improvement.

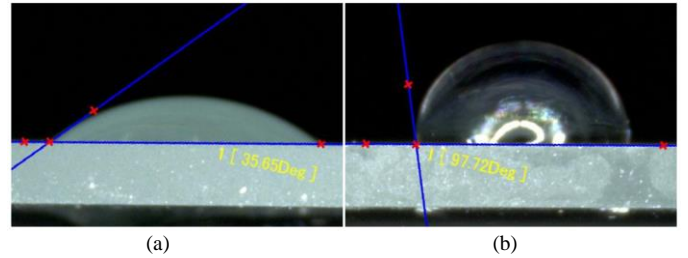


Figure 4. Water contact angle of (a) uncoated, and (b) PMMA/PDMS coated glass slides.

TABLE II. MEASURED EXPERIMENTAL WATER CONTACT ANGLES

Material	WCA
Uncoated glass	36°
PMMA	60°
PMMA/PVC	65°
PMMA/PS	72°
PMMA/PVDF	80°
PMMA/PDMS	98°

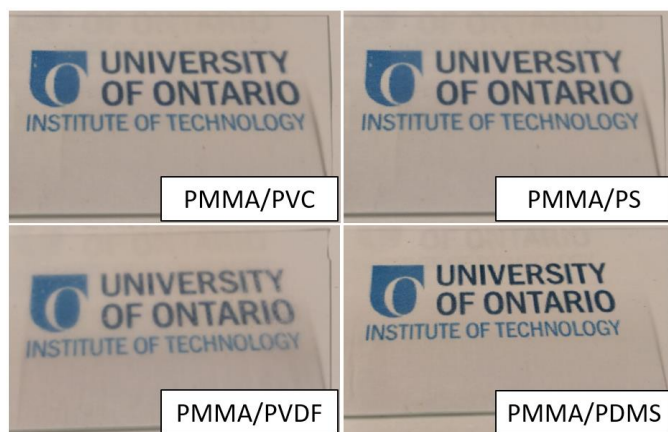


Figure 5. Coating transparency from polymer blends.

The structure of as sprayed PMMA/PDMS is shown in Figure 6 (bottom-right), the transmittance was found to be highest for this combination, and a completely amorphous microstructure was obtained. It suggests that the coating had low surface roughness as small roughnesses would cause light scattering and lower the transmittance. Both polymers did not have high miscibility due to the huge difference in surface tensions, PDMS was therefore distributed evenly by vigorous mixing to create a homogenous suspension. The coating was found to be wetter than all other blends during spray deposition process, probably due to the lower viscosity of solution, resulting in less force required to atomize from the airgun nozzle. The measured WCA of 98° was primarily due to the low surface tension of PDMS, and higher surface roughness will be required by fine-tuning spray parameters to obtain the superhydrophobic state.

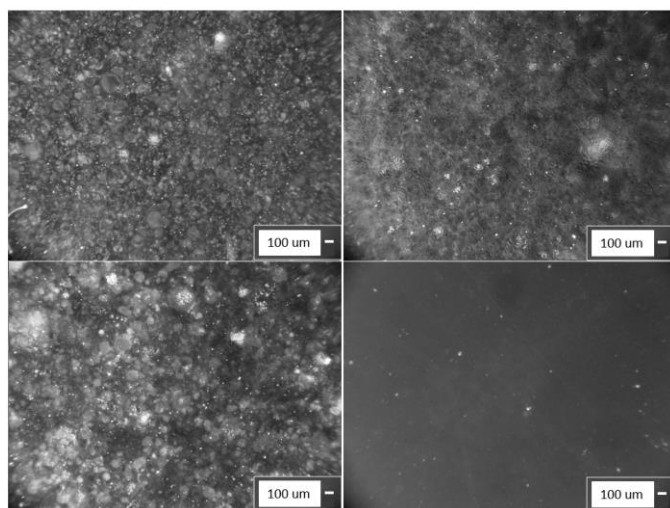


Figure 6. Microstructures of coatings (Top-left) PMMA/PVC, (Top-right) PMMA/PS, (Bottom-left) PMMA/PVDF, (Bottom-right) PMMA/PDMS.

V. CONCLUSIONS

The wetting behavior of polymer blends of PMMA with PVC, PS, PVDF, and PDMS was investigated at 50/50 (wt/wt)

blending ratio. The coatings were fabricated using spray deposition technique with compressed air and spray-gun. Water contact angle (WCA) of glass slide was observed to be 36° and was improved by depositing polymer coatings on top of the glass surface. WCAs increased with decreasing surface tensions of second polymer in the blend. PVC, PS, and PVDF were not sufficient to convert the intrinsically hydrophilic PMMA into hydrophobic coating at the studied blending ratio and spray parameters. The conversion to hydrophobic coating was successful with the use of PDMS, the WCA was measured to be close to 100° with 40° improvement as compared to pure PMMA coating. Transparency of coating was found to decrease with higher miscibility of polymer blends which resulted in homogeneous phase but light scattering within the small deposited particles. Increasing the surface roughness of PMMA/PDMS coating through fine-tuning of spray parameter combinations between pressure, spray distance, spray time and temperature, will potentially result in higher WCA approaching superhydrophobic state.

ACKNOWLEDGMENT

The authors are grateful to NSERC's Discovery Development Grant and NSERC's Chairs in Design Engineering Program for the financial support of this research.

REFERENCES

- [1] M. Ma, R. M. Hill, G. C. Rutledge, "A review of recent results on superhydrophobic materials based on micro-and nanofibers", *Journal of Adhesion Science and Technology*, vol. 22, 1799-1817, 2008.
- [2] Y. Gong, K. Fu, S. Xu, J. Dai, T. R. Hamann, L. Zhang, G. T. Hitz, Z. Fu, Z. Ma, D. W. McOwen, X. Han, L. Hu, E. D. Wachsman, "Lithium-ion conductive ceramic textile: A new architecture for flexible solid-state lithium metal batteries", vol. 21, issue 6, pp. 594-601, 2008.
- [3] L. Y. L. Wu, Q. Zhao, H. Huang, R. J. Lim, "Sol-gel based photochromic coating for solar responsive smart window", *Surface and Coatings Technology*, vol. 320, pp. 601-607, 2017.
- [4] N. Shiraishi, M. Kimura, Y. Ando, "Development of PMMA-based gas sensor and its evaluation using a VOC dilution flow system", *Microelectronic Engineering*, vol. 119, pp. 115-121, 2014.
- [5] K. I. Winey, R. A. Vaia, "Polymer Nanocomposites", *MRS Bulletin*, vol. 32, Issue 4, pp314-322, 2007.
- [6] M. Lewin, A. Mey-Marom, R. Frank, "Surface free energies of polymeric materials, additives and minerals", vol. 16, issue 6, pp. 429-441, 2005.
- [7] T. L. Liu, Z. Chen, C. Kim, "A dynamic Cassie-Baxter model", *Soft Matter*, vol. 11, pp. 1589-1596, 2015.
- [8] J. Tam, G. Palumbo, U. Erb, "Recent advances in superhydrophobic electrodeposits", *Materials*, vol. 9, issue 3, no. 151, 2016.
- [9] Y. F. Shih, M. Y. Chou, H. Y. Lian, L. R. Hsu, S. M. Chen-Wei, "Highly transparent and impact-resistant PMMA nanocomposites reinforced by cellulose nanofibers of pineapple leaves modified by eco-friendly methods", *Express Polymer Letters*, vol. 12, no. 9, pp. 844-854, 2018.
- [10] X. Wang, N. Yang, Q. Xu, C. Mao, X. Hou, J. Shen, "Preparation of a novel superhydrophobic PMMA surface with nanostructure and its blood compatibility", *e-Polymers*, no. 081, 2012.
- [11] H. Liu, J. Huang, Z. Chen, G. Chen, K. Zhang, S.S. Al-Deyab, Y. Lai, "Robust translucent superhydrophobic PDMS/PMMA film by facile one-step spray for self-cleaning and efficient emulsion separation", *Chemical Engineering Journal*, vol. 330, pp. 26-35, 2017.

- [12] R.P.S. Chakradhar, G. Prasad, P. Bera, C. Anandan, "Stable superhydrophobic coatings using PVDF-MWCNT nanocomposite", *Applied Surface Science*, vol. 301, pp. 208-215, 2014.
- [13] C. Peng, S. Xing, Z. Yuan, J. Xiao, C. Wang, J. Zeng, "Preparation and anti-icing of superhydrophobic PVDF coating on a wind turbine blade", *Applied Surface Science*, vol. 259, pp. 764-768, 2012.
- [14] Y. Ma, X. Cao, X. Feng, Y. Ma, H. Zou, "Fabrication of superhydrophobic film from PMMA with intrinsic water contact angle below 90°", *Polymer*, vol. 48, issue 26, pp. 7455-7460, 2007.
- [15] Diversified Enterprises, "Critical Surface Tension and Contact Angle with Water for Various Polymers", 2017.
- [16] Polymer Properties Database, Surface tensions of solid polymers, 2015.

IMPACT OF FIBER MODIFICATION ON THE PROPERTIES OF NATURAL FIBER COMPOSITES

Nabeel Ahmed Syed, Utkarsh, Ghaus Rizvi, and Remon Pop-Iliev

Faculty of Engineering and Applied Science
UOIT-University of Ontario Institute of Technology, Oshawa, Canada
nabeelahmed.syed@uoit.net

Abstract—Utilizing natural fibers as reinforcement in the production of composites has caught researcher's attention in recent years, which is motivated by the need for eco-friendly, efficient, light weight and sustainable substitutes for petroleum-based materials. The most attractive properties of natural fiber composites include relatively low cost of production, light weight, low density, non-toxicity, resistance to abrasion, and CO₂ neutrality. This paper focuses on combining Abaca Fiber (AF) fiber with a Vinyl Ester (VE) thermosetting resin into a composite. As the most critical aspect of natural fiber-based composite processing is the compatibility between the fiber and the polymer matrix, utmost importance needs to be given to modify the fiber and the matrix for achieving better bonding. The presented research work focusses on scrutinizing the influence of three different types of surface treatments of Abaca Fiber and its resultant effect when processed with Vinyl Ester resin to fabricate AF/VE composites using Vacuum Assisted Resin Transfer Molding. The post-modification experimental results revealed that the performance of the resulting composites was significantly enhanced as a result of the applied fiber surface treatment, which notably improved the compatibility between the fiber and the polymer matrix.

Keywords- *Abaca fiber; Vinyl ester; Fiber treatment, Mechanical properties*

I. INTRODUCTION

This paper focuses on utilizing natural eco-friendly, low cost, light weight and renewable materials for useful applications that would work as effectively when compared to materials produced by exhausting traditional resources such as petroleum products [1-2].

The most critical aspect of a natural fiber composite processing is the compatibility between the fiber and the polymer matrix. Utmost importance needs to be given to modify the fiber and the matrix to achieve better bonding of the two.

Abaca fiber is a natural fiber which exhibits good mechanical properties. Abaca, which is commonly referred to

as "Manila Hemp," is extensively produced in the Philippines. The chemical composition of Abaca fiber consists of about 55% of cellulose, 24-28% of hemicellulose, 11-13 % of lignin along with small fragments of ash, pectin, and waxes [3]. Abaca, in distinguished from other natural fibers by its favorable moisture resistance and dye capacity without losing any lustrous characteristics. However, the short coming of this fiber is its hydrophilic nature which makes it difficult to bond with a polymer matrix to perform efficiently. Hence, surface treating the fiber with suitable chemicals is essential to impart some hydrophobicity into the fiber [4].

II. BACKGROUND

Adequate surface treating of fibers improves the compatibility of the fiber and the polymer matrix [5]. Alkali treatment is one of the useful techniques for improving the properties of natural fibers [6]. Saha et al. investigated that alkali treatment on jute fibers resulted in and enhancement of 65% in the tensile strength and 38% in the elongation at break [6]. Gomes et al. indicated that alkali treatment of 10 wt. % increased the mechanical properties of the curaua fibers [7]. Prasad et al. reported an improvement in the mechanical properties of coir/polyester composite after surface treating the coir fiber with 5 wt. % NaOH treatment [8]. Similarly, the acidic treatment with KMnO₄ is also a useful method to improve the adhesion between the fiber and the matrix to increase the mechanical performance of the composite. J.T. Omole et.al reported the enhancement of tensile strength of KMnO₄ treated bagasse fiber/polyester composite up to 35% when compared to the untreated composite [9]. Poly Vinyl Alcohol (PVA) treatment on cellulosic fibers is also adopted in order to improve the compatibility and adhesion between the matrix and the reinforcement. D.G Hepworth et.al investigated the influence of PVA on natural fiber composite and reported that the strength and the stiffness of the PVA treated composite increased after treatment [10].

Vinyl ester resin has a thermosetting matrix, which is a combination of Polyester and epoxy. It exhibits thermal properties and mechanical behavior similar to epoxy while manifesting easier processing and lower cost than polyester resin [11]. Vinyl ester has proved to be an essential matrix which displays water and corrosion resistant behavior.

The Vacuum Assisted Resin Transfer Molding (VARTM) fabrication technique is a widely known method implemented in composite manufacturing industries. This process uses only one-sided molds where the fibers are laid up and covered by a vacuum bag. The infusion of resin takes place through the resin inlet pipes which connects the fiber layup. This technique is often used to produce medium and large-sized composite bodies such as yachts hulls. In addition, lower fabricating cost without compromising the properties of the composite is achieved by utilizing VARTM [12].

The present study concentrates on structurally modifying the Abaca fiber by surface treating with basic (NaOH), acidic (KMnO₄) and polymeric (PVA) chemicals. Consequently, the performance of mechanical properties of the modified AF/VE composite were evaluated.

III. EXPERIMENTAL

The selected Abaca fiber had a density of 1.5g/cm³ and was purchased from Philippines. NaOH, KMnO₄, PVA as chemicals for treatment, Vinyl ester acting as a resin having viscosity of 150 cps and exhibited specific gravity of 1.03, Methyl Ethyl Ketone Peroxide acting as a hardener, Cobalt Napthalate acting as an accelerator and Frekote 700 – NC (releasing agent) was purchased from Sigma Aldrich, USA.

The bundles were carefully striped off from the Abaca fiber with a clipper. The bundles were grouped in 4 categories. First group of bundle fibers was the untreated Abaca. Second, third and fourth group of bundle fibers were treated by immersing in a solution of NaOH, KMnO₄ and PVA respectively. Three different compositions for each chemical treatment were prepared. The fiber immersion time for all the 3 types of treatments was 30 minutes at room temperature. The treated fiber was washed with water thoroughly after the treatment until their pH was 7. These fibers were then air dried for 5 hours and subsequently oven dried for the next 24 hours before subjecting them for testing and characterization. Similarly, Abaca fiber was treated with NaOH, KMnO₄ and PVA by following the above explained procedure. The various compositions of chemicals for the treatment of fiber and their chemical reactions are presented in Table 1. The treated fibers were reinforced to Vinyl ester resin to fabricate the composite by VARTM technique. MEKP (hardener) and Cobalt Napthalate (accelerator) was incorporated at a ratio of 1 weight percentage of the resin. The vacuum pressure throughout the process was maintained at 0.04 MPa. The infusion process as well as curing period was performed at room temperature. The composite was removed from the experimental setup after a curing period of 12 hours.

A. Morphological Analysis

The morphology of Abaca fiber specimens of untreated and 3 types of chemically treated specimens were captured with the aid of CZ/MIRA I LMH which had 1.5nm resolution. Micrographs were studied at various resolutions. Initially, the samples were kept at a temperature of 60°C in the hot air oven

for 24 hours. The samples were then subjected on top of the substrate by an adhesive. Platinum was coated to both unmodified and modified fibers (as it is non-conductive in nature) by the process of vacuum evaporation keeping the acceleration voltage in the range of 5-8 kV.

Table.1. Composition of chemical surface treatment on Abaca

Chemical Treatment	Composition	Specimen Code of VE/AF Composite
Untreated	-	UT
NaOH	1%	1-N
	2%	2-N
	3%	3-N
KMnO ₄	0.2%	0.2-K
	0.4%	0.4-K
	0.6%	0.6-K
PVA	2%	2-P
	4%	4-P
	6%	6-P

B. Tensile Test

All the specimens that were to be tested for tensile properties were first conditioned at 60°C for 24 hours. The tensile testing of untreated and 3 types of chemically treated fibers composites was carried out by the MTS tensile test machine adhering to the standards specified in ASTM D638 with the help of MTS 97 kN load cell which had a crosshead speed of 3mm/min. The tests were carried out at room temperature.

IV RESULTS AND DISCUSSION

A. Morphological Analysis

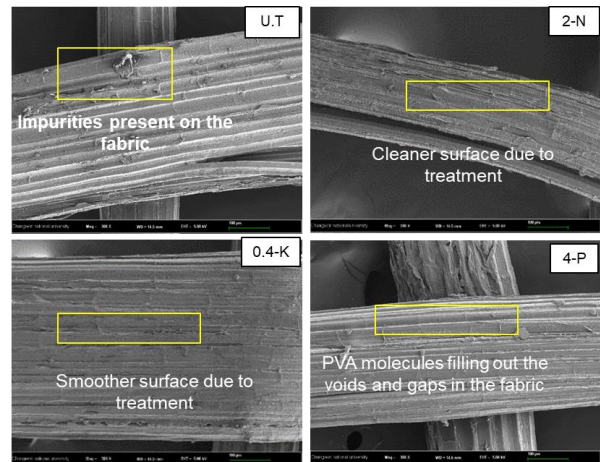


Fig.1 SEM Images of untreated and treated fiber

Pure Abaca and treated Abaca fiber was observed for morphological characteristics via FESEM and the respective micrographs are presented in Fig. 1. The untreated Abaca fiber showed presence of loosely bonded molecules on the surface. In the case of NaOH treated fiber, the micrographs showed that the fiber had become cleaner and smoother with the impurities being washed away due to the chemical treatment. A slight roughness was noticed in 2% NaOH treated Abaca. Similar results were obtained from Abaca fiber treated with 0.4% KMnO_4 . In the case of PVA treatment, the PVA molecules were observed being coated on the fiber surface. The gaps and the voids were reduced with higher concentrations of PVA treatment thereby increasing the friction which is essential for better bonding [13].

B. Tensile Test

The tensile properties of pure, NaOH, KMnO_4 treated and PVA treated AF/VE composites are depicted in Fig. 2. The tensile strength of 2 wt. % NaOH treated composite increased by 57% and it then reduces as the concentration of NaOH increases. After treatment when hemicelluloses are partially removed, the region corresponding to interfibrilous will become less dense resulting in the alignment of the fibrils in the tensile deformation direction thereby increasing the tensile properties [14]. Hence the tensile strength of 1 and 2 wt. % NaOH treated AF/VE increased. However, the fiber treatment involving higher concentrations of NaOH resulted in rupture and loss of structural strength, which eventually showed a decline of tensile properties of the treated composite [15].

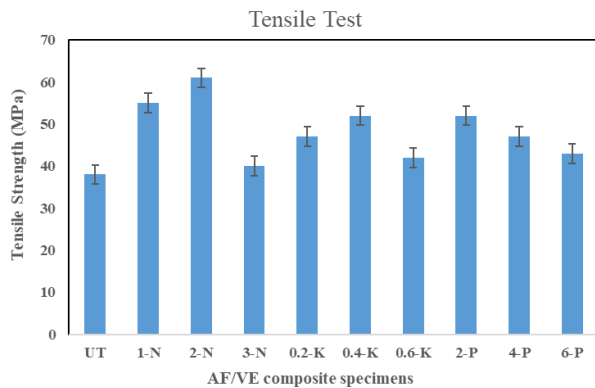


Fig. 2 Tensile properties of AF/VE Composites

The tensile strength of 0.4 wt. % KMnO_4 treated AF/VE composite increased by 35% which might be due to the effect of the treatment on the fiber [9]. It was observed that the tensile strength of KMnO_4 treated AF/VE composites increases up to 0.4 wt. % of KMnO_4 and then it reduces as the concentration of KMnO_4 increases. Higher concentrations of the treatment chemical destroy the structural stability of the fiber thereby in turn lowering the resultant properties of the composite as shown in Fig.2 [9].

The tensile strength of 2 wt. % PVA treated AF/VE composite increased by 36% in comparison with the untreated AF/VE composite as shown in Fig.2. The reason could be due to the improved bonding with the matrix and the reinforcement after the surface treatment of the fiber by PVA [10]. However,

with increasing concentration of PVA, the molecules penetrate and sticks into the gaps between the fibers orientations on the fiber. This results in lower wettability of the fiber with the matrix thereby reducing the mechanical properties as shown in Fig. 2 [13].

IV. CONCLUSION

The influence of surface treatment on Abaca fiber on the properties of AF/VE composite was investigated. Abaca fiber was subjected to NaOH, PVA and KMnO_4 solutions to achieve surface treatment of the fiber via immersion method. The morphological behavior of treated fibers proved that the surface got rougher with the increase in concentration of the surface treating chemicals. The surface treated Abaca/Vinyl ester composite subjected to tensile tests showed significant improvement due to treated AF/VE in comparison with untreated composites. Overall, the modified AF/VE composites can potentially find its applications in the automotive industry preferably in the interior sections because of their improved performance against pure AF/VE composites.

ACKNOWLEDGMENT

The authors are grateful to NSERC's Discovery Development Grant and NSERC's Chairs in Design Engineering Program for the financial support of this research.

REFERENCES

- [1] D. D. Stokke, Q. Wu, G. Han, "Introduction to Wood and Natural Fiber Composites", John Wiley & Sons, West Sussex, UK: 2014.
- [2] N. Suizu, T. Uno, K. Goda, J. Ohgi, "Tensile and impact properties of fully green composites reinforced with mercerized ramie fibers," Journal of Materials Science, vol. 44, pp. 2477-2482, May 2009.
- [3] C. Wong, R. Kiew, G. Argent, O. H. Set, S.K. Lee, Y. Y. Gan, "Assessment of the validity of the sections in Musa (Musaceae) using AFLP," Annals of botany, vol. 90, pp. 231-238, Aug 2002.
- [4] A. K. Bledzki, A. A. Mamun, O. Faruk, "Abaca fibre reinforced PP composites and comparison with jute and flax fibre PP composites," eXPRESS polymer letters, vol. 1, pp. 755-762, Nov 2007.
- [5] A. Roy, S. Chakraborty, S. P. Kundu, R. K. Basak, S. B. Majumder, B. Adhikari, "Improvement in mechanical properties of jute fibres through mild alkali treatment as demonstrated by utilisation of the Weibull distribution model," Bioresource technology, vol. 107, pp. 222-228, Mar 2012.
- [6] P. Saha, S. Manna, S. R. Chowdhury, R. Sen, D. Roy, B. Adhikari, "Enhancement of tensile strength of lignocellulosic jute fibers by alkali-steam treatment," Bioresource technology, vol. 101, pp. 3182-3187, May 2010.
- [7] A. Gomes, T. Matsuo, K. Goda, J. Ohgi, "Development and effect of alkali treatment on tensile properties of curaua fiber green composites," Composites Part A:

Applied Science and Manufacturing, vol. 38, pp. 1811-20, Aug 2007.

- [8] S. V. Prasad, C. Pavithran, P. K. Rohatgi, "Alkali treatment of coir fibres for coir-polyester composites," *Journal of materials science*, vol. 18, pp. 1443-1454, May 1983.
- [9] Omole JT, Dauda B, "Physical and Mechanical Properties of Chemically Treated Bagasse Fibre for Use as Filler in Unsaturated Polyester Composite". *International Journal of Composite Materials* vol 6, pp. 48-54, 2016.
- [10] D. G. Hepworth, D. M. Bruce, J. F. Vincent, G. Jeronimidis, "The manufacture and mechanical testing of thermosetting natural fibre composites," *Journal of materials science*, vol. 35, pp. 293-298, Jan 2000.
- [11] B. D. Agarwal, L. J. Broutman, K. Chandrashekara, "Analysis and performance of fiber composites," John Wiley & Sons, Oct 2017.
- [12] C. Xia, H. Ren, S. Q. Shi, H. Zhang, J. Cheng, L. Cai, K. Chen, H. S. Tan, "Natural fiber composites with EMI shielding function fibered using VARTM and Cu film magnetron sputtering," *Applied Surface Science*, vol. 362, pp. 335-340, Jan 2016.
- [13] H. Ishikawa, H. Takagi, A. N. Nakagaito, M. Yasuzawa, H. Genta, H. Saito, "Effect of surface treatments on the mechanical properties of natural fiber textile composites made by VaRTM method," *Composite Interfaces*, vol. 21, pp. 329-336, May 2014.
- [14] M. Cai, H. Takagi, A.N. Nakagaito, Y. Li, G. I. Waterhouse, "Effect of alkali treatment on interfacial bonding in abaca fiber-reinforced composites," *Composites Part A: Applied Science and Manufacturing*, vol. 90, pp. 589-597, Nov 2016.
- [15] A. El Oudiani, Y. Chaabouni, S. Msahli, F. Sakli, "Crystal transition from cellulose I to cellulose II in NaOH treated *Agave americana* L. fibre," *Carbohydrate Polymers*, vol. 86, pp. 1221-1229, Aug 2011.

On Reliability of Thermally-Sprayed Alumina Dielectric Layers in a Multi-layered Coating-based Heating System

Morvarid Mohammadian Bajgiran, Christian Moreau

Department of Mechanical, Industrial and Aerospace
Engineering (MIAE)
Concordia University
Montreal, QC, Canada
morvarid.mohammadian@mail.concordia.ca
christian.moreau@concordia.ca

Milad Rezvani Rad, André McDonald

Department of Mechanical Engineering
University of Alberta
Edmonton, AB, Canada
milad.rezvani@ualberta.ca
andre.mcdonald@ualberta.ca

Abstract— In coating-based resistive heating systems, for cases in which the substrate is electrically conductive, an intermediary ceramic layer with dielectric properties is required to prevent short-circuiting and leakage of current between the resistive coating heating element and the substrate. Alumina is widely used as an electrical insulator in heating systems thanks to its dielectric properties. A comparative study was conducted with aluminum oxide (Al_2O_3) coatings deposited by using two different thermal spraying processes, namely flame spraying (FS) and suspension plasma spraying (SPS). The morphology and microstructure of the deposited coatings and their mechanical failure were studied by means of a scanning electron microscope (SEM). Phase composition of the samples was examined by using X-ray diffraction (XRD). Porous coating structures were obtained from the FS process, while several coating structures that were dense, cauliflower-like, and porous were observed in the SPS samples. Corundum content in the SPS samples varied from almost 30% to 70% by changing the spraying parameters and the FS samples contained about 60% of corundum. Corundum (alumina α phase) has better stability and hygroscopic properties than the γ phase, which is usually formed under rapid solidification in thermal spray processes. The higher proportion of corundum that can be achieved by the SPS process makes SPS-based alumina coatings a good candidate for use as electrical insulating layers in resistive heating coating systems.

Keywords: Alumina, Dielectric properties, Flame spray, Phase analysis, Suspension plasma spray

I. INTRODUCTION

Thermal-sprayed coatings as resistive heating systems have attracted considerable interest in the recent years. A growing body of literature regarding the fabrication of such heating systems has been documented, starting more than two decades ago with the work done by Younis, *et al.* [1]. Various approaches have been proposed to deposit thermal-sprayed

coatings onto different substrate materials for heating applications. It is well known that the electrical conductivity of the substrate determines the configuration of the heating system. According to a study conducted by Lopera-Valle and McDonald [2], direct deposition of the coating heating element on non-conductive fiber-reinforced polymer composites (FRPC) is possible; while, for cases in which the substrate is electrically conductive, direct deposition of the heating element onto the substrate is not suitable because it leads to short circuiting between the heating element and the conductive substrate. Therefore, an intermediary layer with dielectric properties is required to prevent such undesirable phenomenon, while still allowing for heat transfer through the layers.

Alumina is one of few candidates that has broadly been used for electrical insulation purposes [3 – 5]. Corundum (α - alumina), the only thermodynamically stable form of alumina, possesses favorable properties, namely high chemical stability, high melting point, and high electrical resistance at high temperatures [6]. However, thermally-sprayed alumina coatings principally consist of metastable phases such as γ - or δ - alumina with significantly different properties compared to corundum [5, 7]. Rapid cooling of the molten alumina particles at a lower temperature before the formation of corundum is the main cause of this phase transformation. Then, the reaction of the γ -alumina with water in a humid environment adversely affects the dielectric properties of the material [8]. According to McPherson [9], during the cooling of the alumina particles with diameters less than 10 μm during the thermal spray deposition process, the metastable form of the material is maintained, while larger particles may transform to α -alumina during solidification. Thus, there has been a continuous struggle to retain α -alumina to the greatest possible extent in thermally-sprayed coatings [4], [5], [10].

Although flame-sprayed (FS) alumina is reported to have a porous microstructure [11], suspension plasma-sprayed (SPS) alumina can be produced to have various microstructures, with dense, porous, and columnar features. VanEvery, *et al.* [12] first proposed an explanation for formation of the columnar structure in SPS coatings. It was reported that the plasma drag

forces control the droplet inertia for cases in which powder particles are sufficiently small. This results in the generation of another velocity component that is along the surface of the substrate in addition to the velocity component that is perpendicular to the substrate surface. The peaks of the asperities on the substrate can block the path of the oblique droplet trajectories near the substrate surface and prevent them from reaching the valleys downstream of the asperities during spreading. The microstructure of alumina coatings deposited by using a high enthalpy hybrid water-stabilized plasma torch was investigated by Tesar, *et al.* [10]. Formation of both columnar and well-sintered dense structures for alumina coatings was observed in the coatings that were fabricated by using the suspension plasma spraying technique.

Ganvir, *et al.* [13] investigated plasma-sprayed thermal barrier coatings (TBC) and their microstructures. It was concluded that both particle momentum and substrate roughness have significant effect on the final microstructure of the deposited coatings. Droplet trajectory was significantly affected by lower momentum of the particles and deposition occurred only on the substrate asperities, resulting in the formation of a columnar structure. A comparison study conducted by Müller, *et al.* [14] on the alumina coatings that were produced by suspension plasma spraying and high-velocity suspension flame spraying (HVSFS) processes showed that relatively higher porosity values were obtained from SPS deposition compared to HVSFS method.

Considering the extensive research done on the fabrication of coating-based heating systems with different thermal spray processes and their electrical and heating performances, few detailed comparative studies have been documented on the microstructures that were developed and the phases that were formed during the deposition of the intermediate insulating alumina layer. The present study aims to investigate specifically the reliability aspects related to the morphological and phase transition of alumina coatings deposited by flame spraying and suspension plasma spraying processes.

II. EXPERIMENTAL METHODS

A. Substrate Preparation

The substrates used in this study were obtained from several 6-mm (0.25-inch) thick steel plates (A36/44w mild steel hot rolled flat bar, Metal Supermarkets, Edmonton, AB, Canada). These plates were water jet cut into two different sizes, which were 127-mm (5-inch) long and 25-mm (1-inch) wide rectangular samples that were employed for the heating tests and smaller 25-mm (1-inch) long and 25-mm (1-inch) wide square samples that were utilized for microstructural analyses.

In order to create roughened surfaces for improved adhesion of FS alumina to the substrate, the samples were grit-blasted with #24 alumina grit (Manus Abrasive Systems Inc., Edmonton, AB, Canada) at an air pressure of 586 kPa (85 psig). The samples for deposition of SPS alumina were also grit-blasted by using #80 alumina grits (International Surface Technologies, Montreal, QC, Canada) at an air pressure of 690 kPa (100 psig). The finer alumina grits used for roughening the

samples before the deposition of SPS alumina is mainly due to the smaller particle size of powder in the SPS process compared to the FS process to ensure the required adhesion at the coating-substrate interface. The samples were preheated prior to deposition of the coatings in both processes by passing the torch on the samples with the same spraying parameters that were used for deposition of FS and SPS alumina coatings.

B. Deposition of Electrically Insulating Coatings

Alumina (Al_2O_3 , AMDRY 6060, Oerlikon Metco, Westbury, NY, USA) powder feedstock was fed to an oxy-acetylene flame spray torch (6P-II, Oerlikon Metco, Westbury, NY, USA). The size distribution of the alumina powder particles were 5 to 45 μm ($-45+5 \mu\text{m}$). Argon, as the carrier gas, was used for delivering the powder particles to the torch. A volumetric powder feeder (5MPE, Sulzer Metco, Westbury, NY, USA) was employed to feed the powder to the oxy-acetylene flame spray torch consistently. Detailed spraying parameters of the FS process are mentioned in Table I.

For the deposition of SPS alumina, an alcoholic suspension was utilized as a carrier liquid to convey the fine powder particles (Al_2O_3 , VG-503#, ZiBo V.Gree Trading Co., Zibo, Shandong, China) to the spraying gun. The size distribution of the SPS alumina powders was ($-6+0.2 \mu\text{m}$) with nominal average particle size between 0.45 and 1 μm . A laser diffraction particle size analyzer (Malvern Instruments Ltd., England) was used to characterize the powder size distribution in the prepared suspension. Polyvinylpyrrolidone (PVP360, Sigma-Aldrich, Oakville, ON, Canada) was added to the suspension as dispersant/surfactant to prevent sedimentation of the powder in the suspension. The fine sub-micron alumina powders were dispersed with 10 wt% solid concentration in ethanol and dispersing agent content of 1 wt.%. The prepared suspension was continuously stirred with a magnetic stirrer and was sonicated until uniform dispersion of the powder was obtained. The suspension was then injected into the plasma flame of an Axial III plasma torch (Northwest Mettech, North Vancouver, Canada). Moreover, mechanical stirring of the suspension was done during the deposition process to avoid agglomeration of powder.

Total gas flow rates of 180 and 230 l/min of a mixture of argon, nitrogen, and hydrogen were introduced to the plasma torch. By applying 500 V and passing a current of 180 A, resulted in a plasma power of 90 kW.

TABLE I. FLAME SPRAYING PARAMETERS FOR DEPOSITION OF ALUMINA.

Acetylene flow [m^3/h]	1.3
Oxygen flow [m^3/h]	2.1
Argon pressure [m^3/h]	0.56
Powder flow rate (FMR)	100
Robot speed [mm/s]	250
Stand-off distance (SOD) [mm]	120
Number of passes	5
Number of preheating passes	1

Two different stand-off distance (SOD) values were considered to study their impact on the final microstructure. Continuous flow of compressed air was used to cool down the samples while spraying. The surface temperature of samples was monitored during the deposition of the coating by an infrared (thermographic) camera (A320, FLIR Systems, MA, USA). The three different suspension plasma spraying conditions are presented in Table II.

Given the objective of this study, which was to study the phase composition and mechanical failure of the thermal-sprayed alumina intermediary layer, the details regarding the development and functionality of other layers in the multi-layered coating-based heating system are not within the scope of the present study. That said, fabrication and performance of flame-sprayed nickel-chromium coating, as the heating element layer, and the cold-sprayed dense copper coating, as the electrical connection, which are shown in the following micrographs, can be found in detail in a previous study [15].

C. Characterization

SEM images were taken in secondary and backscattered electron mode by using a scanning electron microscope (Zeiss Sigma 300 VP-FE, Carl Zeiss Canada Ltd., Toronto, ON, Canada, and Hitachi S-3400 N, Tokyo, Japan). To perform microstructural analysis of the coating system, flat coupon samples were cross-sectioned and cold-mounted in an epoxy resin (LECO, Mississauga, ON, Canada). Comprehensive grinding was performed by using 180, 240, 320, 400, 600, 800, and 1200 silicon carbide grit papers (LECO, Mississauga, ON, Canada) and further polishing was conducted by using both 3 and 1 μm diamond slurry suspensions (LECO, Mississauga, ON, Canada). The mounted samples were coated with a thin layer of carbon by using a sputter coater (EM SCD 005, Leica Baltec Instrument, Balzers, Liechtenstein) with a carbon evaporation accessory to prevent the accumulation of static electric charges on the samples surface.

The XRD patterns and phase composition of both FS and SPS alumina coatings were examined by using X-ray diffractometers (Bruker D8-Discovery, Bruker AXS, Madison, WI, USA, and Philips X'Pert Pro, Malvern Panalytical, Westborough, MA, USA).

TABLE II. SUSPENSION PLASMA SPRAYING PARAMETERS FOR DEPOSITION OF ALUMINA.

Spraying conditions	S ₁	S ₂	S ₃
Total Gas Flow (TGF) [l/min]	230	230	180
Argon volume [%]	80	80	45
Nitrogen volume [%]	10	10	45
Hydrogen volume [%]	10	10	10
Solution feed rate [ml/min]	45	45	45
Robot speed [mm/s]	1000	1000	1000
Stand-off distance (SOD) [mm]	50	75	75
Number of passes	30	30	28
Number of preheating passes	6	6	6

The crystallinity and phase purity of the deposited coatings were determined with a copper anode and a graphite monochromator was used to filter *K*-alpha wavelength radiation ($\lambda = 1.54056 \text{ \AA}$). The 2θ diffraction angle was changed from 10° to 90° at a scanning rate of 0.02 degree/min and a dwell time of 1 second. In order to determine the position of the peaks, the X'Pert High Score plus (PW3212) software was employed. Quantitative analysis of the phases that were identified and the resultant crystallite sizes were calculated by using the Rietveld technique.

III. RESULTS AND DISCUSSION

A. Coating Characterization

Images were captured by using a scanning electron microscope (SEM) in secondary mode of feedstock alumina powders and the top surface of the alumina coatings that were deposited by the FS and SPS processes. However, the micrographs from the cross-section of the samples were taken in SEM backscattered mode. Figure 1 presents an image of the alumina powder that was used in the FS process and Figures 2(a) and 2(b) present the top view of the resultant coating. As is shown, considering the 5 to 45 μm size distribution of the alumina powder particles, the deposited coating has relatively large pores on the top surface, some of which are interconnected. These open pores can be a pathway for the penetration of the top heating element, which will lead to a reduction of the effective thickness of the electrically insulating layer.

The SEM image of the sub-micron sized alumina powder used in the SPS process is shown in Figure 3. Even though some large agglomerations were observed in the micrograph of the powder, according to the particle size distribution curve that is shown in Figure 4, it was observed that these large agglomerates were dissolved into the solvent resulting in distribution widths of D10, D50, and D90 of 0.2, 0.5, and 1.5 μm , respectively. Top surface of the SPS alumina-S3 is shown in Figures 5(a) and 5(b). A cauliflower like structure can be observed, which is much denser compared to the alumina top surface deposited by the FS process. Open holes were also detected on the top surface of the SPS alumina coating. The finalized coating system is shown in Figure 6.



Figure 1. Morphology of the alumina powder used in FS process at 500X magnification.

B. Mechanical Failure of Alumina

It is known that the brittleness of alumina is one of the most important factors that can limit the usage of this material in engineering applications [16]. Therefore, the alumina layer in a resistive heating system may be prone to mechanical failures due to thermal stresses at the interfaces with the heating element and the substrate. The microstructure of the coating that was developed by the FS process is shown in Figure 7. In addition to the surface open porosities observed in the top view of the FS alumina, some minor cracks were observed at the interface of the FS alumina layer and the heating element. Further, several cracks were also observed within the FS alumina layer that were likely formed due to the presence of adjacent pores.

Figure 8 shows the cross-sectional microstructure of the SPS alumina-S1 in which detachment of the alumina layer from the top heating element can be observed. It was observed that the porosity increases as the thickness of the alumina coating increases, which was mainly due to the formation of columnar structure at higher thicknesses close to the heating element layer. The generation of compressive stress that may have been induced by the deposition of cold-sprayed copper on the heating element layer, forced the cracks to close while at the interface where the copper layer was not present, the crack was present.

Figure 9 shows the microstructure that was obtained from SPS alumina-S2 condition. The only difference between the spraying conditions of SPS S2 and S1 was the stand-off distance (SOD).

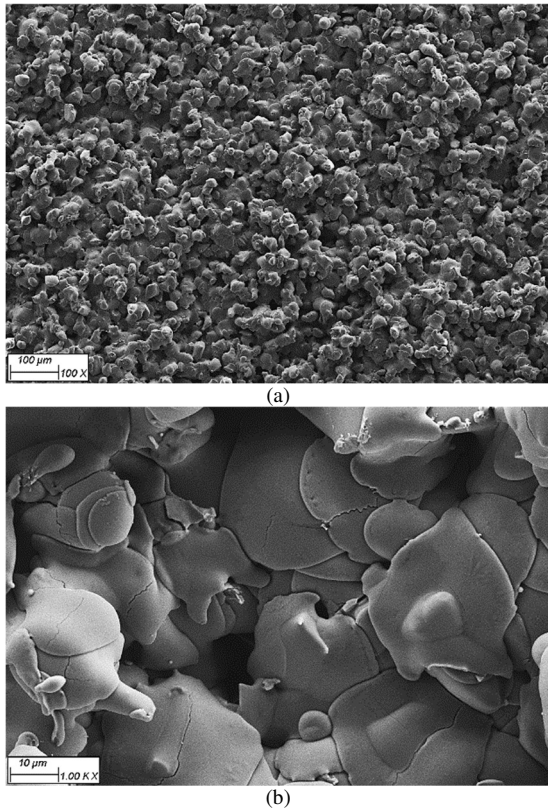


Figure 2. Secondary mode SEM images of the top view of the FS alumina coating at (a)100 X and (b) 1000X magnification.

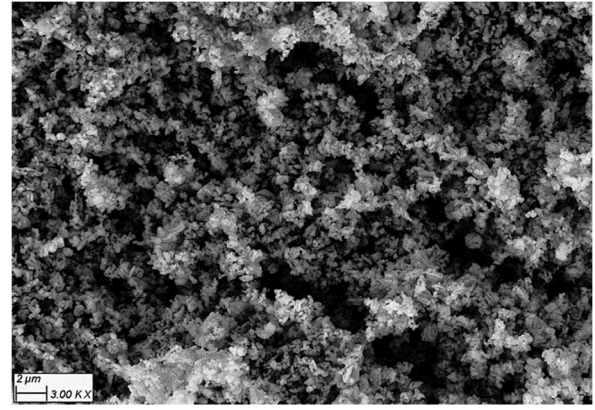


Figure 3. Morphology of the alumina powder used in SPS process at 3000X magnification.

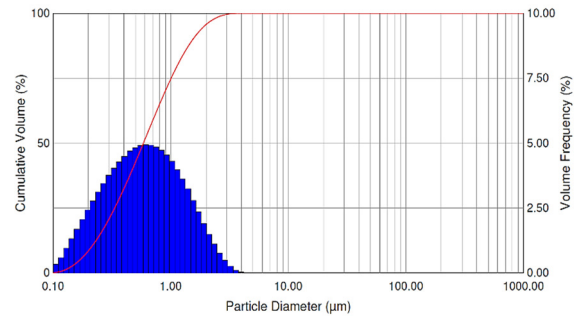


Figure 4. Particle size distribution of 10 wt% alumina suspension.

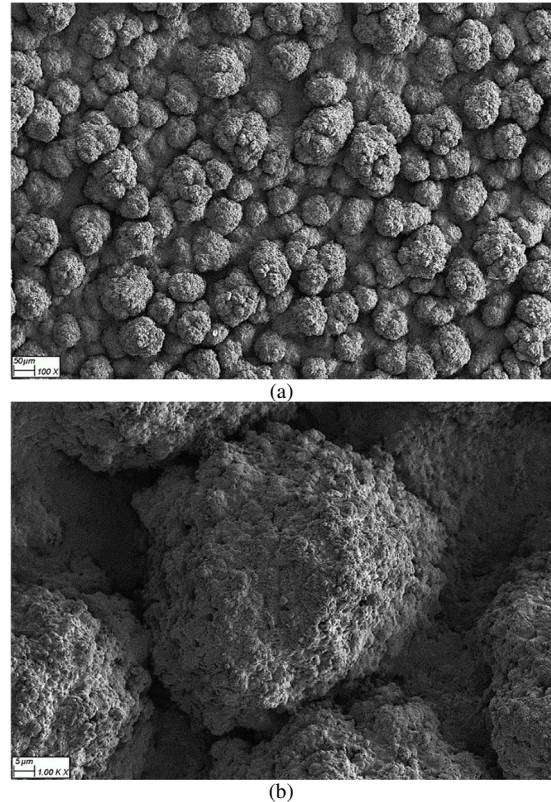


Figure 5. Secondary mode SEM images of the top view of the SPS alumina coating at (a)100 X and (b) 1000X magnification

An increase in the porosity of the alumina layer with the increase of SOD from 50 mm to 75 mm can be observed by comparing the microstructure of the developed SPS alumina coatings in Figures 9 and 10. In the microstructure obtained from the SPS S2 condition, a continuous crack grew along the interface of the heating element and alumina layer, which denotes imperfect adhesion between these layers. Compared to SPS alumina-S1 and S2, the alumina layer deposited with the S3 spraying condition resulted in a dense and fully molten microstructure. The adhesion between the alumina layer and the heating element was satisfactory since limited discontinuity was detected at the interface between the aforementioned layers during the microstructural examination. However, small cracks were observed within the alumina layer that mainly originated from the spaces between the cauliflower-like columns.

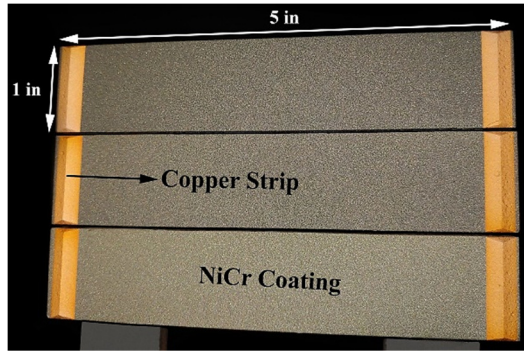


Figure 6. Finalized multi-layered coating systems.

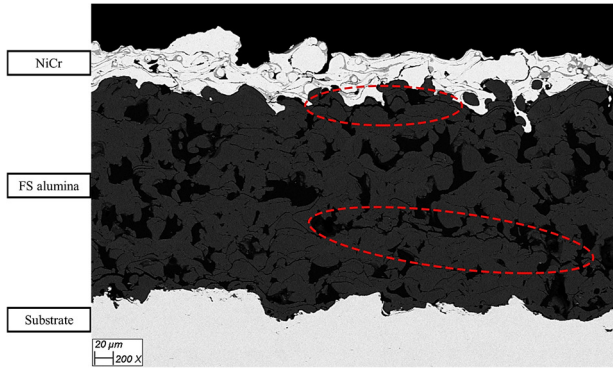


Figure 7. Backscattered SEM image of the cross section of the coating system composed of FS alumina and the heating element layer.

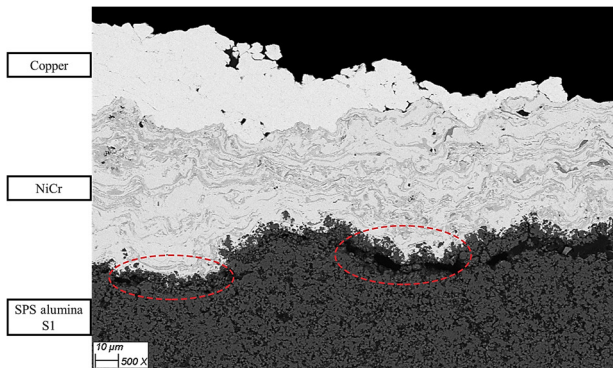


Figure 8. Backscattered SEM image of coating system composed of SPS alumina-S1 with a porous structure and the copper and the heating element layers.

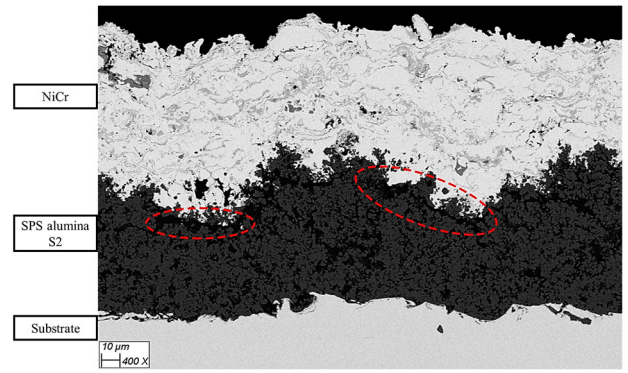


Figure 9. Backscattered SEM images of coating systems composed of the heating element layer and SPS alumina-S2 with a porous structure.

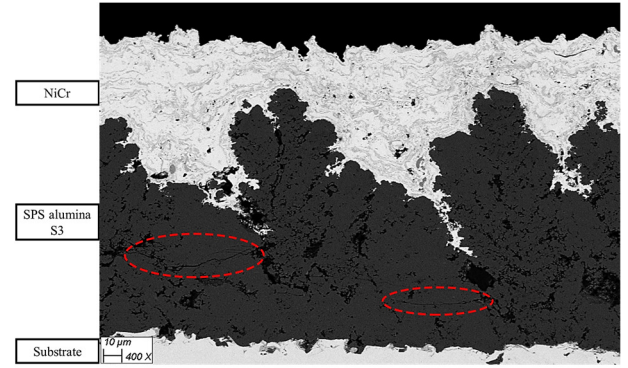


Figure 10. Backscattered SEM images of coating systems composed of the heating element layer and SPS alumina-S3 with a dense structure.

C. Phase Composition

X-ray diffraction (XRD) analyses of the as-sprayed samples are presented in Figure 11. Substantial phase compositional changes were observed in the deposited coatings compared to their initial feedstock. Powders used for both FS and SPS processes were pure α -alumina. However, various proportions of α -alumina and γ -alumina were detected after the spraying process. Quantitative phase analysis of the deposited coatings was conducted using the Rietveld method. A summary of the obtained results is shown in Table III.

FS alumina mainly contained large amounts of the stable α -alumina phase, which indicates incomplete melting of the larger alumina particles during deposition. The α -alumina phase was also dominant in the SPS alumina S1 and S2 conditions. There are two reasons for the high percentage of this phase in the coating. Aside from the incomplete melting of the feedstock during spraying, it is also believed that the low SOD of the SPS process and the continuous heating of the coating during spraying played a significant role in the formation of α -alumina by decreasing the cooling rates. Retention of the initial stable phase could also occur due to the resolidification of alumina particles in flight [17]. On the contrary, for the case of SPS alumina-S3, the higher plasma temperature resulted in complete melting of the powder particles and thus, higher cooling rates, which was the main reason for transition of the alumina phase to the metastable γ -alumina phase.

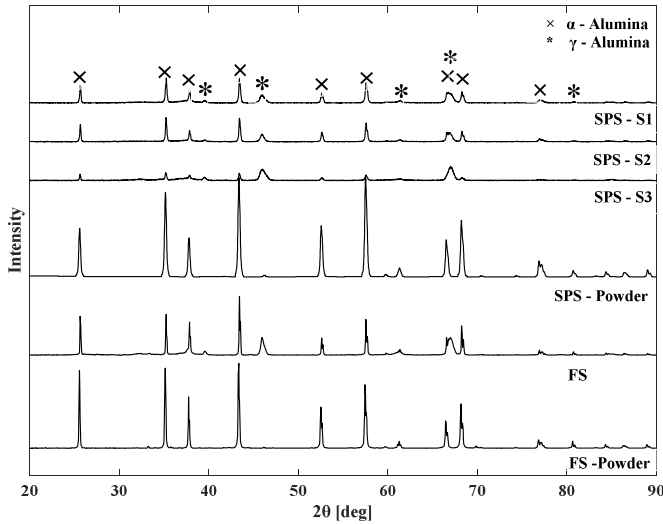


Figure 11. X-ray diffractograms of the as-sprayed alumina coatings.

The mean crystallite size of all samples was also obtained from the Rietveld technique analysis, which is presented in Table III. It can be observed from the XRD patterns that all the SPS alumina samples with smaller crystallite size have broader peaks with lower intensity while the FS sample with its larger crystallite size represent a distinguishable difference. It is known that as the cooling rate of the deposited particles increases, the microstructural length scale of the deposited coatings decreases [18], which is similar to the case when rapid solidification of the γ -alumina results in small crystallite sizes under all four spraying conditions. As it was mentioned previously, one possible reason for the formation of α -alumina is the incomplete melting of the initial powder; therefore, it stands to reason that formation of crystallites larger than the powder initial size is not possible. However, with larger particle sizes of alumina powder for the FS process, and their partial melting, larger crystallite sizes were likely formed.

TABLE III. QUANTITATIVE PHASE ANALYSIS RESULTS OBTAINED FROM THE RIETVELD METHOD.

Sample name	α - Al_2O_3 (%)	α - Al_2O_3 crystallite size (nm)	γ - Al_2O_3 (%)	γ - Al_2O_3 crystallite size (nm)
FS alumina feedstock	100	160	--	--
FS alumina	61	2800	39	19
SPS alumina feedstock	100	50	--	--
SPS alumina (S1)	70	56	30	13
SPS alumina (S2)	70	70	30	19
SPS alumina (S3)	32	47	68	12

IV. CONCLUSIONS

A comparative study was conducted on phase composition and mechanical failure of alumina, an insulating layer in multi-layered coating-based heating systems, deposited by using two different thermal spray processes namely suspension plasma spray and flame spray. The microstructural evaluation of the top view showed cauliflower shape for the SPS samples. Open pores on the top surface were present in both coatings, especially in the FS alumina. Mechanical failure of alumina was observed in the cross-sectional micrographs of all samples. However, their severity and location varied by the type of spraying method and their spraying parameters. The FS samples mainly contained the stable α -alumina, while SPS spraying parameters had a significant effect on the phase composition of the deposited alumina coatings. Their phase composition varied from almost 30% to 70% α -alumina.

ACKNOWLEDGMENT

The authors gratefully acknowledge Dr. Fadhel Ben Ettouil, Research Associate in the Department of Mechanical and Industrial Engineering at Concordia University, for his continued technical support and tireless assistance in preparation of the SPS alumina and HVOF NiCr coatings. The authors also acknowledge financial support from the Natural Science and Engineering Research Council Canada through the Strategic Networks and Discovery Grant programs.

REFERENCES

- [1] H. F. Younis, R. S. Dahbura and J. Lienhard, "Thin Film Resistance Heaters for High Heat Flux Jet-Array Cooling Experiments," ASME-Publications-Htd, vol. 353, pp. 127-134, 1997.
- [2] A. Lopera-Valle and A. McDonald, "Application of flame-sprayed coatings as heating elements for polymer-based composite structures," J. Therm. Spray Technol., vol. 24, (7), pp. 1289-1301, 2015.
- [3] J. Lamarre *et al*, "Performance analysis and modeling of thermally sprayed resistive heaters," J. Therm. Spray Technol., vol. 22, (6), pp. 947-953, 2013.
- [4] F. Toma *et al*, "Comparative study of the electrical properties and characteristics of thermally sprayed alumina and spinel coatings," J. Therm. Spray Technol., vol. 20, (1-2), pp. 195-204, 2011.
- [5] F. Toma *et al*, "Comparison of the Microstructural Characteristics and Electrical Properties of Thermally Sprayed Al_2O_3 Coatings from Aqueous Suspensions and Feedstock Powders," J. Therm. Spray Technol., vol. 21, (3), pp. 480-488, 2012.
- [6] L. M. Berger and C. C. Stahr, "State and Perspective of Thermally Sprayed Ceramic Coatings in the Al_2O_3 - Cr_2O_3 - TiO_2 System," Surface Modification Technologies XXI, pp. 469-478, 2008.

- [7] C. C. Stahr *et al*, "Dependence of the Stabilization of α -Alumina on the Spray Process," J. Therm. Spray Technol., vol. 16, (5-6), pp. 822-830, 2007.
- [8] M. Plummer, "The formation of metastable aluminas at high temperatures," Journal of Applied Chemistry, vol. 8, (1), pp. 35-44, 1958.
- [9] R. McPherson, "On the formation of thermally sprayed alumina coatings," J. Mater. Sci., vol. 15, (12), pp. 3141-3149, 1980.
- [10] T. Tesar *et al*, "Development of suspension plasma sprayed alumina coatings with high enthalpy plasma torch," Surface and Coatings Technology, vol. 325, pp. 277-288, 2017.
- [11] D. Tejero-Martin *et al*, "Beyond traditional coatings, a review on thermal sprayed functional and smart coatings," arXiv Preprint arXiv, 2018.
- [12] K. VanEvery *et al*, "Column formation in suspension plasma-sprayed coatings and resultant thermal properties," J. Therm. Spray Technol., vol. 20, (4), pp. 817-828, 2011.
- [13] A. Ganvir *et al*, "Experimental visualization of microstructure evolution during suspension plasma spraying of thermal barrier coatings," Journal of the European Ceramic Society, vol. 39, (2-3), pp. 470-481, 2019.
- [14] P. Müller, A. Killinger and R. Gadow, "Comparison between high-velocity suspension flame spraying and suspension plasma spraying of alumina," J. Therm. Spray Technol., vol. 21, (6), pp. 1120-1127, 2012.
- [15] M. Rezvani Rad and A. McDonald, "Development of a flame-sprayed coating system to mitigate ice accumulation and freezing damage in carbon steel pipes", in: International Thermal Spray Conference, May 7-10, 2018, Orlando, FL, USA.
- [16] P. Auerkari, Mechanical and Physical Properties of Engineering Alumina Ceramics, Espoo: Technical Research Centre of Finland, 1996.
- [17] P. Müller, A. Killinger and R. Gadow, "Comparison between high-velocity suspension flame spraying and suspension plasma spraying of alumina," J. Therm. Spray Technol., vol. 21, (6), pp. 1120-1127, 2012.
- [18] D. M. Stefanescu and R. Ruxanda, "Fundamentals of solidification," Materials Park, OH: ASM International, pp. 71-92, 2004.

Micro-nano-technologies

Development of BioMEMS 3D Stress/Strain Sensing Mechanism

Mohammed O. Kayed, Amr A. Balbola and Walied A. Moussa
Department of Mechanical Engineering
University of Alberta
Edmonton, Canada
mkayed@ualberta.ca

ABSTRACT

People with joint implants suffer implant loosening over time, which is believed to be due to stress shielding in which the process of fusion of the bone osseous tissues into the implant (osseointegration) is disturbed. To early detect non-painful implant loosening and avoid unnecessary implant-revision surgery, a diagnosis protocol that is capable of monitoring and quantifying the anchorage of prosthesis is needed. Such diagnosis protocol needs full evaluation of the three-dimensional (3D) stress and strain fields at the bone/implant interference. One of the advantages of piezoresistivity in crystalline silicon is the ability to extract all six stress components from a single rosette oriented over the (111) plane. In this study, we propose a bioMEMS 3D stress sensing mechanism that can be integrated into an instrumented implant to provide information regarding the variation of stresses/strains at the implant/bone interface, which will give an indication of behavior of the osseointegration process and predict possible loosening of the joint. The proposed sensing mechanism, formed of an array of sensing elements, can be connected through a wireless RF transmission to communicate with the outside world. The need for the micro size is paramount in order not to affect the function of host material within which the sensing mechanism is implanted and provide redundancy and higher reliability of the sensing mechanism.

Keywords: *Biosensors, MEMS, n-type silicon, stress sensor, piezoresistivity, piezoresistive device, silicon sensor*

Quantifying the Influence of Strain Engineering on the Piezoresistivity Using a Six Elements Sensing Rosette

Amr A. Balbola, Mohamed O. Kayed and Walied A. Moussa

Mechanical Engineering
University of Alberta
Edmonton, Canada

mkayed@ualberta.ca

ABSTRACT

Semiconductor industry has been taking the advantage of strain engineering to enhance the carrier mobility of pMOS and nMOS devices for many years. As the carrier mobility is related with the piezoresistive coefficients, those coefficients were utilized to quantify the influence of strain on the mobility. In this work, a permanent strain is integrated into piezoresistive based stress sensor to enhance its performance. Firstly, a six elements piezoresistive sensing rosette is used to study the effect of biaxial and uniaxial strain on the piezoresistive coefficients. A highly stressing film is deposited onto the substrate to globally produce tensile biaxial strain. In addition, a new local stressor is designed to apply tensile and compressive transverse uniaxial strain. Both designs are fabricated and calibrated to quantify the effect of strain on the piezoresistivity. Ultimately, a new sensing structure is developed to measure the actual strain produced using the local stressor. The outcome is exploited to fabricate a ten elements sensing rosette that is capable of measuring all stress tensor components.

Microfluidic Oil-Water Separation Using a Hydrophobic-Oleophilic Stainless Steel Membrane

Tartela Alkayyali¹, Bradley Haltli^{2,3}, Russell Kerr^{2,3}, and Ali Ahmadi^{1*}

¹Faculty of Sustainable Design Engineering and Departments of ²Biomedical Sciences and ³Chemistry
 University of Prince Edward Island, Charlottetown, Canada

*aahmadi@upei.ca

ABSTRACT

Microfluidics is currently extensively used for droplet generation due to higher monodispersity and throughput achieved compared to other methods. In microfluidics, droplet generation is caused by shear stress or pressure imposed by a continuous phase, usually a type of oil, on a dispersed phase, usually aqueous. For cell encapsulation and drug delivery applications, mineral oil has shown to be slightly toxic to cells. Also, the encapsulated cells require multiple washing steps to remove mineral oil, which is not successful at completely removing the mineral oil. Having trace amounts of mineral oil with sensitive encapsulated cells or in microbeads containing minute drug amounts could be detrimental to their viability and efficiency, respectively. Therefore, on-chip separation of oil and aqueous phases is having increased importance in the stated applications as well as in liquid-liquid extraction and water treatment applications. Deng *et al.* used a microfluidic device with an extraction region in which microbeads move from the oil phase to an aqueous phase based on their laminar interface [1]. A similar approach was implemented by Deng *et al.* based on decreased interfacial tension [2] and changing the surface properties of internal microfluidic channels. Other approaches implemented microfluidic cross-flow filtration to wash away the continuous phase [3]–[5]. Angelescu *et al.* used microchannels and a microfluidic-integrated hydrophobic membrane to achieve liquid-liquid phase separation [6], and Feng *et al.* used a hydrophilic membrane to extract water droplets from a water-oil emulsion [7]. However, the use of intricate microchannels requires the use of highly sophisticated soft lithography microfabrication. Also, these methods do not describe integration with a microfluidic hydrogel droplet generation system to obtain hydrogel droplets in the aqueous phase. In this work, we present a fluidic system in which hydrogel microbeads are transferred from an oil phase to an aqueous phase through the use of a stainless steel membrane (10 μm). The membrane is modified to have hydrophobic and oleophilic properties, aiding the separation of the two phases. The membrane is incorporated in the fluidic system using a 3D-printed separator with modular features (Fig. 1). Vacuum pressure downstream of the membrane is controlled by a vacuum pressure regulator. The separator has curved internal channels to increase resistance at the aqueous phase outlet. Preliminary testing of the system showed the highest separation efficiency (76% (v/v)) by arranging two separators in series, with the downstream pressure of the first and second separators set at -17 kPa and -34 kPa, respectively. Downstream pressure of the aqueous phase is unregulated. The system could be combined with a cell encapsulation setup to determine the effect on cell viability.

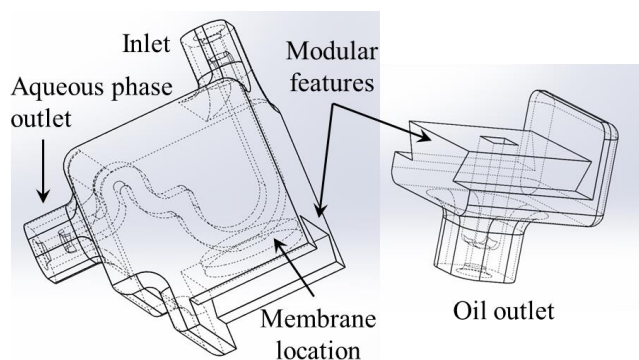


Figure 1: Separator with modular feature to enable the integration of a circular membrane in a fluidic system. Curved channels are used to increase resistance at the aqueous phase outlet.

REFERENCES

- [1] Y. Deng *et al.*, "Rapid purification of cell encapsulated hydrogel beads from oil phase to aqueous phase in a microfluidic device," *Lab Chip*, vol. 11, no. 23, pp. 4117–4121, 2011.
- [2] N. N. Deng, W. Wang, X. J. Ju, R. Xie, and L. Y. Chu, "Spontaneous transfer of droplets across microfluidic laminar interfaces," *Lab Chip*, vol. 16, no. 22, pp. 4326–4332, 2016.
- [3] Y. Y. Chiu, C. K. Huang, and Y. W. Lu, "Enhancement of microfluidic particle separation using cross-flow filters with hydrodynamic focusing," *Biomicrofluidics*, vol. 10, no. 1, 2016.
- [4] L. Lu, R. M. Irwin, M. A. Coloma, J. W. Schertzer, and P. R. Chiarot, "Removal of excess interfacial material from surface-modified emulsions using a microfluidic device with triangular post geometry," *Microfluid. Nanofluidics*, vol. 18, no. 5–6, pp. 1233–1246, 2015.
- [5] J. R. Haliburton, S. C. Kim, I. C. Clark, R. A. Sperling, D. A. Weitz, and A. R. Abate, "Efficient extraction of oil from droplet microfluidic emulsions," *Biomicrofluidics*, vol. 11, no. 3, pp. 1–7, 2017.
- [6] D. E. Angelescu and D. Siess, "Microfluidic phase separation," *Proc. IEEE Sensors*, vol. 2005, pp. 175–178, 2005.
- [7] S. Feng, M. N. Nguyen, and D. W. Inglis, "Microfluidic droplet extraction by hydrophilic membrane," *Micromachines*, vol. 8, no. 11, pp. 1–8, 2017.

A QUANTITATIVE STUDY OF THE DYNAMIC RESPONSE OF SOFT TUBING IN A MICROFLUIDICS CONTEXT FOR PRESSURE-DRIVEN FLOW

Marie Hébert, Jan P. Huissoon and Carolyn L. Ren
Mechanical and Mechatronics Engineering
University of Waterloo
Waterloo, Canada
c3ren@uwaterloo.ca

ABSTRACT

Microfluidics typically uses either a syringe or a pressure pump to drive the flow. In the context of pressure-driven flow, the apparatus requires a connection between the pressurized air output by the pump and the container of the fluid. A soft tubing linking both parts with barbed connections is typically used. The pressure provided at the pump output by the pump software is generally assumed as the pressure provided to drive the fluid. However, the soft tubing introduces its own dynamic from its length and compliance more specifically. Thus, the pressure at the pump outlet specified by the software does not accurately represent the pressure applied at the inlet of the liquid-carrying tubing that goes to the microfluidic chip. The dynamical effects are of particular importance when considering the integration of active control with microfluidics. The transient response affects system behaviour. Both simulations and experiments are used to quantify the tubing dynamics. The relationship is generalized. Hence, online measurements are not required afterwards; the output pressure of the pump software can be processed to more accurately quantify the actual pressure at the liquid tubing inlet.

Keywords: Microfluidics, tubing, dynamic response, pressure-driven flow, active control.

Quantitative Analysis and Predictive Engineering of Self-rolling of Nanomembranes

Jun Song

*Department of Mining and Materials Engineering
McGill University
Montreal, QC, Canada
jun.song2@mcgill.ca*

ABSTRACT

The broad applications of self-rolling technique based on strain-engineered nanomembranes in nano-electromechanical/micro-electromechanical systems (NEMS/MEMS), biomedical devices and optoelectronics have driven extensive research efforts, aiming to control the size and shape of sophisticated non-planar roll-up nanostructures. To achieve this objective, accurate knowledge of the mechanics underlying the rollup process is necessary. The present work presents a quantitative modeling framework for the investigation of effects of lattice mismatch strain anisotropy and surface topographic patterning on the rolling direction and curvature of rolled-up nanomembranes. Our model predictions show the lattice mismatch strain anisotropy and surface topographic patterning can have significant influence on the final rollup direction, shape and curvature, validated through finite element (FE) simulations. Our results explain the considerable discrepancy between experimentally measured and the previous theoretically predicted curvatures. Moreover, guided by our models, novel design strategies based on corner geometry engineering and surface patterning have been proposed to realize unidirectional rollup and controllable helical rolling. Our study offers new mechanistic knowledge towards understanding and predictive engineering of self-rolling of nanomembranes with anisotropic mismatch strain and complicated surface topological geometry.

A NOVEL CONTINUOUS ELECTROPHORETIC PH ELUTION SEPARATION TECHNIQUE

Jeff Farnese, Matthew Courtney, Carolyn L. Ren
Department of Mechanical and Mechatronics Engineering
University of Waterloo
Waterloo, Canada

jfarnese@uwaterloo.ca, mrcourtney@edu.uwaterloo.ca, c3ren@uwaterloo.ca

ABSTRACT

Capillary isoelectric focusing (CIEF) is a common technique that can separate multiple proteins from a sample within a separation capillary. This technique can achieve high resolution separations; however, it is not a continuous process and is difficult to keep adjacent analytes separated during the collection process. A technique that can operate continuously is gradient elution moving boundary electrophoresis (GEMBE). Proteins in the separation capillary are exposed to a constant electrophoretic force, which is a function of the protein's electrophoretic mobility, competing with a dynamically changing hydrodynamic drag force due to the bulk flow of the fluid. This technique will elute a single protein with continuous sample injection once the electrophoretic force becomes dominant. The sensitivity of this technique is limited to the sensitivity of the pressure system employed.

This work presents a novel method for continuously separating proteins by their pI points by dynamically stepping the pH of the solution but keeping a constant flow rate during the elution process. When a sample is loaded, the pH will be tailored to match the pI for a single protein, neutralizing the protein's charge and causing it to elute with the bulk flow of the solution, while the remaining proteins focus at the electrodes. Once the initial protein is eluted and captured downstream, the pH will be stepped to match the next pI of interest and the process is repeated. The advantage of this technique over GEMBE is that pH is a log scale, requiring larger changes in flow rates to produce small changes in pH. This technique could provide similar resolution to GEMBE using a lower resolution pressure control system. COMSOL simulations were done to demonstrate protein separation of single proteins at a time. A device was fabricated from PDMS using standard photolithography techniques to create a microfluidic chip. Experimental results showing the elution of fluorescent pI markers will be presented to compare with the COMSOL predictions. This technique offers an opportunity to achieve high resolution separation and large quantity collection, which could enhance the development of new biopharmaceuticals.

Encapsulation of single cells with Gelatin-Methacryloyl (GelMA) on microfluidic device using a stratified flow with viscosity contrast

Thu H. Nguyen

Waterloo Microfluidics Laboratory
Mechanical and Mechatronics Engineering
Waterloo, Canada

Evelyn K.F. Yim

Regenerative Nanomedicine Lab
Chemical Engineering
Waterloo, Canada

Sarah W. Chan

Regenerative Nanomedicine Lab
Chemical Engineering
Waterloo, Canada

Carolyn L. Ren*

Waterloo Microfluidics Laboratory
Mechanical and Mechatronics Engineering
Waterloo, Canada

ABSTRACT

Encapsulation of stem cells in hydrogel materials, such as gelatin methacryloyl (GelMA), is of particular interest for stem cell therapy as the use of hydrogels as cell carriers offers not only support of stem cell proliferation but also improved cell survival, stability and therapeutic activity upon transplantation. Droplet microfluidics has become a powerful platform for different applications in not only biological and chemical but also cell engineering. Indeed, it enables high-throughput single cell encapsulation in separated droplets for experimentation and analysis studies that would be challenging in batch-based settings or single-phase microfluidics. Generally, the encapsulation of single cells on a droplet microfluidic device is achieved by using inertial cell ordering in micro-channels. This strategy is limited in its range of applications due to practical constraints: high flow rates, specific cell concentrations and the need of sufficient space in a platform footprint for curvature channels, or a very long channel to enhance the equilibrium position for cell ordering. Recently, researchers have found that most applications in biology and biochemistry require the injection of multiple types of reagents and cells/particles into droplets for biological or chemical reactions. Thus, an approach that allows users to encapsulate single cells/particles prepared in one fluid, with another fluid in droplets, is needed. Building on the idea of using stratified flow with viscosity contrast in order to achieve high efficiency of single cell/particle encapsulation, which benefits cell engineering, drug delivery and multiple applications in other fields, this parametric study will investigate parameters that affect the hydrodynamic focusing stream that is formed by stratified flow with viscosity contrast. Flow-focusing geometry with two cross-junctions will be employed to approach the single cell encapsulation. Furthermore, data from this parametric study will provide useful guideline for the single cell encapsulation in hydrogel droplets via stratified flow with viscosity contrast.

ELECTROADHESIVE CARBON NANOTUBE SURFACES

Michael S. H. Boutilier,^{1,2} Changhong Cao,¹ Nigamaa Nayakanti,¹ Sanha Kim,^{1,3}
Seyedeh Mohadeseh Taheri Mousavi,¹ A. John Hart¹

¹ Mechanical Engineering
MIT
Cambridge, MA, USA

² Chemical and Biochemical Engineering
Western University
London, ON, Canada

³ Mechanical Engineering
KAIST
Daejeon, South Korea

ABSTRACT

Dextrous gripping and manipulation is a persistent challenge in robotics, for object sizes ranging from the micro- to macro-scales. In particular, further miniaturization of electronic systems requires pick-and-place handling of ever-smaller integrated circuits and display pixels, which are difficult to handle using mechanical grippers because of the dominance of surface forces over gravity forces at small scales. We recently developed a micro pick-and-place technology using a gripping surface of patterned dielectric-coated carbon nanotube (CNT) forests. Applying voltage to the CNTs induces surface charge in the target object creating a strong attractive force in the “on” state, while the highly porous nature of the CNT forest instills low intrinsic adhesion in the “off” state for reliable object release. In this study, we measure the adhesive forces on millimeter-scale CNT forests in both the normal and shear directions in order to understand the mechanisms of tunable electroadhesive behavior at the interface. We use the results to assess the capabilities and limitations of the technology and inform the design of electroadhesive CNT systems for modern manufacturing applications.

Read Intraocular Pressure Non-Invasively

Angelica Campigotto
Mechanical Engineering
Queen's University
Kingston, ON, Canada

Dr. Robert Campbell, MD
Department of Ophthalmology
Kingston Health Science Centre,
Kingston, ON, Canada

Dr. YongJun Lai
Mechanical Engineering
Queen's University
Kingston ON, Canada

Abstract

Glaucoma affects approximately 64.3 million people around the world and is the primary cause of irreversible blindness. The intraocular pressure is the main factor to evaluate the progression of glaucoma. The current intraocular pressure assessment methods are conducted infrequently and only during office hours. This does not give a complete illustration of the progression of the disease as the intraocular pressure is known to fluctuate throughout the day. A novel approach to monitor the intraocular pressure was generated, in which the intraocular pressure would be directly related to the change in distance of markers placed on a contact lens. As the intraocular pressure increases a soft contact lens will be able to deform at the same rate as the sclera and markers placed on the contact lens will be able to monitor the changing curvature. Based on preliminary testing, the results showed promise for the marked-up contact lens to accurately monitor the changing intraocular pressure.

Keywords-component non-invasive, Intraocular Pressure, Glaucoma

I. INTRODUCTION

Glaucoma is the leading cause of irreversible blindness around the world and is characterized by progressive optic neuropathy. This disease affects approximately 64.3 million people world-wide [1]. Open angle glaucoma is the most common form, which leads to the death of the retinal ganglion cells causing optic nerve deterioration [2] [3]. The human eye is continuously producing internal fluid, known as aqueous humor. The

aqueous humor brings nutrients to the eye and keeps the eye shape for proper light refraction [4]. In a healthy eye, the aqueous humor drains out of the trabecular mesh network; however, in individuals with glaucoma there is a gradual increase in the resistance of the outflow. In response the intraocular pressure (IOP) increases, which causes the increasing damaging force onto the optic nerve. The primary modifiable risk factor for progression of glaucoma is the intraocular pressure (IOP) [5]. A healthy IOP is within the range from 10-21 mmHg [6]. Once the IOP reaches 35 mmHg the probability of having glaucoma is almost 100 % [7]. Increased monitoring of the IOP increases the information of the patients' ocular health to improve the efforts in preventing further vision deterioration with the progression of glaucoma. IOP has been determined to fluctuate throughout the day based on diet, exercise, time of day and more [8]. Current approaches to IOP monitoring are focused to in office testing, performed by a professional, occurring a few times a year [9]. This infrequent sampling may lead to inaccurate results, increased false negatives and a poor IOP control, increasing the patients' probability of vision loss. The need for IOP monitoring at different times of day and at much more frequent intervals has been acknowledged.

Goldmannn applanation tonometry (GAT) is the current standard for monitoring the patient's IOP levels [10]. This in office procedure completed with a device that measures the force required to flatten a specified area of the cornea. The GAT requires local anesthesia and a trained ophthalmologist to perform the test. Generally, patients with glaucoma have their IOP check once every

few months within office hours. This procedure does not consider the diurnal fluctuations and will not properly capture when the elevations in the IOP occur [11] [12]. The need for continuous monitoring of IOP is becoming more prevalent with increased areas of research in this field. To date, devices for continuous IOP monitoring have involved invasive procedures or have a required complex and expensive technologies [13] [14] [15]. Hence, the goal of the research described was to design and test a novel, non-invasive IOP monitoring system using a practical, low-cost design with the potential for wide spread use.

II. PROCEDURE

Using a Purevision Balafilcon soft contact lens and a 0.8mm outer diameter trocar needle, markers were created by puncturing holes out of the contact lens. Three markers were created in a triangle shape, with 2 of the markers located on the same horizontal plane and the third placed lower on the contact lens. **Figure 1** shows a close-up image of the modified contact lens placed onto a cadaver eye.

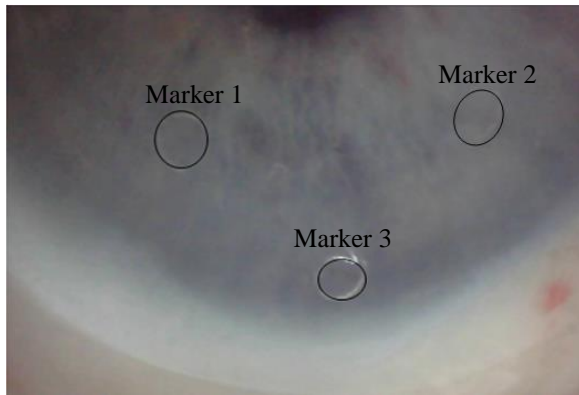


Figure 1: Close up image of the contact lens placed on the cadaver eye with labelled markers.

On one side of the cadaver eye a Millar Mikro Pressure Catheter was inserted through the sclera and into the center. On the opposite side of the sclera a tube connected to a syringe pump was inserted. The syringe pump was used to control the inflow of water into the eye, while the micro-pressure catheter would monitor the real time IOP fluctuations. After the contact lens was then placed on to the cadaver eye, a camera was used to record any changes in the distances of the markers as the IOP was increased. Figure 2, illustrates the experimental set up of the cadaver eye.

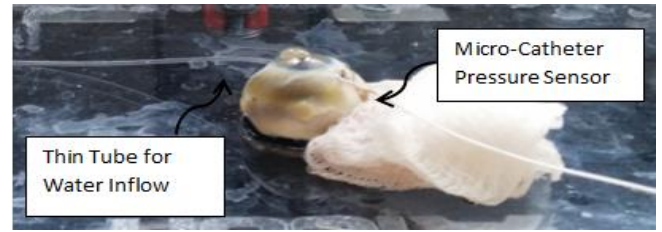


Figure 2: Close up image of the cadaver eye with labels for the inserted thin tube connected to the syringe used to pump water into the eye and the micro-catheter pressure sensor monitoring the IOP.

The pressure was increased from 10 mmHg to 40mmHg. Occasionally the eye would be hydrated using a dropper bottle filled with water to prevent the any effects of dehydration of either the eye or the contact lens. It was assumed that the cadaver eye used had no disorders, such as keratoconus, and had a uniform topography.

III. RESULTS

Three trials were conducted on a single cadaver eye. The markers were analyzed from center to center at various pressures. The cadaver eye used was assumed to be healthy and have a normal tomography with no deformities. Figure 3 demonstrates the change in the marker distances with respect to the IOP. The errors in the change in the marker distance were calculated based on the differences between the three trials completed.

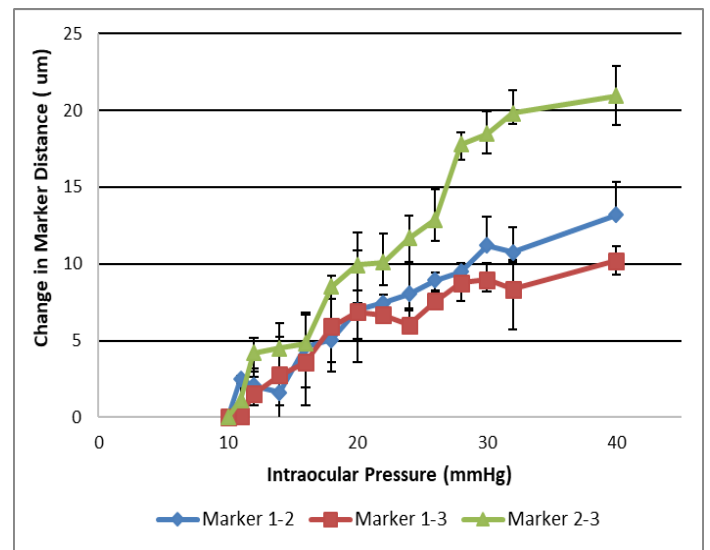


Figure 3: The distance the markers are from each other based on the IOP measured in real time. Three trials were completed, and the average was taken. The regression of squares for marker 1-2 was 0.94, for marker 1-3 was 0.82 and for marker 2-3 was 0.92.

There is a strong linearly increasing relationship between the change in marker distance to the increase in IOP. The distance between marker 2-3 had the largest change at

20.79 μm while the distance change between marker set 1-3 was the least at 10.13 μm . It was expected that the change in distance between the various markers would be similar to each other as the contact lens deforms uniformly. The discrepancies between the change in marker distance may be due to either the angle of the camera with relation to the eye. The angle at which the camera was set would affect the visual tracking between various points on the contact lens. Another reason for these results may be due to the fact that the cadaver eye is not deforming uniformly. It was assumed that the cadaver eye had no disorders or unusual topography, however if this is not the case the deformation will be non-uniform and thus the change in marker distances will be varied.

IV. CONCLUSION

In conclusion the preliminary testing completed with a marked-up contact lens demonstrates promising results of an accurate system to track the fluctuations in the IOP. There is a strong linearly increase relationship between the IOP and the various marker distances; however, some marker sets performed better than others. This may be due to the angle of the camera capturing the data or due to pre-existing irregularities to the cadaver eye that were not accounted for. Future modifications to the lens will include changing the marker shape to that with distinct corners so that the reference point is more accurate for reading.

ACKNOWLEDGMENTS

An acknowledgement to Husam Elkaseh for his assistance with the experimental work.

V. REFERENCES

- [1] Glaucoma Research Foundation, "Glaucoma Facts and Stats," Glaucoma Research Foundation, 5 April 2011. [Online]. Available: <http://www.glaucoma.org/glaucoma/facts-statistics/glaucoma-facts-and-stats.php>. [Accessed 2016].
- [2] H. A. Quigley, "Glaucoma," *The Lancet*, vol. 377, no. 9774, pp. 1367-1377, 2011.
- [3] C. J. Hong and G. E. Trope, "Glaucoma," *Canadian Medical Association Journal*, vol. 187, no. 12, pp. 398-399, 2015.
- [4] Vision Eye Institute, "The aqueous humour," Vision Eye Institute, 26 08 2017. [Online]. Available: <https://visioneyeinstitute.com.au/eyematters/aqueous-humour/>. [Accessed 2018].
- [5] A. G. Actis, E. Versino, B. Brogliatti and T. Rolle, "Risk Factors for Primary Open Angle Glaucoma (POAG) Progression: A Study Ruled in Torino," *The Open Ophthalmology Journal*, vol. 10, no. doi: 10.2174/1874364101610010129, pp. 129-139, 2016.
- [6] Y. Barkana, S. Anis, J. Liebmann, C. Tello and R. Ritch, "Clinical Utility of Intraocular Pressure Monitoring Outside of Normal Office Hours in Patients With Glaucoma," *JAMA Ophthalmology*, vol. 124, no. 6, pp. 793-797, 2006.
- [7] P. M. D. A. R. S. Blika, "The Probability of Having Glaucoma at Different IOP Levels," *ACTA OPTHALMOLOGICA*, vol. 69, no. 5, pp. 565-568, 1991.
- [8] T. Arora, S. Jindal Bali, V. Arora, M. Wadhwani, A. Panda and T. Dada, "Diurnal versus office-hour intraocular pressure fluctuation in primary adult onset glaucoma," *Journal of Optometry*, vol. 8, no. 4, pp. 239-243, 2014.
- [9] K. S. K. S. e. a. Gautam N, "Postural and diurnal fluctuations in intraocular pressure across the spectrum of glaucoma," *British Journal of Ophthalmology*, vol. 100, no. 4, pp. 537-541, 2016.
- [10] N. Arend, C. Hirneiss and M. Kernt, "Differences in the measurement results of Goldmann applanation tonometry with and without fluorescein," *Der Ophthalmologe*, vol. 111, no. 3, pp. 241-246, 2014.
- [11] M. Yaniv Barkana, M. Sarah Anis, M. Jeffrey Liebmann, M. Celso Tello and M. Robert Ritch, "Clinical Utility of Intraocular Pressure Monitoring Outside of Normal Office Hours in Patients With Glaucoma," *JAMA Ophthalmology*, vol. 124, no. 6, pp. 793-797, 2006.
- [12] L. John H. K., "Diurnal Measurement of Intraocular Pressure," *Journal of Glaucoma*, vol. 10, no. 5, pp. 39-41, 2001.
- [13] I. E. Araci, B. Su, S. R. Quake and Y. Mandel, "An implantable microfluidic device for self-monitoring of intraocular pressure," *Nature Medicine*, vol. 20, no. 10.1038/nm.3621, pp. 1074-1078, 2014.
- [14] J. O. Lee, H. Park, J. Du, A. Balakrishna, O. Chen, D. Sretavan and H. Choo, "A microscale optical implant for continuous in vivo monitoring of intraocular pressure," *Microsystems & Nanoengineering*, vol. 3, no. 10.1038/micronano.2017.57, p. 17057, 2017.
- [15] G. E. Dunbar, B. Y. Shen and A. A. Aref, "The Sensimed Triggerfish contact lens sensor: efficacy, safety, and patient perspectives," *Clinical Ophthalmology*, vol. 11, no. doi: 10.2147/OPHTH.S109708, pp. 875-882, 2017.
- [16] Y.-C. Tham, X. Li, T. Y. Wong, H. A. Quigley, T. Aung and C.-Y. Cheng, "Global Prevalence of Glaucoma and Projections of Glaucoma Burden through 2040," *American Academy of Ophthalmology*, vol. 121, no. 11, pp. 2081-2090, 2014.

DESIGN OF A HIGH DENSITY MEMS MEMBRANE-ROD BASED HYDROPHONE

Tara Ahmadi
 Electrical and Computer Engineering
 University of Windsor
 Ontario, Canada

Mohammed Jalal Ahamed
 Mechanical, Automotive and
 Materials Engineering
 University of Windsor
 Ontario, Canada

Shahpour Alirezaee
 Electrical and Computer Engineering
 University of Windsor
 Ontario, Canada

Abstract— This paper presents a MEMS based hydrophone that uses a flexible micro-membrane and rod assembly as a sensing element for greater sensitivity. Our design consists of a collection of arrays, composed of a circular base and a resonating cilium (rod) running through the center to detect acoustic frequencies. In this work, we present finite element simulation results for optimization of membrane and rod geometry for achieving higher sensitivity and wider operating frequencies.

Keywords—MEMS hydrophone, MEMS membrane, Eigen frequency and MEMS resonator

I. INTRODUCTION

The word hydrophone directly translates to ‘water-sound’; naturally what the hydrophone does is pick up sound underwater. A system to detect acoustic waves underwater would be used for monitoring underwater acoustic activities and as well applications in underwater security. Traditionally there are two types of hydrophones: passive and active. In passive hydrophones the acoustic wave is generated elsewhere in the source and is detected by the onboard receiver. With the recent advancement in microelectromechanical systems (MEMS) based fabrication and sensing technologies [1], various miniatures of passive and active MEMS hydrophones are gaining popularity.

A MEMS based hydrophone is a miniaturized hydrophone utilizing micro scale sensing components and conventional semiconductor based fabrication methods. Various geometric designs for the MEMS acoustic sensor have been explored for underwater applications. By combining bionic sensing principles, MEMS hydrophones have been demonstrated with piezo resistive principle [2]. MEMS bionic vector hydrophones with beam microstructures showed promising bandwidth and sensitivity performance [3]. Cup-shaped MEMS microstructures were utilized for 20 Hz to 1 kHz bandwidth vector hydrophones [4]. MEMS diaphragm hydrophones integrated with ring MOSFET transducer showed an advanced approach to increase the device’s sensitivity [5]. MEMS hydrophones with four-beam [6] and T-shaped geometric structures [7] were demonstrated. In this paper, we present a design where incoming acoustic vibrations are detected by a membrane interfaced with a rod to increase sensitivity, precision and operating frequency range.

Our MEMS hydrophone design is inspired by the lateral line seen in aquatic vertebrates, which has a cell that works similarly to the cilium which stands upon other cells that take in the vibrations the cilium experiences. Our design is composed of three main parts: the membrane, the rod (cilium) that runs through the center of the electrode membrane, and the still surface beneath the membrane (Figure 1). We aim to achieve greater sensitivity and accuracy by the membrane-rod interface, which will receive the vibrations of acoustic sound that pass through a cilium (rod) at the center of the membrane. This paper presents finite element model results for geometry optimization of a membrane-rod MEMS hydrophone with greater sensitivity, density, and frequency bandwidth.

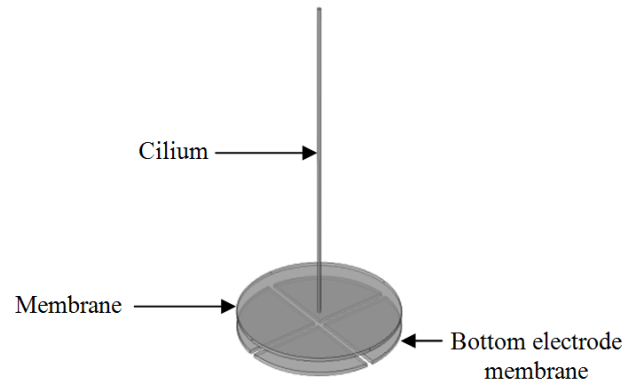


Figure 1. Basic structure of the membrane-based MEMS hydrophone.

II. MODEL APPROACH

The proposed hydrophone design consists of a circular membrane, attached to a longer, thinner cylinder (cilium) at the membrane center as shown in Figure 1. The cilium is rigidly connected to the center of the membrane, and the outer edge of the membrane is fixed; all else is free to move in this model. The electrode array at the bottom will be connected to sensors that detect the capacitive changes when the upper membrane attached to the cilium, deforms. To detect motion in this model, a fixed membrane of the same shape and size is placed directly beneath the membrane with the cilium attached to it; it is called the bottom electrode membrane. The motion of the membrane attached to the rod relative to the bottom electrode membrane beneath it will cause a change in capacitance, in the space

between them. The data on the change in capacitance will then be used to determine the amount of change in motion, and therefore the frequency of sound will be determined. The fabrication will be based on two double layer bonded silicon-on-insulator (SOI) wafers. The top SOI will host the membrane and bottom SOI will host the sense electrodes. The advantage of this design is that there is no need for piezoelectric materials and straight forward SOI based fabrication is used, which is well established. The classical four-beam microstructure design available in the literature can be seen in figure 2. The four beams are all joined at a single end, each at a 90-degree separation from the two neighboring beams. A main rod runs through the center point where these four beams meet, perpendicular to the imaginary plane the four beams share; this main rod detects sound and transfers it to the beams. Each beam has its own piezo resistor, made of a semiconductor whose resistance correlates to how mechanically strained it is; this variance of resistance gives us information on the acoustic sound vibrations that cause the strain, from both horizontal displacement and angular rotation. The T-model [7] is similar to the 4 beam microstructure models as it uses the motion along each beam to detect the frequency of the acoustic sound that is the cause of the motion. In our design we use a flat, circular membrane and a cilium/rod that runs down the center of the membrane.

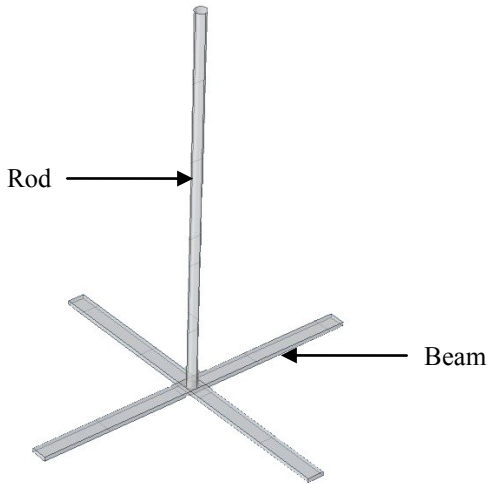


Figure 2. Classical four-beam-rod MEMS hydrophone design.

The cilium and membrane simultaneously pick up acoustic disturbance. These disturbances cause the membrane to vibrate, which changes the distance between the membrane and the bottom electrode membrane beneath it. The displacement can be picked up by the changes in capacitance. The bottom electrode and the membrane with the cilium attached to it create a capacitor. Information from the capacitance change between the two membranes will be used to differentiate the acoustic frequency being picked up by the MEMS hydrophone. Eigenfrequencies play a large role in this model, as each relevant vibrational shape/mode is unique to a certain eigenfrequency, so long as the dimensions (cilium length, cilium diameter, membrane thickness, membrane diameter) are

not changed. The device will be operated at the resonance frequencies to maximize the displacement and signal sensitivity.

III. RESULTS AND DISCUSSIONS

Finite element simulation software COMSOL Multiphysics [8] was used to do eigenfrequency modeling on the geometry presented in Figure 1. Each set of dimensions for this model, as in values for membrane thickness, membrane diameter, cilium thickness and cilium length, create unique vibrational shapes at each eigenfrequency. By investigating the variation of the rod diameter and rod length or, membrane diameter and membrane thickness, the correlation to the value of eigenfrequency observed in those dimensions can be observed. Analyzing these effects will provide an understanding on the best parameters to fabricate a hydrophone MEMS device to detect low and high frequencies with wider bandwidth. The frequency each hydrophone will detect will be the frequency it experiences resonance in and is the mode shape of interest, one resonance mode shape can be seen in Figure 3. The mode shape is a dome, and the colors on the surface plot display the amount of displacement that particular part of the hydrophone has experienced; regions that are purple have experienced the highest displacement, regions that are light green have experienced the least displacement relative to their resting locations. In this mode, the membrane shows the highest possible displacement relative to its original location, particularly at the center of the membrane.

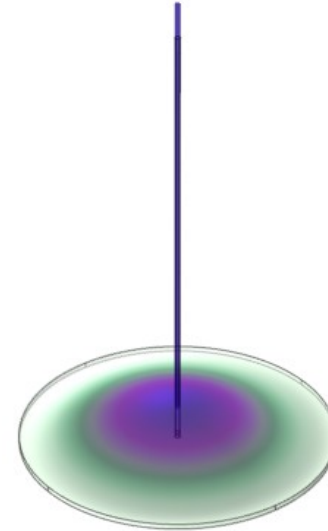


Figure 3. Surface plot showing mode shape of interest.

The further away the membrane attached to the cilium is from the bottom electrode membrane, the higher the capacitance between those two membranes will be. Other mode shapes also exist, however the deformation that occurs in other mode shapes provide less information as they would not cause as large of a change in capacitance. The membrane itself must be moving to cause a capacitance variance large enough to be registered by the sensors.

Figure 4 shows the effects of varying the membrane diameter and membrane thickness on the eigenfrequency at which the vibrational mode of interest is displayed. Our simulation results show that there is a rapid increase in the eigenfrequency as the membrane diameter is lowered while the membrane thickness is increased. This gives the design more flexibility at a range of 650 to 750 μm for the membrane diameter and 14 to 22 μm for membrane thickness.

Varying the membrane diameter 650 to 1150 μm in increments of 50 μm and the membrane thickness 10 to 22 μm in increments of 2 μm will give hydrophones that have the vibrational mode of interest at frequencies ranging roughly 100 kHz to 900 kHz to make an array of sensors that can work in this large bandwidth. This is a wide range for acoustic detection and makes varying the membrane size a viable method of varying the MEMS hydrophone. The values for cilium diameter and cilium length were both held constant throughout this sweep; cilium diameter is 12 μm and cilium length is 1000 μm . However, to detect frequencies in the lower range the membrane needs to be thinner and larger.

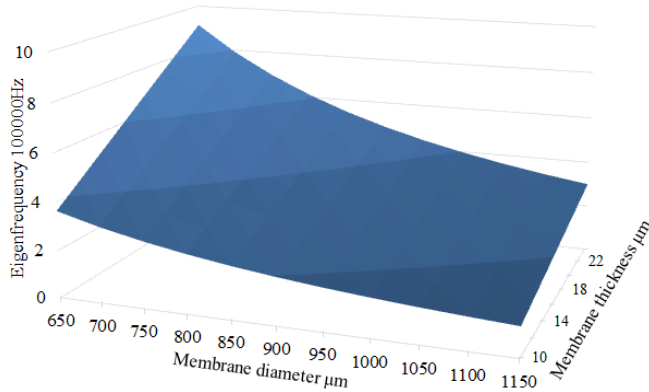


Figure 4. Effects of membrane diameter and thickness on eigenfrequency.

Figure 5 shows the effects of varying cilium diameter and cilium length on the eigenfrequencies where the dome-shaped vibrational mode of interest is seen. Figure 5 shows that the eigenfrequency at the vibrational mode decreases linearly as the cilium length and cilium thickness increase, both when they increase alone and when they are increased simultaneously. When the cilium thickness and cilium diameter are increased together, the rate of decrease of the eigenfrequency is greater than if one is an isolated variable.

Varying the cilium dimensions does not change the eigenfrequency as rapidly as changing the membrane dimensions does, and this creates a smaller range of frequencies of sound detected. However, changing the dimensions of the cilium will be beneficial as the space the hydrophones will occupy on a membrane surface plate will be constant, as it is connected to a base structure at the edges of the membrane, and vibrates freely everywhere else. This is important considering that there will be many of these hydrophones in one location, with a number of different groups of hydrophones, and each group will share the same

dimensions and all hydrophones in that group will be used to detect the same frequencies.

For the change displayed in figure 5, the cilium diameter range is 12 μm to 24 μm , swept in increments of 2 μm , and the cilium length range is 950 μm to 1200 μm , swept in increments of 50 μm , which gives frequency detection in the range 250 kHz to 320 kHz. The membrane diameter was held constant at 700 μm and the membrane thickness was held constant at 12 μm during both these sweeps.

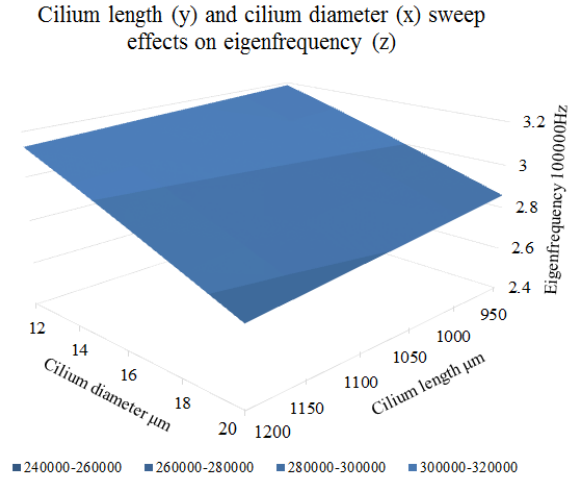


Figure 5. Effects of cilium diameter and length on eigenfrequency.

IV. CONCLUSIONS

This paper contributes to the development of MEMS based high resolution underwater acoustic monitoring. Semiconductor based micro-fabrication processes mimic bionic sensory systems at a microscopic-scale, with high sensitivity. Inspired by the aquatic vertebrates' lateral lines, the MEMS acoustic vibration sensing system for picking up sounds underwater is designed. A collection of small MEMS devices with a circular base and a resonating cilium (rod) running through the center will detect acoustic frequencies. These small devices have the same basic design but slightly changing dimensions across the array, to detect the desired range of frequencies. Our circular membrane acts as the base of the MEMS device, as opposed to separate pillars, is poised to provide greater sensitivity in sound detection and more underwater information to work with. Our results show resonant frequencies in the range of 100 kHz to 900 kHz. Finite element analysis is used to iterate the design of membrane and rod as well as numerically predict the device performance. The fabrication and experimental validation is in progress. The results presented in this paper will further assist in the development of high density and highly sensitive MEMS hydrophones for wide bandwidth.

CORRESPONDING AUTHORS

Prof. Ahamed – jahamed@uwindsor.ca

Prof. Alirezaee- s.alirezaee@uwindsor.ca

ACKNOWLEDGMENT

MJA would like to thank Natural Sciences and Engineering Research Council of Canada (NSERC) and University of Windsor Faculty of Engineering for financial supports. The authors would also like to thank CMC Microsystems for software support.

REFERENCES

- [1] M. J. Ahamed, D. Senkal, A. A. Trusov and A. M. Shkel, "Study of High Aspect Ratio NLD Plasma Etching and Postprocessing of Fused Silica and Borosilicate Glass," in *Journal of Microelectromechanical Systems*, vol. 24, no. 4, pp. 790-800, Aug. 2015.
- [2] G. Zhang, P. Zhao, and W. Zhang. "Resonant frequency of the silicon micro-structure of MEMS vector hydrophone in fluid-structure interaction", *AIP Advances* 2015 5:4.
- [3] C. Xue, S. Chen, W. Zhang, B. Zhang, Gu. Zhang, Hui Qiao, Design, fabrication, and preliminary characterization of a novel MEMS bionic vector hydrophone, *Microelectronics Journal*, Volume 38, Issues 10–11, 2007, Pages 1021-1026
- [4] W. Xua, Y. Liua, G. Zhangb, R. Wang, C. Xue, W. Zhang, and J. Liu, Development of cup-shaped micro-electromechanical systems-based vector hydrophone, *Journal of Applied Physics* 2016 120:12
- [5] P. Amiri and Z. Kordrostami, "Sensitivity Enhancement of MEMS Diaphragm Hydrophones Using an Integrated Ring MOSFET Transducer," in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 11, pp. 2121-2130, Nov. 2018.
- [6] Smitha G Prabhu, Nagabhushana S, "Design and Simulation Of Under Water Acoustic Mems Sensor", *COMSOL Conference*, Pune 2015
- [7] S. Hande, "Design and Simulation of T-Shaped Hydrophone Utilizing COMSOL Multiphysics," *International Journal of Engineering Science and Computing*, vol. 7, no. 12, pp. 15862-15865, 2017.
- [8] COMSOL Multiphysics® v. 5.3. www.comsol.com. COMSOL AB, Stockholm, Sweden.

Nuclear Materials

Orientation dependent irradiation hardening in pure Zr studied by nanoindentation, TEM, and crystal plasticity finite element modeling

Qiang Wang, Christopher Cochrane, Travis Skippon,
Zhouyao Wang, Mark R. Daymond*

Department of Mechanical and Materials Engineering
Queen's University
Kingston, Canada

* E-mail address: daymond@queensu.ca

Hamid Abdolvand

Department of Mechanical and Materials Engineering
Western University
London, Canada

ABSTRACT

Ten grains in pure Zr samples with an orientation coverage of 0° to 90° between basal pole and surface normal were studied by nanoindentation before and after ion irradiation. Irradiation to ~ 0.3 dpa was achieved by Zr-ions using a series of energies at 250°C , generating relatively uniform irradiation layer approximately $5.5\text{ }\mu\text{m}$ thick. The effect of irradiation of hardening and orientation dependence of hardness was investigated. A maximum increment of hardness due to irradiation occurs when the angle between sample surface normal and the $\langle c \rangle$ axis of the grain is $\sim 60^\circ$. The corresponding deformation mechanisms were studied using Transmission Electron Microscope (TEM). The anisotropic irradiation hardening was found to be primarily related to the activation and propagation of prismatic $\langle a \rangle$ slip. A crystal plasticity finite element model (CPFEM) of nanoindentation was employed to simulate experimental results. CRSS values of all slip systems and $\{10\text{-}12\} \langle -1011 \rangle$ tension twinning before and after irradiation were obtained by tuning the model inputs to fit simulation results with experiment. The sequence of irradiation induced increment of CRSS in the deformation systems was found to be prismatic slip (150 %), basal slip (40 %), pyramidal slip (20 %) and tension twinning (0 %).

Plastic deformation evolution at a crack tip in zirconium foil

Zhouyao Wang, Mark Daymond
Mechanical and Materials Engineering Department
Queen`s University
Kingston, ON, Canada
zhouyao.wang@queensu.ca

ABSTRACT

Understanding fracture mechanisms of zirconium alloys has significant importance in preventing abrupt failure and improving reliability in the fuel channel of CANDU power reactors. In our study, plastic deformation was investigated at the crack tip of zirconium foils with a center cracked tension configuration through a macroscopic strain analysis and a microscopic dislocation analysis. By using digital image correlation and quasi-in-situ electron backscattered diffraction (EBSD) characterization, the strain evolution and geometrically necessary dislocation (GND) evolution was captured.

The results show that as the load was increased, the plastic zone extended in size; even when large-scale yielding occurred close to the crack-tip, a J-dominated zone still exists in this small-scale specimen which is operating under non-plane strain conditions. The strain distribution as well as the dislocation distribution in this single parameter dominated zone follows the power-law singularity predicted by the first term of the asymptotic strain solution and higher order terms can be neglected. The concentrated plastic deformation at the crack tip was contributed mostly by screw $\langle a \rangle$ dislocations and first order pyramidal $\langle c+a \rangle$ dislocations.

On the State of Thermal Residual Stresses in Zirconium: Modeling and Experimentation

Abdulla Alawadi, Hamidreza Abdolvand

Mechanical and Materials Engineering
Western University
London, Canada
aalawadi@uwo.ca

ABSTRACT

Zirconium alloys are widely used for manufacturing of nuclear reactors components, such as pressure tubes. In CANDU reactors, these tubes are the primary pressure boundary between the hot water coolant and the “cold” calandria tubes. During fabrication, pressure tubes are heat treated at a high temperature and then cooled down to room temperature. This will result in the development of thermal residual stresses. To investigate the levels of thermal residual stresses, in this study, a commercially pure zirconium that has been heat treated then air cooled is studied in detail. Three dimensional synchrotron X-ray diffraction (3D-XRD) is used to measure the orientations, grain volumes, and the state of the residual stresses of more than 11000 grains. Furthermore, a weighted Voronoi tessellation technique was used to import the “as measured” microstructure into a crystal plasticity finite element (CPFE) model to simulate the heat treatment process. The results from both the 3D-XRD and CPFE model show that the residual stresses developed after the heat treatment process are significant and that grain size and neighborhood are factors that affect the magnitude of the developed stresses.

A non-local crystal plasticity finite element model for simulating irradiation growth in α -zirconium polycrystals

Omid Sedaghat, Hamidreza Abdolvand
Mechanical and Materials Engineering Department
Western University
London, Canada
osedagha@uwo.ca

ABSTRACT

α -zirconium, with hexagonal close packed (HCP) crystal structure, has been widely used in the core of various nuclear reactors due to its low neutron absorption cross section and good corrosion resistance. In a nuclear reactor, zirconium alloys are exposed to an intensive neutron flux resulting in a phenomenon known as irradiation damage, which affects the deformation mechanism of the alloy over the service time. Understanding the interaction between dislocation slip and irradiation induced damage is a great challenge. In this study, a dislocation based non-local crystal plasticity finite element (CPFE) model is developed to simulate the effects of crystal anisotropy and polycrystal microstructure on the localized stresses that develop as a result of neutron irradiation. After discussing the required steps toward modeling a non-local CPFE model, irradiation growth and the resulting anisotropy in growth strain will be studied in detail; It is shown that the model is capable of simulating the irradiation induced hardening and softening and the associated localized stress fields that develop at the vicinity of grain boundaries.

Synhrotron X-ray diffraction for extracting μ -level full stress tensors in three dimensions

Karim Louca

Department of Mechanical & Materials Engineering
Western University
London, ON, Canada
klouca@uwo.ca

Hamidreza Abdolvand

Department of Mechanical & Materials Engineering
Western University
London, ON, Canada
hamid.abdolvand@uwo.ca

ABSTRACT

Understanding deformation mechanisms on micro-, meso- and nanoscale levels are vital in predicting service life of engineering components. Due to low symmetry crystal structure in hexagonal close packed (HCP) metals, deformation mechanisms can vary based on a number of factors, including the state of the applied stress or materials texture. Microstructural characterization techniques, like scanning electron microscope (SEM), have been used to investigate materials on grain and sub-grain levels and progressions in these techniques were witnessed when electron backscatter diffraction (EBSD) and high resolution (HR)-EBSD were implemented. However, most of these techniques focus on investigating the surface of the material rather than the bulk and the retrieved microstructural information is simply in two dimensions, including grain shape, volume (i.e. area), and stress tensors. In this study, a novel technique, known as three-dimensional X-ray diffraction (3D-XRD), has been used to study individual grains in three dimensions. This research focuses on comparing post-processed data acquired from 3D-XRD with those measured using EBSD on a Zirconium specimen.

Depth dependence of indentation stress of helium implanted Inconel X-750

Loabat Shojaei-Kavan
Mechanical and Materials Engineering
Western University
London, Canada
lshojaei@uwo.ca

Robert. J. Klassen
Mechanical and Materials Engineering
Western University
London, Canada
rjklasse@uwo.ca

Pyramidal constant average strain rate nano-indentation tests were performed on samples of the aged Inconel X-750 CANDU annulus gas spacer alloy to understand the effect of accumulated helium on the operative deformation mechanisms. Multi-energy He^+ implantations were performed, at 300°C , to achieve uniform helium concentrations of 100, 1000, and 5000 appm over a depth of about $2.5\text{ }\mu\text{m}$ into the X-750 samples. Nano-indentation tests were conducted under constant \dot{P}/P conditions to invoke constant average indentation strain rates of $\dot{\epsilon}_{ind} = 0.0025, 0.025, \text{ and } 0.05\text{ sec}^{-1}$. At any given indentation depth, between 200-1600 nm we observed that the average indentation stress σ_{ind} increases when $\dot{\epsilon}_{ind}$ or the helium concentration are increased or when the indentation depth is small. The strain rate sensitivity m of σ_{ind} was assessed and the apparent activation volume $V^* \approx 3\sqrt{3} kT / \sigma_{ind} m$ was calculated. We observed that, for shallow indentations of 200 nm depth, m decreased from 0.03 to 0.02 while V^* increased from 3.8 to 5.5 b^3 when the accumulated helium content increased from 100 to 5000 appm. Similar analyses were performed at the larger indentation depths up to 1600 nm to indicate the effect of accumulated helium on the length-scale dependence of the operative deformation mechanisms of Inconel X-750.

Robotics, Mechatronics and Automation

Nonlinear Model Predictive Control of Path Following by Waterloo's Autonomous Car

Mohammad Hassanpour
Systems Design Engineering
University of Waterloo
Waterloo, Canada
m66hassa@uwaterloo.ca

John McPhee
Systems Design Engineering
University of Waterloo
Waterloo, Canada
mcphee@uwaterloo.ca

Krzysztof Czarnecki
Electrical and Computer Engineering
University of Waterloo
Waterloo, Canada
k2czarne@uwaterloo.ca

ABSTRACT

Nonlinear Model Predictive Control (NMPC) is an optimization-based method that can benefit from current and future sensor information. Autonomous vehicles are usually equipped with numerous sensors to have a better understanding of their surroundings and detect other vehicles, pedestrians, road signs, and obstacles. Therefore, the gathered data can be exploited by the controller to provide the passengers with a safe and smooth ride. Having an accurate model is vital for developing NMPCs. The high-fidelity model of the University of Waterloo autonomous vehicle (Autonomoose) has been developed using MapleSim [1,2]. Figure 1 shows the vehicle equipped with the Vehicle Measurement System (VMS) to gather experimental data against which the high-fidelity model was validated. A control-oriented model of the vehicle was extracted from the high-fidelity model, with some simplifying assumptions, and implemented in the heart of the controller to predict the behaviour of the vehicle. NMPC will make sure the vehicle follows the reference trajectory decided by the path planner modules and stays within an acceptable range from the center of the path. The optimization problem is formulated to achieve improvements in safety and comfort. Hard constraints are introduced to make the system robust to any uncertainty and disturbance. Finally, parameters in the NMPC were calibrated through numerous model-in-the-loop simulations. Figure 2 shows the Unreal environment used to simulate the vehicle behaviour interacting with traffic. The vehicle performs well in the simulated environment; the next step of this research is to deploy the controller on the car (using ROS) to determine the robustness of the NMPC.



Figure 1 (Left): Autonomoose equipped with VMS to gather experimental data
Figure 2 (Right): Autonomoose model in the Unreal environment

ACKNOWLEDGEMENT

The authors would like to thank Maplesoft, the Natural Sciences and Engineering Research Council of Canada, and the Ontario Centres of Excellence for funding this research.

REFERENCES

1. Van Gennip, M.D. and McPhee, J., "Parameter Identification for Combined Slip Tire Models using Vehicle Measurement System," SAE Technical Paper 2018-01-1339, 2018, doi:10.4271/2018-01-1339
2. Hosking, B.A. and McPhee, J., "Powertrain Modeling and Model Predictive Longitudinal Dynamics Control for Hybrid Electric Vehicles," SAE Technical Paper 2018-01-0996, 2018, doi:10.4271/2018-01-0996.

Pneumatic Hyperelastic Actuators for Grasping Soft Organic Objects

Alexandre Galley, George K. Knopf
 Department of Mechanical and Materials Engineering
 The University of Western Ontario
 London, Ontario Canada
 gkknopf@uwo.ca

ABSTRACT

Pneumatically driven soft robotic grippers [1,2] can elastically deform to grasp delicate and irregularly shaped organic objects with minimal surface damage. Traditional robotic end-effectors consist of two or more rigid fingers that apply point forces on the target object which tend to bruise or damage organic surfaces. In contrast, soft grippers are essentially air driven actuators fabricated from hyperelastic materials such as synthetic rubbers or silicone polymers [3]. The stress-strain behavior of these elastomeric materials is, however, very non-linear making it difficult to predict the actuator deformation and the touch forces generated at the object surface. In this research, a pneumatically driven hyperelastic polydimethylsiloxane (PDMS) actuator [4] for delicately grasping non-rigid organically shaped objects is developed (Fig. 1a), validated through computational simulation (Fig. 1b), and experimentally investigated (Fig. 1c). The primary application is to gently grasp and hold various horticultural products such as mushrooms, grapes, strawberries or tomatoes during automated harvesting. Each curved pneumatic actuator in the gripper assembly is based on the PneuNet's principle of inflatable elastomeric chambers [1,2] and designed to apply a distributed force over a region on the object's surface during the grasping action. Only the wall of the elastomeric chamber that makes direct contact with the object is unrestrained in the assembly and allowed to expand during air inflation. COMSOL Multiphysics software is used to predict the impact of key design parameters on the gripper's performance during inflation. The physical actuators of the gripper are fabricated using soft molding techniques and experimentally tested in a laboratory environment. The comparative assessment between simulations and experiments demonstrate that proposed pneumatic PDMS actuators can be used to create robotic grippers that can gently grasp horticultural products without damaging the surface. Furthermore, the analysis of the hyperelastic actuators provides additional insight on how to design simple but effective air driven soft robotic systems.

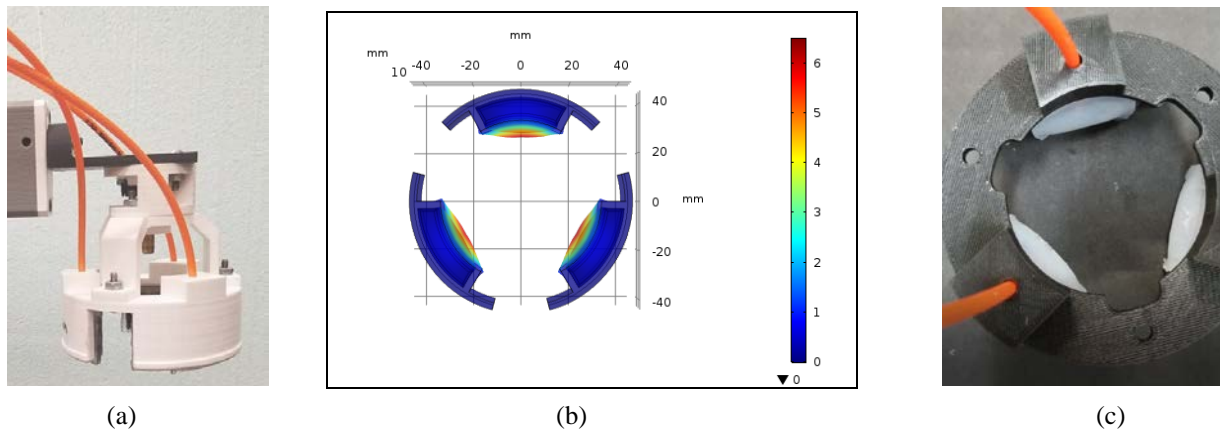


Figure 1. (a) Robotic end-effector with the elastomeric actuators inserted into gripper assembly. (b) COMSOL Multiphysics stress-strain simulation and force analysis. (c) Three inflated hyperelastic actuators.

REFERENCES

- [1] Mosadegh, B. et al. (2014) "Pneumatic networks for soft robotics that actuate rapidly," *Adv. Funct. Materials* 24:2163-2170.
- [2] Shintake, J. et al. (2018) "Soft robotic grippers," *Advanced Materials* 30:1707035
- [3] Bower, A.F. (2014) "Hyperelasticity- time independent behavior of rubbers and foams subjected to large strains," *Applied Mechanics of Solids*, 1-19.
- [4] Kim, T.K. et al. (2011) "Measurement of nonlinear mechanical properties of PDMS elastomer," *Microelectr Eng.* 88:1982-1985.

A PAYLOAD-QUADROTOR SYSTEM MODELLING ANALYSIS

Hugh H.T. Liu

Institute for Aerospace Studies
University of Toronto
Toronto, ON, Canada
liu@utias.utoronto.ca

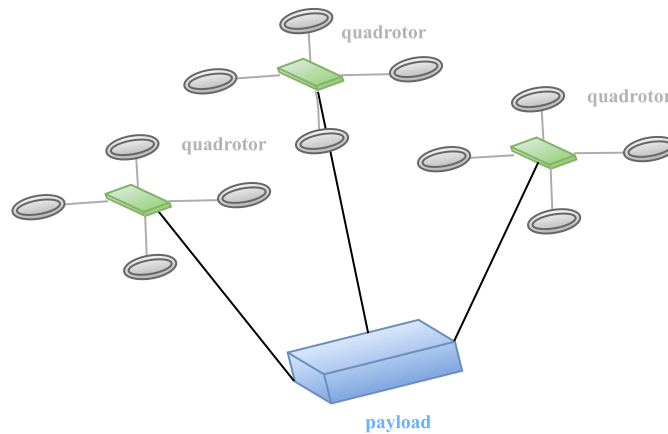
Longhao Qian

Institute for Aerospace Studies
University of Toronto
Toronto, ON, Canada
Longhao.qian@mail.utoronto.ca

ABSTRACT (*HEADING 1*)

Particle delivery is considered one promising application deploying unmanned aerial vehicles. It brings interesting study topics in aerial robotics, from sensing and perception, navigation and guidance, dynamics and control, to autonomy and mission planning. Instead of custom design of drones for different size and weight of payload, one alternative solution being investigated is to use multiple standard drones to carry payload. It offers modular design and flexibility to the variety of payload. In addition, it provides redundancy for additional benefit in terms of reliability and robustness in operation.

In our research lab, a multi-quadrotor (drone) system is under development to carry tethered payload. It is a research and development platform to study drone delivery problems. In this presentation, the equations of motion of such a configuration is treated as multi-body dynamics. Modelling and analysis are performed via Kane's method. Different configurations and corresponding control problems will be addressed. For example, when the UAV-payload system is attached through tethering system, the connecting point to the centre of mass or off-centered point will make significant difference. The presentation will highlight these different conditions and impact.



Noise and Thermal Drift Reduction in MEMS Gyroscopes With Predictive Error Modeling and Compensation Techniques

Matthew Shamon
MicroNano Mechatronics Laboratory
University of Windsor
Windsor, Ontario
shamonm@uwindsor.ca

Mohammed Jalal Ahamed
MicroNano Mechatronics Laboratory
University of Windsor
Windsor, Ontario
jahamed@uwindsor.ca

I. ABSTRACT

Recent advancement in microsystems based fabrication is allowing the development of next generation 2D and 3D microelectromechanical systems (MEMS) based gyroscopes with exceptional sensor performance characteristics [1], and resulting into many new applications [2]. However, it is yet to replace the traditional fiber optic or quartz gyroscopes in high-precision navigation and aerospace applications because of the performance degradation of MEMS gyroscopes due to the drift and high noise. One approach to reduce the bias drift error is by improving the sensor performance by innovative sensor designs and use of robust materials. However, such techniques often demand complex and expensive fabrication methods. Numerous post-fabrication noise and drift reduction method exists that are either applied electronically at hardware level or applied during the signal processing stages. In this research, we use a predictive physics based error modeling technique that can accurately estimate the bias error and temperature drift for MEMS vibratory gyroscopes, which can be implemented to compensate for errors to achieve accurate angular rate signal. Our predictive model can find temperature compensated frequency drift and thermal rate error in the gyroscope response.

MEMS vibratory gyroscopes employ a small resonating mass to detect and quantify the angular motion of an object. Semiconductor based MEMS gyroscopes mostly utilize silicon as the resonator material due to its fabrication compatibility, low-cost and wafer-scale fabrication. However, silicon has poor thermal stability due to the large and linear temperature sensitivity of silicon which results into changes in its physical characteristics with temperature. Additionally, temperature impacts on thermo-mechanical stress on the packaging that also induces error. In this paper, we present a model that can accurately predict the temperature drift error in the gyroscope rate response by predicting the changes in elastic and spring constants of a vibratory MEMS resonator due the changes in the temperature. Previously, we developed a model that can be used to accurately predict the changes in elastic constants of silicon MEMS resonator with temperature [3]. The changes in elastic and spring constants impact on the gyroscope rate response, here we extended our previous model to predict the gyro response. Experiments were performed to validate our model. In the experiments, we placed a gyroscope in a controlled rate table and recorded null measurement to quantify noise over long time to characterize short-term random drift and long-term temperature drift. We utilized probability density function and Allan variance to characterize gyroscope errors. We compared our predicted error model with experimentally measured error. Our predictive model can be implemented to compensate the error and achieve temperature and bias stability for further improvement in gyroscope performance. The work presented in this research will be useful for further development of highly stable MEMS gyroscopes for high precision applications.

II. REFERENCES

- [1] A. M. Shkel, D. Senkal and M. Ahamed, "Method of fabricating micro-glassblown gyroscopes". United States of America Patent US9702728B2, 19 01 2017.
- [2] J. Jaekel and M. J. Ahamed, "An inertial navigation system with acoustic obstacle detection for pedestrian applications," 2017 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL), Kauai, HI, 2017, pp. 109-112.
- [3] P. Rajai, M. Straeten, J. Liu, G. Xereas and M. J. Ahamed, "Modeling of temperature frequency-compensation of doped silicon MEMS resonator," 2018 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL), Moltrasio, 2018, pp. 1-2.

A Point Cloud Generation Scheme

Shuwei Qiu

Department of Electrical and Computer Engineering
The University of Western Ontario
London, ON, Canada
Email: sqiu47@uwo.ca

Mehrdad R. Kermani

Department of Electrical and Computer Engineering
The University of Western Ontario
London, ON, Canada
Email: mkermani@eng.uwo.ca

Abstract—For many tasks in robotics, it would be preferred if geometric information of the target can be obtained in advance. A scheme for obtaining the point cloud of the target using single-touch interactive segmentation and a stereo matching technique based on Phase-Only Correlation (POC) enhanced with image pyramid is developed in this paper. To test the proposed approach, objects with different shapes were used: a 3D printed ball and an electric drill. Finally, a grasp planned is briefly introduced as a potential application of the proposed method.

Index Terms—Point Cloud Generation; Single-Touch Object Segmentation; Stereo Matching

I. INTRODUCTION

As robotic technologies dramatically evolve, robots are expected to do more complex tasks than simple pick and place tasks. During the procedure of many complex tasks, robots need to manipulate different objects. Generally, obtaining the geometric information of the target is the first step. There are many ways to get the geometric information of objects and using cameras is one of them. By making use of the information acquired from cameras, the point cloud of the target can be generated so that the geometric information of the target can be estimated. This paper introduces a scheme for point cloud generation by using a stereo camera on the robotic hand (as shown in Fig. 1).

In order to get the point cloud information of the object of interest, image segmentation and stereo matching are required. Image segmentation is for cropping the target object from the scene so that only the points belong to the target are used to generate the point cloud. To precisely partition the target from the scene, the location of the target is required. In order to provide the rough location of the target while minimizing human interaction, only one seed point is required during the segmentation phase of the proposed scheme. Such image segmentation with only one selected seed point is called single-touch interactive segmentation. The related works of single-touch interactive segmentation are provided in Section II-A.

After segmenting the target object out of one image from a pair of stereo images (usually the left image), the next step

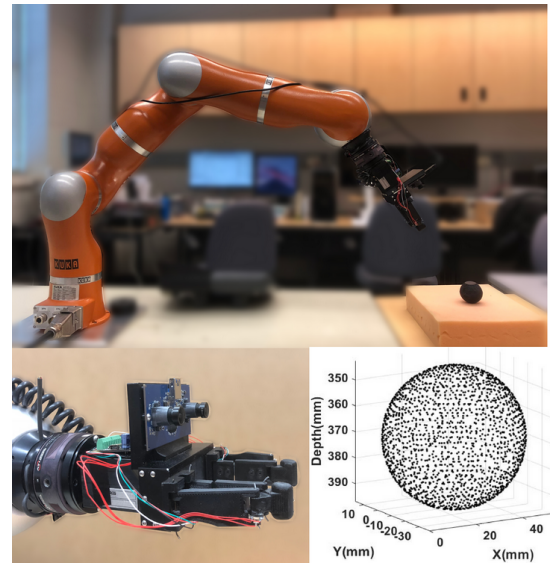


Fig. 1. Point cloud generation for a 3D-printed ball

is to find the corresponding points for the points that belong to the object in the other image (the right image). Such a procedure needs to employ a technique called stereo matching. The related works of stereo matching are introduced in Section II-B.

After finding the matched points, the depth information can be recovered through triangulation. In other words, the corresponding 3D points can be formed from the matched points. There are some minor differences between a pair of stereo images because the locations of the two cameras of a stereo camera are different. Moreover, the depth information can be used to generate a point cloud of the target. However, from one pair of stereo images, only a partial point cloud of the target can be generated. Besides, the geometric information of the target is hard to be extracted from a partial point cloud. Hence, other techniques are necessary to be applied to estimate a complete point cloud from one or more partial point clouds. In this work, ellipsoid fitting and point cloud stitching are employed. The details about ellipsoid fitting and

point cloud stitching will be provided in Section III-C.

The contributions of this paper are as follows:

- This work proposed a scheme for generating a point cloud of the target object
- Traditional flood-fill algorithm is improved to be iterative in this work for object segmentation

The rest of the paper is organized as follows: Section II briefly reviews related works in single-touch interactive segmentation and stereo matching. Section III introduces the proposed method. Section IV provides some experiment results of the proposed method. Section V introduces a potential application of the proposed approach. Finally, section VI concludes this work.

II. RELATED WORKS

A. Single-Touch Interactive Segmentation

Over the past decade, some effective single-touch interactive segmentation approaches are proposed. TouchCut is an interactive segmentation using a single finger touch to identify the object of interest. It is under the level set [1] framework with a new model that takes account of the edge map, colour distribution of foreground and background as well as the geometrical cue obtained from the user touch [2].

There is another object segmentation using single-touch interaction under the framework of graph-cut [3]. It firstly detects the salient edges of the input image. Then, the salience edges are used to determine the range of the target by using a random walk technique. However, since this approach determined the contour of an object from its edges, the accuracy of this approach depends on edge detection. Therefore, it is not reliable when dealing with the object with much texture.

Similar to these prior works, a single finger touch interaction is employed in our work. However, unlike these prior works, our work employs a salience detection technique to highlight the central region. This salience detection method uses the colour contrast between the target and the background [4].

B. Stereo Matching

Stereo matching is to find the corresponding points in a pair of stereo images that are different projections of the same 3D points. The first step of stereo matching is to build a descriptor for reference points using other points in the image. The job of descriptors is to describe some features of the reference point. According to the number of points that are used to build descriptors, The methods for stereo matching can be divided into two groups: local methods and global methods.

Local methods are the methods using a small number of points surrounding the point of interest to build descriptors for an interesting point. Among all the methods for building local descriptors, Phase-Only Correlation (POC) is a powerful one. POC is an image matching technique that uses the phase information obtained from the Discrete Fourier Transform (DFT) [5]. The cross-phase spectrum between two

input images is obtained with the help of DFT. And the 2D POC function is the 2D Inverse Discrete Fourier Transform (2D IDFT) of the cross-phase spectrum [6]. POC function can be used to evaluate the similarity between two images. Specifically, if the two images are similar, their POC function would have a sharp, distinct peak, and the position of the peak is the displacement between the two images [6].

Global methods are the methods using a large number of points or even all the points in an image to build a global descriptor or model of the image and to minimize a global cost function [7]. Graph cuts are widely employed in global methods for stereo matching. The first global method based on graph cuts for stereo matching was proposed in [8]. This method reformulated the stereo matching problem as a problem of finding the maximum-flow in a graph. A general constraint called local coherence was proposed in this method, and it was applied in all directions [8].

III. POINT CLOUD GENERATION METHOD

This section provides the details about the system used to get the point cloud of the object of interest in our work. The system consists of three parts: single-touch interactive segmentation, POC-based stereo matching and point cloud stitching.

A. Single-Touch Interactive Segmentation

Single-touch interactive segmentation is to partition the object of interest from the left image of a pair of stereo images. Moreover, there is only one point on the object selected by users to identify the object. The purpose of segmentation is to make sure the reference points only come from the object so that there will be no other parts except the object in the final point cloud. The procedure of the proposed single-touch interactive segmentation method is as follows.

Firstly, dividing the image into different regions using simple linear interactive clustering (SLIC) [9] algorithm (see Fig. 2(a)). SLIC is an algorithm that groups pixels into perceptually meaningful regions called superpixels [9]. The pixels that belong to one superpixel have similar attributes, such as colours. The attributes of each super-pixel are defined as the average attributes of all the pixels belonging to this super-pixel. It is noteworthy that the SLIC algorithm is used in a Lab colour space instead of RGB colour space. Thus, the attributes of each super-pixel are l, a, b, x, y , where l is the channel for lightness, a and b are the channels for the colour opponents green-red and blue-yellow, and x and y are the Cartesian coordinates.

Secondly, computing global contrast for each super-pixel r_k as the salience value using this formula,

$$S(r_k) = \sum_{r_i \neq r_k} \frac{\sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2}}{\sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}} \quad (1)$$

where the denominator is the Euclidean distance between two super-pixels, and the numerator is the color distance in a Lab color space [4].

Then, filtering all the super-pixels using a threshold distance to make the algorithm focus on the regions around the

selected point (as shown in Fig. 2(b)). Specifically, only the super-pixels whose distance to the selected point are less than the threshold distance are considered in the following steps.

Next, prominent super-pixels are determined. The method employed in this step is inspired by a method called key region focusing [4]. The purpose of this step is to highlight the super-pixels which are close to the selected point so that the number of super-pixels could be further reduced. To this end, an adaptive value T_α is determined as follow

$$T_\alpha = \frac{\alpha}{K} \sum_{r_i \in R} S(r_i) + (1 - \alpha)S(r_s) \quad (2)$$

where R is the set of all super-pixels considered during this step, $K = |R|$, α is a parameter that matters to the number of highlighted super-pixels, and $S(r_s)$ is the saliency value of the super-pixel containing the selected point. After calculating T_α , the super-pixels whose saliency value are less than T_α are deleted.

Finally, getting rid of the residual background. In order to do this, the background points need to be determined at first. Since people tend to select the centroid when they indicate the object of interest [2], the selected point is regarded as the centroid of the target. Then, the Euclidean distance between all the remaining pixels and the centroid are calculated as well as the average distance. The pixels whose distance to the centroid are further than this average distance are considered as background pixels (as shown in Fig. 2(c)). Next, delete all the background pixels by using iterative flood fill algorithm proposed by our work. Flood fill algorithm is a fill tool of paint programs to fill connected, similarly-colour areas. The purpose of the iterative flood fill algorithm is to delete all the potential background pixels. In each iteration, firstly, a background pixel is chosen, and the region connected to this pixel is determined using the flood fill algorithm. Then, the pixels belonging to this region are deleted from the set of all background pixels, and the image is updated by erasing this region. Next, check if all the background pixels are deleted. If they are, the iteration would be ended. If they are not, then the iteration would continue.

After the final step, the object of interest is partitioned from the left image of a pair of stereo images, and the next phase is to find the corresponding point for each point chosen from the object.

B. Phase-Only Correlation Based Stereo Matching

Our stereo matching approach contains two parts, image pyramid and 2D POC. And the details are provided in the next sections.

1) *Image Pyramid*: The image pyramid is a technique to generate a sequence of reduced resolution images for an original image. Specifically, the original image is repeatedly filtered and subsampled. The purpose of employing an image pyramid in our work is twofold. On the one hand, the image pyramid is used to obtain the initial guess of the matching points. Specifically, after multiple times of resolution reduction, the difference between a pair of stereo images could be neglect. In our work, the original image is filtered and

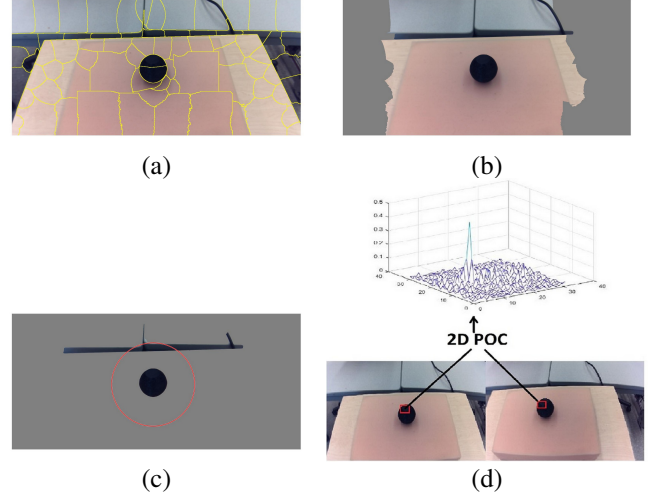


Fig. 2. Point Cloud Generation. (a) The super-pixel regions after applying SLIC algorithm (b) The image after thresholding (c) The prominent region and background pixel determination. The pixels outside of the red circle are regarded as background pixels. (d) 2D Phase-Only Correlation

subsampled four times. Thus, we assume the coordinates of matching points are the same as the coordinates of reference points in the coarsest images. So, the coordinates of reference points on the image with the lowest resolution are the initial guess of matching points. On the other hand, the image pyramid is used to apply the coarse-to-fine strategy to improve matching accuracy. As mentioned before, the coarsest image provides the initial guess of matching points. Moreover, from the second coarsest image to the original image, 2D POC is used to correct the matching result obtained from the last coarser image. In other words, for each reference points, the initial guess of its matching point is corrected several times.

2) *2D Phase-Only Correlation*: The purpose of 2D POC is correcting the matching points from the last coarser image. Given two points q and p , q is a reference point, and p is the matching position of q obtained from the last coarser image. Then, two image blocks, f and g , are partitioned from a pair of stereo images. And the centers of f and g are q and p , respectively. Let F and G denote the Discrete Fourier Transforms of f and g . The cross-phase spectrum R is defined as $R = \frac{F\bar{G}}{|F\bar{G}|}$, where \bar{G} is the complex conjugate of G . The POC function r is the Inverse Discrete Fourier Transform of R . If these two image block, f and g , are similar, there would be a distinct sharp peak of their POC function (as shown in Fig. 2(d)). The position of the peak is the displacement between f and g . Therefore, the peak position is used to correct the previous matching points in our work.

C. Complete Point Cloud Generation

1) *Ellipsoid fitting*: Ellipsoid fitting is a technique using discrete data points to reconstruct a 3D ellipsoidal surface. In order to do an ellipsoid fitting to the point cloud, a set of points on the surface of the target need to be extracted. Delaunay triangulation is applied to the point cloud to obtain the convex hull enclosing the point set on the object surface. A subset of the point cloud is selected using the convex hull.

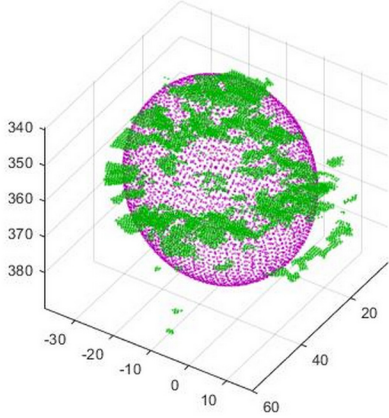


Fig. 3. The comparison between the complete point cloud of a 3D printed ball estimated by ellipsoid fitting and the partial point cloud

An ellipsoid is fitted to the point subset using linear least squares. A new point cloud is generated using the vertices of the ellipsoid model. An example of using ellipsoid fitting to generate a complete point cloud of a 3D printed ball (shown in Fig. 2) is presented in Fig. 3. As we can see, the point cloud of this ball has a much higher quality compared with the original partial point cloud. However, the disadvantage of this method is also apparent. Ellipsoid fitting is only suitable for round or ellipsoidal objects. For the objects with a complex shape (e.g. an electric drill), a more advanced technique is necessary to be applied.

2) *Point Cloud Stitching*: Point cloud stitching is a technique that merges multiple partial point clouds of the same object to reconstruct a complete point cloud of the object. Generally, point cloud stitching contains two steps: point cloud registration and point cloud merging. Point cloud registration is to find the transformation matrix between two or more point clouds that could best align these point clouds. If these point clouds belong to the same object, but from a different perspective, a complete point cloud of the object can be estimated from multiple partial point clouds by applying point cloud registration to find the transformation matrices between these partial point clouds and applying these matrices to combine them. In this work, one of the most famous methods for point cloud registration, Iterative Closest Point (ICP) [10][11], is employed.

IV. EXPERIMENT

The details about the experiment are provided in this section. The stereo camera used in our work is Dual Lens USB3.0 Camera Module from Kayeton Technology Co., Ltd, and the model is KYT-U100-960R301. The size of this camera is 74mm x 30mm. The resolution of this camera is 2560 x 720. During the experiment, the proposed method was tested by using three objects with different shapes: a 3D printed ball (as shown in Fig. 2) and an electric drill (as shown in Fig. 4). The reason why these three objects were chosen is that the complexities of the shape of these object

are different. Specifically, the shape of the ball is the simplest and the shape of the drill is the most complex among these three objects. With the objects whose shape are from different complexity levels, the limits of the proposed method can be found.

A. The Test With A 3D Printed Ball

First, the proposed method was tested by using a 3D printed ball. As it can be seen from Fig. 2, the ball can be nicely segmented from the original scene. Since the ball is round, the ellipsoid fitting technique can be used for this object. The complete point cloud of this ball generated from ellipsoid fitting can be seen in Fig. 3. It is evident that the quality of the complete point cloud is better than the partial point cloud, which means that the ellipsoid fitting can be used to generate a complete point cloud for the objects with round shapes.

B. The Test With An Electric Drill

Finally, the proposed approach was tested with an object whose shape is more complicated than a ball, an electric drill. Because the shape of the mushroom model is not round, the ellipsoid fitting is not suitable anymore, so that point cloud stitching is required to generate a complete point cloud of the model. To do point cloud stitching, multiple partial point cloud needs to be acquired. As described in Section III, single-touch object segmentation was performed before stereo matching. The result is shown in Fig. 4. Similar to the test using a mushroom model, the electric drill can be segmented from the scene with some remains of the background. The result of the stereo matching is shown in



Fig. 4. Some results of the single-touch object segmentation for an electric drill

Fig. 5, from which it can be seen that the shape of the drill is obvious.

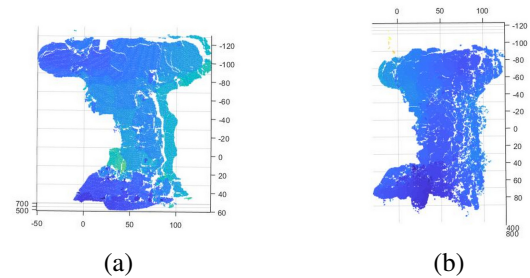


Fig. 5. Some results of the stereo matching for an electric drill

However, the result of point cloud stitching is not satisfactory (as is shown in Fig. 6). The possible reasons are that the accuracy of the stereo matching needs to be improved

and, most importantly, that a more accurate point cloud registration method is required.

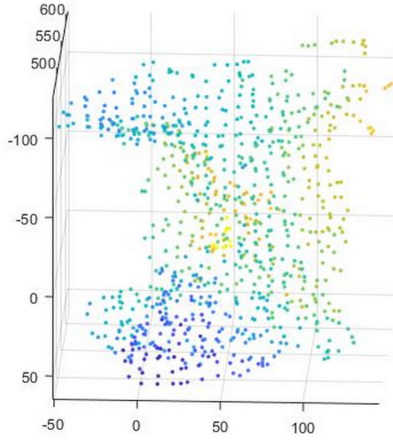


Fig. 6. The result of point cloud stitching for an electric drill

V. APPLICATION

A complete point cloud is quite useful in many areas. Since the research area of the authors is robotic grasping, an application of this point cloud generation method in the area of robotic grasping is briefly introduced in this section.

To control a robot to grasp an object, a common way is to analyze the contact point for each finger on the surface of the target from its geometric information which can be extracted from a point cloud of the target. To obtain some contact points from the point cloud, a grasp planner using hand-object geometric fitting [12] was implemented. The idea of this planner is building a score map for a point cloud of the target, and finding some contact points based on the score map with the kinematic parameters of a robotic hand. The robotic hand that was considered in this work is an under-actuated hand [13]. The potential contact points of the ball are shown in Fig. 7 wherein the blues dots denote the contact points.

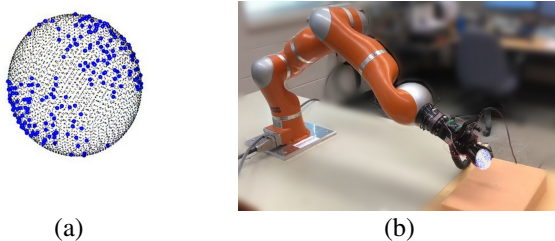


Fig. 7. The potential contact points of the 3D printed ball

VI. CONCLUSIONS

In this paper, a point cloud generation method is proposed. This method consists of three parts: single-touch object segmentation, Phase-Only Correlation (POC) based stereo matching, and ellipsoid fitting/point cloud stitching. In the

part of single-touch object segmentation, the simple linear interactive clustering (SLIC) algorithm and a method called key region focusing are employed to determine the region of the foreground. To get rid of background pixels, the flood fill algorithm is improved to be iterative in this work. POC-based stereo matching is to build a partial point cloud from a pair of stereo images. Moreover, the job of ellipsoid fitting and point cloud stitching is to estimate a complete point cloud of the target from one or more partial point clouds. It is noteworthy that ellipsoid fitting is only suitable for objects with a round shape. To test the proposed approach, a 3D printed ball and an electric drill were used. The result of single-touch object segmentation is satisfactory, even though there are still some remains of the background in the final object image. Besides, according to the result, the ellipsoid fitting can be used to estimate a clean and complete point cloud based on a partial point cloud. However, the accuracy of the stereo matching method still needs to be improved. Finally, a grasp planner is briefly introduced as an application of the proposed approach.

REFERENCES

- [1] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, "A variational level set approach to multiphase motion," *Journal of computational physics*, vol. 127, no. 1, pp. 179–195, 1996.
- [2] T. Wang, B. Han, and J. Collomosse, "Touchcut: Fast image and video segmentation using single-touch interaction," *Computer Vision and Image Understanding*, vol. 120, pp. 14–30, 2014.
- [3] V. Gopalakrishnan, A. Purwar, S. Lokkoju, R. Kumar, and K. N. Iyer, "Interactive object segmentation using single touch," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 281–285.
- [4] J. Wang, C. Zhang, Y. Zhou, Y. Wei, and Y. Liu, "Global contrast of superpixels based salient region detection," *Computational Visual Media*, pp. 130–137, 2012.
- [5] S. Sakai, K. Ito, T. Aoki, and H. Unten, "Accurate and dense wide-baseline stereo matching using sw-poc," in *Pattern Recognition (ACPR), 2011 First Asian Conference on*. IEEE, 2011, pp. 335–339.
- [6] T. Takahashi, T. Kawano, K. Ito, T. Aoki, and S. Kondo, "Performance evaluation of a geometric correction method for multi-projector display using sift and phase-only correlation," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 1189–1192.
- [7] A. Zureiki, M. Devy, and R. Chatila, "Stereo matching and graph cuts," in *Stereo Vision*. InTech, 2008.
- [8] S. Roy and I. J. Cox, "A maximum-flow formulation of the n-camera stereo correspondence problem," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 492–499.
- [9] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk *et al.*, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [10] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor Fusion IV: Control Paradigms and Data Structures*, vol. 1611. International Society for Optics and Photonics, 1992, pp. 586–607.
- [11] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International journal of computer vision*, vol. 13, no. 2, pp. 119–152, 1994.
- [12] P. Song, Z. Fu, and L. Liu, "Grasp planning via hand-object geometric fitting," *The Visual Computer*, vol. 34, no. 2, pp. 257–270, 2018.
- [13] M. Abdeetel and M. R. Kermani, "Development and grasp analysis of a sensorized underactuated finger," in *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*, 2017, p. In Press.

MODEL PREDICTIVE CONTROL, IDENTIFICATION AND IMPLEMENTATION DESIGN FOR A PENDUBOT: A MICROCONTROLLER-BASED ARCHITECTURE

Ricardo Bautista-Quintero
Department of Mechanical Engineering
University of New Brunswick
Fredericton, NB, Canada
r.bautista@unb.ca

Rickey Dubay
Department of Mechanical Engineering
University of New Brunswick
Fredericton, NB, Canada
dubayr@unb.ca

Abstract— Pendubot systems have been recognized as a complex nonlinear mechanical testbed and clearly effective for both academic and research purposes. The multiple concepts of control and identification covered by analyzing such testbed are broad. However, literature in model predictive control (MPC) implemented on underactuated systems such as a pendubot, inverted pendulum, inertia wheel, biped robots have been poorly addressed when computer resource-constrained microcontrollers are used. By implementing such advanced model-based controller the authors' objective is to close the gap between the advanced theory of this strategy and the practical implementation. The analysis presented in this paper shows relevant information about how model-based control developers can make a smooth transition from academia to industry. The low-cost hardware and open software presented in this paper allow exploring new features and recent contributions of MPC with industrial perspective.

Keywords- *MPC; underactuated mechanical system; resource-constrained microcontroller; low-cost practical implementation.*

I. INTRODUCTION

Over the last decades, an unprecedented fast expansion of the global markets has driven the creation of new forms of technology. The aggressive competition under new global rules demands better quality products. For instance, the emergence of embedded computer technology in the 1960s opened new paradigms of control. Since then, this low-cost technology not only improved quality but also increased the profitability of products and processes. The computer power and miniaturization of embedded controllers have been a key factor in accelerating mass production of certain products. Dedicated software for these embedded computers has matured enough to develop advanced forms of control. Based on the control core embedded on these computers some features as energy consumption, precision and reliability clearly enhance quality while

reducing cost. In this context, both industry and academia have quickly incorporated this form of control since MPC became cost-effective using embedded hardware. However, software support for industrial applications became predominantly close-architecture (black-box) due to intellectual property issues. In contrast, academia is focused on exploring MPC on theoretical concepts and validation based on numerical simulations. There is a gap between industrial needs since the short life cycle of the product implies a rapid integration of hardware and software tools.

A. MPC the gap between academia and industry

MPC deals with a dynamic optimization strategy that uses an efficient mechanism to achieve a control objective by means of predicting the dynamic behavior of the plant [1]. The prediction can be designed to avoid possible violations of both input and output signals [2].

The history of these predictive control tools started on Matrix Dynamic Control (DMC) developed by C. R. Cutler and B. L. Ramaker during the 1980s. This predictive form of optimal control improved the performance and profitability for the refinery industry [3]. Since that economic success, more companies started to develop new algorithms and software tools. The commercial introduction of MPC to industrial processes with minimum knowledge of the plant model and the control strategy simplified the implementation-tuning cycle. Some integrated products evolved to the market, for instance: [4], [5], [6], [7], [8]. A company AspenTech improved DMC after important developments of Shell industries [9]. Simultaneously, Shell developed their own improvements of DMC. Another company Profimatic developed a consolidated tool that established a new era of common implementation use of MPC [5]. Similarly, Honeywell released an algorithm based on Robust Multivariable Predictive Control Technology (RMPC) [10]. Companies like MDC Technology, Inc. and Yokogawa further improved DMC and created software tools to ease implementation for different applications [11]. Based on these platforms and the fusion of

Identify applicable sponsor/s here.

different companies today MPC technology has matured by using advanced software tools. Some examples of the most predominant technologies in this field are MATLAB MPC toolbox [8], Pavilion8 software from Rockwell Automation [6], System 800xA from ABB [4], DeltaV PredictPro multivariable MPC from Emerson Process Management [12].

Industrial developers can substitute multiple decentralized controllers with complex interactions into a global multivariable system that guarantee a performance index by using these modern tools. Some industries deal with intricate dynamic behavior in which identification techniques provide information to design a model-based controller. The accuracy of the modeling identification determines the accuracy of the prediction controller that determines performance and constraint management.

Despite some attractive features of industrial MPC controllers, industry slowly migrates towards such technology on a large scale. This might indicate that most developers remain unaware on how to exploit capabilities of such tools due to the commercial close-architecture policies. Additionally, there exist small scale but non-trivial control applications where the commercial tools are by far an unpractical choice due to the cost. Particularly, when the control system is implemented on a low-cost embedded microcontroller there is no support. In this context, the main contribution of this paper is meant to provide a comprehensive theoretical and practical method for integrating an efficient implementation on low-cost hardware and open-source software for small scale systems that are challenging to control.

A form of MPC based on General Predictive Control (GPC) algorithm is implemented on a Pendubot nonlinear underactuated electro-mechanical system. Both, low-cost and open architecture features presented in this paper may provide an accelerated and customized understanding of dealing with highly complex industrial close-architecture tools mentioned previously.

B. Pendubot: a suitable academic testbed

A pendubot is a mechanical system which consists in two rigid links connected with a rotational joint. One of these joints is connected to the other end to an actuator. This mechanical system is considered underactuated since it has two degrees of freedom but only one of them is actuated. This underactuation restriction introduces important challenges to control position of the two links with only one input actuation [13], [14], [15]. Academic researchers in the field of control commonly explore diverse forms of approaches for this challenging nonlinear system [15], [16], [17]. The common approaches explored in this system are control, identification and swinging-up strategies. Effective teaching in concepts as robustness, noise rejection, optimal analysis, and many more topics implies important challenges in the academic world. In this context, academia relies on commercial versions of pendubots along with general purpose computer for visualization and data acquisition and control

boards in order to execute the control in hard real-time [18]. Some features of the hardware and software remain unknown for academics and consequently, important concepts that link theory and practice are not clearly understood by students. Instead, this paper suggests the use of low-cost hardware and open-source in order to integrate three basic approaches: parameter identification, control and swing-up strategy.

This paper has the following structure. Section II describes the mathematical model of the pendubot dynamic system. Section III provides the synthesis of the General Predictive Controller and its implications for resource-constrained embedded computer implementation platform. Section IV shows the embedded control implementation, from the identification algorithm, the control strategy, and swinging-up sequence. Section V summarizes the contribution and presents the conclusions.

II. DYNAMIC MODEL DESCRIPTION

The pendubot mechanical system shown in Figure 1 has a DC motor in the based (actuator) and it is directly connected to Link 1. The Link 1 rotates and the angular position is defined as q_1 . The second link is connected to Link 1 by a rotational joint and the angular position is defined as q_2 as shown in Figure 1.

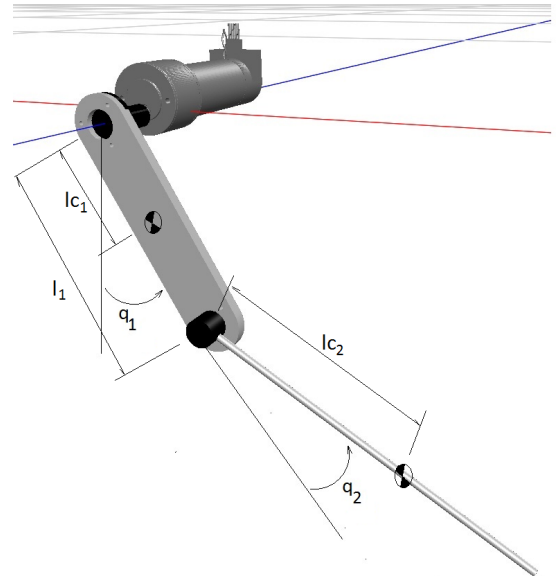


Figure 1. Pendubot underactuated dynamic system

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \boldsymbol{\tau} \quad (1)$$

where $\mathbf{q} = [q_1 \ q_2]^T$ is the vector of generalized coordinates, $\boldsymbol{\tau} = [\tau_1 \ \tau_2]^T$ is the vector of generalized torque/forces and \mathcal{L} is the difference between the kinetic energy and the potential energy, see more details in [19]. The dynamic structure of the pendubot is modeled as a 2-DOF serial link mechanism formulated in the following equation:

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) + \mathbf{f}(\dot{\mathbf{q}}) = \boldsymbol{\tau} \quad (2)$$

Parameter	Symbol	Value	Units
Mass link1	m_1	0.163	kg
Mass link2	m_2	0.065	kg
Length link1	l_1	0.15	m
Length link2	l_2	0.26	m
Length centre mass link1	l_{c1}	0.105	m
Length centre mass link2	l_{c2}	0.13	m
Gravity constant	g	9.81	m/s^2
Inertia moment link1	I_1	1.2e-3	$kg\ m^2$
Inertia moment link2	I_2	5e-4	$kg\ m^2$
Coefficient 1	c_1	0.0737	$N\ m/V$
Coefficient 1	b_1	0.0261	$kg\ m/rad/s$
Coefficient 2	b_2	0.0001	$kg\ m/rad/s$

TABLE I. The design parameters

where: $\mathbf{q} \in \mathbb{R}^{2 \times 1}$ angular position vector, $\dot{\mathbf{q}} \in \mathbb{R}^{2 \times 1}$ angular velocity vector, $\ddot{\mathbf{q}} \in \mathbb{R}^{2 \times 1}$ angular acceleration vector. $\mathbf{M}(\mathbf{q}) \in \mathbb{R}^{2 \times 2}$ inertia matrix, $\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) \in \mathbb{R}^{2 \times 2}$ Coriolis matrix, $\mathbf{g}(\mathbf{q}) \in \mathbb{R}^{2 \times 1}$ gravity vector, and $\mathbf{f}(\dot{\mathbf{q}}) \in \mathbb{R}^{2 \times 1}$ vector that represents viscous friction torques.

$$\mathbf{M}(\mathbf{q}) = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \quad (3)$$

where,

$$\begin{aligned} m_{11} &= m_1 l_{c1}^2 + m_2 [l_1^2 + l_{c2}^2 + 2l_1 l_{c2} \cos(q_2)] + I_1 + I_2 \\ m_{12} &= m_2 [l_1 l_{c2} \cos(q_2)] + I_2 \\ m_{21} &= m_{12} \\ m_{22} &= m_2 l_{c2}^2 + I_2 \end{aligned}$$

$$\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \quad (4)$$

where,

$$\begin{aligned} c_{11} &= -2m_2 l_1 l_{c2} \sin(q_2) \dot{q}_2 \\ c_{12} &= -m_2 l_1 l_{c2} \sin(q_2) \dot{q}_2 \\ c_{21} &= m_2 l_1 l_{c2} \sin(q_2) \dot{q}_1 \\ c_{22} &= 0 \end{aligned}$$

$$\mathbf{g}(\mathbf{q}) = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \quad (5)$$

where,

$$\begin{aligned} g_1 &= m_1 l_{c1} g \sin(q_1) + m_2 l_{c2} g \sin(q_1 + q_2) \\ g_2 &= m_2 l_{c2} g \sin(q_1 + q_2) \end{aligned}$$

$$\mathbf{f}(\dot{\mathbf{q}}) = \begin{bmatrix} b_1 \dot{q}_1 \\ b_2 \dot{q}_2 \end{bmatrix} \quad (6)$$

and

$$\boldsymbol{\tau} = \begin{bmatrix} c_1 u_1 \\ 0 \end{bmatrix} \quad (7)$$

Note that τ_1 is the motor torque, this torque is produced by the applied voltage u_1 and $\tau_2 = 0$ since there is no actuator for the link 2. The parameters of the system are presented in Table I.

A. Simplified Parameter Identification method

Since the system in Equation (2) is linear in parameters then the model can be re-written as a function that depends on positions and their derivatives, such as:

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) + \mathbf{f}(\dot{\mathbf{q}}) = \mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})\boldsymbol{\Phi} \quad (8)$$

where $\boldsymbol{\Phi} = [\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7, \phi_8, \phi_9]^T$ can be defined by:

$$\boldsymbol{\Phi} = \begin{bmatrix} m_1 l_{c1}^2 + m_2 l_1^2 + m_2 l_{c2}^2 + I_1 + I_2 \\ l_1 m_2 l_{c2} \\ m_2 l_{c2}^2 + I_2 \\ g(l_{c1} m_1 + m_2 l_1) \\ g m_2 l_{c2} \\ b_1 \\ b_2 \end{bmatrix} \quad (9)$$

Based on Equation (2) and the lumped parameters proposed in Equation (9), $\mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ corresponds to Equation (10) which is a matrix $\in \mathbb{R}^{2 \times 7}$.

$$\mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}) = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} & Y_{14} & Y_{15} & Y_{16} & 0 \\ 0 & Y_{22} & Y_{23} & 0 & Y_{25} & 0 & Y_{27} \end{bmatrix} \quad (10)$$

where,

$$\begin{aligned} Y_{11} &= \ddot{q}_1 \\ Y_{12} &= \cos(q_2)(2\ddot{q}_1 + \ddot{q}_2) - \sin(q_2)\dot{q}_2(2\dot{q}_1 + \dot{q}_2) \\ Y_{13} &= \ddot{q}_2 \\ Y_{14} &= \sin(q_1) \\ Y_{15} &= \sin(q_1 + q_2) \\ Y_{16} &= \dot{q}_1 \\ Y_{22} &= \ddot{q}_1 \cos(q_2) + \sin(q_2)\dot{q}_1^2 \\ Y_{23} &= \ddot{q}_1 + \ddot{q}_2 \\ Y_{25} &= \sin(q_1 + q_2) \\ Y_{27} &= \dot{q}_2 \end{aligned}$$

An estimation of the real values $\boldsymbol{\Phi}$ is defined as $\hat{\boldsymbol{\Phi}}$ and it is obtained by applying a standard regression algorithm. The procedure consists on providing a voltage signal to the motor (as it is shown in Figure 3) and recording the value of each corresponding matrix element $\mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$. After 10 seconds of recording (1000 values for each element) the input and output is used to obtain the parameters of the system by using an algorithm based on least minimum square (LMS). This numerical approximation of the parameters ($\boldsymbol{\Phi}$) is obtained by using the following equation (11).

$$\hat{\boldsymbol{\Phi}} = (\mathbf{Y}_i^T \mathbf{Y}_i + \lambda \mathbf{I})^{-1} \mathbf{Y}_i^T \boldsymbol{\tau}_i \quad (11)$$

where $\mathbf{I} \in \mathbb{R}^{7 \times 7}$ identity matrix, λ is a constant regularization parameter (when $\mathbf{Y}_i^T \mathbf{Y}_i$ is ill-conditioned), $i = 1 \dots N_s$ (N_s is the number of the samples), ΔT is the sample time, $N_s = T_{final}/\Delta T$, and $\mathbf{Y}_i \in \mathbb{R}^{2N_s \times 7}$ and $\boldsymbol{\tau}_i \in \mathbb{R}^{2N_s \times 1}$.

Since the matrix $\mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ includes the first the second derivative of the angular positions (q_1 and q_2) the embedded computer (microcontroller) must be able to estimate these variables. Particularly, numerical estimation of the second derivative is constrained by the noise during data acquisition. This problem

was previously studied and solve by using an equivalent filtered regression model, see [20] for more details.

The following section shows the control synthesis design to the pendubot.

III. GPC CONTROL SYNTHESIS

General Predictive Control (GPC) has been widely accepted in both academia and industry since it can handle unstable open-loop plants and non-minimum phase features [1]. The GPC algorithm is based on obtaining a sequence of future control signals in a fashion to minimize the multi-stage cost function defined over a prediction horizon. The predictions should be generated accurately and based on a real dynamic system in order to supply a proper output signal. GPC control algorithms have been combined with adaptive algorithms which are meant to provide an accurate discrete model either off-line or on-line execution [21]. Description and implementation of different adaptive predictive control are being explored in [22]. As previously described in [23] a closed-loop control identification system based on GPC. In previous publications [24], [1], the CARIMA model has been used to describe the input-output relation of a system in order to design a GPC controller [21]. However, for state-space representation a different approach is used, this is based on increasing the degree of the model plant. The state-space representation with augmented state consists on making an additional internal state as can be observed in the following equation:

$$\begin{aligned} \begin{bmatrix} x(k+1) \\ u(k) \end{bmatrix} &= \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \begin{bmatrix} x(k) \\ u(k-1) \end{bmatrix} + \begin{bmatrix} B \\ I \end{bmatrix} \Delta u(k) \\ y(k) &= [C \quad D] \Delta u(k) \begin{bmatrix} x(k) \\ u(k-1) \end{bmatrix} + D\Delta u(k) + d(k) \end{aligned} \quad (12)$$

Similar to the CARIMA approach presented in [1] for SISO plants the GPC state-space formulation has a prediction as follows:

$$\hat{y} = G\Delta u + R_{\text{free}} \quad (13)$$

The G matrix is described according to the augmented state-space plant described previously.

$$G = \begin{bmatrix} CB & 0 & \dots & 0 \\ CAB & CB & \dots & 0 \\ CA^2B & CAB & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ CA^{N_2-1}B & CA^{N_2-2}B & \dots & CA^{N_2-N_u}B \end{bmatrix} \quad (14)$$

And the free response is defined as the following equation:

$$R_{\text{free}} = \begin{bmatrix} CA \\ CA^2 \\ CA^3 \\ \vdots \\ CA^{N_2} \end{bmatrix} x(k) \quad (15)$$

According to the following criteria the control signal u that minimizes the difference between the prediction \hat{y} and the actual output y and can be written in as (for details see [21]):

$$K_{\text{gpc}} = \Gamma(GG^T + \lambda I)^{-1}G^T(SP - R_{\text{free}}) \quad (16)$$

where $\Gamma = [1, 0, 0, \dots, 0] \in \mathbb{R}^{1 \times N_2}$

Note that N_2 is the prediction horizon and N_u the control horizon.

A block diagram of the controller proposed can be seen in Figure 2, where the simulation is presented in the following section.

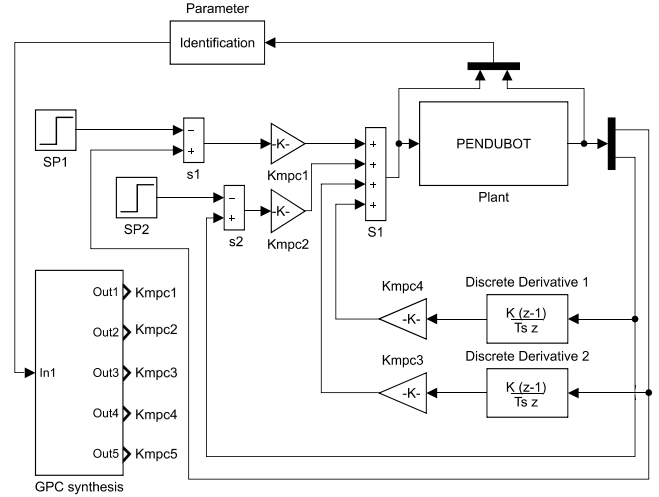


Figure 2. Block Diagram: pendubot control synthesis and ID.

The following section integrates the model, identification, and control synthesis equations presented in the previous and current section in order to show the details on the implementation with the embedded computer.

IV. EMBEDDED CONTROL IMPLEMENTATION

The integration of the control implementation is described in this section by linking the previous theory presented with the set of software structures developed by a microcontroller technology platform and a custom low-cost hardware. The control core for this application is based on an advanced RISC Machines (ARM) processor. A 32-bit microcontroller LPC1768 (cortex-M3) with a CPU frequencies of up to 100 MHz, a set of built-in peripherals such as serial communication, I/O ports, PWM signal generators, ADC converters, quadrature encoder interface among other features (more details in [25]). A C/C++ compiler used for this board is cloud-based available for free access [26]. Each position q_1 and q_2 is acquired by an incremental rotational encoder which provides two pulses signals (phase A and phase B). The relative angular position of the shaft of an incremental sensor is given by counting either phase A or phase B pulses. An increment in the counting is giving when the edge of phase A leads the edge of phase B.

By counting the pulses CW or CCW, a digital estimation of the rotation of the link 1 and link 2. Resolution of encoders and mechanically attached to the links define how accurate is the position estimation. The position q_1 is estimated by using the following expression:

$$q_1 = N_p 2\pi / Res_{p1} \quad (17)$$

where N_p is the number of the pulses (edges) of the encoder attached to the pendulum, $Res_{p1} = 5900$ is the encoder resolution in pulses per revolution (ppr). The same formula applies for the encoder attached to link 2, with a resolution of $Res_{p2} = 4000$ ppr. Once the positions of both car and rod are obtained the velocities are estimated by using the backward differentiation algorithm, see Equation (18).

$$\dot{q}_1[k] = (q_1[k] - q_1[k-1]) / T_s; \quad (18)$$

where $q_1[k-1]$ is the previous sample position and T_s is the sample time (10 ms). Similarly, \dot{q}_2 is estimated likewise equation (18).

The motor activation is performed by pulse width modulation (PWM) with relatively high frequency (more than 20 KHz). This high-frequency commutation allows the transfer of energy to the motor more efficiently since the power stage (commonly based on transistors) acts as a switch between saturation and cutoff stages. The actuator used for the experimental part is a DC gearmotor Pittman GM9236S015 with a gear-box ratio of 5.9:1.

A. Parameter Identification experimental results

An estimation of the parameters of the underactuated system presented in this paper generates an accurate regression based on the input-output map. The equation (11) obtains a numerical approximation of the lumped parameters given in (9). The input signal $u(t)$ is proposed in the equation (19).

$$u(t) = 1.2 \sin(\pi t) + 0.4 \sin(3\pi t) \quad (19)$$

A block diagram of the parameter identification strategy can be seen in Figure 3. In figure 3, the proportional control gain

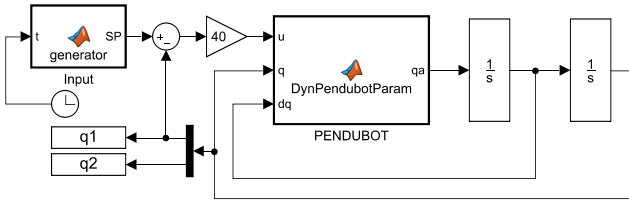


Figure 3. Parameter Identification q_1 in closed-loop and q_2 in open-loop

$k_p = 40V/rad$ induce a control motion over q_1 and thus q_2 . The parameters estimated are shown in Table II.

Figure (4) shows a comparison between the real experimental signal of the position q_2 and the signal constructed by the model using the identified parameters shown in Table II.

Symbol	Theoretical estimated	Experimental estimated
ϕ_1	0.0050	0.0013
ϕ_2	0.0013	0.0016
ϕ_3	0.0016	0.0021
ϕ_4	0.2725	0.3241
ϕ_5	0.0829	0.0836
ϕ_6	0.3511	0.3617
ϕ_7	0.0010	0.0008

TABLE II. Identified parameters vs. real lumped values

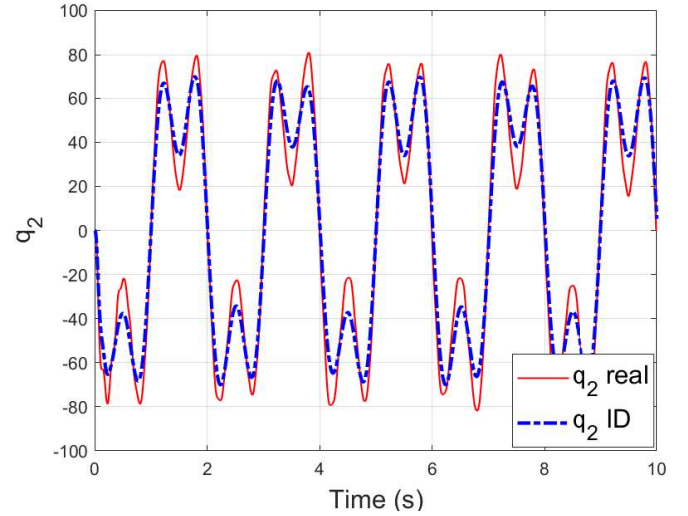


Figure 4. Comparison between real q_2 position and the q_2 simulation signal based on the identified model.

B. Control implementation

According to figure 1 there are 3 visible equilibrium points (EP), 1) an stable EP ($q_1 = 0, q_2 = 0$), two unstable ($q_1 = 0, q_2 = \pi$) and ($q_1 = \pi, q_2 = 0$). The code implemented is capable to transit from the stable EP towards the each unstable EP. Once q_1 and q_2 are closed to the unstable EP the implementation of the GPC control strategy is triggered. The first step to calculate the GPC gains is summarized as follows:

- Obtain $\ddot{\mathbf{q}} = \mathbf{M}^{-1}[\boldsymbol{\tau} - \mathbf{C}(\dot{\mathbf{q}}, \mathbf{q})\dot{\mathbf{q}} - \mathbf{G}(\mathbf{q}) - \mathbf{f}(\dot{\mathbf{q}})]$
- Define the space-state $\mathbf{x} = [q_1, q_2, \dot{q}_1, \dot{q}_2]^T$
- Define $\dot{\mathbf{x}} = \mathbf{f}_x$, thus, $\mathbf{f}_x = [\dot{q}_1, \dot{q}_2, \ddot{q}_1, \ddot{q}_2]^T$
- Define the unstable equilibrium points, (e.g., $q_1 = 0, q_2 = \pi$)
- Obtain the Jacobian matrix $\mathbf{A}_c = \mathbf{J}(\mathbf{f}_x, \mathbf{x})$ and evaluate at the equilibrium point previously selected.
- Obtain the Jacobian matrix $\mathbf{B}_c = \mathbf{J}(\mathbf{f}_x, \boldsymbol{\tau})$ and evaluate at the equilibrium point previously selected.
- Define C_c and D_d according the state-space previously described, (e.g., $\mathbf{C}_c \in \mathbb{R}^{2 \times 4}$ and $\mathbf{D}_c \in \mathbb{R}^{2 \times 2}$)

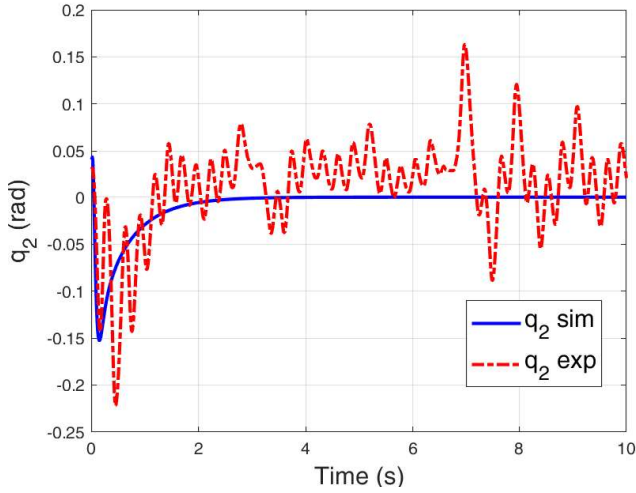


Figure 5. Comparison between simulation and experimental data of q_2 .

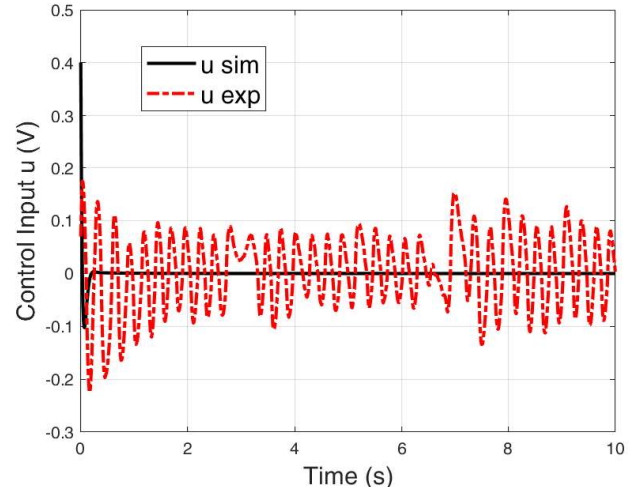


Figure 6. Comparison between simulation and experimental signal u .

- Define a sample time (e.g., $\Delta T = 0.01$ s)
- Obtain the equivalent discrete model by means of any transformation method, for instance: zero-order-hold: (e.g., $[A_d, B_d] = c2d(A_c, B_c, 'zoh')$)
- Ad must be full rank in order to ensure the system is controllable.
- Obtain matrix G according to equation (14)
- Obtain vector R_{free} according to equation (15)
- Define prediction and control horizon according to performance required, e.g., see examples: [1] and [21]
- Calculate the GPC control gains by means of equation (16), in this case, $P = 190$, $N_u = 8$, and $\lambda = 0.08$ so that $K_{GPC} = [-26.37; -26.91; -5.43; -3.60; 1.44]^T$

The following Figures (Figure 5 and 6) shows a comparison between simulation and experimental results by using a prediction horizon of $P = 190$ and control horizon of $N_u = 8$, with $\lambda = 0.08$ and control sample time is $\Delta T = 10ms$.

According to Figure 5 the simulation shows in position q_2 a maximum angular displacement of 0.15 rad (8.6 degrees) away of the (opened-loop) unstable equilibrium point while real experimental signal shows a very close response (maximum deviation from equilibrium is approximately 0.2 rad). This small difference between simulation and experimentation is caused by external noise and parametric discrepancies between the model and real plant. The lack of actuation on the second link demands a control strategy able to provide the minimum amount of energy to the first link and control the position of the second link. In the same context, the experimental control signal u remains under 0.5 Volts. The fluctuations observed are due to the 20Khz PWM signal provided by the ARM microcontroller that are partially smoothed by the load.

C. Swing-up strategy

The GPC control presented is activated once q_1 is approximate π while q_2 approximately 0, due to the fact the system is underactuated, there is not an arbitrary tracking trajectory for link 2. Thus a closed loop signal between link 1 with a proportional controller is established and the following parameters were chosen to approximate towards unstable equilibrium point. The following control parameters were chosen numerically and then implemented in the real device.

$$u(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5 \quad (20)$$

where, $a_0 = 0$, $a_1 = 0$, $a_2 = 0$, $a_3 = 122.69$, $a_4 = -289.83$, $a_5 = 182.57$, $K_p = 40$ and the switching conditions are $\pi - 0.1 < q_1 < \pi + 0.1$, and $-0.1 < q_2 < 0.1$. Transition between the swing-up and GPC control strategies are performed in a single sample time. The high-level programming used for this implementation allows to verify the switching condition and execute the code accordingly.

V. CONCLUSIONS

This paper shows a case study on using low-cost components to execute advanced schemes such as Model Predictive Control. The identification method was based on a standard linear regression technique in order to find a set of lumped parameters for model-based techniques. The combination of GPC control and identification approach provides adaptive features on the closed-loop control performance. The results obtained suggest that a relatively simple set-up open firmware is able to explore the functionality of an advanced form of control on a non-trivial dynamic plant. Additionally, it is suggested that when a resource-constrained computer system is used for portable applications and adaptive closed-loop control is required this approach can be a practical and feasible choice. We suggest that MPC industrial developers that commonly implement this model-based technique as a black-box may consider this

novel reinterpretation of how to implement a General Predictive Control as an open-guide resource. For academics, the main contribution of this paper is the methodology of integration both identification of nonlinear model with a predictive control (GPC) with low cost components in order to experience a smooth transition between the engineering school and more complex industrial applications. Many other publications deal with this process separately and more general. The objective of this contribution is meant to close the gap between the MPC family controller and off-the-shelf experimental components.

ACKNOWLEDGMENT

We want to acknowledge the Mitacs Elevate Program and the company Eigen Innovations Inc. for supporting the research and providing grant funding that made possible this work.

REFERENCES

- [1] E. F. Camacho and A. C. Bordons, *Model Predictive Control*, Springer, 2007.
- [2] J. Maciejowski, *Predictive Control with Constraints*, Pearson, Prentice Hall, 2000.
- [3] C. R. Cutler and B. C. Ramaker, "Dynamic matrix control," in *Automatic Control Conference*, 1980.
- [4] E. Gallestey, M. Lundh, T. Alloway, R. Martini, M. Stalder, and R. Santini, "Advancing system 800xa," Technical report, ABB, Jan 2014.
- [5] S. K. Lahiri, *Multivariable Predictive Control: Applications in Industry*, John Wiley and Sons, 2017.
- [6] R. Automation, "Model predictive control rockwell automation model predictive control delivers results," Tech. rep., Rockwell Automation, 2012.
- [7] J. R. Koelsch, "Model-based predictive control: No longer a black art," Tech. rep., Automation World, August 2013.
- [8] N. Khaled and B. Pattel, *Practical Design and Application of Model Predictive Control*, Elsevier, 2018.
- [9] R. Golightly, "The aspen dmc3 difference," Tech. rep., Aspen Technology, 2016.
- [10] E. F. Camacho and A. C. Bordons, *Model Predictive Control in the Process Industry*, Springer, 1995.
- [11] S. Qin and T. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, 2003.
- [12] E. P. Management, "Emerson process management," Tech. rep., Control Engineering, 2004.
- [13] R. Tedrake, "Underactuated robotics: Learning, planning, and control for efficient and agile machines," Course Notes for MIT 6.832.
- [14] M. W. Spong, J. K. Holm, and L. Dongjun, "Passivity-based control of bipedal locomotion," *IEEE Robotics Automation Magazine*, vol. 14 (2), June 2007, pp. 30–40.
- [15] M. W. Spong, "Partial feedback linearization of underactuated mechanical systems," in *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems '94*, Sep 1994, vol. 1, pp. 314–321.
- [16] R. Bautista-Quintero and M. J. Pont, "Implementation of h-infinity control algorithms for sensor-constrained mechatronic systems using low-cost microcontrollers," *IEEE Transactions on Industrial Informatics*, vol. 4 (3), Aug 2008, pp. 175–184.
- [17] E. R. Westervelt, J. W. Grizzle, and D. E. Koditschek, "Hybrid zero dynamics of planar biped walkers," *IEEE Transactions on Automatic Control*, vol. 48 (1), Jan 2003, pp. 42–56.
- [18] G. Butazzo, *Hard Real-Time Computing Systems Predictable Scheduling Algorithms and Applications*, Springer, 2002.
- [19] R. Kelly, D. V. Santibáñez, and P. J. A. Loria, *Control of Robot Manipulators in Joint Space*, Springer, 2005.
- [20] M. Gautier and P. Poignet, "Extended kalman filtering and weighted least squares dynamic identification of robot," 2000, vol. 33, pp. 935–940.
- [21] J. A. Rossiter, *Model Based Predictive control*, CRC press, 2003.
- [22] M. Awrynzuk, "On improving accuracy of computationally efficient non-linear predictive control based on neural models," *Chemical Engineering Science*, 2011.
- [23] R. Bautista-Quintero, R. Dubay, J. A. Carretero, and M. Diaz-Rodriguez, "Close-loop control identification of an inverted pendulum based on parameter linear regressor and generalised predictive control+ integral compensator," in *The 14th IFToMM World Congress*, 2015, pp. 318 – 322.
- [24] D. W. Clarke, C. Mohtadi, and P. Tuffs, "Generalized predictive control, part i the basic algorithm," *Automatica*, 1987.
- [25] X. Perry, *Designing Embedded Systems and the Internet of Things (IoT) with the ARM mbed*, Wiley, 2018.
- [26] MBED, "Mbed c/c++," <https://os.mbed.com/compiler/>, 2010.

DESIGN OF THE MECHANICAL TRANSMISSION FOR 5 DOF ROBOT WITH DISTRIBUTIVE ACTIVE SEMI-ACTIVE ACTUATION*

Sergey Pisetskiy

Department of Electrical and Computer Engineering
The University of Western Ontario
London, ON, Canada
Email: spisetsk@uwo.ca

Mehrdad R. Kermani

Department of Electrical and Computer Engineering
The University of Western Ontario
London, ON, Canada
Email: mkermani@eng.uwo.ca

Abstract—The Magneto-Rheological Distributive Active Semi-Active (MR DASA) actuation approach is based on antagonistic work of a pair of clutches at each joint of a robot. All the clutches in the DASA system are rotated by a single motor. To transmit the rotation from the motor shaft to the input of each clutch, a mechanical transmission is used.

This paper presents the mechanical design of the transmission for the five degrees of freedom manipulator with a single motor located at the base of the robot. After an introduction of the antagonistic MR DASA actuation approach, the details of the mechanical design of the transmission are presented. The paper concludes with the preliminary test results of the 1st joint of the robot.

I. INTRODUCTION

Mechanical actuation is the essential part of every robot. The mechanical aspects of the implementation of robot joints actuation determine the shape and performance of the whole system.

There are a number of well known approaches to configure the actuation system for the robot joint. The most desirable arrangement is considered to be a direct-drive approach [1]. In such configuration, the absence of transmission and reduction elements ensures excellent controllability and mechanical simplicity of an articulation.

However, in many cases the actuators tend to be too heavy and bulky for mounting them at the joints directly. Relocating an actuator away from a joint benefits in reducing the total inertia of a link. In order to transfer the motion from the active component to the joint, a proper transmission system needs to be implemented. In some cases the transmission can also perform the function of a reduction system.

A recent trend concerning the development of compliant robots introduced new requirements for the actuation and



Fig. 1. Visualisation of the 5-DOF robot.

transmission systems. Reduction of the inertia of the links plays an important role in achieving a robot intrinsic compliance and safety for an operator. In turn, it leads to a desire to relocate the actuators away from robot links and to introduce an effective system to transmit the motion.

A number of different configurations were proposed for compliant actuation of the joints with a remotely located power source. Possible solutions included a belt transmission for MR clutch actuation [2], [3], a system of tendons and pulleys [4], [5], a system of Bowden cables [6], air distribution system for pneumatic muscle actuation [7], etc.

In this paper the mechanical transmission for Distributed Active Semi-Active actuation for 5 degrees of freedom (5-DOF) manipulator is presented. The DASA technology [8]

*This work was supported in part by Canada Foundation for Innovation (CFI) and Natural Sciences and Engineering Research Council (NSERC) of Canada under grant No.25031 and RGPIN-346166.

based on magneto-rheological technology is chosen as the main approach to achieve the intrinsic compliance of the robot. The separation of the active component and semi-active components helps to reduce the overall inertia of the manipulator. The unique properties of MR clutches permit further improvements to compliance and to the safety of the robotic system.

The contributions of this work are as follows:

- Mechanical design of the transmission system that delivers a bi-directional rotational motion from a single power source to 5 separate revolute joints of the articulated manipulator.
- A mechanical arrangement of the spur gears that allows two parts of the Link #1 lower frame to act as a single united structure, while the MR clutch stator inside the frame is rigidly connected to the structure outside the frame.

II. OVERVIEW

A. Transmission design requirements

The main characteristics of the developed transmission are based on the manipulator design requirements. The robot is required to have 5 degrees of freedom in an articulated configuration, the work-space with the outreach of 900 mm, and 10 kg targeted payload at full extension of the arm.

There is a single electrical motor located at the base of the manipulator for actuating all five joints. Each of the joints utilizes a similar system of two magneto-rheological clutches coupled in an antagonistic configuration. All the robot links are connected to the MR clutch stators, and thereby directly actuated with the required joint torques listed in Table I.

TABLE I
TORQUES OF MR CLUTCHES AT THE ROBOT JOINTS

Joint	1	2	3	4	5
Torque [Nm]	15	200	80	15	15

Therefore, the transmission system of the robot is designed in order to deliver the bi-directional rotational motion to every joint from the single motor in the base. The components of the system are able to transmit the required torques. The design of the transmission allows the manipulator to work within its configuration limits with minimum friction.

B. DASA actuation approach

DASA or Distributive Active Semi-Active approach represents an actuation where a single power source is used to generate a constant rotational motion (see Fig. 2), and a number of clutches (magneto-rheological clutches) control the delivery of the motion to the robot links. To produce both forward and backward movement of the link, the two clutches are used at each joint. Two MR clutches spinning in opposite directions form an antagonistic pair that allows bi-directional actuation. Due to the unique properties of the MR clutch, the transmitted torque and the speed of rotation can be controlled

independently from the power source (a motor). Additionally, the clutches are able to filter unwanted perturbations from the power source and the transmission.

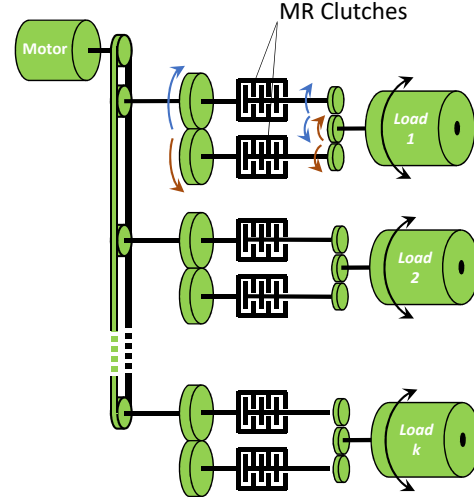


Fig. 2. Distributive Active Semi-Active actuation approach.

The mechanical transmission in the DASA system is designed to deliver bi-directional rotation to the joint. Multiple joints are actuated from a single motor located usually at the base of the robot.

C. Basic principles of MR clutch

An MR clutch is one of the key components of the proposed manipulator. The clutch ensures the compliant behavior of the robot by disengaging the reflected inertia of the motor and transmission from the actuated link.

MR clutch is able to transmit the exact amount of torque to the robot link. This is achieved through the precise control of the viscosity of the magneto-rheological fluid between inputting and outputting components of the clutch. The viscosity can be changed to the required value by applying a magnetic field with determined intensity. In the presence of the magnetic field, the micron-sized particles in the MR fluid form columns aligned in the direction of the field. The density of the created columns and their strength determine the yield stress of the fluid.

During operation of the clutch, the displacement of the moving part relative to the stationary part of the clutch causes shearing in MR fluid. The total force transmitted from input to output component can be found by multiplying the fluid yield strength by the surface area of the output (or the input) component.

An example of an MR clutch that is able to transmit rotational movement from the outer component (rotor as an input) to the inner component (Stator as an output) is shown in Fig. 3.

The clutch in Fig. 3 consists of a single ferromagnetic disk mounted on the rotor driven by the input shaft. The ferromagnetic stator with the electromagnetic coil is mounted

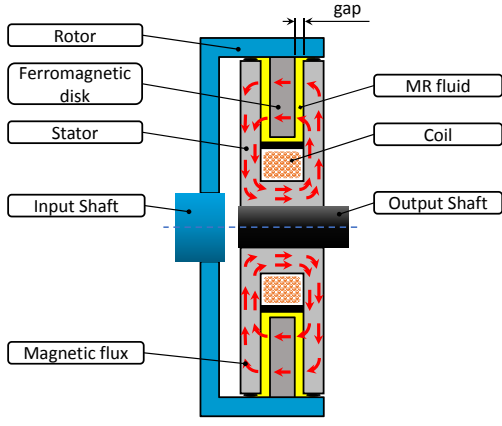


Fig. 3. Section view of a single-disk MR clutch.

on the output shaft. The gap between the stator and the rotor disk is filled with magneto-rheological fluid. The coil generates a magnetic field in order to change the viscosity (yield stress) of the MR fluid in the gap. The shearing of the columns in the fluid due to movement of the rotor disk (input) relative to the stator (output) causes the stress that transmits the torque. The amount of torque transmitted by each surface element of the disk is given by,

$$dT = 2\pi r^2 \tau dr \quad (1)$$

where r is the distance from shaft axis to the element (radius), τ is the MR fluid shear stress calculated with the use of Bingham visco-plastic model as follows [9],

$$\tau = \tau_y(B) + \eta \omega r / l_f, \tau > \tau_y \quad (2)$$

where $\tau_y(B)$ is the magnetic field dependent yield stress, η is the Newtonian viscosity of the MR fluid, ω is the angular velocity of the rotor disk relative to the stator, l_f is the fluid gap size.

By integrating (1) along the whole surface of the disk, the total torque transmitted by the clutch can be found.

The MR clutch can include more than one disk to achieve the required torque capacity. In this case the number of disks should be taken into account for the total torque calculation. The details of the construction of MR clutches construction and their features can be found in [10], [11], [12]. And the comparison of various MR clutch configurations can be found in [10], [13].

III. MECHANICAL DESIGN OF THE MANIPULATOR

A. Manipulator transmission

The transmission system is designed to transmit the rotational motion generated by one motor to 5 joints of the robot. As is shown on the kinematic diagram in Fig. 4, the single motor at the bottom of the robot base provides the motion to the joints through a serial transmission that consists of shafts, belts, and bevel gears.

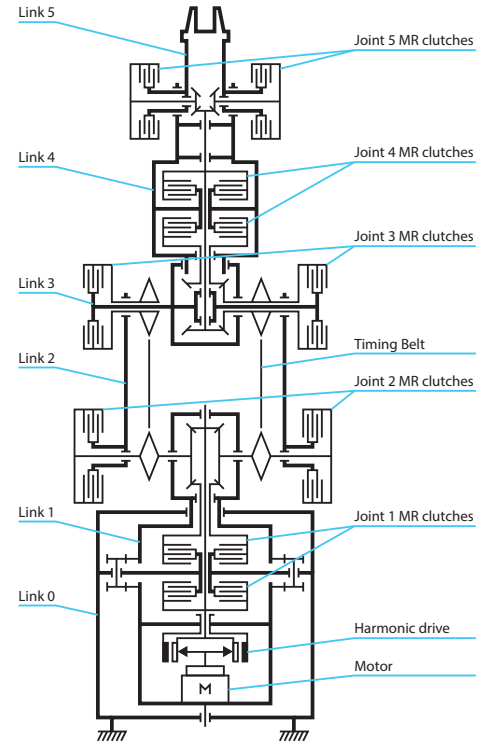


Fig. 4. Kinematic diagram of the 5-DOF robot transmission.

The rotation speed of the motor is reduced by the harmonic drive (HD drive) gearbox. From the output of the HD drive, the rotation is passed directly to the rotor of the bottom MR clutch and through the long inner shaft to the top bevel gear in the 2nd joint.

The opposite rotating motion is delivered to the MR clutch pair at each joint as follows:

1st joint - The rotor of the bottom clutch is directly connected to the HD drive output. The rotor of the top clutch is connected to the hollow shaft coming from the bottom bevel gear of the 2nd joint bevel gear set.

2nd joint - Rotors of both MR clutches are connected to the large lateral bevel gears of the bevel gear set.

3rd joint - Rotors of both MR clutches are connected with the rotors of the 2nd joint through the timing belt transmission.

4th joint - Rotors of both MR clutches are connected to the miter gears set located in the 3rd joint. The connection is realized through the pair of coaxial shafts. The bottom MR clutch rotor is driven by the top miter gear via the outer hollow shaft. The top MR clutch rotor is driven by the bottom miter gear via the inner shaft.

5th joint - Rotors of both clutches are driven in opposite directions by the set of three miter gears. The middle miter gear receives rotation from the shaft connected to the top clutch rotor in the 4th joint. Lateral miter gears are connected to the rotors of the MR clutches.

The robot transmission can be modelled based on relative angular velocity of rotor and stator for each clutch similar to what is done by A. S. Shafer and M. R. Kermani in [3] for 3-

DOF robot. Clutch rotor-stator velocities depend on the speed of the motor and angular velocities of the robot links (see Fig. 5). The relationship for angular velocities derived from the robot transmission scheme is given by

$$\begin{bmatrix} \dot{\theta}_1^\pm \\ \dot{\theta}_2^\pm \\ \dot{\theta}_3^\pm \\ \dot{\theta}_4^\pm \\ \dot{\theta}_5^\pm \end{bmatrix} = \pm \begin{bmatrix} n_{g/1} \\ n_{g/2} \\ n_{g/3} \\ n_{g/4} \\ n_{g/5} \end{bmatrix} \dot{\theta}_g + \begin{bmatrix} n_{1/1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \dot{q}_1 + \begin{bmatrix} 0 \\ -n_{2/2} \\ -n_{2/3} \\ -n_{2/4} \\ -n_{2/5} \end{bmatrix} \dot{q}_2 + \begin{bmatrix} 0 \\ 0 \\ -n_{3/3} \\ -n_{3/4} \\ -n_{3/5} \end{bmatrix} \dot{q}_3 + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -n_{4/5} \\ +n_{4/5} \end{bmatrix} \dot{q}_4 + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -n_{5/5} \end{bmatrix} \dot{q}_5 \quad (3)$$

where $\dot{\theta}_i^+$ and $\dot{\theta}_i^-$ are relative angular velocities between rotor and stator of pair of clutches at joint i , \dot{q}_i is the angular velocity of joint i , and $n_{i/j}$ is a gear ratio between joints i and j .

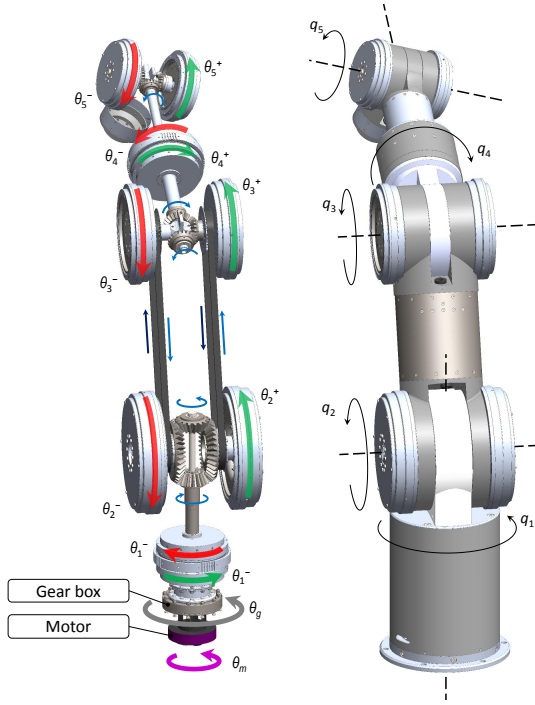


Fig. 5. Transmission of the manipulator.

B. Design of the robot base and Joint #1

The manipulator base (Fig. 6) is the heaviest part of the robot. The base represents link #0 and carries all the load from the rest of the links. It is designed to be mounted on a horizontal surface and has holes in the bottom flange for proper fixation with bolts. The inner space of the base is used to carry the lower frame of Link #1 with motor, gearbox, and two MR clutches of the 1st joint. The lower frame consists of two parts: top and bottom. The bottom part carries an electrical motor and a gearbox. The top part carries the rotor of an encoder.

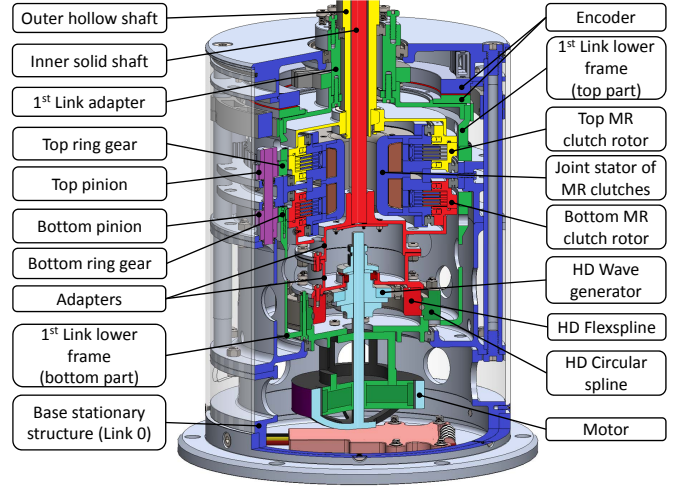


Fig. 6. Manipulator base zonal section view.

The electrical brushless motor Hacker Q80-13XS has maximum 6520 RPM and peak power of 2.7 kW. The motor is mounted just below the harmonic drive gearbox SHD-25-100-2SH. The shaft of the motor is directly connected to the wave generator of the HD drive, while the stationary circular spline is fixed to the structure of the Link #1 frame. The outputting flexspline is driving the rotor of the bottom MR clutch and the inner solid shaft, both connected to the gearbox via two cylindrical adapters. The harmonic drive has a 1:100 reduction ratio and 184 N-m limit for momentary peak torque.

The pair of MR clutches above the gearbox are responsible for the 1st joint actuation - the rotation of link #1 around the vertical axis. Each clutch can transmit torque up to 20 N-m. In order to optimize the weight and used space, both MR clutches share a common magnetic core with two separate windings. The joint core is fixed to the stationary frame of the robot base. Such fixation does not allow simple uniting of the top and the bottom parts of the Link #1 frame. Instead of a rigid connection, the spur gear arrangement is used to transmit the motor reaction torque from the bottom part to the top one. The gear arrangement ensures the rotation of both frame parts as one unit and at the same time allows the joint core of the clutch pair to be stationary fixed to the base structure.

The spur gear arrangement consists of two ring gears and six pinion pairs. One spur ring gear is located on the bottom part of the Link #1 frame. It is in mesh with bottom pinions of six pinion pairs mounted on the base. The top pinion of each pair is in mesh with the second ring gear. The second ring gear is mounted on the top part of the Link #1 frame. Through the pinion pairs, the rotation of the bottom ring is mechanically transmitted to the top ring gear. As a result, two parts of the Link #1 lower frame act as a single united structure.

As it was mentioned above, the rotor of the bottom MR clutch is actuated directly by the circular spline of the harmonic drive. At the same time, the rotor of the top clutch is driven in an opposite direction by the hollow shaft coming from the set of bevel gears above (see Fig. 7). Top and bottom

MR clutches work as an antagonistic pair, and by energizing one or the other, the 1-st link of the robot can be rotated in opposite directions.

To measure the angular position of the first joint, an absolute encoder Zettlex INC-9-150-171001 is mounted at the top of the manipulator base. The stator of the encoder is fixed on the top plate of the stationary base, while the rotor is mounted on the top part of the Link #1 frame.

C. Design of joint #2

The lower and upper frames of Link #1 are connected with the hollow adapter as shown in Fig. 7. The upper frame carries a set of bevel gears, the belt pulleys, an encoder, and a pair of MR clutches for Joint #2. The top bevel gear receives the rotation from the inner solid shaft. It is in mesh with the two large bevel gears on the sides and spins them in opposite directions with the ratio 1:2. Each side bevel gear is connected with the timing belt pulley and with the clutch rotor of the second joint.

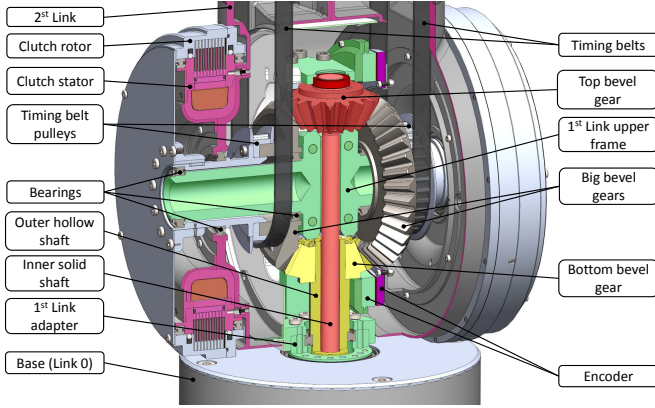


Fig. 7. Joint #2 zonal section view.

Each MR clutch is designed to transmit up to 200 N·m torque in order to lift the robot links with the required payload. The small bottom bevel gear is in mesh with the side bevels. The gear is driving the hollow shaft connected to the top MR clutch of Joint #1 (see Fig. 6). By means of the bevel gear system, the top clutch of Joint #1 is driven in opposite direction relative to the bottom clutch. Likewise, the large MR clutches of Joint #2 are actuated in opposite directions to each other to be able to work in antagonistic pair.

The absolute encoder Zettlex INC-9-175-171001 is used to measure the angular position of Joint #2. The stator of the encoder is fixed to the upper frame of Link #1, and the rotor is attached to the structure of Link #2.

The pulleys are used in timing belt transmission for actuation of the upper joints. Two belts are used to transmit motion in both directions to create antagonistic pair in the Joint #3.

D. Design of joint #3

The motion from the timing belts is received by two pulleys connected to the MR clutches of Joint #3 (Fig. 8). The rotors of the clutches are rotated in opposite directions to create an

antagonistic pair of actuators. The pulleys and the clutches are mounted on the frame of the 3-rd link. One of the pulleys (shown on the right side in Fig. 8) is connected to the miter gear. The set of miter gears is used for transmitting rotational motion at a 90-degree angle needed for the next joint (Joint #4) in the robot. Two receiving miter gears are mounted on concentric shafts and turn in opposite directions. Both shafts in the concentric shaft pair are hollow. The inner shaft has a cavity in order to pass cables from Link #3 to the upper links.

The absolute encoder Zettlex INC-4-100-161001 is used to measure the angular position of the third joint. The stator of the encoder is attached to the structure of Link #2, and the rotor is fixed to the frame of Link #3.

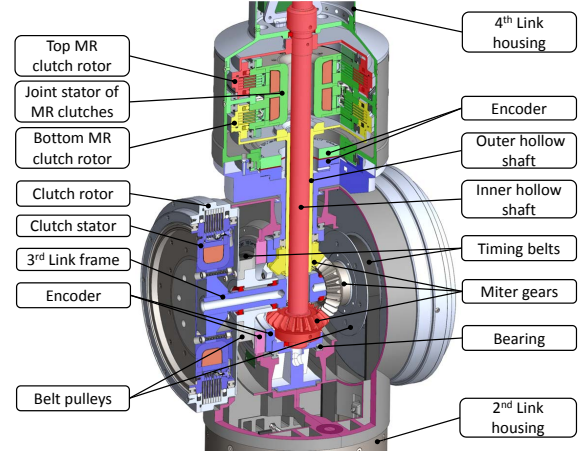


Fig. 8. Joint #3 and #4 zonal section view.

E. Design of joint #4

As can be seen in Fig. 8, the "roll" rotation of Joint #4 is perpendicular to the "pitch" rotation of Joint #3. A set of two concentric shafts is used to transmit the contra-rotating motion from the miter gear set to the MR clutches. The rotor of the top clutch receives rotation from the inner shaft, and the rotor of the bottom clutch is driven by the outer shaft. The pair of MR clutches in Joint #4 and in Joint #1 have identical design. Both clutches share the common stator attached to the housing of the link. By rotating in opposite directions, the rotors realize the antagonistic actuation approach.

The absolute encoder INC-4-75-151001 is used to measure the angular position of the 4th joint. The rotor of the encoder is mounted on Link #3, and the stator is attached to the structure of Link #4.

F. Design of joint #5

Joint #5 has the smallest miter gear set designed to receive rotation from the shaft to the bottom gear, and transmit it at a 90-degree angle with concentric lateral gears (see Fig. 9). The gears are attached to the MR clutch rotors that spin in opposite directions to each other. The stators of the clutches are attached to the structure of Link #5.

The smallest absolute encoder Zettlex INC-4-75-151001 is used to measure the angular position of the last joint. The

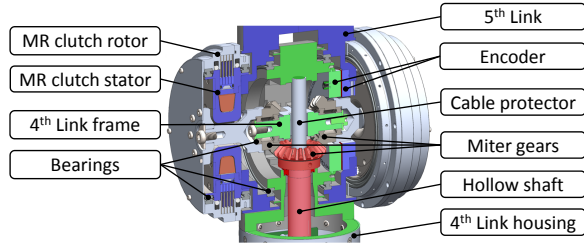


Fig. 9. Joint #5 zonal section view.

stator of the encoder is fixed to the frame of Link #4, and the rotor is attached to the structure of Link #5.

IV. ASSEMBLY AND TESTS

A. Preliminary tests of the robot base with the 1st joint

The robot base with Joint #1 and Link #1 is assembled as shown in Fig. 10. The integrity of the structure, the mechanical transmission performance, and Joint #1 actuation efficiency is successfully tested.

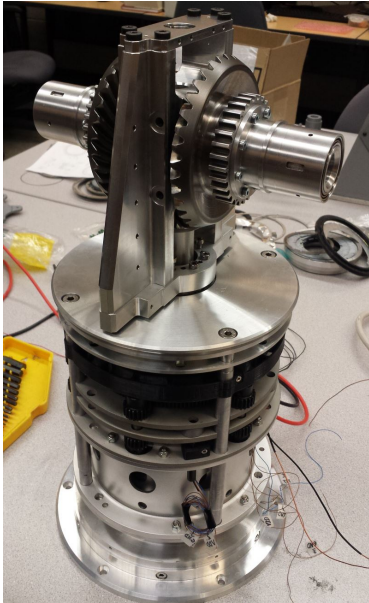


Fig. 10. Assembled base of the robot.

The maximum torque generated by the Joint #1 MR clutches is measured at 3A current. For the top clutch, the maximum transmitted torque reached 17 N·m. For the bottom clutch, the maximum torque observed is 17.3 N·m.

The preliminary frequency response test demonstrated high bandwidth reaching 55 Hz with open loop actuation control scheme.

The tests are performed using a static load cell Transducer Techniques SBO-1K. A motor driver AM AZ12A8 is used to supply current to the MR clutch coil. The real-time controller board dSPACE DS 1103 is used to control the test setup.

V. CONCLUSION

The design of the transmission for the 5-DOF manipulator, presented in this paper, demonstrates the possibility of building a complex robot that can be actuated by a single motor. The utilization of bevel gears and miter gears makes it possible to transmit rotation at a 90-degree angle and to attain contra-rotating motion to realize the antagonistic actuation approach. Timing belts and coaxial shafts are used to deliver the motion to the robot joints. The spur gear arrangement is used in the 1st joint to transmit the torque and motion between the parts of the frame. Such an arrangement facilitates the rigid fixation of the MR clutch stator inside the frame to the base structure outside of the frame.

Feasibility of the presented design is proved with assembly of the most complicated part of the robot - the base with the 1st joint. Preliminary tests of the prototype demonstrate that the performance of the mechanical transmission is sufficient to actuate the upper links as they are designed to move under optimal conditions.

Further work is planned to assemble and test the rest of the robot joints.

REFERENCES

- [1] A. AlbuSchffer et al., "The DLR lightweight robot: design and control concepts for robots in human environments", *Industrial Robot: An International Journal*, Vol. 34 Issue: 5, pp. 376-385
- [2] T. Kikuchi, K. Oda and J. Furusho, "Development of Leg-Robot for Simulation of Spastic Movement with Compact MR Fluid Clutch," in *Robotics and Automation (ICRA), 2009 IEEE International Conference on*. IEEE, 2009, pp. 1903-1908.
- [3] A. S. Shafer and M. R. Kermani, "Development of high performance intrinsically safe 3-DOF robot," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 619-624.
- [4] J. Viau et al., "Projected PID Controller for Tendon-Driven Manipulators Actuated by Magneto-Rheological Clutches," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 5954-5959.
- [5] D. B. Camarillo et al., Mechanics modeling of tendon-driven continuum manipulators, *Robotics, IEEE Transactions on*, vol. 24, 2008, pp. 1262-1273.
- [6] A. Schiele et al., Bowden cable actuator for force-feedback exoskeletons, *Intelligent Robots and Systems (iROS), 2006 IEEE/RSJ International Conference on*. IEEE, 2006, pp. 3599-3604.
- [7] D. Shin, I. Sardellitti and O. Khatib "A Hybrid Actuation Approach for Human-Friendly Robot Design," in *Robotics and Automation (ICRA), 2009 IEEE International Conference on*. IEEE, 2009, pp. 4369-4374.
- [8] A. S. Shafer and M. R. Kermani, "On the feasibility and suitability of mr fluid clutches in human-friendly manipulators," *Mechatronics, IEEE/ASME Transactions on*, vol. 16, no. 6, 2011, pp. 1073-1082.
- [9] R. W. Phillips, "Engineering applications of fluids with a variable yield stress," Ph.D. dissertation, University of California, Berkeley, 1969.
- [10] W. Li, P. Yadmellat, and M. R. Kermani, "Design optimization and comparison of magneto-rheological actuators," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5050-5055.
- [11] M. Moghani and M. R. Kermani, "Design and development of a hybrid Magneto-Rheological clutch for safe robotic applications," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 3083-3088.
- [12] P. Fauteux, M. Lauria, B. Heintz, and F. Michaud, "Dual-differential rheological actuator for high-performance physical robotic interaction," *Robotics, IEEE Transactions on*, vol. 26, no. 4, 2010, pp. 607-618.
- [13] A. S. Shafer and M. R. Kermani, "Design and validation of a magneto-rheological clutch for practical control applications in human-friendly manipulation," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 4266-4271.

On-Line Crack Detection in High Pressure Pipelines Using Distributed PZT Sensors

Khaled M. Al-Arife, Abdullah H. Abdullah

Department of Mechanical Engineering
Abu Dhabi University
Abu Dhabi University, UAE
Khaled.alarife@adu.ac.ae

Abstract— A technique for monitoring the structural health of high pressure pipelines using distributed piezoelectric (PZT) strips is described in this paper. The networking of these physical sensing and related computational units represents an industrial application of Internet of Things (IoT) for structural health monitoring (SHM). An array PZT strips captures the vibration signals as the fluid flows through the pipe and transforms it to a measureable voltage signal to the data acquisition instruments. To experimentally test the SHM system under controlled conditions, tests are performed to detect the presence of surface cracks by using a miniaturized vibro-motor with 100 Hz that was placed on the top of the pipe with temporary fixtures. Each measurement test included multiple measurements at different locations along the pipe. The PZT strips were radially arranged with 72° as a separation angle between each two adjacent strips over the 360° . The results showed a 4X increase in the amplitude of the voltage fluctuations for the cracked pipe when compared to undamaged healthy pipe. Furthermore, the number of oscillations increased with the presence of a crack. This experimental finding is interpreted into crack detection indicators, and then linked to the structural health.

Keywords—Structural health monitoring; piezoelectric strips; crack detection; smart sensors

I. INTRODUCTION

Industrial fluidic pipelines are high risk because they operate under high pressures, exposed to highly corrosive environments, and transport toxic fluids (liquids or gases). In addition to the negative impact on the surrounding environment, pipeline failures may lead to high shutdown and maintenance costs [1]. The structural health of a pipeline is often observed by monitoring the pressure dynamics because surface cracks and structural flaws create vibrations that alter the dynamical behavior. On-line detection of small and often invisible structural cracks is essential for the safe operation of these fluid transport systems and prevention of fluid leakage.

Fiber optic sensors (FOS) [2] have been used to the line's integrity over time. Essentially the technique looks for changes to the FOS communication signal that may occur due

to any physical damage in the pipeline structure. The FOS is an immune method from electric leakages and has high signal bandwidth [2,3]. However, the effectiveness of this method is associated with fully developed cracks that produce significant leakages. Alternative structural health monitoring (SHM) methods for pipeline structures include the use of acoustic emission in detecting fluidic leakages [4,5] and Eddy current testing (ECT). The acoustic emission technique relies on the energy waves that could be generated during leakage flows in their sonic and ultrasonic frequencies [5,6,7]. In contrast, ECT relates the structural changes that are associated with crack generation to the variations in the structural magnetic fields [8,9]. The changes in the structural magnetic fields are recorded by orthogonal coiled sensors and then correlate them to the electrical impedance of the sensor [9,10]. Even though this technique showed good results, its performance is linked to the type of pipe's material.

Recently, piezoelectric (PZT) transducers have been proposed as an embedded sensor for determining structural cracks in pipelines [13,14,15,16]. The approach is based on monitoring the structural impedance changes when triggered with excitation signals and then linking them to the existence of structural cracks [13,14]. This method has proven to be effective with high frequency excitation signals above 30 kHz [13]. The PZT transducers are also used to detect cracks by monitoring changes in the amplitudes of its electric outputs and linked them to the crack formation [17,18].

Recent advances in communication and computation enabled networking these sensing devices with computational processes for rapid data transfer. This connectivity is known as the Internet of Things (IoT) [19]. The application of IoT in the SHM is anticipated to promote data processing and decision making to be faster and more accurate [19,20,21]. The IoT is used in this paper to integrate the PZT stripes with PC, data acquisition unit, and decision maker to generate more reliable online crack detection procedure in the high pressure pipelines, Fig. 1.

The implementation of a distributed array of piezoelectric sensors, called a PZT strips, for detecting and locating cracks is discussed in this paper. Crack presence and location is determined by looking for changes in the amplitude and frequency of the piezo-voltages, and then linking them to the

existence and location of the structural cracks. This method showed success even in low frequency and low powered excitation signals. Basically, the used signal is in the range of 100Hz from low power vibrator. The excitation signal was tracked as an energy signal with the piezoelectric strips. The high accuracy and less complexity make this Internet of Things (IoT) based crack detection method a very good candidate for the SHM of wide range of static and dynamic structures.

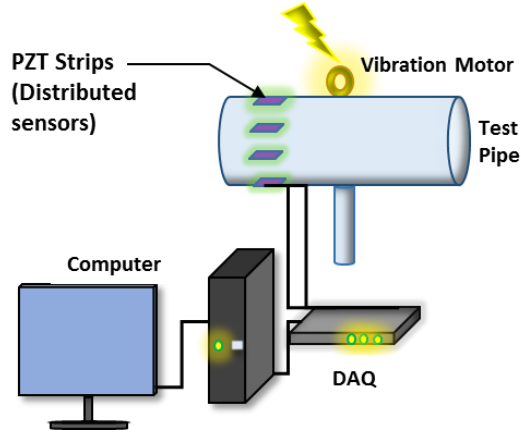


Figure 1. Schematic drawing of the experimental SHM process.

II. MATERIALS AND EXPERIMENTAL SETUP

A. Materials

The test specimen for this study was an ASTM A525 galvanized steel pipe with a reduced diameter to emulate the branched pipelines often used in fluid transport systems. Length and outer diameter of the main pipe are 1.00m and 0.112m, respectively. The branch pipe is 0.5m and 0.03m in length and diameter, respectively. The branch pipe was arc welded to an opening in the main pipe. This type of piping was selected due to its applicability, weld-ability, and ease of deformability. The physical properties of the used piping are listed in Table I.

TABLE I. ASTM A525 GALVANIZED STEEL PIPE DIMENSIONS AND PHYSICAL PROPERTIES.

Physical Properties	Main Pipe (1)	Branch Pipe (2)
Length	1.00 m	0.50 m
Outer Diameter	112 mm	30 mm
Wall Thickness	1.9 mm	3.1 mm
Modulus of Elasticity	200 GPa	200 GPa
Shear Modulus	80 GPa	80 GPa
Density	7850 kg/m ³	7850 kg/m ³
Poisson's Ratio	0.29	0.29

The controlled excitation signal for the experiment and analysis is generated by a low frequency and low power

micro-vibro-motor. The motor had a rated operating voltage of 1.5 V, operating current of 70 mA, and produced a repetitive mechanical vibration with a frequency of 100Hz.

The distributed sensor signals are generated by five piezoelectric sensing strips in each of individual tests. The piezo-strips were purchased from Piezo Systems, Inc. (MA, USA). Each strip is a bi-mount piezoelectric bending sensor that has physical dimensions of 54mm, 6.3mm, and 1.8mm in total length, width and thickness respectively, Fig. 2. Each of the piezo-strips is connected to the data acquisition device (National Instruments -DAQ) and connected to a computer running LabVIEW software.

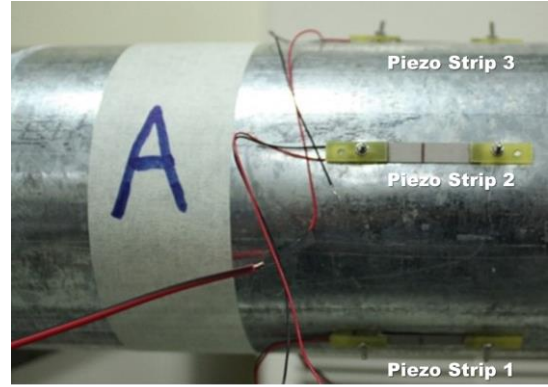


Figure 2. Photograph of the piezo strips at test location A.

B. Experimental Setup

The experimental setup considered two specific locations as shown in Fig. 3: A (cracked side of the pipe) and B (healthy side of the pipe). Tests were done in a series of four trials that included a control test with no input signal and no cracks present. A second trial involved an input excitation signal on a pipe with no structural cracks. The third trial involved no excitation but a planned crack on the pipe. Finally, the fourth and last trial considered both the excitation signal and planned crack. Each of these sessions was done in location A and then in location B. On the other hand, the excitation vibration signal was always sent from the same location X for excited sessions to understand the effect of the measurement location and the impact of the existence of the planned crack on the piezo-voltage.

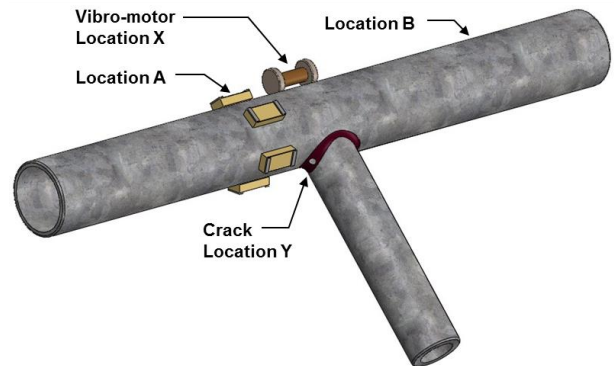


Figure 3. Schematic drawing of the experimental setup.

The test sessions have used five piezo-strips that were distributed on the pipe's circumference with 72° as a

separation angle to ensure a balanced distribution. The main theme behind selecting locations A and B is to understand how the piezo strips will receive the propagating strain energy that were driven by the excitation signal in the healthy and cracked pipes. This can be realized by applying the current sessions. In different words, these different sessions and their assigned locations may make the trouble shooting processes faster and more accurate.

Physically, the detection location A was located at 16cm before pipe's branching in the crack's side, while location B was 16cm away from the branch in the healthy side of the pipe. It is important to note that, the input vibration source was placed at the same point in all of the excited sessions; it is indicated with location X, Fig. 3. The test arrangements are described in Fig. 4. The piezo strips are connected to the DAQ and to the PC for providing the LABVIEW with streams of the piezo-voltages. All of the hardware connections are made to enable all of the sensors to operate in the same time. Once operated, the strips sent their readings to the computer. Finally, the strips' signals that are sent to the computer are processed by LABVIEW to present the piezo-strips' outputs numerically and graphically.

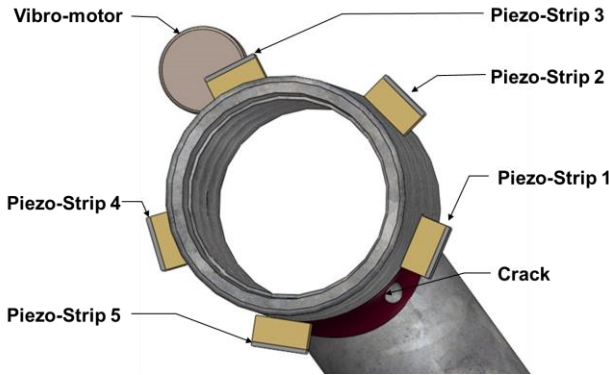


Figure 4. Schematic drawing of the experimental setup used for the tests.

III. RESULTS AND DISCUSSION

A. Experimental Results

The effectiveness of using low frequency mechanical excitation signals in determining the existence and location of structural cracks is demonstrated by conducting four different sets of experiments. These experimental tests investigate the PZT strip responses to both healthy (i.e. undamaged) and cracked pipes when exposed to the same mechanical excitation. For the tests the fault in the cracked pipe was introduced at the welded joint that connects the main pipe to the reduced diameter branch (Fig. 5). The approximate diameter of the crack is 6mm. Damage near the welded joint is realistic because of the excessive thermomechanical stresses that can be expected due to the difference in the thermomechanical responses between the welded joint and the rest of the piping structures.

The first set of experimental tests involved exposing the pipe structure to mechanical excitation at location X (Fig. 4) and then monitoring the signal flow using circumferentially

arranged PZT strips at location A. The measurements were initially recorded by the PZT strip 1 at location A (A1) for both the healthy and cracked/damaged pipes as shown in Fig. 6. When excited by the vibration motor, the measured response by the PZT strip on the healthy pipe showed a fast buildup with a steady state in less than 1.25 seconds and oscillations with magnitudes $> 1\text{mV}$ (blue line). In contrast, the same inputs applied to the cracked pipe showed a rapid built-up in approximately 0.5 seconds leading to oscillations with an amplitude of 4mV (red line).

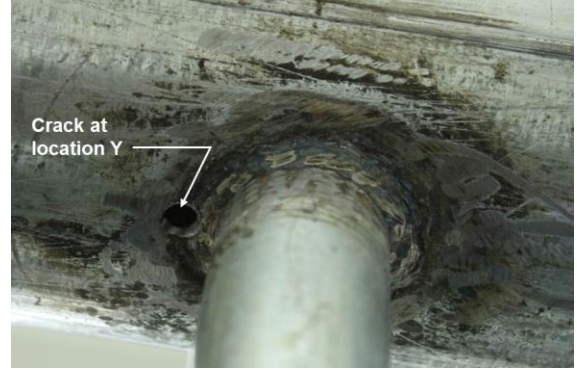


Figure 5. Photograph of the cracked pipe.

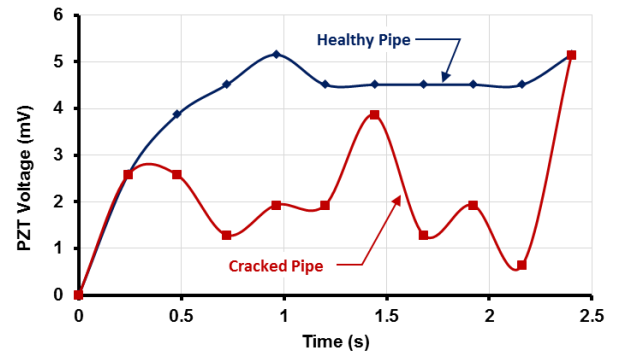


Figure 6. Piezo-strip voltage response at location A1 for the healthy pipe (blue line) and cracked/faulty pipe (red line).

The second set of tests explored the difference in response between the healthy and cracked pipes when excitation signals are applied at the same location A1 (Fig. 7). The initial test is performed on the healthy structure without an input excitation signal (dashed line). The detected piezo voltage showed smooth built-up voltage with maximum amplitude of less than 2 mV . The impact of applying a mechanical signal to the healthy structure is explored by exposing the same healthy pipe to a low frequency mechanical vibration at 100Hz . Once the signal was sent, the PZT strips detected the highest voltage amplitude among this set of tests, where it built up smoothly over 1 second (blue line). For the damaged or cracked pipe without mechanical excitation, the PZT voltage built up rapidly in approx. 0.25 sec and then exhibited oscillations with varying amplitude (green line). Similarly, the response of the cracked pipe to mechanical excitation was observed (red line).

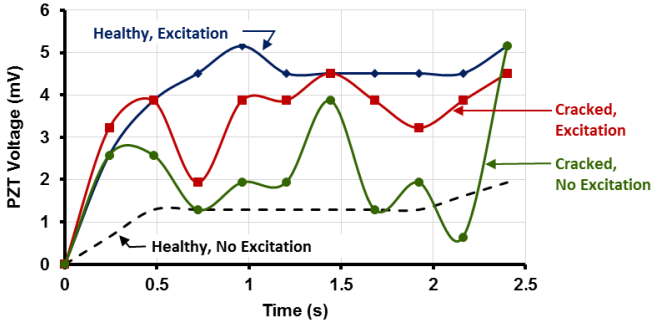


Figure 7: Piezo-strip voltage response for healthy pipe without excitation (black dashed line); healthy pipe with excitation (blue line); cracked pipe without excitation (green line); and cracked pipe with excitation (red line).

The third series of tests examined the effect of the PZT sensor's location by measuring the differences on undamaged (Fig. 3: A) and damaged (Fig. 3: B) sides of the same pipe when mechanically excited (Fig. 8). The recorded voltage showed rapid buildup for both cases but the PZT sensor near the crack exhibited sudden fluctuations after ~ 0.25 sec. It is important to note that the maximum voltage amplitude of the undamaged side is higher than that of the faulty side which implies that the cracks may have dissipated, or damped, part of the received energy in the signal. It is possible that the excitation signal was momentarily stored in the form of elastic strain in the pipe's material before changing its flow direction around the crack. Changing the signal direction and using the available structural stiffness are accompanied with the material's damping which is inherent in the material's grains contributed in dissipating part of the excitation energy.

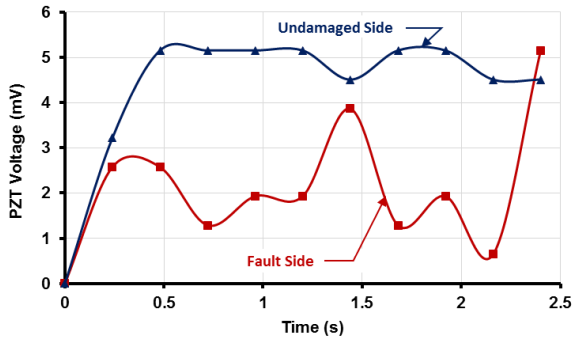


Figure 8: Piezo-strip voltage response for undamaged side (blue line with solid circles) and fault side (red line) of the cracked pipe.

The fourth and final experimental test investigated the impact of the strip's location on the cracked pipe's circumference. Five PZT strips were distributed on the pipe's circumference at 16 cm distance from the branch at the cracked side (i.e. location A). This arrangement is to link the cracks radial location to the responses of the PZT strips. Furthermore, the five piezo strips were distributed on the pipe's circumference equally with a 72° as a separation angle. PZT strip number 3 was the closest strip to the signal generator and the PZT strip number 1 was the nearest strip to the crack. When the excitation signal was sent, the vibration signal is tracked on the pipe's circumference with two different pathways. The first flow direction (or pathway) is from the

signal generator to Strip 3 then Strip 2 then Strip 1, while the second pathway is from Strip 3, to Strip 4, then to Strip 5.

Figure 9 shows the recorded voltages of the five circumferential piezo strips for the first signals pathway (Generator→Strip 3→Strip 2→Strip 1). At Strip 3 (red line), the response oscillates about the zero voltage. This implies that the excitation signal triggers this strip and then moves to the next PZT sensor or Strip 2. Similarly, the response of Strip 2 (green line) also oscillates about the zero voltages. Then the signal moves to the Strip 1 (black line) which is located nearest one to the crack with the highest oscillations. Note that the PZT voltages are negative because the PZT strips can bend either upward or downward (one direction positive and the other negative).

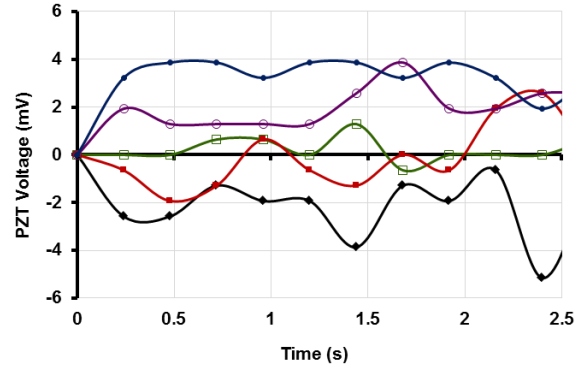


Figure 9: Piezo-strip voltage response for the cracked pipe at location A. Strip 1: black line; Strip 2: green line; Strip 3: red line; Strip 4: purple line; Strip 5: blue line.

The second signal pathway is also shown in Fig. 9 (Generator→Strip 3→Strip 4→Strip 5). Once more, the response for Strip 3 (red line) shows oscillations about the zero voltage. The signal along this pathway is transferred to Strip 4 (purple line) where it oscillates away from the zero voltage with smaller amplitude. After Strip 4, the signal goes to Strip 5 (blue line) which produces the lowest oscillations and high voltages. The piezo-strip voltage characteristics in both of the signal pathways suggest that both the amplitudes and centers of oscillations provide a clear link to the fault's (i.e. crack) location.

B. Discussion

The recorded oscillations in the piezo-strip voltages for the cracked pipe suggest that the mechanical vibration signal sent from position X is influenced by the existence of the crack and this is observed by the changes in the flow pathway characteristics. Essentially the energy in the vibration signal flows smoothly for the healthy structure but in the presence of a crack oscillation occur because as the signal reaches a physical gap in the solid material structure the pathway is altered. This change in the direction of signal flow arises, in part, when the signal's energy is momentarily stored in the structural material as elastic energy and after reaching the capacity-limit of its stain energy, it is released in the form of pulsed strains. These strains are detected by the network of PZT strips, captured by the data acquisition device, and then

transmitted to computer software for signal analysis (e.g. spectral analysis, pattern recognition).

IV. CONCLUSIONS

The application of a PZT strip network in detecting and locating cracks along high pressure pipelines has been explored. Specifically, five PZT strips in two pathway configurations were examined to determine vibration signal flow through a faulty (i.e. cracked) pipe. The investigative study included four separate stages: (i) the existence of crack was demonstrated by monitoring the piezo-strip voltage peaks and oscillations where the cracked pipe showed significant differences in the amplitude and oscillatory behavior when compared to health, undamaged pipe; (ii) the impact of using vibration excitation signal where the cracked structure exhibited oscillating voltages with and without mechanical vibration signal (Note: the oscillating response reflects the impact of the surrounding noise signals on the structure dynamics); (iii) locating the crack on the axial axis by linking the voltage oscillations to the distance from the structural crack; and (iv) locating the crack on the pipe's circumference as demonstrated by monitoring voltage changes in terms of voltage oscillation and amplitude values (i.e. crack on side with highest amplitudes and greatest signal fluctuations). Once captured and recorded the PZT sensor data can be analyzed in an effort to determine the overall health of the pipeline system. The proposed methodology makes monitoring loaded structures easier and less costly. Once any change in the PZT sensor voltage dynamics is observed, it can be electronically e-linked to the health conditions of the structures.

ACKNOWLEDGMENT

This work has been supported by the research office & sponsored programs of Abu Dhabi University.

REFERENCES

- [1] M. A. Murad, "The Development of structural health monitoring (SHM) procedures for the structural integrity and maintenance repair of offshore Ageing pipelines," Proceedings of the 4th European-American workshop on reliability of NDE, Berlin, Germany, pp. 1-15, 2009.
- [2] P. Rajeev, J. Kodikara, W.K.Chui, T. Kuen, "Distributed Optical Fibre Sensors and Their Applications in Pipeline Monitoring," Key Engineering Materials vol. 558, pp. 424-434, 2013.
- [3] S. Z. Yan and L. S. Chyan, "Performance enhancement of BOTDR fiber optic sensor for oil and gas pipeline monitoring," Optical Fiber Technology, vol. 16, no. 2, pp. 100-109, 2010.
- [4] S. Pan, Z. Xu, D. Li, D. Lu, "Research on Detection and Location of Fluid-Filled Pipeline Leakage Based on Acoustic Emission Technology," Sensors, vol. 18, 3628, pp. 1-17, 2018
- [5] A. Anastasopoulos, D. Kourousis, and K. Bollas, "Acoustic emission leak detection of liquid filled buried pipeline," Journal of Acoustic Emission, vol. 27, pp. 27-39, 2009.
- [6] D. Ozevin and J. Harding, "Novel leak localization in pressurized pipeline networks using acoustic emission and geometric connectivity," International Journal of Pressure Vessels and Piping, vol. 92, pp. 63-69, 2012.
- [7] A. Mostafapour and S. Davoodi, "Analysis of leakage in high pressure pipe using acoustic emission method," Applied Acoustics, vol. 74, no. 3, pp. 335-342, 2013.
- [8] D. Rifai, A. N. Abdalla, R. Razali, K. Ali, M.A. Faraj, "An Eddy Current Testing Platform System for Pipe Defect Inspection Based on an Optimized Eddy Current Technique Probe Design," Sensors vol. 17, 3, pp. 1-24, 2017.
- [9] C. Camerini, J. M. A. Rebello, L. Braga, R. Santos, T. Chady, G. Psuj, G. Pereira, "In-Line Inspection Tool with Eddy Current Instrumentation for Fatigue Crack Detection," Sensors, vol. 18, 2161, pp.1-12, 2018
- [10] J.B. Nestleroth, R.J. Davis, "Application of eddy currents induced by permanent magnets for pipeline inspection," NDT & E International, vol. 40, Issue 1, pp. 77-84, 2007
- [11] A. Mohan, S. Poobal, "Crack detection using image processing: A critical review and analysis," Alexandria Engineering Journal, vol. 57, Issue 2, pp. 787-798, 2018
- [12] S.-H. Peng, H.-S. Kim, H.-D. Nam, "Metal Crack Detection in X-ray Images Based on Local Brightness Variation and Multi-scale Analysis," International journal on information vol. 15, pp. 1961-1972, 2012.
- [13] G. Park, H. H. Cudney, and D. J. Inman, "Feasibility of using impedance-based damage assessment for pipeline structures," Earthquake Engineering and Structural Dynamics, vol. 30, no. 10, pp. 1463-1474, 2001.
- [14] G. Du, Q. Kong, T. Lai, G. Song, "Feasibility Study on Crack Detection of Pipelines Using Piezoceramic Transducers," International Journal of Distributed Sensor Networks vol. 2013, Article ID 631715, pp. 1-7, 2013.
- [15] D. Mateescu, Y. Han, A. Misra, "Dynamics of Structures with Piezoelectric Sensors and Actuators for Structural Health Monitoring," Key Engineering Materials Vol. 347, pp 493-498, 2007.
- [16] B. Chomettea, A. Fernandes, J.-J. Sinouc, "Cracks detection using active modal damping and piezoelectric components," Shock and Vibration, vol. 20, 619-631, 2013
- [17] M. Gresil, L. Yu, V. Giurgiutiu, "Fatigue crack detection in thick steel structures with piezoelectric wafer active sensors," Proc. of SPIE of the SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring, 2011, San Diego, California, United States vol. 7983, pp. 79832Y-1 - 79832Y-15, 2011.
- [18] Julie L. Briand, "Experimental validation of novel structural health monitoring strategy for bolted pipe joints" Dalhousie University, M.Sc dissertation, 2010.
- [19] P.G. Benardos, G.-C. Vosniakos, "Internet of Things and Industrial Applications for Precision Machining," Solid State Phenomena, 1662-9779, vol. 261, pp 440-447 2017.
- [20] C. Mendes, R. Osaki, C. da Costa, "Internet of Things in Automated Production," EJERS, European Journal of Engineering Research and Science, vol. 2(10), pp. 13-16, 2017.
- [21] L. Nagalakshmi, K.S. Ansari, N. Gowtham, J. Raman, "Solar Cell Crack Identification using Internet of Things (IoT)," Proceedings of the International Journal of Advanced Research in Computer and Communication Engineering, vol. 6(3), pp. 869-872, 2017.

Plant Irrigation Water Sprinkler Robot

Adel Younis

Assistant Professor, Dept. of Mechanical Engineering
Australian College of Kuwait
Kuwait, Kuwait
Email: a.younis@ack.edu.kw
(Corresponding Author)

Dalal AlSalahi, Ahmad AlKhudher, Hamoud AlBaijan,
Eissa AlFailakawi and Mohamed Ahmed
B.Tech. Candidates, Dept. of Mechanical Engineering
Australian College of Kuwait
Kuwait, Kuwait

Abstract—This paper encompasses the different aspects associated with the design of an irrigation robot which could be used as a replacement for complex irrigation systems. It provides a design that is more financially beneficial as well as the requirements needed for the creation of this product. The objectives of this report are to create an environmentally friendly robot with a reasonable cost which can automatically irrigate plants in a specified area without the need of direct interference from individuals, is user friendly and can function indoors as well as outdoors. Commencing with an introduction, which encompasses the objective of designing an irrigation robot and moving to a theoretical background of irrigation systems and robots; this report aims to provide detailed information on each aspect to reach an understanding in depth of the requirements for the creation of this project. Governing equations and calculations are included in order to provide a numerical explanation of the path leading to the implementation of this design, as well as to eliminate errors which may occur in the design phase of this project. Finally, the water irrigation robot was designed, manufactured, tested and proved to be working efficiently.

Keywords—component; mechatronics; irrigation; robot

I. INTRODUCTION

Generally speaking, with the busy lives and frequently hectic situations which many individuals face on an almost daily basis or simply because of forgetting, people tend to forget to water their plants. In addition, installing an irrigation system for large areas such as plant production units tend to usually be expensive and difficult to uninstall once installed. This report presents the user with a solution, which is cost-efficient and is completely autonomous for both situations. The report displays the design of the robot, which is made up of a mobile robot, solar panel, Arduino and its respective required components such as moisture sensors. This project's robot automatically measures the plant's water needs based on the soil water content and waters the plants according to their respective soil needs. The robot is portable, allowing it to be used in several different locations, only requiring the change in

the track of motion required through programming. Furthermore, it uses solar panels to charge the battery, which will run an electric motor and the pump, making it environmentally friendly.

II. LITERATURE REVIEW

A. Irrigation system

Irrigation system is the principle of providing the amount of water needed for watering plants and agricultural fields. It, helps to increase the production of agricultural crops. As a matter of fact, the irrigation system enhances the soils in dry areas by means of nurturing the soil using water provided by mechanical and hydraulic devices, as well as in areas that rainfall rate is low. Irrigation system has been beneficial to the residential and commercial properties, in such way that reduces consumption of the usage of excessive amounts of water and saves money. The irrigation system has many values in growing more pastures and crops. Such values are: the water can maximize the benefits of fertilizer by irrigating the grounds which have plants to have best fertilization for plant growth; the other value is, by using this system, the impact of water stress can be avoided in order to produce higher quality crops/pastures. Furthermore, watering the garden with a hose may have the effect of excessive water entering the soil, which will result in damaging roots and plants due to reduced absorption of nutrients and the soil will be compacted. Vice versa, when using the irrigation system, the amount of the supplied water will be known and that will help to preserve nutrients and to minimize the soil compaction.

Historically speaking, the first idea of irrigation system has been practiced in Egypt and in Mesopotamia as far back as the 6th millennium BCE, where the natural rainfall was insufficient to assist the crops.

Archaeological investigation has found evidence from 1800 BCE of the ancient Egypt, in the twelfth dynasty of pharaoh Amenemhet, Faiyum Oasis, which is a natural lake that swelled annually by the effect of the annual flooding of the Nile, was used as a reservoir for storing the overflowing of the water for usage during the dry seasons.

Among the oldest known irrigation methods that still have been used nowadays where it can be found in the Middle East,

Asia and North Africa, called the Qanats. This method was developed in ancient Persia in about 800 BCE. The system consists of a network of vertical wells and gently sloped tunnels driven in the slopes and steep hills to take advantage of groundwater.

One of the most complex irrigation systems of the ancient world was the one in the reign of King Pandukabhaya, dating back to about 300 BCE, Sri Lanka. Although the underground canals exist, building the first completely artificial reservoirs in order to store water, were carried out by the Sinhalese. The Sinhalese were often called 'masters of irrigation' because of their engineering superiority in the irrigation sector. During the time between 1153 - 1186 CE in the reign of King Parakrama Bahu, the system was widely restored and extended as well.

Li Bing, Chinese hydrologist and irrigation engineer who devised the Dujiangyan Irrigation System in 256 BCE. The function of the system is to irrigate a wide area that still provides water today.

In 1441, a Korean engineer of the Joseon Dynasty, Jang Yeong-sil, invented the world's first rain gauge. It was installed in irrigation tanks as part of a national system for measuring and collecting rainwater for agricultural implementation.

B. Robot Design

Any machine, especially if programmable and capable of performing a specified set of actions, can be defined as a robot. The embedment of the control system for a robot can be external like remote controls or internal and their appearances vary from looking like humans to simple form like a mini car but they all carry a task.

Robots range from being autonomous to being semi-autonomous and they vary in shape and function from humanoids to medical, industrial, drones or even microscopic sized robots. These robots and more, serve to proliferate and facilitate many functions, which are undesirable to be carried out or dangerous too. The branch, which deals with such technology, is known as robotics and it involves all aspects such as designing, construction, operation and even control systems of robots.

The term robot itself comes from *robota*, a Czech word which means forced labor and the word robot was used in naming a fictional character which was a humanoid robot in the play *R.U.R* written by Karel Capek, a Czech writer, in 1920. His brother, Josef Capek, though, invented the actual word. The first real robot to be created was developed by American-born British neurophysiologist, robotician and cybernetician, William Grey Walter, in 1948. Furthermore, Unimate, the first commercially programmable robot that was built in 1954 by George Dovel and was used in General Motors assembly.

Since ancient times, the concept of automated devices can be found in many different forms and configurations of automated devices, some resembling animals and others humans, which were used mainly as means of entertainment. However, as we move towards the industrial age, the usage of automated devices or machines has become of a more practical use in the industry.

The reason for such development is mostly due to the fact that electricity has been introduced by that time, the industrial

age, and was widely used, unlike in ancient times, even though there have been several accounts of discovering some form electricity usage in ancient times, there were not so common. The industrial age brought about the usage of small power suppliers like portable motors and batteries, which left thinkers with a wider range of possibilities to create new concepts, one of which came in towards the beginning of the 20th century, namely, humanoid machines. Nowadays, in the 2000s, the development of robots has reached to levels of human-like intelligence and appearance thanks to artificial intelligence and even having the first humanoid robot citizen in Saudi Arabia who's fully functional and is capable of withholding conversations on an intellectual level.

Earlier accounts of this concept can be found all through history, though they reside mostly in the realms of mythology and unasserted claims of the past. Starting with Greek mythology, the three-legged table which can move as if having a mind of its own, which was created by Hephaestus, the god of fire (in Greek mythology). Also, from the ancient legend of Cadmus, who was said to turn dragons' teeth into soldiers, the concept of creating an intelligent form can be found. Moving to the time of BC, 4th century BC to be specific, a steam propelled mechanical bird, which was created by Archytas of Tarentum, is yet another great example of automated machines of the past.

C. Irrigation Robot

In May 2003, "A Robotic Plant Care System" by Kevin Sikorski [9] a system, which irrigated plants in the Intel Lab have been developed. The system consists of laser range finders to locate pot plants in the lab environment. Unfortunately, the system had several limitations such as the system was very expensive, the approximate cost of only the major components of the system accounted to approximately \$9000 (Prices according to Active Media pricelist June 2002). Moreover, the laser range finders failed to work sometimes as the Lab contained surfaces that were hostile to laser reflections like the glass doors. The pot plants were considered to be circles when detected by the range finder, with an assumption that only circular base pots were used, but in fact the shape of the pots can vary from being square to a hexagon to a circular one.

Another irrigation system "PotPet: Pet-like Flowerpot Robot" by Ayumi Kawakami, Koji Tsukada, Keisuke Kambara and Itiro Siio developed in 2011 [2], which is a flowerpot-type robot called PotPet. This system requires each potted plant to be equipped with all the sensors in order to stimulate the robot when irrigation is needed. Needing sensors in each potted plant increasing the cost of the system with each newly added pot, which drawback the system. Moreover, the system doesn't actually water the plants in fact it just grabs their attention and reminds the user to water the plants making the watering operation manual rather than completely autonomous.

Another system with the same purpose is ‘A Smart System for Garden Watering using Wireless Sensor Networks’ by Constantinos Marios Angelopoulos, Sotiris Nikolettseas and Georgios Constantinos Theofanopoulos that was also developed in 2011[5]. This system waters the potted plants by analyzing their soil moisture using sensors and waters them using the attached valves. The problem with this system is that it is not portable; the valves are always attached to the potted plant spoiling the beauty of the environment. Moreover, each potted plant is equipped with separate sensors and valves for watering, which increases the cost of the system as the number of potted plants increase.

In 2012, an irrigation robot has been designed with the use of a temperature-humidity sensing module to water plants. The system is fully automated irrigation system, which uses wavelength communication to communicate between the robot and the sensing module. This robot is completely portable and is equipped with a Radio Frequency Identification (RFID) module, a microcontroller, an on-board water reservoir and an attached water pump. It is capable of sensing the watering needs of the plants, locating them and finally watering them autonomously without any human intervention. Mobilization of the robot to the potted plant is achieved by using a predefined path. For identification, an RFID tag is attached to each potted plant.

III. PROJECT DISCRIPTION

Irrigation systems as mentioned in the literature review section, have developed so much to the point of where robots have been used for watering plants. However, the robots have only been limited to being used indoors in a specified area, while this project promises to deliver a robot, which can be used outdoors in a considerably larger area. The robot being designed in this project consists of a $450 \times 400 \times 150$ mm sized robot, having its top as a flat plate with a tank which the size of was calculated using equation (1). In addition, tank treads were used instead of wheels, though the robots, which were for indoors have regular shaped wheels. This is because when used outdoors, the ground in which a robot operates on top of would not be completely flat, as is the case with the area indoors. Furthermore, the robot has a solar panel, which will be used to charge the batteries needed to empower the robot's motion and general functionalities. As for the sprinklers in used, lawn sprinklers (garden sprinklers) are to be used in order to reduce water wastage and to eliminate the possibility of harming the plants and to enhance the efficiency of plant fertilization. As for how the robot will operate, it will commence with the robot sensor arm sensing the moisture of the location it's on, then the robot acts accordingly to irrigate the surrounding area. Otherwise, the robot will move to locations as shown in Fig. 1, sensing the moisture in the specified locations.

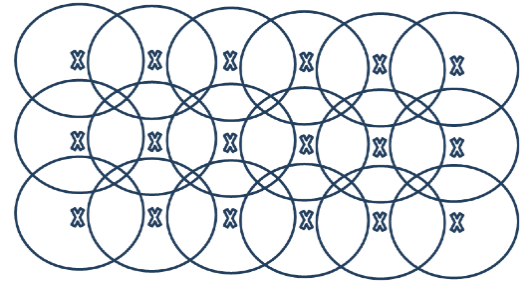


Figure 1. Irrigation water distributions

A. Robot Design

A robot can be defined, as previously stated, as any machine, which can perform a specified set of actions. When it comes to designing a robot, the first thing that comes to mind is the purpose or the objective behind a robot's creation, in other words, the tasks required to be carried out by the said robot. These specifications depend on many factors such as the application required, the field of operation for the robot and the marginal scale of severity of that application. Based on the said factors, the materials needed for the design can be selected in order to satisfy the requirements of the design.

This paper concentrates on the application of a robot as a substitute for irrigation systems. To be able to satisfy such a requirement, it was determined that a car-base-sized robot, with tank wheels instead of regular car wheels of a total dimension of $40 \times 15 \times 45$ cm is to be used, which is based on choosing a reasonably sized robot in order to be able to be used outdoors while having enough capacity to carry the function of irrigation of plants. Tank wheels were used instead of regular wheels to facilitate the movement of the robot in uneven grounds and in muddy areas without affecting the movability of the robot. Above the plate, a tank of size $35 \times 15 \times 25$ cm was chosen based on the maximum allowable weight to be carried by the robot of 40 kg and this tank can hold up to 30 liters of water. The choice for the tank size was determined by the fact that a solar panel is to be added alongside other components such as a robotic arm and an Arduino board, so the maximum allowable tank size was reduced in order to have the total weight of the objects on the robot below the specified 40 kg. Equation (1) was used to calculate the specified tank size.

Moving to the materials used, aluminum was selected for the car-sized-plate base of the robot, as well as the wheels for aluminum's ability to resist corrosion, since the robot will have allot of contact with corrosive elements such as water and the salty soil which would accelerate the speed of corrosion in metals. Furthermore, aluminum is lighter but fairly strong, making it easier to carry around. As for the movement of the robot, a motor with a power of 60 W is used. This motor gives a speed of 0.5 /s and has a load capacity of 40 kg/m³.

$$V = A_T h_T \quad \dots\dots (1)$$

where

V : Tank volume
 A_T : Tank cross section area
 h_T : Height of the tank

B. Sprinkler and Pump Specification

A sprinkler can be defined as a device, which sprays water. The choosing of sprinklers, as the design, depends on the application required. For example, there are types of sprinklers specified for firefighting, others for irrigation and each of type has several functionalities depending on the usage such as the area of coverage. To be able to choose a sprinkler, several calculations should be taken into consideration. Commencing with the velocity, the height of the tank needs to be determined in order to use equation (2) to calculate the velocity and gravity is also needed. Following that, the area of the nozzle from the tank to the sprinkle is to be determined using equation (3). After determining both the velocity and area, equation (4) is used to determine the volume flow rate, which is produced, finding one of the factors needed to choose a sprinkler. Another factor to be considered while choosing a sprinkler is the pressure supplied through the nozzle. To make the calculation simpler, equation (5) is used to calculate the pressure where g is the gravity and h is the height of the sprinkler above the base of the tank. The pressure is to be determined in kilo Pascals (kPa) and is then added to the Bernoulli equation (6) along with velocity in order to determine the h_a which is the power head of the pump, by which the pump needed can be chosen for the specified sprinkler.

$$V = \sqrt{2gh} \quad \dots\dots(2)$$

V : Velocity m/sec
 g : Gravity 9.81 m/sec
 h : head metres

$$A_n = \frac{\pi}{4} D_n^2 \quad \dots\dots(3)$$

A_n : Area of the nozzle
 D : Diameter of nozzle

$$Q_n = A_n V \quad \dots\dots(4)$$

Q_n : Flow rate at the nozzle section

$$P = \rho gh \quad \dots\dots(5)$$

P : Pressure in kPa
 ρ : Density of water 1000 kg/m³





Figure 1. Sprinkler types

The steps to follow to select a centrifugal pump are:
The flow rate needs to be determined. In order to do so, the largest water consumer needs to be determined, for example, in a house, the largest water consumer would be a bathtub due to its volume, requiring about 10 gallons per minute (gpm) to fill up which is equivalent to 0.6 liters per second (L/s). In other cases, such as production lines of plantation, the amount required would be much higher, around 10 times more, making it about 100 gpm or 6.3 L/s. After determining the flow rate, another factor must be determined as it is also vital in the selection of the pump, the static head, which depends on the height of the tank or in other cases, the distance between the uppermost point in a tank and the discharge pipe. In the case of this project, the total height of the tank is taken due to the fact that the sprinkler is placed above the tank and not below it. There are cases, in which the distance travelled by the fluid will be large enough to require the calculation of friction, which depend on the size and length of the pipe in which the fluid travels, as well as the flow rate. However, due to the short distance to be travelled by the fluid in this project, losses will be minimal and may be neglected. Finally, the selection of the pump will require the usage of the total head which is usually the sum of the friction and static heads, however, the friction head, as previously stated, will be neglected due to the very low losses in the distance travelled by the fluid and the flow required.

The amount of water dispensed is calculated so that it covers a certain radius as shown in Figure I. The pump pressure and the flow rate controls the area (radius) needs to be covered.

Table I. pressure and nozzle measurement for micro sprinkler

180° and 360° spray irrigation micro sprinkler						
Product name	Photo	Product code	Size	Qty/Bag(pcs)	Qty /Ctn (pcs)	Ctn size(cm)
360° sprinkler		ER-W48	360° sprinkler Working pressure: 1bar Flow rate: 120L/H Spray width: 30cm	1000	50000	54°35'28
180° sprinkler		DC-M22	180° sprinkler Working pressure: 1bar Flow rate: 90L/H Spray width: 30-50cm	1000	50000	54°35'28
						30

$$\frac{P_1}{\rho g} + \frac{V_1^2}{2g} + Z_1 + h_p = \frac{P_2}{\rho g} + \frac{V_2^2}{2g} + Z_2 + h_L \quad \dots\dots(6)$$

$$\dot{m} = \rho \cdot \dot{V} \quad \dots\dots(7)$$

\dot{m} : Mass flow rate of water in kg/s
 \dot{V} : Volume flow rate of water in m^3/s

$$\dot{W} = \dot{m} \cdot g \cdot h_v \dots\dots(8)$$

\dot{W} : Power of pump in Watt

h_v : Velocity head in m ($\frac{v^2}{2g}$)

C. Pump Performance

A pump is a device, which is used to move fluid by increasing the pressure through a suction method, increasing pressure and speed of a fluid. In order to choose the most appropriate pump, using the required or calculated flow rate and pressure head, the pump performance curve is used to select the needed pump. The pump chosen for this project is a centrifugal pump. To be more precise, an enclosed impellor type centrifugal pump, which is submerged in the tank. This type of pumps must be submerged in water as to disallow cavitation caused by sucking air. As for how it functions, a negative, low pressure, through priming, to allow suction of water, displaces the water on the suction side. The rotation of the impeller by the shaft; increases the pressure and kinetic energy of the fluid on the exit, increasing both velocity and pressure at the exit. Choosing a centrifugal pump, over other types of pumps, is beneficial in terms of having a steady and even flow, having a low initial cost and a high efficiency.

D. Solar Panels

Sustainability is one of the most important factors that must be considered in this project. Sustainability certifies to reduce cost, save the environment and raise social awareness, so after careful consideration of the several ways in which sustainability can be applied to this project, solar panels were found to be the most suitable.

A solar panel is a set of solar cells packed in order and connected together which absorb the solar energy and convert it into electrical power. There are several types of solar panels made up of different materials and sizes. Each different material has different characteristic and function, as well as different levels of solar absorption and electrical output. For example, monocrystalline solar panel, which is made of silicone, has about 20 % efficiency of converting that percentage of absorbed energy to electrical power which is considered as being of a high performance and efficiency for solar panels. Another type is thin film which is also silicon but has an amorphous cell structure and it has an efficiency of between 7 and 10% making it less efficient than the monocrystalline but it is allot cheaper to produce and is easier to produce as well as being flexible with weather changes. In this project, Monocrystalline solar panels were chosen because of it is more efficient than others. The size of the solar panel depends on the application as well as the type of solar panel.

For example, a villa would need its roof covered with thin film solar panels to produce electricity for the whole villa while it would require half the roof to produce electricity for the villa if monocrystalline type is used. Finally, the number and size of solar cells needed to operate all the machines (such as recharging the battery and acting as a power supply for the Arduino) can be determined using the following steps:

- Determine power consumption demands
 1. Calculate total Watt-hours per day for each appliance used:
 KWh needed from the solar panel array= daily KWh/Peak Sun-Hours
 2. Calculate total Watt-hours per day needed from the PV modules:
 Total watts needed from the solar panels= KWh needed from the solar panel array $\times 1000$
- Calculate the number of PV panels for the system:
 Number of panels= Total watts needed from the solar panels/ solar panel size selected

E. Robot Program

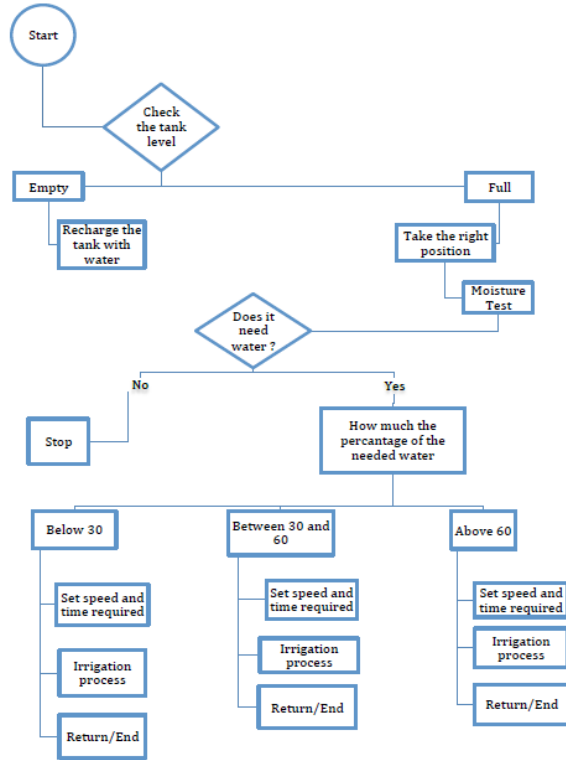


Figure 2. Robot program flowchart

The robot, no matter how fitting the materials used, design and power it has is of high quality, a robot is functionless without either having a previously installed command or being given a set of commands to carry out on the spot. Robots can receive commands in many forms, some come instantaneous like remote controlled robots which function with the press of a button and keep functioning as long as that button is pressed, while others would have a set of previously installed

commands and would be programmed to carry these commands out. The latter method would not require human interference for the robot to carry out the installed commands and would carry the required tasks out according to the programs installed in it. This project uses the programming method to make the robot function in the manner required. **The location coordinates of the field needs to be watered is entered (programmed) and the robot will move accordingly. The robot moves in straight line and then turns either left or right and keep moving in straight line and so on until all area of the field is watered. The next step in this project is to link the robot with a global positioning system (GPS) so that the robot receives the coordinates from satellite and moves accordingly.**

There are several programs, which can be used to program a robot such as MATLAB and Arduino. The program chosen is the Arduino due to its facility of usage and the availability of components such as the Arduino board and the sensors required. The first set of commands to be programmed is the movement of the robot. Following the movement of the robot, a robotic arm which has a temperature and moisture detector will have the command of measuring the moisture in the soil to check the required areas to be irrigated in order to reduce water wastage and to optimize the irrigation process. The sensors take analogue reading and those readings are then transformed through the Arduino board into a digital reading in order to fit with the digital commands programmed, allowing the robot to act according to the readings obtained. In addition, a sensor in the water tank is used to keep a constant check-up on the water level. Finally, based on the readings received from the sensors, the codes programmed will command the robot to activate the sprinklers where they are needed, thereby completing the irrigation process.

IV. RESULTS AND DISCUSSION

As previously stated, certain factors need to be considered and calculated in order to implement and determine certain factors of this robot design. This chapter encompasses the results and discussion of these results to facilitate the understanding of each calculation.




A. Tank size

In order to calculate the size of the tank, certain factors need to be kept in mind. The first factor is the mass. The robot can withstand 40 kg without having any issues, which will be the reference point for the tank calculation. Considering the other components that will be installed such as solar panels, motor, pump and such, 10 kg were reduced from the total allowable mass, leaving 30 kg of water to be used for irrigation. The base area of the tank was agreed up on to be the same as the plate of the robot base, which is 0.4 x 0.3 m, leaving the only factor needed to be calculated as the height. Rearranging equation (1), the height will be calculated through dividing the mass by density multiplied by the area, finding the height of the tank to be 13 cm or 0.13 m, having the dimension of the tank be 0.4 x 0.3 x 0.17 m.

B. Sprinkler

Unlike the case with the tank, the sprinkler can be determined with much simpler calculation, then choosing from the table II from which the pressure and volume flow rate required for the sprinkler were determined to be 1 bar and 90 L/H. The type of sprinkler used is 180 degrees. This is so that the electric components and the robot don't get damaged.

Table II: Sprinkler specification

PERFORMANCE FLOW RATE VS. PRESSURE						
	 360° – Spray Nozzle Color: Red, 0.075"		 180° – Spray Nozzle Color: Green, 0.060"		 90° – Spray Nozzle Color: Blue, 0.040"	
	Flow Rate (GPM)	RADIUS (ft)	Flow Rate (GPM)	RADIUS (ft)	Flow Rate (GPM)	RADIUS (ft)
15 PSI	0.45	3.5	0.31	3.5	0.15	3.5
20 PSI	0.52	4.0	0.36	4.0	0.18	4.0
25 PSI	0.57	4.5	0.41	4.5	0.21	4.5
30 PSI	0.63	5.0	0.45	5.0	0.23	5.0

C. Pump Performance

In order to choose the most suitable pump, the pressure head required must be calculated. In this project's case, the required pressure would be needed to supply sprinklers with water; hence, the calculations carried were based on the sprinklers used, to determine the most suitable pump. Commencing with calculating the velocity output of the sprinkler using the specified volume flow rate of the sprinkler, 90 liters per hour (2.5×10^{-5} m³/s), as well as the diameter, 3.5 mm, the velocity was determined to be 2.599 m/s (Appendix A, Sprinkler Calculation). There are other factors, which are known, such as the potential head, which is 0.31m high, taking the bottom of the tank as the datum line. In addition, the pressure output is almost atmospheric; therefore, the pressure was neglected. Following that, the losses in the pipe were determined, using Reynold's number formula, as well as the roughness and moody chart, combined with the velocity previously calculated, to determine the losses in the pipe, finally calculating the required pressure head, which was found to be 1.788 m, using Bernoulli's equation. After calculating the required pressured head, the power required was similarly calculated, finding a required power of 1.754 Watts.

D. Solar Panel

The solar panel is the easiest to calculate as it depends on the amount of hours at which the sun is hot in Kuwait multiplied by the Kilo-Watt-Hour Which is the power, giving the required total to select the number of solar panels, based on the desired size preference of the user.

E. Sensor Operations

a) Tank level sensor:

The tank level sensor is used to keep a measure on the amount of water inside the tank. This sensor will keep a reading on the level of water and send feedback to the

Arduino. The range set is based on the experimental values from the experiment carried to check how long it takes for the tank to be emptied with set pump and sprinklers, in addition to recording the values of when the tank is filled and when it is emptied, as the maximum and minimum values, emptied.

b) *Moisture sensor:*

The moisture sensor is a sensor that is used in this project to measure the moisture in the soil to determine how much water is present in the soil. The maximum or minimum values of the sensor were determined by an experiment which included using a sample of the soil the robot will be used on to determine the maximum and minimum water required for the grass.

The moisture sensor gets inserted in the soil by the robot arm to a depth equal to 2 cm. the soil moisture is measured at least at three locations and these data gets averaged and based on the average soil moisture, the robot either orders watering or not. An experiment was carried out to test the moisture sensor. The obtained experimental data is presented in Table III. When the soil is dry the readings of the soil moisture were high and as soon as the soil gets watered, the reading s of the soil moisture starts to decrease because of the surface tension that makes water stick to the soil particles. The robot is equipped with moisture sensor to determine soil moisture so that it gives a command to the robots to start watering if the value of the soil moisture is high. In the experiment, when the soil was dry the values of soil moisture were around 900 g/m³ and as soon as the soil was watered, the values went down to around 300g/m³

The robot was programmed to start watering if the values yielded by the moisture sensor were 500 g/m³ or above.

Soil moisture sensor readings (g/m ³)
900
970
950
935
925
390
375
325
300



Figure 3. Moisture sensor

V. STRESS ANALYSIS

Stress analysis was carried out to avoid failure in the structure (main body) of the robot. Specifically, it was carried out on the top plate body cover and on the beam holder.

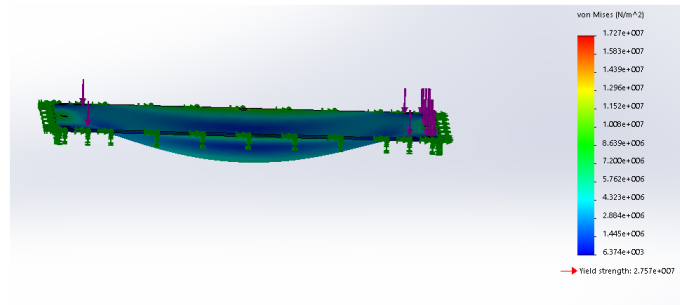


Figure 4. Stress Analysis top plate body cover

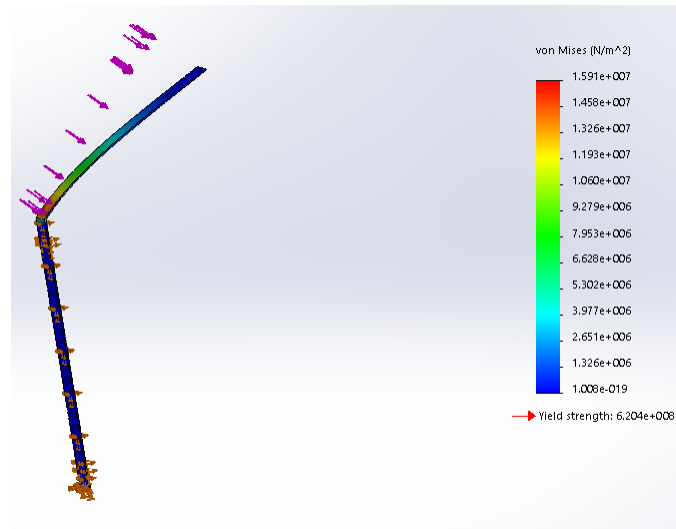


Figure 5. Stress Analysis beam holder

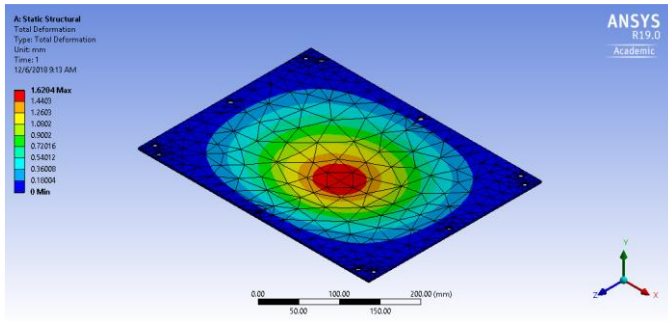


Figure 6. Stress Analysis deformation

Fig. 4 shows the stress analysis for the body plate which holds the tank and the system, as well as the beam holder (Fig. 5) which holds the solar panel. Fig. 6 shows the amount of deformation which is happening in the middle where the loads are concentrated.



Figure 7. Designed and tested irrigation robot

VI. CONCLUSION

In conclusion, this report summarizes the different aspects of irrigation systems, as well as robots and how this design has come to be. In addition, it journeys through the different methods, equations and calculations necessary to be followed in order to achieve this design, in a manner which is detailed to facilitate the creation of this product in the smoothest possible manner. Furthermore, a set of sample calculations, data and results were supplied in order to give a numerical example of the possible outcomes, though they may be changed in the future, however all the sample calculations were carried out with the correct set of equations needed, having the only possible variable in future changes, being the entries of numbers estimated to ensure the correctness and functionality of the above-suggested design. Finally, it is determined that the design is ready to be implemented into practical manner, based on the results obtained.

REFERENCES

- [1] ["Amenemhet III". *Britannica Concise*. Archived from the original on 2007-05-10. Retrieved 2007-01-10.
- [2] Ayumi Kawakami, Koji Tsukada, Keisuke Kambara and Itiro Sio, "PotPet: Pet-like Flowerpot Robot", *Tangible and Embedded Interaction 2011*, Pages 263-264 ACM New York, NY, USA, 2011.
- [3] Baek Seok-gi (1987). *Jang Yeong-sil 장영실*. Woongjin Wiin Jeon-gi Woongjin Publishing Co., Ltd.China – history. *Encyclopædia Britannica*, 1994 edition.
- [4] [B.C. Wolverton, Anne Johnson, and Keith Bounds, "Interior Landscape Plants for Indoor Air Pollution Abatement: Final Report", National Aeronautics and Space Administration (NASA-TM-101768) Science and Technology Laboratory, Stennis Space Center, 1989.
- [5] Constantinos Marios Angelopoulos, Sotiris Nikolettseas, Georgios Constantinos Theofanopoulos, "A Smart System for Garden Watering using Wireless Sensor Networks", *MobiWac '11 Proceedings of the 9th ACM international symposium on Mobility management and wireless access* Pages 167-170 ACM New York, NY, USA, 2011.
- [6] Currie, Adam (1999). "The History of Robotics". Archived from the original on 18 July 2006. Retrieved 10 September 2007 *Deborah Levine Gera (2003)*. Ancient Greek Ideas on Speech, Language, and Civilization
- [7] *De Silva, Sena (1998)*. "Reservoirs of Sri Lanka and their fisheries". *UN Food and Agriculture Organization*. Retrieved 2007-01-10.
- [8] E.J. Van Henten, J. Hemming, B.A.J. Van Tuijl, J.G. Kornet, J. Meuleman, J. Bontsema and E.A. Van Os; "An Autonomous Robot for Harvesting Cucumbers in Greenhouses"; *Autonomous Robots*; Volume 13 Issue 3, November 2002.
- [9] [9] Kevin Sikorski, "Thesis- A Robotic PlantCare System", University of Washington, Intel Research, 2003.
- [10] [10] Oxford English Dictionary. Retrieved November 27, 2016. Ivan Margolius, 'The Robot of Prague', Newsletter, The Friends of Czech Heritage no. 17, Autumn 2017, pp. 3 - 6. <https://czechfriends.net/images/RobotsMargoliusJul2017.pdf>
- [11] Pearce, Jeremy. "George C. Devol, Inventor of Robot Arm, Dies at 99", *The New York Times*, August 15, 2011. Retrieved February 7, 2012. "William Godwin (1876). "Lives of the Necromancers"
- [12] "Qanat Irrigation Systems and Homegardens (Iran)". *Globally Important Agriculture Heritage Systems*. *UN Food and Agriculture Organization*. Retrieved 2007-01-10.

Kalman Filtering and PID Control of an Inverted Pendulum Robot

Claire Bourque*, Andrew Lee*, Elyse Hill*, S. Andrew Gadsden*, Alex Bardelcik* and Mohammad Biglarbegian*

*Mechanical Engineering

University of Guelph, Guelph, Ontario, Canada

Emails: cbourque, alee32, ehill06, gadsden, abardelc, mbiglarb@uoguelph.ca

Abstract—The inverted pendulum is a classical controls theory problem that is unstable and nonlinear. In this paper, the state space equations for this system were derived and then linearized using small angle approximations. Using a PID controller with the Kalman and Unscented Kalman filters, the system was simulated in Matlab and then programmed into an inverted pendulum robot built for experimentation. The performance of these filters and controllers were then compared. The results found that for both the simulations and the actual implementation, the Kalman filter produced better state estimates and allowed the system to resist interference and further improve system stability.

I. INTRODUCTION

In this paper, the simulation and implementation of a PID controller and Kalman Filter in an inverted pendulum robot was performed. The purpose of this experiment was to control the inverted pendulum robot with the Kalman Filter and prove the Kalman Filter improves the system response of the robot. The system was first simulated using a PID controller, a Kalman Filter, and an Unscented Kalman Filter using Matlab. The mechanical system was first modeled using the Lagrange method and then the state space equations were used to simulate the system. To simulate the Kalman Filter, the system and measurement noise had to be found. The measurement noise could be found in the sensor manual while the system noise was approximated using the measurement noise. Next, using an Arduino, the Kalman filter was used to balance the movements of an Inverted Pendulum Robot. The sensors used to record the angle and speed of the robot were an ADXL345 accelerometer and a ITG 3200 gyroscope. The first sensor measured the acceleration of the system while the second measured the angular displacement and angular velocity. Using these sensors, the angular displacement of the robot was measured and controlled based on a setpoint of 0° from the vertical axis.

II. LITERATURE REVIEW

A. Inverted Pendulum Robot and PID Controller

An inverted pendulum is a classical controls problem that involves a nonlinear, unstable system with one input signal and several output signals. As such, PID controllers are often used to control inverted pendulum robots because they optimally model Single-Input-Single-Output (SISO) systems. In cases

where there are multiple inputs to control, a Multiple-Input-Multiple-Output system is implemented because one PID-controller is not enough [1]. PID control is commonly implemented on inverted pendulum systems and has been found to be simple, effective, and robust according to the works of Cole et. al [2], Sondhia et. al [1], and Wang et. Al [3]. The experimental set up used in this paper was based on the work done by Cole et. al [2], the team that built the robot implemented in this experiment and provided a summary on the construction and programming of the segway-bot using PID control. In conjunction with this, the tutorial provided by Arduino that provides an example of simple programming for PID control was used to refine the system.

B. Inverted Pendulum Robot and Kalman Filter

Several articles on the application of the Kalman filter (KF), extended Kalman filter (EKF), and unscented Kalman filter (UKF) to the control of an inverted pendulum were reviewed to understand the effect of filters on the system. The KF has been shown to improve control of an inverted pendulum through the reduction of noise and improvement of the robustness of the system [4][5]. Further, the filter tends to be applied to the angle measurements of the system rather than the position error [5]. The EKF also improves control and stability of the system, evidenced by its ability to achieve results quickly and accurately [6]. Unlike the KF, the EKF can be used to estimate states in a non-linear system [8]. During the prediction stage, the EKF uses the Jacobian of the non-linear state equations to calculate the a priori state estimates. [6]. However, linearizing the system can result in lost information, which is why the UKF is often chosen for more complex systems. By using only the non-linear equations, the UKF is able to improve final estimation results for inverted pendulums, making it an optimal filter for the inherently non-linear system [7].

III. MECHANICAL MODEL

In order to model the system, the state space model of the system was derived. This was accomplished by modeling the mechanical system using the Lagrange method. The parameters of the system are described in Table I. A free body diagram of a wheel and the body of the robot can be seen Figure 1. Using the free body diagram, the discrete state space equations were derived. The experimental setup used for this

TABLE I
LIST OF NOMENCLATURE

Symbol	Meaning
θ	Angle of rod with respect to vertical
φ	Angle of wheel with respect to vertical
z	Distance to center of gravity of rod
m_w	Mass of wheel
m_b	Mass of rod (body of robot)
r	Radius of wheel
g	Gravity
w	Thickness of robot body
J_1	Moment of inertia of the wheels
J_2	Coupled moment of inertia
J_3	Moment of inertia of the body of the robot
τ	Torque or input of the motor

experiment was built in-house, and is shown in Figure 2 to illustrate further the freed body diagram.

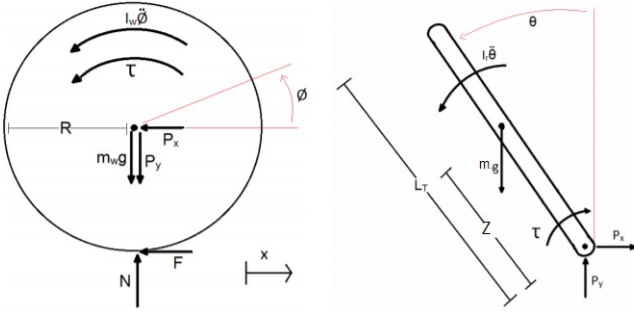


Fig. 1. Free Body Diagram of System

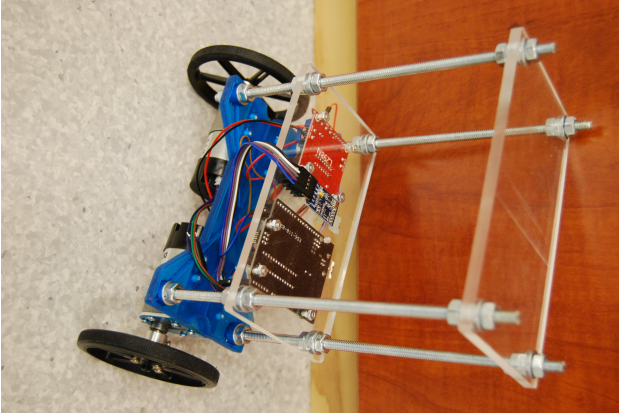


Fig. 2. Experimental Setup Used to Simulate an Inverted Pendulum

$$\frac{d}{dt} \begin{bmatrix} \varphi \\ \theta \\ \dot{\varphi} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & \frac{-J_2 g m_b z}{J_1 J_3 - J_2^2} T & 1 & 0 \\ 0 & \frac{J_1 g m_b z}{J_1 J_3 - J_2^2} T & 0 & 1 \end{bmatrix} \begin{bmatrix} \varphi \\ \theta \\ \dot{\varphi} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{J_3 + J_2}{J_1 J_3 - J_2^2} T \\ \frac{J_1 - J_2}{J_1 J_3 - J_2^2} T \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \tau \\ \tau \end{bmatrix}$$

$$C_s = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

In the above state space representation, J_1 , J_2 , and J_3 are represented by the following equations:

$$J_1 = (m_b + m_w)r^2 + I_w$$

$$J_2 = m_b r z$$

$$J_3 = m_b z^2 + I_b$$

where:

$$I_w = m_w r^2$$

$$I_b = \frac{1}{12} m_b (2 \times z + 2 \times w)^2$$

IV. SYSTEM SIMULATION

A. PID Controller

Following derivation of the state space equations, the system was simulated with a PID controller to determine if the state space equations accurately modeled the inverted pendulum system. The values found in Table II were used in the simulation and represent actual measurements taken from the inverted pendulum robot. However, it is important to note that this only compares the linearization of the model but not its accuracy. To evaluate the model properly, the simulated results should be compared with the experimental results. However, in this paper, the simulation was used to tune a PID controller initially before implementation on the experimental setup. The system

TABLE II
LIST OF PARAMETERS

Parameter	Value	Description
T	0.01	Period of system, in seconds
m_w	0.1	Mass of wheel, in kilograms
m_b	1.82	Mass of rod (body of robot), in kilograms
r	0.045	Radius of wheel, in meters
g	9.81	Gravity, in meters per seconds squared
w	0.09	Thickness of robot body, in meters
z	0.08	Distance to center of gravity of rod, in meters
τ	1	Torque or input of the motor, in newtons
t	0:T:5	Arbitrary time interval, in seconds
x	$[0 \ 0.1 \ 0 \ 0]'$	Initialization of x

was initially calibrated and tuned without the incorporation of noise in order to clearly model ideal conditions. Both the linear and nonlinear state space equations were used in the model in order to ensure that they approximated the same values. Initially, the model displayed a typical dynamic response, oscillating around the setpoint ($\theta = 0^\circ$) until it eventually achieves steady state after several oscillations. The gains were tuned to minimize steady state error (Table III), resulting in an ideal inverted pendulum response where the angle of the body decreases until a steady state value of $\theta \approx 0^\circ$ is achieved. The value of K_i is equivalent to zero in this case because the alterations to K_p and K_d already minimized the steady state error, meaning K_i was not needed. This position would be maintained for the duration of the simulation as no noise or

other destabilizing occurrences would be present (see Figure 3a).

Measurement and system noise were then added to the system to create a more realistic simulation (Figure 3b). These noise values were assumed to be Gaussian and white. Measurement noise stemmed from the errors produced through the position tracking provided by the gyroscope and accelerometer sensors implemented on the inverted pendulum robot. The values for measurement noise were found within the specification manuals for the sensors used for the inverted pendulum robot. System noise stemmed from any possible interference that could affect the system measurements. This includes things such as friction of the table and wheels, loose bolts in the apparatus, or unwanted vibrations in the system. These values were approximated since it is very difficult to completely understand the total amount of system noise the robot experiences. The measurement noise values were used to approximate the system noise values. This experiment assumed most of the noise would stem from the sensors, thus the system noise was approximated to several factors less than the measurement noise. The covariance matrices for system (Q) and measurement (R) noise can be found in Table IV.

TABLE III
GAIN VALUES

Gain	Value
K_p	1.1
K_i	0
K_d	0.2

TABLE IV
COVARIANCE MATRICES

System	Measurement
$Q = \begin{bmatrix} 6 \times 10^{-7} & 0 & 0 & 0 \\ 0 & 6 \times 10^{-7} & 0 & 0 \\ 0 & 0 & 6 \times 10^{-7} & 0 \\ 0 & 0 & 0 & 6 \times 10^{-7} \end{bmatrix}$	$R = \begin{bmatrix} 6 \times 10^{-4} & 0 & 0 & 0 \\ 0 & 6 \times 10^{-4} & 0 & 0 \\ 0 & 0 & 6 \times 10^{-4} & 0 \\ 0 & 0 & 0 & 6 \times 10^{-4} \end{bmatrix}$

B. Kalman Filters applied to PID Controller

In order to linearize the system model, small angle approximations were made. Both the linear and non-linear system models were simulated and the results were recorded over a ten second interval to confirm that both models behaved similarly for small angles. Thus, the linearized model was assumed to be the true model and was used to calculate the *a priori* state estimates and *a priori* state covariance in the Kalman Filter. Using the non-linear equations in the EKF would not significantly improve the accuracy of the state estimates for small angles. Since the physical system is only expected to operate within small angles, the EKF was not implemented. If this assumption was not made, then the EKF would have to be applied in place of the KF since when large angles are encountered, the system would behave nonlinearly. This issue did not apply to the UKF, which uses the nonlinear equations

without linearizing them. The discretized, nonlinear state space equations can be found below:

$$\frac{d}{dt} \begin{bmatrix} \varphi \\ \theta \\ \dot{\varphi} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \varphi T & 0 \\ 0 & 1 & 0 & \theta T \\ 0 & \frac{J_2 \sin(\theta) \cos \theta g m z + J_2 \dot{\theta}^2}{(J_2 \cos(\theta))^2 - J_1 J_3} T & 1 & \frac{J_2 J_3 \sin(\theta) \dot{\theta}^2}{(J_2 \cos(\theta))^2 - J_1 J_3} T \\ 0 & \frac{\sin(\theta) J_1 m_b z + (1 + J_2^2 g \cos(\theta) \sin(\theta) \dot{\theta}^2)}{(J_2 \cos(\theta))^2 - J_1 J_3} T & 0 & \frac{1 + J_2^2 g \sin(\theta) \cos(\theta) \dot{\theta}^2}{(J_2 \cos(\theta))^2 - J_1 J_3} T + 1 \end{bmatrix} \begin{bmatrix} \varphi \\ \theta \\ \dot{\varphi} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{J_3 + J_2 \cos(\theta)}{J_1 J_3 - J_2^2} \\ \frac{-J_1 - J_2 \cos(\theta)}{J_1 J_3 - J_2^2} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \tau \\ \tau \end{bmatrix}$$

$$C_s = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Figure 4 displays the filters applied to the system and Table V displays the root-mean squared errors of the filters. Application of the KF revealed that the filter was able to accurately track the system, never deviating very much from the true approximations. The UKF varied in response when applied. While on some trials it would appear to track the system perfectly, on others its values would explode, resulting in incorrect approximations. From the reviewed theory, the UKF should improve the system approximation. This suggests that there is an issue in the code used to approximate with the UKF. However, given the linear approximation made on the system, it is reasonable that the KF would be the best estimator since it is optimal for linear systems. Since the

TABLE V
RMSE VALUES

State	KF (rad^2)	UKF (rad^2)
θ	2.44×10^{-3}	2.53
φ	5.22×10^{-4}	1.55×10^{-3}
$\dot{\theta}$	1.65×10^{-3}	3.94×10^{-1}
$\dot{\varphi}$	1.86×10^{-2}	2.58×10^{-2}

implementation on the real system involves improved body angle control by applying the KF, a final simulation comparing the measurement data was also conducted. This can be seen in Figure 5. The efficacy of the Kalman filter was measured by tracking body body angle error of the PID controller where the setpoint was 0° . In one scenario, the sensor measurement was used to calculate the body angle error of the PID controller. The other scenario used the Kalman estimate to calculate the the body angle error. The RMSE value of the PID error decreases, as seen in Table VI, illustrating that applying the KF to the inverted pendulum system should produce a better response. The UKF's response when applied in both cases was unexpected. Because it is unstable, it makes sense that it is unable to approximate the measurements in Fig. 5a. When

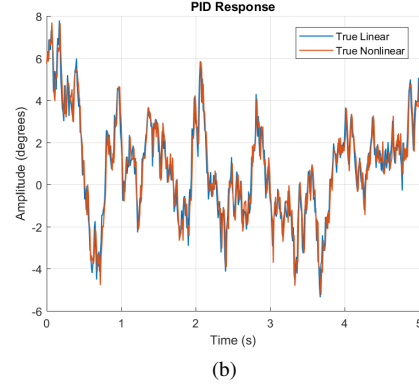
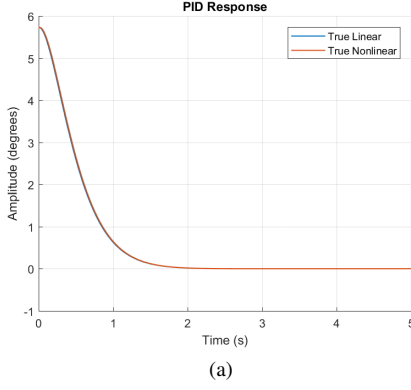


Fig. 3. Initial Simulation Results: (a) Ideal PID Response (without noise), (b) Ideal PID Response (with noise).

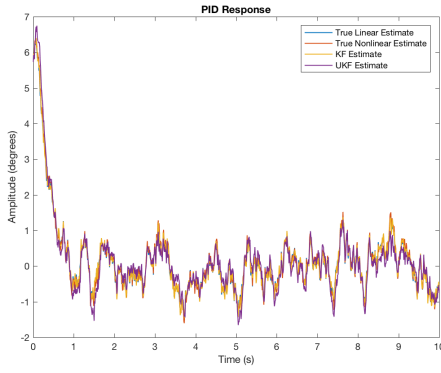


Fig. 4. System with Filters Applied

the KF was used in the PID controller (Fig. 5b), the UKF was able to produce more accurate estimates. However, it is currently unknown why it has this effect.

TABLE VI
RMSE VALUES FOR MEASUREMENT COMPARISON

Without Kalman Error	$4.78 \times 10^{-2^\circ}$
With Kalman Error	$3.66 \times 10^{-2^\circ}$

V. SYSTEM IMPLEMENTATION

A. System Model Accuracy vs. Kalman Filter Efficacy

In order to tune and implement the KF into the inverted pendulum robot, the value of variables in the initial state space equations were changed. Eq. 1 represents the system matrix constant (F) which multiplied the angular position of the body of the robot. From the initial measurements in Table II, this factor was equal to 75.6. Eq. 2 represents the input matrix constant (B) and has a value of 0.0015. Initial simulation of the system model with these values was problematic, resulting in an inaccurate system model that poorly balanced. After implementing this model in Arduino, it was found that the

errors occurring in the system were the result of using incorrect measurement values.

$$F = \frac{J_1 g m_b z}{J_1 J_3 - J_2^2} \quad (1)$$

$$B = \frac{J_1 - J_2}{J_1 J_3 - J_2^2} \quad (2)$$

Figure 6a uses the incorrect system model. When compared with the true states, the incorrect system model produces estimates with high error. Aforementioned, the initial values in Table II ranged from being measured and approximated. This caused some inaccuracies in the model constants F and B that adversely affected the system. The inaccurate estimates caused the PID controller to produce the wrong torque needed to control the system, resulting in an unstable system that does not balance. The instability of the system indicates the importance of using proper parameter measurement values in the system model. The inertia of the pendulum system as well as the effect of torque on that system are vital to its control. Thus, accurate models are necessary to ensure the system is stable. The system model variables were then adjusted through trial and error until the correct system response was achieved. Figure 6b illustrates the system response with the adjusted system model. Using values which accurately model the system produces more accurate body angle errors for the PID controller.

B. Kalman Filter Process

As described, the KF is a predictor-corrector estimation strategy that yields optimal estimates to linear systems and measurements. A nonlinear form of the KF is known as the extended KF, and essentially first-order Taylor series approximations are used to create Jacobians or linear matrices. The process is similar between the KF and EKF. The KF process is described as follows. The first two equations below represent the prediction stage and predict the state estimates and covariances, respectively.

$$\hat{x}_{k+1|k} = F\hat{x}_{k|k} + Bu_k \quad (3)$$

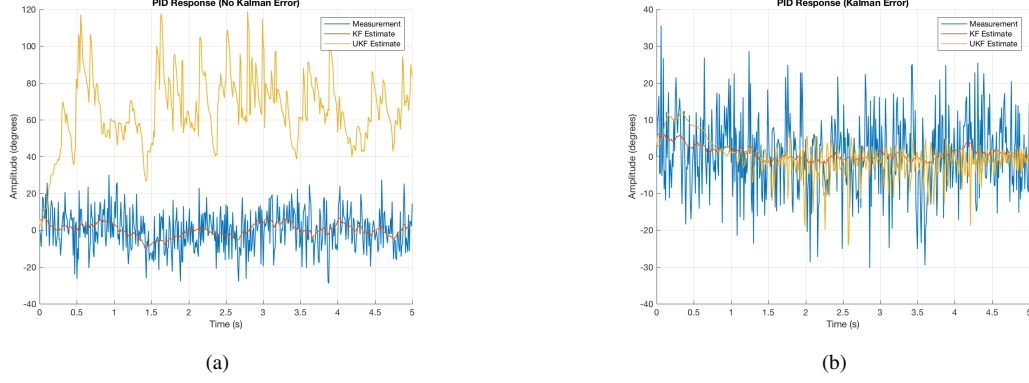


Fig. 5. Simulation Comparison to Measurement Values: (a) without Kalman Error; (b) with Kalman Error.

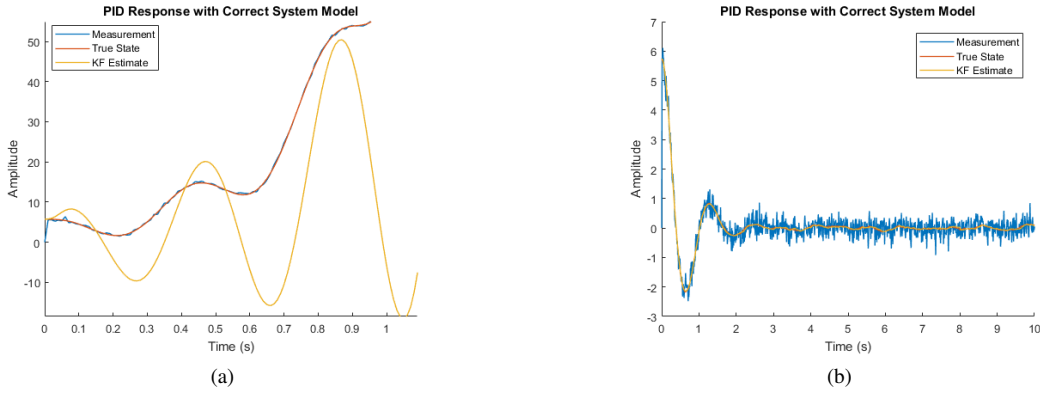


Fig. 6. Simulated System Models: (a) with Poor System Measurements; (b) with Adjusted System Measurements.

$$P_{k+1|k} = F P_{k|k} F^T + Q_k \quad (4)$$

The next section is referred to as the update stage. The following equations are used iteratively with the prediction stage to calculate the updated state estimates and covariances, respectively. The first equation is used to calculate an innovation covariance used to generate the Kalman gain. The H term refers to the measurement matrix. Note that the updated state estimates are used by the PID controller to generate the control signal based on the error (difference between the desired and estimated states).

$$S_{k+1} = H P_{k+1|k} H^T + R_k \quad (5)$$

$$K_{k+1} = P_{k+1|k} H^T S_{k+1}^{-1} \quad (6)$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}(z_{k+1} - H\hat{x}_{k+1|k}) \quad (7)$$

$$P_{k+1|k+1} = (I - K_{k+1}H)P_{k+1|k} \quad (8)$$

C. Kalman Filter System Tuning

The Kalman filter was initially tuned in the Arduino assuming a zero measurement covariance. The system constants were tuned until the Kalman estimates matched the complementary filter estimates. Initially, F was assumed to be 75.6 based

on measurements of the physical system and B was assumed to be 0.015 based on the simulation. These system constants overestimated the effect of the torque on the body angle. Thus, the Kalman estimates produced much higher angles than the complementary filter as seen in Figure 7a. F was tuned for the KF to a value of 61 while B was tuned to a value of 0.001. After the model was corrected, the KF estimates matched the complementary filter estimates (Figure 7b).

D. Covariance Matrix

After tuning the system matrices, the covariance matrices were tuned to further improve the estimates. The measurement covariance matrix is not constant for all angles and thus needed to be calculated for different angles. The calculated values were taken in 5° increments and can be found in Table VII. At small angles ($< 10^\circ$), the covariance values were relatively similar in magnitude. At angles $\geq 10^\circ$, however, the magnitude increases substantially.

E. PID Tuning

The PID controller was retuned for the inverted pendulum robot. The gains for the implementation were tuned in the same order as in the Matlab simulations: The proportional gain was tuned first while setting all other gains to zero.

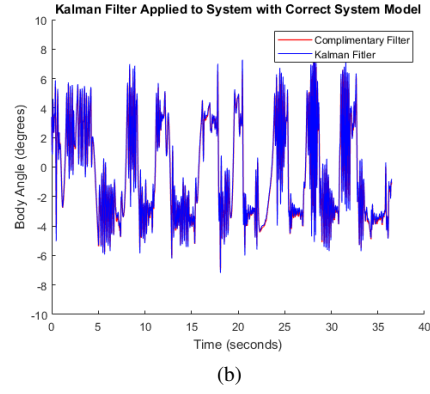
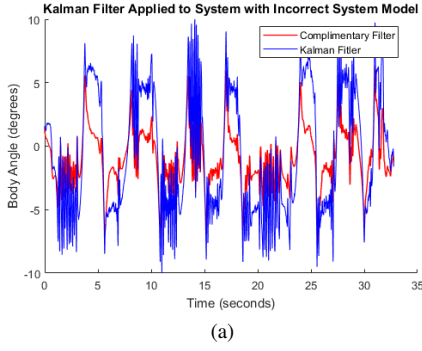


Fig. 7. Implemented System Models: (a) with Poor System Measurements; (b) with Adjusted System Measurements.

TABLE VII
COVARIANCE VALUES

Degrees	Covariance
0	$R = 4.7801 \times 10^{-4}$
5	$R = 5.7013 \times 10^{-4}$
10	$R = 1.1378 \times 10^{-3}$

Then, the derivative gain was slowly adjusted to smooth the oscillations. Finally, integral gains were omitted because the system is unstable and never expected to settle completely with 0 steady state error. The PID was tuned initially on the complementary filter, which combines sensor data from both the accelerometer and gyroscope. Listing 1 displays the system update equation can be seen for the complementary filter. Using both the sensor data from the accelerometer and gyroscope, a full picture of how the system reacted while a controller or filter was maintaining its upright position can be seen. The KF was then applied and also had to be tuned until the parameter values in Table VIII were achieved. These parameters created a stable system. The variables' greatest change in magnitude came from tuning the system noise and the value for the input matrix constant.

Listing 1. Code Equation for Complementary Filter

```
C1= 0.98
C2 = 0.02
Complementary(k)= C1*(Complementary(k-1)
+ gyroscope(k)*T) + C2*(accelerometer(k));
```

VI. EXPERIMENTAL RESULTS

The PID controller was implemented using the complimentary filter estimates as input followed by using the Kalman filter estimates as the input (represented in Figures 8a and 8b, respectfully). The initial body angle offset was set to approximately 4° with a setpoint of 0° which represents a perfectly upright robot. The PID controller was allowed to run until the inverted pendulum robot reached steady state. The

TABLE VIII
INITIAL AND TUNED PARAMETERS

Initial Parameters	Tuned Parameters
$Q = \begin{bmatrix} 6 \times 10^{-7} & 0 \\ 0 & 6 \times 10^{-7} \end{bmatrix}$	$Q = \begin{bmatrix} 2 \times 10^{-7} & 0 \\ 0 & 2 \times 10^{-7} \end{bmatrix}$
$R = 6 \times 10^{-4}$	$R = 2 \times 10^{-5}$
System Constant = 75.6	System Constant = 61
$B = 0.015$	$B = 0.001$
$K_p = 70$	$K_p = 12$
$K_d = 0.25$	$K_d = 0.1$

RMSE values of both trials were calculated and are displayed in Table IX.

TABLE IX
PID CONTROLLER RMSE VALUES

PID Controller Input	RMSE of Error from Setpoint
Complimentary Filter Estimate	68.4210°
Kalman Filter Estimate	65.0810°

VII. DISCUSSION

A. Experimental Analysis

The RMSE of the error from the setpoint was 68.412° and 65.081° for the complimentary filter and Kalman filter respectively. In addition, the complimentary filter estimates produced a steady error of approximately 0.15° while the Kalman filter had virtually no steady state error. Finally, the settling time using the complimentary filter was approximately 7.8 seconds while the settling time using the Kalman filter was approximately 4.5 seconds. The Kalman filter appears to offer an improvement, over the complimentary filter, as predicted in the simulated results. However, the results are very comparable when considering that there was high variability in the experimental setup. Some of the main contributors of variability were the power and USB cables. The cables produced an uncertain amount of positive torque and damping when they dragged behind the robot. The Kalman filter estimates were

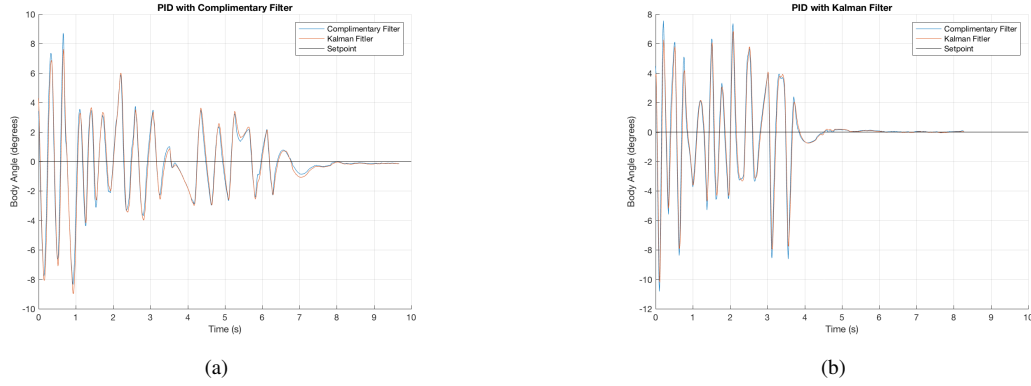


Fig. 8. PID Controller Results: (a) Using Complimentary Filter Estimate as Input; (b) Using Kalman Filter Estimate as Input.

relatively consistent with the complimentary filter estimates and offered only marginal improvement when used in the PID controller as shown in Fig. 8. Thus it is reasonable to expect that the Kalman filter would perform comparably with the complimentary filter. While the Kalman filter is able to smooth out some of the noise, it appears to lag slightly behind the complimentary filter during rapid oscillations. Thus, when the angular velocity is positive, the Kalman estimates tend to be lower than the complimentary filter estimates. The opposite is true for negative angular velocities. The PID response shows several oscillations before finally being able to settle close to the setpoint. The magnitude of the oscillations do not monotonically decrease as expected in a second order system model. There are several contributing factors for the additional oscillations. The motor driver was not consistent in its output and often failed to output the correct torque based on the setpoint error. In addition, there was significant backlash in the motor-wheel system which was not modeled into the controller or system matrix of the Kalman filter. The backlash would often cause several degrees of movement before torque was actually applied to the system.

B. System Problems

There were several physical limitations with the system that greatly impacted the results. One issue was that the wires connected to the power supply created a great deal of interference with the system. In ideal situations, the controller or filter should be the only stabilizing factor on the robot. In practice, however, as the robot travelled away from the power supply, the connecting wires would act like a spring and become tense before pulling the robot backwards, causing a reaction force on the robot that artificially controlled its movements. Another issue stemmed from the motors. There were discrepancies with the performance between the individual motors. The motor response of the left motor was lacking compared to the right motor. The left motor was occasionally unresponsive which made the robot travel in circles instead of in a forward and backward motion. The unresponsiveness could be due to the amount of current entering the right motor compared to the left motor. Another problem with controlling the robot was that the

motor could not output small values of torque to correct small body angle errors. This is because the DC motors required a certain threshold of voltage in order to operate. For small angle errors, the PID controller would output a small voltage to correct the body angle. However, the motors outputted zero torque because the voltage was below the required threshold. Thus, the motors would not activate until a certain body angle error was achieved. One final issue with the robot was its sensitivity to disturbances. This was due to the robot's physical design which caused it to have a small moment of inertia and a low center of gravity. To improve the controllability of the robot, more weight should be added higher up on the body of the robot.

C. Future Work

Several additions can be made to the inverted pendulum robot to improve performance. One addition would involve incorporating more controllers, such as for the position or the yaw of the system. A linear-quadratic regulator controller could also be added in place of a PID controller to alter system performance. Additionally, the estimation techniques need to be refined as well. The UKF needs to be corrected in order to correctly approximate the system. Fixing the UKF will allow for incorporation of other estimation techniques, such as the interacting multiple model, that could produce more refined results. Finally, several physical aspects of the system should be adjusted based on the limitations mentioned in the previous section. This would include making the robot powered by a battery to allow the wires to be removed and redesigning the system to increase its mass and its the moment of inertia

VIII. CONCLUSIONS

Experimentation showed that implementation of a Kalman filter with a tuned PID controller improved the response of an inverted pendulum robot. The simulated and implemented results also further demonstrated that an accurate system model is important to good overall system performance. If any estimates are inaccurate, the system model will not correctly correspond with the actual system, resulting in inaccuracies and possibly instabilities. Additional experimentation should

examine effects of different controllers and estimators in the system. In future experiments, several modifications can be made to the inverted pendulum robot in order to improve performance such as removing the wires from the system, correcting the unresponsive motor, and increasing the inertia of the system.

REFERENCES

- [1] S. Sondhia, R. R. Pillai, S. S. Hedge, S. Chakole and V. Vora, "*Development of Self Balancing Robot with PID Control*," International Journal of Robotics, vol. 7, no. 1, pp. 1-6, 2017.
- [2] N. Cole, A. Cataford, S. Hockin, A. Gaiero and S. Rossi, "*Self Balancing Bot*," ENGG*4480 W17, pp. 1-28, 2017.
- [3] J. Wang, "*Simulation studies of inverted pendulum based on PID controllers*," Simulation Modeling Practice and Theory, vol. 19, no. 1, pp. 440-449, 2011.
- [4] W. Zhang, G. Li and W. Lirong, "*Research on the Control Method of Inverted Pendulum Based on Kalman Filter*," in 2014 IEEE 12th International Conference on Dependable, Autonomic and Secure Computing, Dalian, 2014.
- [5] H. Lee and S. Jung, "*Balancing and navigation control of a mobile inverted pendulum robot using sensor fusion of low cost sensors*," Mechatronics, vol. 22, no. 1, pp. 95-105, 2012.
- [6] S. Hoseini and J. Poshtan, "*Fault tolerant control applied on an inverted pendulum by using extended Kalman filter*," in 2007 International Conference on Intelligent and Advanced Systems, Kuala Lumpur 2007.
- [7] M. Zheng, K. Ikeda and T. Shimomura, "*Parameter Estimation of Rotary Inverted Pendulum based on Unscented Kalman Filter*," in 2006 SICE-ICASE International Joint Conference, Tokushima, 2006.
- [8] S. Bonadies, N. Smith, N. Niewoehner, A. S. Lee, A. M. Lefcourt, and S. A. Gadsden, "*Development of Proportional-Integral-Derivative and Fuzzy Control Strategies for Navigation in Agricultural Environments*," ASME Journal of Dynamic Systems, Measurement, and Control, Vol. 140, No. 6, 2017.
- [9] J. Kim, K. Chang, B. Schwarz, A. S. Lee, M. Al-Shabi, and S. A. Gadsden, "*Dynamic Model and Motion Control of a Robotic Manipulator*," Journal of Robotics, Networking and Artificial Life, Vol. 4, No. 2, pages 138-141, 2017.
- [10] S. A. Gadsden, S. R. Habibi, and T. Kirubarajan, "*Kalman and Smooth Variable Structure Filters for Robust Estimation*," IEEE Transactions on Aerospace and Electronic Systems, Vol. 50, No. 2, pages 1038-1050, 2014.
- [11] S. A. Gadsden and A. S. Lee, "*Advances of the Smooth Variable Structure Filter: Square-Root and Two-Pass Formulations*," Journal of Applied Remote Sensing, Vol. 11, No. 1, 2017.

Solid Mechanics

NON-FOURIER HEAT CONDUCTION AND HEAT WAVE INTERFERENCE IN CNT REINFORCED COMPOSITES

Aminallah Pourasghar

Department of Mechanical Engineering
University of Alberta
Edmonton, Canada
pourasgh@ualberta.ca

Zengtao Chen

Department of Mechanical Engineering
University of Alberta
Edmonton, Canada
Zengtao.chen@ualberta.ca

ABSTRACT

We present a framework to investigate the effects of transient heat conduction on heat wave interference and the overshooting phenomenon in carbon nanotube (CNT)-reinforced composites. Different kinds of heat conduction models such as Fourier, hyperbolic or single-phase-lag (SPL) and dual-phase lag (DPL) models are considered while the media is subjected to a heat pulse or periodic thermal shock. This paper contributes to the existing literature in three aspects: (1) using the differential quadrature method (DQM) to solve the nonlinear DPL heat conduction equations, while the materials and properties are geometry- and temperature-dependent; (2) showing the effects of the volume fraction of CNTs and their distribution on the transient heat conduction and heat wave interference; (3) demonstrating heat wave interference and the high-temperature gradient as the origins of the overshooting phenomenon. In the numerical algorithm based on the DQM, the temporal domain is marched block by block for the time history of temperature. Each block contains several time levels, and the numerical results at these time levels are obtained simultaneously. Thus, the numerical solution at the $(n+1)^{\text{th}}$ time level depends on the solutions at its previous levels all the way from the 1st to the n -th levels. The accuracy of the present solution is confirmed by comparing it with available results from the literature. A detailed numerical study is conducted to illustrate the effects of heat wave speed and heat flux, carbon nanotube volume fraction and heat wave interference on the time variation of temperature and heat flux.

A New Anisotropic Elastic Metamaterial

Xiaodong Wang
Department of Mechanical Engineering
University of Alberta, Edmonton, Alberta, Canada T6G 1H9
Xiaodong.wang@ualberta.ca

This talk presents a new anisotropic elastic metamaterial, which consists of a series of properly arranged periodic representative cells. The unique structural feature of its representative cell endows the model with special wave filtering abilities, broad bandgaps in the low frequency ranges and isolation of transverse waves. Numerical simulations are conducted to investigate wave propagation in this metamaterial and the effect of the component material properties on its dynamic behaviour.

Effect of nonlinear material viscosity on the dynamic performance of dielectric elastomer oscillators

Yuanping Li, Liying Jiang

Department of mechanical and Materials Engineering

Western University

London, Ontario, Canada N6A 5B9

ljjiang55@uwo.ca

ABSTRACT

As a category of soft electroactive materials, dielectric elastomers (DEs) are capable of producing large voltage-induced deformation, which makes them desirable materials for the application of tunable oscillators and resonators. Nevertheless, the dynamic performance of such DE-based vibrational devices are not only strongly affected by external loads but also significantly affected by the material viscoelasticity of DEs. By adopting the state-of-art modeling framework of finite-deformation viscoelasticity theory, and theory of polymer dynamics, this work first investigates the effect of deformation-dependent viscosity (nonlinear viscosity) on the frequency tuning process of a viscoelastic DE membrane oscillator. A comparison between the effect of the nonlinear and linear material viscosity on the frequency tuning process under AC voltage is presented. Moreover, the vibrational behavior of the oscillator under harmonic excitation is also considered. From the simulation results, it is concluded that not only the inelastic deformation changes the resonant frequency of the oscillator but also the relaxation time significantly affects the frequency tuning. When the oscillator is subjected to AC or DC voltage, the nonlinear material viscosity strongly influences the frequency tuning process, especially for its transient response. The modeling framework and simulation results in this work are expected to provide guidelines for the optimal design of dielectric elastomer oscillators and resonators.

Manipulating the architecture of beams for high toughness and strength

Ahmed Dalaq, Francois Barthelat

Department of Mechanical Engineering, McGill University
817 Sherbrooke Street West,
Montreal, Quebec H3A 2K6, Canada
ahmed.dalaq@mail.mcgill.ca

ABSTRACT

Strength and toughness are usually mutually exclusive properties in engineering materials, but recent work has demonstrated that this limitation can be overcome by segmenting monolithic materials into blocks of well-defined geometries and arrangements. The “architected” systems offer a wealth of design parameters which can be manipulated with precision to induce sliding, rotation, and interlocking of unit blocks. In this work we present models and experiments for the flexural response of linearly segmented architected beams. We use finite element (FE) model to explore deformation and failure modes of simple beams made of confined cubes, and to predict and optimize overall strength and toughness as function of number of blocks, blocks geometry and coefficient of friction between the blocks. The geometry of the blocks was then enriched systematically using different degrees of polynomials. We found that most of the geometric enrichments improve toughness, by improving the mechanisms of sliding and jamming at the interfaces between individual blocks. In particular, toughness and strength are proportional to the average curvature of the block faces except for shapes that fall under minimal surfaces. We finally fabricated and tested architected beams made of glass, with the best identified design from the modeling section. The architected glass was six times tougher than the monolithic glass (reaching up to 111 times in terms of work of fracture) and preserved about 60% of its strength. Material architecture and segmentation provide new approaches to designing materials and structures. Here we show that architecture can be manipulated with precision to obtain optimum combinations of strength and toughness.

Investigating the accuracy of various many-body potentials in estimating mechanical properties of a nanoscale Aluminum

Sabir Subedi

Graduate Student,
 Dept. of Mechanical Engineering
 Memorial University of Newfoundland
 sabirs@mun.ca

Sam Nakhla

Assistant Professor
 Mechanical Engineering, Faculty of Engineering
 Emergency Medicine, Faculty of Medicine
 Memorial University of Newfoundland

ABSTRACT

Molecular dynamics was used to estimate the mechanical properties of nanoscale aluminum from an atomistic uniaxial tension test. The simulation procedures followed in current research work are proven efficient for a nanoscale iron [1]. Previous research [1-3] has tested mechanical properties using molecular dynamics for various materials such as Iron, Molybdenum, Silicon, Titanium, Zirconium, Magnesium. Results from these studies agreed that the potential could significantly affect the predicted results. However, there is no comparison for the accuracy of various many-body potentials for nanoscale Aluminum. Hence, the objective of the present study was to evaluate and compare accuracies amongst the various many-body potentials in estimating the mechanical properties of nanoscale Aluminum. The various many-body potentials used in this study are available in NIST [4] interatomic potentials repository and suitable for use in Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [5]. Overall, the atomistic tension test was carried out for nine Embedded atom model (EAM), three Modified EAM (MEAM) and two Reactive Force Field (ReaxFF) potentials. A sample set of results obtained from the simulation and the comparison with experimental data to determine the most accurate potential are provided in table 1. Remaining set of data and results are to be presented later. From table 1, EAM potentials predicted the most accurate results. In contrast, the MEAM and ReaxFF potentials significantly over and under predicted the elastic modulus and Poisson ratio at room temperature. Therefore, these findings demonstrate the importance in the investigation the accuracy of the given potential before applying it to simulations outside of its initial purpose of parameterization.

Table 1: Sample data for the comparison of the mechanical properties of highly pure Aluminum with Literature [19] for various many-body potentials.

Potentials name		Poisson Ratio		Elastic Modulus (GPa)	
		Present Study	Error relative to pure Aluminum ($\nu = 0.35$) [9]	Present Study	Error relative to pure Aluminum ($E = 66.6$ GPa) [9]
EAM	Sheng <i>et. al.</i> [6]	0.345	-1.43	65.41	-1.79
MEAM	Pascuet <i>et. al.</i> [7]	0.349	-0.288	58.45	-12.23
ReaxFF	Hong <i>et. al.</i> [8]	0.407	16.28	40.57	-39.07

Keywords: Molecular dynamics simulation, potentials, Modulus, Poisson Ratio

References:

- [1] L. S. Morrissey, S. M. Handrigan, S. Subedi, & S. Nakhla, "Atomistic uniaxial tension tests: investigating various many-body potentials for their ability to produce accurate stress-strain curves using molecular dynamics simulations," *Molecular Simulation*, pp. 1-8, December 2018.
- [2] S. A. Etesami, M. Laradji, & E. Asadi, "Interatomic potentials transferability in predicting temperature dependency of elastic constants for titanium, zirconium and magnesium," *Modelling and Simulation in Materials Science and Engineering*, vol. 27, January 2019.
- [3] Y. V. Lysogorskiy, T. Hammerschmidt, J. Janssen, J. Neugebauer & R. Drautz, "Transferability of interatomic potentials for molybdenum and silicon" *Modelling and Simulation in Materials Science and Engineering*, vol. 27, January 2019.
- [4] Interatomic Potentials Repository, <https://www.ctcms.nist.gov/potentials/system/Al/>, (Accessed on 13th February, 2018)
- [5] LAMMPS Molecular Dynamics Simulator, <https://lammps.sandia.gov/>, (Accessed on 13th February, 2018)
- [6] H. W. Sheng, M. J. Kramer, A. Cadieu, T. Fujita & M. W. Chen, "Highly optimized embedded-atom-method potentials for fourteen fcc metals," *Physical Review B*, vol. 83, pp.134118, April 2011.
- [7] M. I. Pascuet & J. R. Fernández, "Atomic interaction of the MEAM type for the study of intermetallics in the Al-U alloy," *Journal of Nuclear Materials*, vol. 467, pp. 229-239, December 2015.
- [8] S. Hong & A. C. Van Duin, "Atomistic-scale analysis of carbon coating and its effect on the oxidation of aluminum nanoparticles by reaxff-molecular dynamics simulations," *The Journal of Physical Chemistry C*, vol. 120, pp. 9464-9474, April 2016.
- [9] L. F. Mondolfo (2013). *Aluminum alloys: structure and properties*. Elsevier. (Accessed on 13th February, 2018).

Investigating the effect of porosity and temperature on modulus of nanoscale Aluminium

Sabir Subedi
 Graduate Student

Dept. of Mechanical Engineering
 Memorial University of Newfoundland
 sabirs@mun.ca

Sam Nakhla
 Assistant Professor

Mechanical Engineering, Faculty of Engineering
 Emergency Medicine, Faculty of Medicine
 Memorial University of Newfoundland

ABSTRACT

There are many experimental studies which have [1-6] evaluated the effect of porosity on elastic modulus for various metals. However, many of these studies do not actually perform the uniaxial tension tests that are commonly done on the macroscale [7]. Moreover, there is limited research on the mechanical properties of Aluminum from molecular dynamics. As such, the objective of the current research was to simulate the effect of porosity on the modulus of FCC structured nanoscale Aluminum. Previous work on this topic demonstrated that the EAM potentials, Mishin *et. al.* [8] and Sheng *et. al.* [9], were the most accurate potential in estimating the modulus of nanoscale Aluminum. Therefore, in the current study these interatomic potentials are used to simulate a molecular dynamics tensile test under uniaxial loading for a nanoscale Aluminum with porosity to estimate the modulus at room temperature. Porosities ranging from 1-5% were studied. Results showed that the modulus of the aluminum is inversely proportional to porosity percentage. Similarly, the effect of temperature in modulus of nanoscale Aluminum was investigated. A uniaxial tensile test was simulated in MD using the most accurate EAM [8-9] potentials at different temperatures (0-900 K) for a nanoscale Aluminum. Like effect of porosity, the modulus of the Aluminum decreases with increase in temperature. Also, a comparison of MD predictions of elastic modulus with increasing temperature was made to experimental data found in literature. Furthermore, the combined effect of the porosity and temperature were investigated for nanoscale Aluminum. Results demonstrated the reduction in the material's modulus due to porosity at different temperature for best accurate EAM potentials, i.e. Mishin *et. al.* [8] and Sheng *et. al.* [9].

In this abstract, a sample set of results, figure 1, is provided which showed the reduction in modulus increased from 10% to 25% with increasing porosity. Remaining set of data and results are to be presented later. In the present study, dislocation and grain boundary were not taken in consideration for a single crystal of Aluminum. Hence, results may be over predicting when compared to bulk polycrystalline experimental data [1-6].

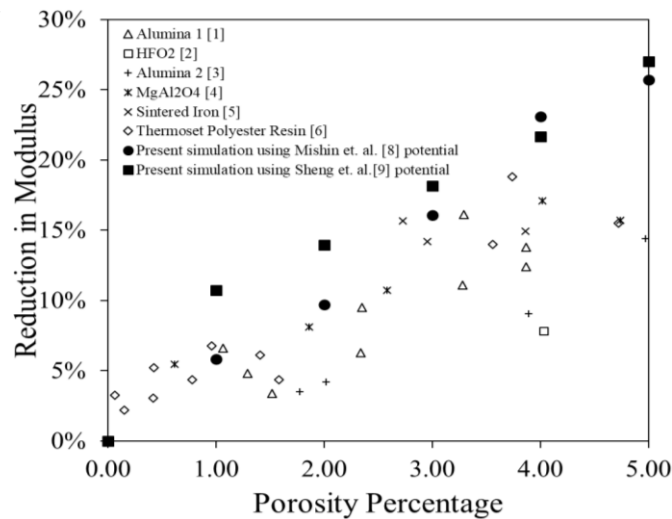


Figure 1. Sample set of results: comparisons of the results obtained from present simulation for various porosity percentage in Aluminum with the available experimental [1-6].

Keyword: Molecular Dynamics, Modulus, Porosity, Temperature

Reference:

- [1] M. Asmani, C. Kermel, A. Leriche, & M. Ourak, "Influence of porosity on Young's modulus and Poisson's ratio in alumina ceramics," *Journal of the European ceramic society*, vol. 21, pp. 1081-1086, August 2001.
- [2] S. L. Dole, O. Hunter Jr, & C. J. Wooge, "Elastic properties of monoclinic hafnium oxide at room temperature," *Journal of the American Ceramic Society*, vol. 60, pp. 488-490, November 1977.
- [3] F. P. Knudsen, "Effect of porosity on Young's modulus of alumina," *Journal of the American Ceramic Society*, vol. 45, pp. 94-95, February 1962.
- [4] D. F. Porter, J. S. Reed, & D. L. III, "Elastic moduli of refractory spinels," *Journal of the American Ceramic Society*, vol. 60, pp. 345-349, July 1977.
- [5] J. P. Panakkal, H. Willems, & W. Arnold, "Nondestructive evaluation of elastic parameters of sintered iron powder compacts," *Journal of Materials Science*, vol. 25, pp.1397-1402, February 1990.
- [6] K. K. Phani, & R. N. Mukerjee, R. N., "Elastic properties of porous thermosetting polymers," *Journal of materials science*, vol. 22(10), pp. 3453-3458, October 1987.
- [7] W. J. Chang, "Molecular-dynamics study of mechanical properties of nanoscale copper with vacancies under static and cyclic loading," *Microelectronic Engineering*, vol. 65, pp. 239-246, January 2003.
- [8] Y. Mishin, D. Farkas, M. J. Mehl & D. A. Papaconstantopoulos, "Interatomic potentials for monoatomic metals from experimental data and ab initio calculations," *Physical Review B*, vol. 59, pp. 3393, February 1999.
- [9] H. W. Sheng, M. J. Kramer, A. Cadien, T. Fujita & M. W. Chen, "Highly optimized embedded-atom-method potentials for fourteen fcc metals," *Physical Review B*, vol. 83, pp. 134118, April 2011.

Molecular Dynamics Simulations of Nanoindentation – The Effect of Force-Field Choice on the Predicted Elastic Modulus

Douglas Pratt

Graduate Student

Department of Engineering and Applied Science
 Memorial University of Newfoundland
 dougp@mun.ca

Dr. Sam Nakhla

Assistant Professor

Department of Engineering and Applied Science
 Memorial University of Newfoundland

ABSTRACT

Nanoindentation is an extremely useful, non-destructive technique for measuring the mechanical properties of metals, polymers, biomaterials and thin-film structures [1]. As its name suggests, nanoindentation is performed on a very small length-scale, making its results dependent on the indented material's microstructure. Consequently, Molecular Dynamics (MD) is an excellent method for simulating nanoindentation because atomistic models can achieve near-experimental sizes [2]. When designing a nanoindentation simulation there are many factors to consider, such as sample dimensions, temperature, indenter material and indenter size. One of the most important factors however to is the choice of force-field potential type, which describes how atom-atom interactions are calculated. Different force-fields result in significantly different values for elastic modulus and other material properties. Common force-field types include the embedded-atom method (EAM) and the modified embedded-atom method (MEAM), which model atom interactions based on electron density [3-4]. Another type is the ReaxFF potential, which uses a bond-order based method to predict the system's energy [5]. The objective of this research is to predict the modulus of Aluminum from nanoindentation simulations with various force-fields of the three mentioned and compare their accuracy to experimental data. The predicted elastic moduli are calculated using the Hertz approximation in the elastic region and by the Oliver and Pharr method [6]. Figure 1 shows sample results for the nanoindentation of aluminum using an EAM force-field [7]. Both methods overestimate the predicted elastic modulus. EAM is the most accurate compared to pure aluminum (70 GPa), ReaxFF is the least accurate and MEAM has average accuracy. This research investigates the possible reasons for overestimating the moduli and differences associated to force-field types.

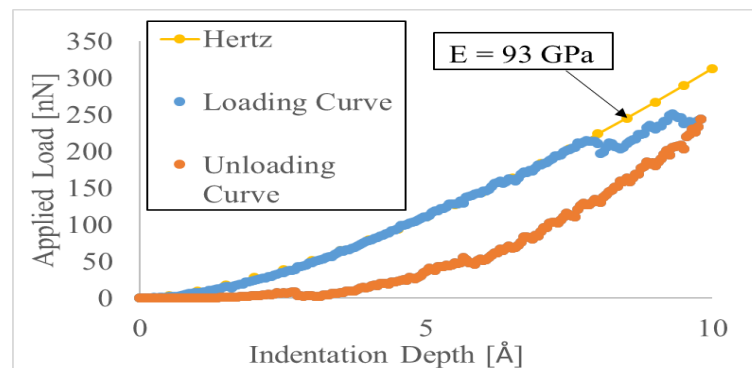


Figure 1 – Nanoindentation simulation of pure aluminum using an EAM force-field.

References

- [1] D. A. Lucca, K. Herrmann and M. J. Klopstein, "Nanoindentation: Measuring methods and applications," *CIRP Ann. Manuf. Technol.*, vol. 59, (2), pp. 803-819, 2010. . DOI: 10.1016/j.cirp.2010.05.009.
- [2] E. T. Lilleodden *et al.*, "Atomistic simulations of elastic deformation and dislocation nucleation during nanoindentation," *J. Mech. Phys. Solids*, vol. 51, (5), pp. 901-920, 2003. . DOI: 10.1016/S0022-5096(02)00119-9.
- [3] M. S. Daw and M. I. Baskes, "Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals," *Physical Review B*, vol. 29, (12), pp. 6443-6453, 1984. . DOI: 10.1103/PhysRevB.29.6443.
- [4] M. I. Baskes, "Modified embedded-atom potentials for cubic materials and impurities," *Physical Review, B: Condensed Matter*, vol. 46, (5), 1992. . DOI: 10.1103/PhysRevB.46.2727.
- [5] S. Plimpton, "Fast Parallel Algorithms for Short-Range Molecular Dynamics," *Journal of Computational Physics*, vol. 117, (1), pp. 1-19, 1995. . DOI: 10.1006/jcph.1995.1039.
- [6] W. C. Oliver and G. M. Pharr, "Measurement of hardness and elastic modulus by instrumented indentation: Advances in understanding and refinements to methodology," *J. Mater. Res.*, vol. 19, (1), pp. 3-20, 2004. . DOI: 10.1557/jmr.2004.19.1.3.
- [7] Y. Mishin *et al.*, "Interatomic potentials for monoatomic metals from experimental data and ab initio calculations," *Physical Review B - Condensed Matter and Materials Physics*, vol. 59, (5), pp. 3393-3407, 1999. . DOI: 10.1103/PhysRevB.59.3393.

Dynamics of a Mono-stable Ring-based Vibratory Energy Harvester

Ibrahim F. Gebrel

Department of Mechanical and Materials Engineering
Western University, London, Ontario
London, Ontario, Canada
igebrel@uwo.ca

Ligang Wang

Department of Mathematica, College of Science
Harbin Engineering University
Harbin, China
wangligang@hrbeu.edu.cn

Samuel F. Asokanthan

Department of Mechanical and Materials Engineering
Western University, London, Ontario
London, Ontario, Canada
sasokant@uwo.ca

ABSTRACT

Exploitation of nonlinear dynamic system phenomena in the design of vibration-based harvesters have recently received much attention. The nonlinearity is often brought to the system via external coordinate-dependent nonlinear forces. A mono-stable and Bi-stable tube structure based energy harvester that utilizes magnetic force is proposed by Mann and Owens [1]. Masana and Daqaq investigated the influence of the potential shape and the excitation level on the performance of mono-stable and bi-stable clamped-clamped beam structure vibration energy harvester [2]. These two studies form the basis for the present study on a novel energy harvester that is based on a flexible ring structure. The present study focusses on interaction of nonlinear magnetic forces that act on a flexible ring structure. Investigation of the performance of such novel design, namely the mono-stable ring structure with nonlinear magnetic force, has been demonstrated via numerical simulation. Mathematical model for the ring structure developed by Asokanthan and Cho [3] as well as a model to generate nonlinear magnetic force that acts on the ring structure is formulated. Galerkin's procedure is employed to reduce the equations of motion to a set of nonlinear ordinary differential equations. The reduced equations represent a Duffing type equation in the presence of external magnetic field. The highly sensitive ring second flexural mode when combined with the nonlinear magnetic force results in an ideal combination that yields increased frequency range, and can be considered as novel in the field of vibration-based energy harvesters. For the purposes of increasing device efficiency, operation of ring-based harvester with nonlinear external excitation frequency close to the system resonant frequency is considered essential. The form of the system potential energy is examined to ensure that the mono-stable behavior exists in the proposed model. Restoring force which is the derivative of the potential energy with respect to the displacement is also examined for the purpose of identifying the softening and hardening behavior of the system.

Numerical predictions of time response, frequency response and phase diagram when the system is subjected to ambient harmonic excitation have been performed. Examination of power variation with time and extracted current vs amplitude of excitation is performed. The results reveal that the steady state value of the power is 0.08586 *Watt* while steady state values of the generated current and displacement are 0.03052 *A* and 0.005827 *m*, respectively. In addition, frequency response of the system is investigated to characterize the behavior due to varying frequencies of the input excitation. The obtained output power from the current ring design is comparable to the bi-stable tube or bi-stable beam harvesters. This analysis gives an insight into the dynamics and power generation of this new class of harvesters.

References

- [1] Mann, B.P., and Owens, B.A., 2010 "Investigations of a Nonlinear Energy Harvester with a Bi-stable Potential Well," *Journal of Sound and Vibration*, Vol.329, No.9, pp.1215-1226.
- [2] Masana, R., and Daqaq, M. F., 2009, "Relative Performance of a Vibratory Energy Harvester in Mono- and Bi-stable Potentials," *Journal of Sound and Vibration*, Vol. 330, No.24, pp. 6036–6052.
- [3] Asokanthan, S.F., and Cho, J., 2006, "Dynamic Stability of Ring – Based Angular Rate Sensors," *Journal of Sound and Vibration*, Vol. 295, No. 3-5, pp. 571-583.

Time Response of The Cylindrical Fused Silica Resonator Subjected to Different Input Angular Velocity

Yiming Luo, Mohamed Bognash, Ibrahim F. Gebrel
 and Samuel F. Asokanthan*
 Department of Mechanical and Materials Engineering,
 The University of Western Ontario,
 London, Ontario, Canada
 yluo373@uwo.ca; sasokant@uwo.ca*

Tianliang Qu
 College of Advanced Interdisciplinary Studies,
 National University of Defense Technology,
 Changsha, Hunan, China
 qutianliang@nudt.edu.cn

ABSTRACT

In this paper, time response analysis of a fabricated cylindrical fused silica resonator (named as Ge02) via simulation is investigated. The resonator Ge02 has been introduced in a previous study [1], and for the present study, diameter of bottom holes are selected to be 3mm, as shown in Figure 1. After fabrication, the vibratory characteristics of the resonator have been measured experimentally via a laser Doppler vibrometer under air atmospheric pressure. The measured second flexural mode natural frequency and Q factor for the stationary, respectively, are 5003 Hz and 3822.

The equations of motion for a ring-based gyroscope derived by Asokanthan et. al [2] for the purposes of investigating the stochastic response of ring-based gyroscopes

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ddot{q} + \begin{bmatrix} 2\zeta\omega_{01} & -2\Omega\gamma \\ 2\Omega\gamma & 2\zeta\omega_{02} \end{bmatrix} \dot{q} + \begin{bmatrix} \kappa_1 + \kappa_2\Omega^2 & -\dot{\Omega}\gamma \\ \dot{\Omega}\gamma & \kappa_1 + \kappa_2\Omega^2 \end{bmatrix} q = \begin{bmatrix} f \cos \omega t \\ 0 \end{bmatrix} \quad (1)$$

is employed for the numerical predictions in the present study. In Eq. (1), f and ω denote, respectively, the external driving force and natural frequency while Ω represents the input angular velocity. The parameter ζ represents the damping ratio and detailed expressions for the parameters κ_1 , κ_2 , and γ can be found in [2]. The second mode natural frequencies of Ge02 are ω_{01} and ω_{02} under rotation. In the present study, driving force magnitude f is set at 1.8N and ω is tuned continuously at the resulting lower natural frequency ω_{01} of the resonator under rotation. Numerical simulation of time response with vibratory characteristics of Ge02 along the sensing axis in the range of Ω from 0 to 100 deg/s is performed. The time response of vibratory amplitude of Ge02 under input angular velocity of 100 deg/s has been shown in Figure 2, and the resonant stable amplitude and the time taken for Ge02 to reach the stable amplitude are depicted in Figure 3.

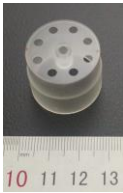


Figure 1. The fabricated cylindrical fused silica resonator.

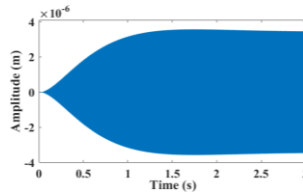


Figure 2. Time response of vibratory amplitude of resonator Ge02.

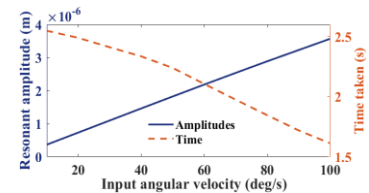


Figure 3. The time taken and resonant amplitude for Ge02 under different input angular velocity.

The results illustrate that the sensing resonant amplitude increases with increasing Ω while the time taken for Ge02 to reach resonant amplitude decreases. The study in this paper is expected to help in the device design, in particular, in selection of the driving scheme and to achieve minimal start-up time for the CVG.

REFERENCES

- [1] Pan Y, Wang D, Wang Y, et al. Monolithic cylindrical fused silica resonators with high Q factors[J]. Sensors, 2016, 16(8): 1185.
- [2] Asokanthan S F, Arghavan S, Bognash M. Stability of Ring-Type MEMS Gyroscopes Subjected to Stochastic Angular Speed Fluctuation[J]. Journal of Vibration and Acoustics, 2017, 139(4): 040904.

3D PRINTED ARCHITECTED SHELLULAR METAMATERIALS

J. Shi

*AM³L Laboratory, Department of Bioresource
Engineering
McGill University
Island of Montreal, QC H9X 3V9, Canada*

*A.H. Akbarzadeh**

*AM³L Laboratory, Departments of Bioresource
Engineering and Mechanical Engineering
McGill University
Island of Montreal, QC H9X 3V9, Canada
Email: hamid.akbarzaeh@mcgill.ca

A. Mirabolghasemi

*AM³L Laboratory, Department of Bioresource
Engineering
McGill University
Island of Montreal, QC H9X 3V9, Canada*

G. Desharnais

Axis Prototypes, Montreal, QC, QC H1P 3C1, Canada

ABSTRACT

Shellular materials derived from triply periodic minimal surfaces (TPMS) with zero mean curvature reveal high specific stiffness and surface-to-volume ratio, pore connectivity, less stress concentration, and enhanced energy absorption capacity, making them appropriate for applications in medicine, photovoltaic, and electrochromic devices. In this paper, we explore numerically and experimentally the topology-property relationship of shellular materials with perforated periodic metasurfaces. The Schwarz Primitive (P) is selected as a shellular material, while different periodic perforation patterns, e.g., circular and elliptical holes with different aspect ratios and sizes, are mapped to the surface in order to develop perforated shellular metamaterials. The shellular metamaterials are 3D printed by selective laser sintering (SLS) and their mechanical properties are tested to provide inputs for numerical simulation and for validating the numerical predictions.

In order to compare the mechanical performance of this new class of advanced materials with conventional P shellulars and lattices, three different 4×4×4 shellular materials (i.e. P shellular without holes and perforated P shellular with circular and elliptical holes) and octet truss lattices with the same relative density are 3D printed. The compression tests are conducted on 3D printed samples to determine their stress-strain curves and to serve as a validation scheme for later numerical simulation. It can be found that conventional P shellulars are stretching dominated, while introducing metasurfaces decreases their stiffness. In addition, all three P shellulars are more resilient than stretching-dominated octet truss. More specifically, compared with the conventional P shellulars, the strain at the onset of densification is increased for perforated P shellulars and their plateau stress is also decreased, which are important factors for evaluating the energy absorption performance of advanced materials. After validation with the experimental results, comprehensive numerical studies are conducted by finite element analysis. The influence of several geometrical parameters, i.e., number, shape, and location of holes and relative density of P shellulars, on Young's modulus, deformation mode, yielding stress, densification strain, and energy absorption efficiency, are investigated. The purpose of this work is to optimize the mechanical properties of P shellulars with periodic perforation metasurfaces, showing great potential on lightweight architectural design, energy harvesting, and porous bone implants.

Time-fractional dual-phase-lag heat conduction and transient crack problem in viscoelastic materials under thermal shock

Wenzhi Yang

Department of Mechanical Engineering,
University of Alberta
Edmonton, Canada

Zengtao Chen

Department of Mechanical Engineering,
University of Alberta
Edmonton, Canada

ABSTRACT

In this study, a theoretical model is built to investigate the crack problem in a half infinite viscoelastic material under a thermal shock. To consider the generalized lagging behaviors in heat conduction and remove some non-physical results, the time-fractional dual-phase-lag (DPL) model is employed. With the increasing application of polymer composites, the viscoelasticity is considered in deriving the transient thermal stresses. The Fourier and Laplace transform, coupled with the singular integral equations, are used to solve the governing, partial differential equations (PDEs) numerically. Finally, the temperature distribution around cracks and stress intensity factors (SIFs) are shown graphically with comparing the results predicted by Fourier's law. The numerical results show the fractional DPL model would be more accurate compared to the conventional DPL model. Significant difference in SIFs exists between viscoelastic and elastic materials.

Modelling of Creep Recovery in Carbon Fiber Reinforced Polymer Composites

Spencer Cabel

Composite Materials and Structures Research Group &
Department of Mechanical Engineering
University of Manitoba
Winnipeg, MB R3T 2N2, Canada
umcabels@myumanitoba.ca

Raghavan Jayaraman

Composite Materials and Structures Research Group &
Department of Mechanical Engineering
University of Manitoba
Winnipeg, MB R3T 2N2, Canada
Raghavan.Jayaraman@umanitoba.ca

ABSTRACT

Composite materials such as carbon fiber reinforced polymers are used in a multitude of applications under varying service conditions. Due to the viscoelasticity of the polymer matrix, the composite tends to experience creep under constant load and creep recovery upon unloading. Past research has focused mainly on the creep; however published work on the creep recovery has been limited. The understanding and successful modelling of the creep recovery portion is of importance in the deformation and fracture of polymer composites subjected to cyclic creep conditions. Hence, creep recovery of a carbon/epoxy composite is focused in this study. Creep recovery after tensile creep has been studied using a dynamic mechanical analyzer over a range of temperatures and applied loads and the data is suitably modeled. The effect of creep recovery on cyclic creep behavior of a polymer composite subjected to cyclic loading is presented and discussed.

Application of the Theory of Configurational Mechanics in Soft Dielectrics

Jianyou Zhou and Liying Jiang
Department of Mechanical and Materials Engineering
Western University
London, Canada
ljjiang55@uwo.ca

ABSTRACT

Introduced by Eshelby, the theory of configurational mechanics studies the evolution of the defects in the material when subjected to external loads. Since its introduction, the theory of configurational mechanics has been applied to analyze a wide range of problems in metals, rubber-like materials and composites, most of which are crack-related. In this work, the theory of configurational mechanics is further developed to tackle some unsettled issues in a new category of rubber-like materials, namely, soft dielectrics. Capable of producing exceptionally large deformation under electrical stimuli, soft dielectrics have extensive applications, such as artificial muscles, actuators, oscillators and energy harvesters. In these applications, soft dielectrics are designed to undergo cyclic deformation for a long period of time. Therefore, one critical issue of using soft dielectrics lies in the prediction of their durability and fatigue life. However, as discussed in the literature, it is rather difficult to evaluate the fatigue life of soft dielectrics since they are usually subjected to electromechanical coupled field and both the electrical load and mechanical load may contribute to their fatigue damage. From this perspective, the configurational stress tensor shows promise of being a predictor for the fatigue life of soft dielectrics since it includes all the energetic properties that govern the evolution of the defects in the material, which is regarded as the mechanism behind their fatigue damage. In this work, the configurational stress tensor for viscoelastic soft dielectrics under electromechanical coupled field is developed and adopted as a fatigue life predictor. The developed predictor is able to adopt most of the hyperelastic relations for rubber-like materials and thermodynamics evolution equations for viscoelastic solids. With the developed predictor, factors that may induce fatigue damage in soft dielectrics under electromechanical coupled field are investigated. The modeling framework and simulation results are expected to provide useful guidelines for further study on the fatigue life of soft dielectrics under different loading conditions and other soft materials that commonly subjected to coupled load. Another critical issue for the application of soft dielectrics is to predict their dielectric strength or electrical breakdown field. Studies have shown that the dielectric strength of soft dielectrics changes with their deformation along with other factors. On the other hand, a robust dielectric strength predictor that accounts for these factors is lacking in the literature. Knowing that electrical breakdown is linked to the deformation of the defects in the material, the configurational stress tensor is also further developed as a predictor for the electrical breakdown field of soft dielectrics. The developed dielectric strength predictor can capture experimental data measured from different experiment set-ups and explain different experimental phenomena. Based on the theory of configurational mechanics, this work is expected to provide helpful solutions to the two main issues of using soft dielectrics. As there may be other possible applications of the theory of configurational mechanics in soft dielectric, this work could be still far away from presenting a complete overview. Nevertheless, the first purpose of this work is to draw attention to the configurational stress which could serve as a powerful tool to solve complicated problems emerging in soft dielectrics and offer guidelines for their optimal design.

Multiscale synergistic damage mechanics methodology for predicting progressive failure of composite structures

Chandra Veer Singh

Department of Materials Science & Engineering, University of Toronto
184 College St., Suite 140
Toronto, ON M5S 3E4, Canada

Email: chandraveer.singh@utoronto.ca

ABSTRACT

Composite materials are being increasingly used for their high stiffness to weight properties in several structural applications, in particular aerospace applications. However, these materials are susceptible to damage during service, especially due to hierarchical nature of the material, micro-cracking and evolution of damage leading to complete failure is a threat and limitation is sophisticated application such as aerospace structures. In order to reduce investment in detail experimental characterizations required for certification process, extensive research has been performed in developing physics based novel computational methods to predict the mechanical properties, damage tolerance and durability of the composites. While, prediction of stiffness and strength has been tackled in somewhat classical theories of composites, prediction of damage initiation and evolution is still a challenging problem. In this work we will present the concepts and developments underpinning the Synergistic Damage Mechanics (SDM) approach [1] which is capable of predicting damage initiation and damage evolution in a multiscale manner, while accounting for matrix cracking at microscale and its propagation through crack multiplication leading to macroscale damage and degradation in properties.

This talk will present the progress of multi-year model development of a multiscale synergistic damage mechanics approach as a comprehensive tool for industrial design and damage tolerance predictions for practical composite structures. This approach can predict damage initiation, its progression and degradation in stiffness properties for multidirectional laminated structures under multiaxial loading conditions with good reliability. We will illustrate the SDM approach and its implementation in ABAQUS and ANSYS, for analyzing variety of structural problems such as wind turbine rotors, helicopter blades, and automotive panel under impact [2-4]. Finally, current limitations and suggestions for future directions will be discussed.

References:

- [1] Talreja, R., & Singh, C.V. (2012). Damage and failure of composite materials. Cambridge University Press., ISBN:9780521819428
- [2] Montesano, J., Chu, H., & Singh, C.V. (2016). Development of a physics-based multi-scale progressive damage model for assessing the durability of wind turbine blades. *Composite Structures*, 141, 50-62.
- [3] Berton, T., Haldar, S., Montesano, J., & Singh, C.V. (2018). Time-dependent damage analysis for viscoelastic-viscoplastic structural laminates under biaxial loading. *Composite Structures*, 203, 60-70.
- [4] Berton, T., Najafi, F., & Singh, C.V. (2019). Development and implementation of a multi-scale model for matrix micro-cracking prediction in composite structures subjected to low velocity impact. *Composites Part B: Engineering*, 168, 140-151.

The Fracture Toughness of Elastic-Plastic 3D Nanolattices

Sahar Choukir

Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Canada
sahar.choukir@mail.utoronto.ca

Chandra Veer Singh

Department of Material Science & Engineering
University of Toronto
Toronto, Canada
chandraveer.singh@utoronto.ca

ABSTRACT

3D hierarchical nanolattices are instrumental in creating engineering materials with both high strength and toughness. Cellular structures are known to have high elastic moduli and specific strengths, but much less is known about their resistance to fracture. Numerical modelling of the fracture toughness of 3D nanolattices remains limited and the effect of different topologies, as well as relative densities on the fracture toughness is still unexplored. Previous studies have shown that continuum fracture mechanics can predict the failure in nano-architected materials. This work reports the stress intensity factors (SIFs: K_I , K_{II}) under tension (Mode I) and shear (Mode II) of centrally cracked plate made of metallic nanolattices using the finite element method. Seven different topologies that exhibit both isotropic and anisotropic elasticity and plasticity have been studied to investigate the relationship between the topology and fracture toughness behavior. Plane strain assumptions required for application of linear elastic fracture mechanics (LEFM) to the lattices were investigated. We present a general numerical framework for computing SIFs under mode I and mode II loading conditions, taking into account the unit cell anisotropy. A 3D finite element stress analysis is carried in Abaqus. Their fracture resistance was determined using the J-integral for a range of lattice orientations (100,110,111). Effects of thickness, crack growth range and anisotropy on fracture toughness parameters were determined. The dependence of mode I and mode II fracture toughness upon relative density is determined for each lattice, and the fracture envelope is obtained in combined mode I–mode II stress intensity factor space. Finite element simulations reveal that the toughness is sensitive to the type of lattice and could be optimized by modifying the topology. The results present a qualitative basis for evaluating the fracture properties of nanolattices.

Vibration method for very high cycle fatigue in aluminum alloy

M.L. Sentissi, A. Ross, M. Brochu

Dep. Mechanical Engineering
Polytechnique Montreal
Montreal (Quebec) Canada
annie.ross@polymtl.ca

E. Pessard

Laboratoire Angevin de mécanique, procédés et innovation
Arts et métiers ParisTech
Angers, France

INTRODUCTION

Assessing very high cycle fatigue (VHCF) life of materials is critical to the design of numerous mechanical systems. Aluminum alloys used for aircraft engines or for civil engineering structures are different, but in both cases they must endure high numbers of vibration cycles. In some applications, high order bending can occur [1]. VHCF testing is a long and fastidious process that can be accelerated by using high frequency vibration equipment instead of standard cycling machine testing [2]. The present work aims to present an approach for fast uniaxial and biaxial fatigue testing. The long term objective of this research project is to understand the fatigue behavior of aluminum alloys used to manufacture structural components.

EXPERIMENTAL SETUP AND METHODOLOGY

The experiments are performed using a 6kN vibration shaker from Sentek Dynamics, which can provide 100g accelerations and frequencies up to 4500 Hz. Thin plate-like specimens are installed in clamped-free configuration for out-of-plane base excitation [3]. Different sample shapes have been considered, including rectangular, kite and cruciform shapes, as in Figure 1(a). Specimen thickness is 1,5mm to 3mm and in-plane dimensions vary from 60mm to 640mm, depending on the geometry and on the target excitation frequency (i.e. ranging between 100 and 2200 Hz).

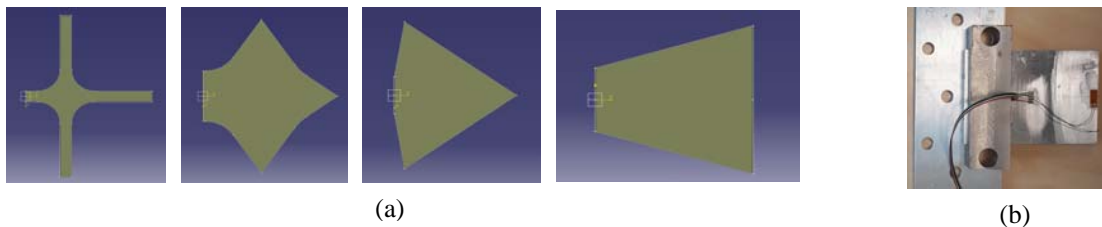


Figure 1. Various configurations for VHCF samples (a) and strain gage mounted at location fatigue testing (b).

Finite element modeling is used to design the specimens, determine the vibration modes required to obtain uniaxial and biaxial stress fields, and therefore fix the required excitation frequency and amplitude, as well as the location of the fatigue zone. Specimens are subjected to biaxial bending. Noncontact vibration measurements are performed using a Polytech vibrometer. Calibration of the actual stress amplitude in the fatigue zone with respect to the velocity measured with the vibrometer is done beforehand using a strain gage as in Figure 1(b).

PRELIMINARY RESULTS

Cruciform-like specimens were designed to provide appropriate low- and high-frequency bending deformations at selected modes. Small changes in specimen shape allow for generating either local uniaxial or local biaxial stress states at the center of the cross. The damping ratio of the entire setup can be as low as 0,08%, which means the resonance peak is extremely narrow and frequency control must be very precise in order to reach sufficient stress levels and properly assess fatigue life.

REFERENCES

- [1] George, T. (2004). *International Journal of Fatigue*, 26(5), 477-486.
- [2] Papakyriacou *et al.* (2001). *Materials Science And Engineering: A*, 308(1-2), 143-152.
- [3] Ellyson, B. (2016). Étude du comportement en fatigue vibratoire [...]. Masters thesis. Polytechnique Montreal (Canada).

Using lime/hemp concrete as a multi-fonction material

E.R. Fotsing, A. Ross
Dep. Mechanical Engineering
Polytechnique Montreal
Montreal (Quebec) Canada
annie.ross@polymtl.ca

T. Lecompte
Laboratoire d'ingénierie des Matériaux de Bretagne
Université Bretagne Sud
Lorient, France

INTRODUCTION

Using bio-sourced or bio-inspired materials is becoming a reasonable solution for reducing our environmental imprint. However, sustainable ecological materials must also provide good mechanical properties, in addition to thermal, hygroscopic or acoustic protection, or other functions. Lime/hemp concrete is a mixture of plant based aggregates and hydraulic aerated binder that can be used as a functionalized filler material [1]. Due to the multi-scale porosity of the material, acoustic properties can be estimated using a double porosity model [2]. In the present experimental work, various lime/hemp configurations were tested in order to better understand the relationship between the material parameters and its acoustic properties.

SAMPLES AND EXPERIMENTS

The lime/hemp composite material was compression molded to produce cylindrical specimens. The diameter was 30mm and the lengths were 30mm, 60mm and 90mm. The compaction pressure was varied in order to obtain a theoretical porosity ranging between 66% and 89%. Figure 1 illustrates the aspect of resulting specimens.

Acoustic testing was performed in a 30 mm diameter impedance tube equipped with two microphones (Fig.2) in absorption configuration [ASTM 2012], to measure the absorption coefficient. Alternatively, the tube was equipped with four microphones (Fig.3) in transmission configuration, to measure the transmission loss (TL) of the material.



Fig.1: Example of low-density (a) and high-density (b) lime/concrete samples [3]

RESULTS

Using the Olny and Boutin model [4] as well as microscope observation, it was shown that microscopic and mesoscopic pores in the composite material are effectively represented and acoustic absorption is determined appropriately. Broadband absorption typical of porous materials is accompanied by a relatively low frequency resonance-like peak. The location of the peak, however, varies significantly with the various material parameters. Stacked multi-layer samples allow producing two overlapping resonance-like peaks, thus offering high absorption coefficients ranging between 0.8 and 1.0 over a wide frequency band (such as from 500 to 1500 Hz). Considering the low density of the materials (around 400 kg/m³), this is considered good acoustic performance.

REFERENCES

- [1] V. Cerezo (2005). Propriétés mécaniques, thermiques et acoustiques [...]. Ph.D. thesis. Ecole doctorale MEGA, Lyon, France.
- [2] P. Gle et al. (2012). *Construction and building materials*, 37, pp.801-811.
- [3] E.R. Fotsing et al. (2017). 2nd International conference on Bio-based building materials & 1st Conference on Ecological valorisation of granular and fibrous materials. June 21-23 2017, Clermont-Ferrand, France.
- [4] X. Olny, C. Boutin (2003). Acoustic wave propagation in double porosity media. *J Acoust Soc Am*, 114(1), pp.73-89.

MULTISCALE ANALYSIS OF FERROELECTRIC CELLULAR METAMATERIALS

J. Shi

AM³L Laboratory, Department of Bioresource
 Engineering
 McGill University
 Island of Montreal, QC H9X 3V9 Canada

A.H. Akbarzadeh*

AM³L Laboratory, Departments of Bioresource and
 Mechanical Engineering
 Island of Montreal, QC H9X 3V9 Canada
 *Email: hamid.akbarzaeh@mcgill.ca

Abstract — In this paper, asymptotic homogenization (AH) is implemented for predicting the effective thermo-electro-mechanical properties of 2D extruded ferroelectric cellular metamaterials. The effect of pore microarchitecture (relative density and cell topology) and polarization direction on ferroelectric properties of periodic ferroelectric cellular metamaterials is explored. The pore topology is determined by Fourier series expansion. Alternative pore microarchitectures are considered by tailoring the shape parameters, scaling factor, and rotation angle of the constitutive pore. It is found that the piezoelectric and pyroelectric figures of merit of ferroelectric cellular metamaterials can be significantly improved compared to the commonly used honeycomb cellular materials if an appropriate microarchitecture is selected for the pore. For example, piezoelectric charge coefficient (d_h) for an architected transversely polarized ferroelectric cellular metamaterial with a solid volume fraction of 0.4 can be 350% higher than the corresponding figures of merit of ferroelectric honeycomb.

Keywords - *Ferroelectric cellular metamaterial; Asymptotic homogenization; Piezoelectricity; Pyroelectricity*

I. INTRODUCTION

By tailoring their periodically repeated microarchitecture, metamaterials deliver many exotic properties beyond those found in natural materials or chemically synthesized substances, e.g., electromagnetic metamaterials [1], optical metamaterials [2], acoustic metamaterials [3] and mechanical metamaterials [4]. If these metamaterials are made of smart materials (e.g., piezoelectric, piezomagnetic, or magnetostrictive materials) in the form of optimized cellular architectures, they can open a new venue for designing lightweight advanced multifunctional materials responsive to arbitrary multiphysical stimuli [5, 17].

Advanced 3D printing technology [6] has recently provided reliable routes for manufacturing of ferroelectric metamaterials with complex geometries in micro and nano scales. Based on the interconnection of the inclusions/pores, four types of composite/cellular ferroelectric metamaterials can be defined: (1) 0-3 type (inclusions/pores are enclosed by the matrix), (2)

1-3 type (inclusions/pores connect in one direction and the matrix connects in all three directions), (3) 2-2 type (both inclusions/pores and matrix exhibit connectivity in two dimensions), and (4) 3-3 type (both inclusions/pores and matrix exhibit connectivity in all three dimensions).

Introducing pores into ferroelectric materials has shown remarkable improvement in figures of merit, especially for hydrophone applications. For 0-3 type of ferroelectric cellular metamaterials, alternative pores, e.g. flat-cuboidal, spherical and short-cylindrical, showed different effective electro-mechanical properties [7]. The electro-mechanical properties of 1-3 ferroelectric cellular materials can be improved by alternative parameters, e.g. pore shape [8], pore aspect ratio [9], and polarization direction [9-11]. For longitudinally-polarized ferroelectric cellular materials, the electro-mechanical properties were insensitive to the change of pore shape and aspect ratio; opposite behavior has been found for transversely polarized porous piezoelectric materials [8, 10, 11]. Compared to 1-3 type of ferroelectric cellular materials, 3-3 type of ferroelectric cellular materials have shown great improvement in ferroelectric figures of merit at the expense of deteriorating mechanical performance [12].

Both analytical and numerical methods have been proposed to obtain the effective properties of architected cellular metamaterials. Several analytical mean field methods [13-15], e.g. dilute, self-consistent, and Mori-Tanaka, were proposed to predict the effective electro-mechanical properties of ferroelectric materials. However, the local fluctuations of the field quantities were not taken into account in these analytical micromechanical models. Asymptotic homogenization (analytical method) [16, 17] and finite element method (numerical method) [8-11] have been widely used for predicting the effective properties of materials with periodic microstructures. Although the finite element method is straightforward to follow, the asymptotic homogenization method has a robust mathematical basis and provides closed-form expressions for accurate theoretical predictions.

In this paper, two kinds of 1-3 ferroelectric cellular metamaterials with different types of pore topologies are modeled and multiscale asymptotic homogenization is proposed to predict their thermo-electro-mechanical properties. The role of the microarchitectural features and polarization

direction of constitutive cells on their effective properties and figures of merit are also studied. In order to show the improvement of ferroelectric cellular metamaterials, their performance is compared to conventional honeycomb piezoelectric cellular materials. The optimum pore topologies with maximum ferroelectric properties and figures of merit are also presented, which are beneficial for ferroelectric architecture design for hydrophone and IR detection device.

II. MODELLING OF FERROELECTRIC CELLULAR METAMATERIALS

A. Formulation of Cell Geometry

As shown in Figs. 1(a) and 1(b), according to the polarization direction (axis 3) of the constitutive ferroelectric materials, the 1-3 ferroelectric cellular metamaterials can be further classified as a longitudinally polarized (i.e. polarized along the pore axis, type I) and transversely polarized (i.e. polarized orthogonal to the pore axis, type II) ferroelectric cellular metamaterials.

The following Fourier series expansion is used to generate pores of a general two-dimensional (2D) topology [18]:

$$x = r(\theta) \cos(\theta), y = r(\theta) \sin(\theta) \quad (1)$$

where $r(\theta) = r_0[1 + c_1 \cos(\theta) + c_2 \sin(\theta)]$ and $\alpha, \beta = 1, 2, \dots$.

In Eq. (1), α and β control the symmetry of the pore topology. For example, when $\alpha = 2$ and $\beta = 4$, the pore is two-fold symmetry and when $\alpha = 3$ and $\beta = 6$, the pore is three-fold symmetry. In addition, c_1 and c_2 determine the specific pore shape. Assuming that the representative volume element (RVE) of cell is a unit-length square, the volume fraction of the solid for ferroelectric cellular metamaterials made of a single material is expressed by:

$$f_v = 1 - \frac{\pi(2 + c_1^2 + c_2^2)}{2} r_0^2 \quad (2)$$

where f_v is the volume fraction of the solid.

By rotating (θ_r as rotation angle) and scaling (s as scaling factor or aspect ratio) of the pore, a wider range of pore microarchitecture can be achieved to further optimize the thermo-electro-mechanical properties of ferroelectric cellular metamaterials. Fig. 2(a) presents 25 different shapes of four-fold symmetrical pores with $\alpha = 4$, $\beta = 8$, $\theta_r = 0$, $s = 1$ and $f_v = 0.8$, while c_1 and c_2 vary between -0.4 and 0.4 to show their effects on the pore topology. Fig. 2(b) scales down ($s = 2$) and rotates ($\omega_r = 45^\circ$ counterclockwise) pores presented in Fig. 2(a). Fig. 2(c) shows three-fold symmetrical pores with $\alpha = 3$, $\beta = 6$, $s = 1$, $\omega_r = 0$ and $f_v = 0.8$ to reveal how shape parameters α and β tune the pore microarchitecture in cellular metamaterials. To showcase the unprecedented thermo-electro-mechanical properties of architected ferroelectric cellular metamaterials, we compare their effective properties with conventional 2D extruded ferroelectric honeycombs.

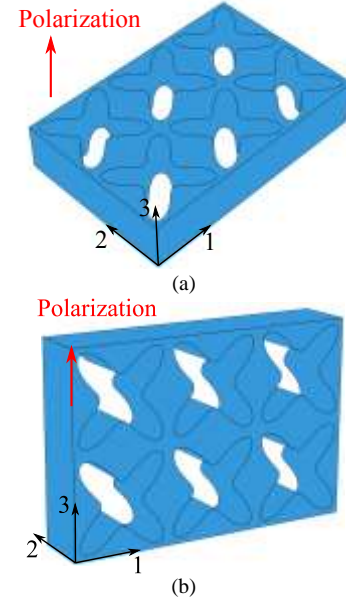


Figure 1. Two types of 1-3 cellular, 2D extruded, ferroelectric metamaterials: (a) Type I (longitudinal polarization); (b) Type II (transverse polarization).

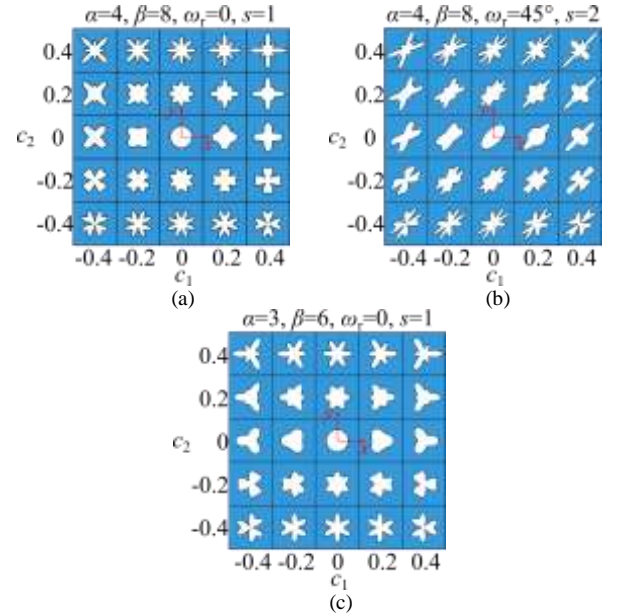


Figure 2. Cross section of 1-3 cellular, 2D extruded, ferroelectric metamaterials with: (a) $\alpha = 4$, $\beta = 8$, $\theta_r = 0$ and $s = 1$, (b) Single material with $\alpha = 4$, $\beta = 8$, $\omega_r = 45^\circ$ and $s = 1$, and (c) Single material with $\alpha = 3$, $\beta = 6$, $\omega_r = 0$ and $s = 1$.

B. Constitutive and Governing Equations of Piezoelectric Materials

Herein, we assume that the mechanical and electric fields are coupled while temperature field is not affected by the mechanical and electric fields; such analysis is so-called uncoupled thermo-piezoelectricity [17, 19]. In linear elasticity, the linear constitutive relation of thermo-piezoelectric materials can be expressed as:

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl} + e_{kij}(-E_j) - \lambda_{ij}\theta \quad (3a)$$

$$D_i = e_{ikl}\varepsilon_{kl} - \kappa_{ij}(-E_j) - p_i\theta \quad (3a)$$

where σ_{ij} , D_i , ε_{kl} , E_j and θ ($i, j, k, l=1,2,3$) are stress tensor, electric displacement vector, strain tensor, electric field vector, and temperature change from the reference temperature, respectively; C_{ijkl} , e_{kij} , κ_{ij} , λ_{ij} and p_i are respectively, elastic, piezoelectric, dielectric, thermal-stress and primary pyroelectric coefficient vectors.

In order to assess the performance of devices made of ferroelectric materials, three figures of merit are here discussed, i.e., the piezoelectric charge coefficient (d_h), the hydrostatic figure of merit ($d_h \times g_h$), and voltage responsivity (F_v) [20].

Hydrostatic charge coefficient: The hydrostatic (or piezoelectric) charge coefficient, $d_h = d_{33} + d_{31} + d_{32}$, assesses the conversion of mechanical loads (under hydrostatic loading) to electrical signals. Therefore, higher d_h will enhance hydrophone device sensitivity to detect sound.

Hydrostatic figure of merit: The hydrostatic figure of merit, $d_h \times g_h$, defines the ability of a hydrophone device to identify and distinguish a signal from the ambient background noise. Here, the hydrostatic voltage coefficient, g_h , is given by d_h / κ_{33}^σ , where $\kappa^\sigma = \kappa + \mathbf{e}^T \mathbf{C}^{-1} \mathbf{e}$ and a larger $d_h \times g_h$ means a higher signal-to-noise ratio.

Voltage responsivity: For heat and infrared (IR) detection, which are based on the generation of maximum current or voltage for a given thermal energy input, a higher voltage responsivity (F_v) improves the device sensitivity to thermal fluctuation:

$$F_v = \frac{P_3}{c_E \cdot \kappa_{33}^\sigma} = \frac{P_3}{\rho \cdot c_p \cdot \kappa_{33}^\sigma} \quad (4)$$

C. Asymptotic Homogenization for Periodic Ferroelectric Cellular Metamaterials

As shown in Fig. 3, in order to predict the overall multiphysical properties of heterogeneous cellular metamaterials, asymptotic homogenization analyzes their representative volume elements (RVE) to obtain a macroscopic homogenous medium equivalent to the original heterogeneous microstructure. More specifically, there are two distinct scales in this periodic structures. The first one is the macroscopic scale with a coordinate of $\{O, x_1, x_2, x_3\}$ and the other one is the microscopic scale with a coordinate of $\{O, y_1, y_2, y_3\}$; η ($\ll 1$) is a magnification factor and $y_i = x_i / \eta$.

Based on asymptotic expansion and periodic boundary conditions, we can express the effective stiffness tensor $\bar{\mathbf{C}}$, piezoelectric tensor $\bar{\mathbf{e}}$, dielectric permittivity tensor $\bar{\mathbf{\kappa}}$, thermal stress tensor $\bar{\boldsymbol{\lambda}}$ and pyroelectric coefficient $\bar{\mathbf{p}}$ as

$$\begin{aligned} \bar{C}_{ijmn} &= \langle C_{ijmn}^* + C_{ijkl}^* MU_{k,l}^{mn} + e_{ijl}^* NU_{,l}^{mn} \rangle \\ \bar{e}_{ijn} &= \langle e_{ijn}^* + C_{ijkl}^* M \Phi_{k,l}^n + e_{ijl}^* N \Phi_{,l}^n \rangle \\ \bar{\kappa}_{in} &= \langle \kappa_{in}^* - e_{ikl}^* M \Phi_{k,l}^n + \kappa_{il}^* N \Phi_{,l}^n \rangle \\ \bar{\lambda}_{ij} &= \langle \lambda_{ij}^* - C_{ijkl}^* M \Theta_{k,l} - e_{ijl}^* N \Theta_{,l} \rangle \\ \bar{p}_i &= \langle p_i^* - e_{ikl}^* M \Theta_{k,l} + \kappa_{il}^* N \Theta_{,l} \rangle \end{aligned} \quad (5)$$

where $\langle (\cdot) \rangle = \frac{1}{V_\Omega} \int_\Omega (\cdot) dV$ and V_Ω is the volume of the RVE;

MU_k^{mn} , $M\Phi_k^n$, $M\Theta_k$, NU^{mn} , $N\Phi^n$ and $N\Theta$ are microscopic fluctuations and can be obtained by solving the three local problems [17].

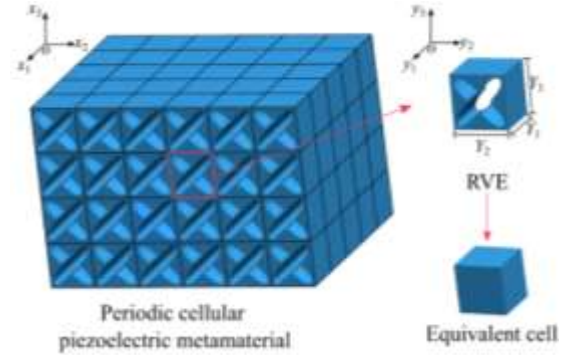


Figure 3. Schematic figure of Homogenization.

III. RESULT AND DISCUSSION

In the following subsections, BaTiO_3 is selected as the ferroelectric material, which is lead-free environmentally-compatible material. Its piezoelectric properties of the constituent materials are given in Table 1, whereas their associated thermal properties are given in Table 3.

TABLE I. PIEZOELECTRIC PROPERTIES OF MATERIAL CONSTITUENTS OF FERROELECTRIC CELLULAR METAMATERIALS: \mathbf{C} IN ($\times 10^{10}$) PA, \mathbf{E} IN ($\times 10^{-2}$) C/M², $\mathbf{\kappa}$ IN ($\times 10^{-11}$) C/(VM²) [21].

Properties	Epoxy	BaTiO ₃
C_{11}	8	150
C_{12}	4.4	66
C_{13}	4.4	66
C_{33}	8	146
C_{44}	1.8	4.4
e_{31}	0	-4.3
e_{33}	0	17.5
e_{42}	0	11.4
κ_{11}	3.7	987.2
κ_{33}	3.7	1116

TABLE II. THERMAL PROPERTIES OF MATERIAL CONSTITUENTS OF FERROELECTRIC CELLULAR METAMATERIALS P IN G/CM³, Δ IN ($\times 10^{-6}$ K⁻¹) PA, Γ IN VM⁻¹K⁻¹, K IN WM⁻¹K⁻¹, C_p IN J/(GK) [2,22].

Properties	Epoxy	BaTiO ₃
P	1.15	5.7
Δ_{11}	60	8.53
Δ_{33}	60	1.99
P_3	0	13300
K	0.7	2.9
C_p	1.10	0.43

In order to understand the relationship between thermo-electro-mechanical properties and the volume fraction of solid phase or porosity, Fig. 4 presents three piezoelectric related properties and one pyroelectric related properties for type I (red area) and type II (blue area) metamaterials with f_v varying from 0.3 to 1. For considered metamaterials, c_1 and c_2 are selected as -0.4, -0.2, 0, 0.2 or 0.4 while two groups of α and β , i.e. $\alpha = 4, \beta = 8$ and $\alpha = 3, \beta = 6$, are chosen for generating the microarchitectures. The pore shapes are further tailored by rotating and scaling the pores for the following four sets: $\theta_r = 0$ and $s = 1$ (original shape), $\theta_r = 0$ and $s = 0.5$ (scaled pore), $\theta_r = 45^\circ$, $s = 1$ (rotated pore), and $\theta_r = 45^\circ$ and $s = 0.5$ (rotated and scaled pore). Omitting inadmissible pore architectures, 1482 different cellular ferroelectric metamaterials are analyzed. Figure 4 shows that type II metamaterials provide a wider range of material properties for design selection. As shown in Fig. 4(a), κ_{33} of both type I and type II metamaterials show an increasing trend with the increase of volume fraction; the variation is linear with the volume fraction for type I metamaterials, while nonlinear for type II metamaterials. In addition, the value of κ_{33} of both type I metamaterials is always higher than that of type II metamaterials. In Fig. 4(b) and (c), although d_h and $d_h \times g_h$ of type I metamaterials almost remain unchanged with different volume fractions, d_h and $d_h \times g_h$ of type II metamaterials show a significant improvement by increasing porosity. Moreover, pore topologies also have great influence on d_h and $d_h \times g_h$ of type II metamaterials. For example, for type II metamaterials with $\alpha = 4, \beta = 8, c_1 = 0.2, c_2 = -0.2, \theta_r = 0, s = 1$, and $f_v = 0.4$, d_h is 375% higher than that of type I honeycomb ferroelectric materials with $f_v = 0.4$, and is 27% higher than that of type II honeycomb with $f_v = 0.4$. Similar as d_h and $d_h \times g_h$, F_V (Fig. 4(d)) also decreases with an increase of volume fraction. However, type I and type II metamaterials have similar F_V and type II honeycomb shows the lowest F_V . Generally, introducing pores to 1-3 cellular metamaterials, especially type II metamaterials, will obviously improve the $d_h, d_h \times g_h$ and F_V , which is due to the significant decrease of κ_{33} (Figs. 4(a)).

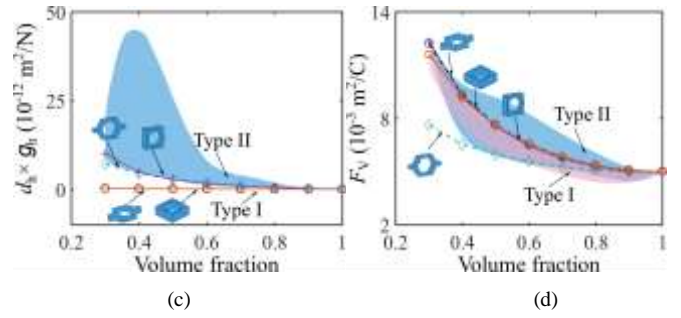
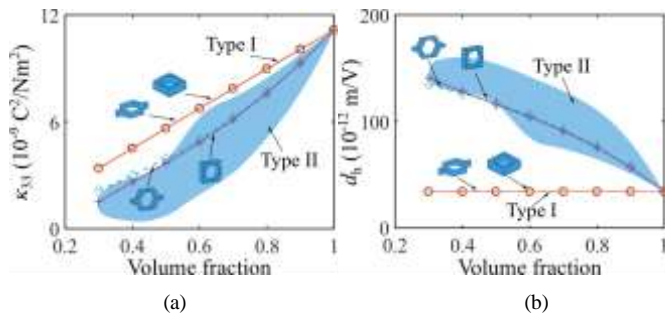


Figure 4. Property and figures of merit design charts for type I and type II cellular ferroelectric metamaterials with solid volume fraction f_v varying from 0.3 to 1.

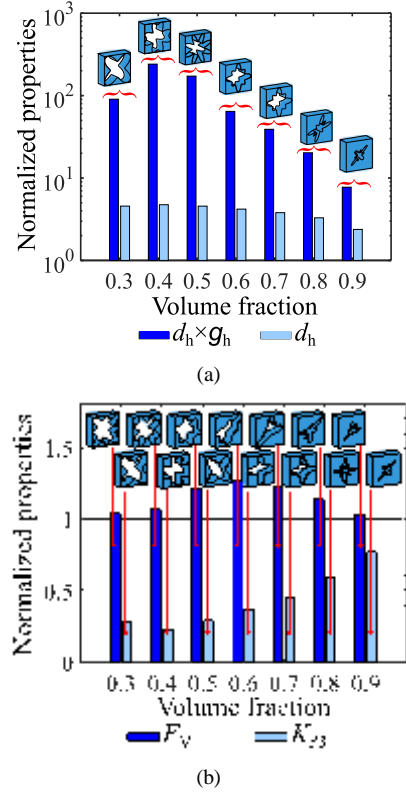


Figure 5. Normalized maximum figures of merit of Type I and type II ferroelectric cellular metamaterials.

In order to provide quantitative comparison between the architected cellular ferroelectric metamaterials and conventional honeycomb ferroelectric materials, Fig. 5 presents normalized optimum figures of merit shown in Fig. 4, where maximum values of $d_h, d_h \times g_h, F_V$, and K_{23} (thermal conductivity along polarization direction) at different volume fractions are normalized by type I honeycomb ferroelectric materials with the same solid volume fraction. As shown in Fig. 5(a), the microarchitectures with optimum d_h and $d_h \times g_h$ vary for different volume fractions of BaTiO₃. The improvement, however, decreases when the volume fraction increases. For example, when f_v is 0.4, d_h is increased by 375%, while 138% for $f_v = 0.9$. In Fig. 5(b), it can be seen that there is a slight improvement for F_V (appropriate for IR detection applications), increasing by 3.4% ~ 26.5%, and significant

decrease for K_{33} by 23.18% ~ 77.24% which is appropriate for thermal insulation applications.

IV. CONCLUSION

This work conducts a strategic study on the role of pore microarchitecture and polarization direction on the effective thermo-electro-mechanical properties of 1-3 type of architected ferroelectric cellular metamaterials. In particular, the multiscale asymptotic homogenization used for predicting the effective thermo-electro-mechanical properties of ferroelectric materials are proposed. The influence of pore topologies on two kinds of 1-3 cellular ferroelectric metamaterials is also presented. The results show that introducing pores into ferroelectric materials leads to a significant improvement of d_h , $d_h \times g_h$, and F_v , resulting from the decrease of κ_{33} . In addition, by tailoring the pore topologies, the values of d_h , $d_h \times g_h$, and F_v can be further modified, which provides a reliable design motif for tuning the ferroelectric properties.

ACKNOWLEDGMENT

A.H. Akbarzadeh acknowledges the financial support by Fonds de Recherche du Québec – Nature et technologies (Établissement de nouveaux chercheurs universitaires - New university researchers start-up program) [Grant number : 252976]. J. Shi also acknowledges the financial support provided by the China Scholarship Council [File No. 201706220086] and McGill University.

REFERENCES

- [1] Lagarkov A N, Semenenko V N, Chistyayev V A, et al. Resonance properties of bi-helix media at microwaves. *Electromagnetics*, 1997, 17(3):213-23.
- [2] Shalaev V M, Cai W, Chettiar U K, et al. Negative index of refraction in optical metamaterials. *Optics Letters*, 2005, 30(24):3356-3358.
- [3] Li J, Chan C T. Double-negative acoustic metamaterial. *Physical Review E*, 2004, 70(2):055602.
- [4] Rafsanjani A, Akbarzadeh A H, Pasini D. Snapping Mechanical Metamaterials under Tension. *Advanced Materials*, 2015, 27(39): 5931-5935.
- [5] Shimada T, Le V L, Nagano K, et al. Hierarchical ferroelectric and ferrotoroidic polarizations coexistent in nano-metamaterials. *Science Report*, 2015, 5:14653.
- [6] Bodkhe S, Turcot G, Gosselin F P, et al. One-step solvent evaporation-assisted 3D printing of piezoelectric PVDF nanocomposite structures. *Acs Applied Materials & Interfaces*, 2017, 9(24):20833.
- [7] Bravo-Castillero J, Rodríguez-Ramos R, Guinovart-Díaz R, et al. Analytical formulae for electromechanical effective properties of 3–1 longitudinally porous piezoelectric materials. *Acta Materialia*, 2009, 57(3):795-803.
- [8] Della C N, Shu D. The performance of 1–3 piezoelectric composites with a porous non-piezoelectric matrix. *Acta Materialia*, 2008, 56(4):754-761.
- [9] Iyer S, Venkatesh T A. Electromechanical response of (0-3) porous piezoelectric materials: Effects of porosity shape. *Journal of Applied Physics*, 2011, 110(3):034109.
- [10] Kar-Gupta R, Venkatesh T A. Electromechanical response of porous piezoelectric materials: Effects of porosity distribution. *Applied Physics Letters*, 2007, 93(1): 062904-1.
- [11] Kar-Gupta R, Venkatesh T A. Electromechanical response of porous piezoelectric materials. *Applied Physics Letters*, 2006, 54(15):4063-4078.
- [12] Challagulla K S, Venkatesh T A. Electromechanical response of piezoelectric foams. *Acta Materialia*, 2012, 60(5):2111-2127.
- [13] Wang, B. Three-dimensional analysis of an ellipsoidal inclusion in a piezoelectric material. *International Journal of Solids & Structures*, 1992, 29: 293–308.
- [14] Dunn M L, Taya M. Micromechanics predictions of the effective electroelastic moduli of piezoelectric composites. *International Journal of Solids & Structures*, 1993, 30(2): 161–175.
- [15] Chen, T. Piezoelectric properties of multiphase fibrous composites: some theoretical results. *Journal of the Mechanics and Physics of Solids*, 1993, 41(11): 1781–1794.
- [16] Medeiros R D, Rodríguez-Ramos R, Guinovart-Díaz R, et al. Numerical and analytical analyses for active fiber composite piezoelectric composite materials. *Journal of Intelligent Material Systems & Structures*, 2015, 26(1):101-118.
- [17] Shi J , Akbarzadeh A H. Architected Cellular Piezoelectric Metamaterials: Thermo-Electro-Mechanical Properties[J]. *Acta Materialia*, 2018.
- [18] Overvelde J T B, Shan S, Bertoldi K. Compaction through buckling in 2D periodic, soft and porous structures: effect of pore shape. *Advanced Materials*, 2012, 24(17):2337-2342.
- [19] Akbarzadeh A H, Chen Z T. Hygrothermal stresses in one-dimensional functionally graded piezoelectric media in constant magnetic field. *Composite Structures*, 2013, 97(2):317-331.
- [20] Zhang Y, Xie M, Roscow J, et al. Enhanced pyroelectric and piezoelectric properties of PZT with aligned porosity for energy harvesting applications. *Journal of Materials Chemistry A: Materials for Energy & Sustainability*, 2017, 5(14):6569-6580.
- [21] Fantoni F, Bacigalupo A, Paggi M. Multi-field asymptotic homogenization of thermo-piezoelectric materials with periodic microstructure. *International Journal of Solids & Structures*, 2017, 120: 31-56.
- [22] Yao K, Biscani F, Belouettar S, et al. Multi-coating inhomogeneities approach for the effective thermoelectro-elastic properties of piezoelectric composite materials. *Composite Structures*, 2010, 92(4):964-972.

VERIFICATION AND VALIDATION OF A FINITE ELEMENT MODEL FOR STRUCTURAL HEALTH MONITORING OF ALUMINUM STRUCTURES USING GUIDED WAVE ULTRASONIC TESTING

Muhammad Tariq Vatao*, Remon Pop-Iliev, and Ghaus Rizvi

Faculty of Engineering and Applied Science

UOIT-University of Ontario Institute of Technology, Oshawa, Canada

Email: muhammadtariq.vatao@uoit.net

Abstract— Guided Wave Ultrasonic Testing (GWUT) has become a promising technique for non-destructive testing in recent years. Guided waves propagate in omni-direction in the form of longitudinal and transverse waves and offer inspecting advantages by travelling through the whole geometry of objects, which makes them a useful tool for structural health monitoring. While the experimental correlation of the related key variables has been previously studied and established, more verification and validation studies are needed to quantify the confidence and predictive accuracy of model-based calculations. In this context, this paper focuses on studies of wave propagation during GWUT through experiments versus results obtained from a two-dimensional finite element (FE)-based model analysis. Two transducers were used in the experimental setup which were attached on an aluminum plate at some distance. Out of the two transducers, one was used as an actuator transducer excited by a Gaussian-modulated sinusoidal pulse. The other transducer was used as a sensor. The FE analysis was performed using COMSOL Multiphysics®. The comparative study of the results for guided wave propagation in a plate obtained from the FE analysis and the ones obtained through experiments demonstrated good agreement. The verified FE model is used to predict wave propagation at different distances between the sensor and the actuator. The results comparison and the discrepancy in the predicted wave propagation behavior are presented and discussed in this paper.

Keywords-GWUT; Structural health monitoring; Finite element; wave propagation; Gaussian-modulated

I. INTRODUCTION

Guided waves ultrasonic testing (GWUT) based damage detection technique is commonly used for damage detection and structural health monitoring [1]. These waves propagate along the wave guide in the form of transverse and longitudinal waves [2]. Conventional ultrasonic method including pulse echo is restricted to investigate the simple geometries and the zones close to the transducer availability [3]. Whereas, guided waves can travel longer distances and can inspect the whole structure from a single point [4].

GWUT usually performs by two transducers, one as an actuator and other as a sensor. However, a single transducer

can emit and receive waves at the same time [5]. Unlike conventional ultrasonic probes, piezoelectric transducers are light weight, inexpensive and scan large structures with minimal invasion [6]. However, due to the inertial resistance of piezoelectric transducer, it is difficult to perfectly transform electronic signal into mechanical vibration. To overcome this problem, most of the researchers have used Gaussian-modulated sinusoidal pulse or Hann window instead of using simple sinusoidal pulse to provide gradual increment of the amplitude [7].

In this paper, experiments have been done to observe the behavior of guided waves in aluminum plate. Gaussian-modulated pulse is used to excite the transducer. Finite element model has been made to simulate guided waves propagation. The numerical model is validated with the experimental results. The comparison of wave propagation pattern observed in the experiments and predicted by FE model are also presented.

II. EXPERIMENTAL

A plate of aluminum alloy 6061-T6 is utilized in the experiments. The dimension of the plate was 1.2 m by 1 m with a thickness of 2 mm. Two piezoelectric transducers were rigidly attached on the plate surface by using cyanoacrylate as a coupling agent for efficient transmission of ultrasonic waves from transducer. The thickness and the diameter of the transducers are 1mm and 10mm with natural frequency of 200 kHz. One of the transducer was utilized to excite the guided waves and the other one was used as a receiver. Both transducers were positioned at the middle part of the plate to keep away the disturbance arises from reflected waves as appeared in Fig. 1. The experiments were performed by placing the sensor transducer at the distance 50 mm and 10 mm from actuator transducer.

A Gaussian-modulated sinusoidal pulse was used to excite the transducer. The frequency of the pulse was set to 200 kHz to excite the transducer at the resonant frequency. The input signal was generated by combining the Gaussian window with the sine wave as mentioned in Eq. 1.

$$x(t) = e^{-0.5\left(\frac{t-t_m}{t_0}\right)^2} \sin(\omega t) \quad (1)$$

Where t_0 is the standard deviation, and t_m is the mean of Gaussian function. The signal was generated by function generator (GW-Instek AFG-2225) and transferred to the actuator to excite it on 5 count tone burst as shown in the Fig. 2. Upon receiving the signal, actuator has converted the signal in to mechanical vibration which propagated towards sensor transducer in the form of guided waves. The waves received at the sensor transducer was investigated with an oscilloscope as shown in Fig 4.

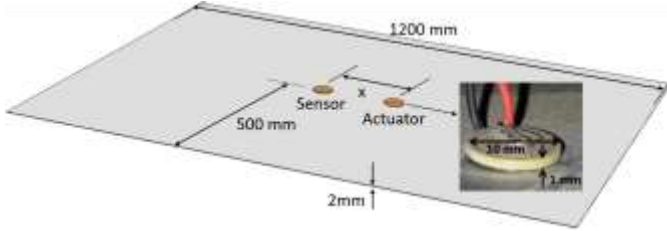


Fig. 1 Experimental setup illustration. Inset is showing the bonding of the piezoelectric transducer with the aluminum plate..

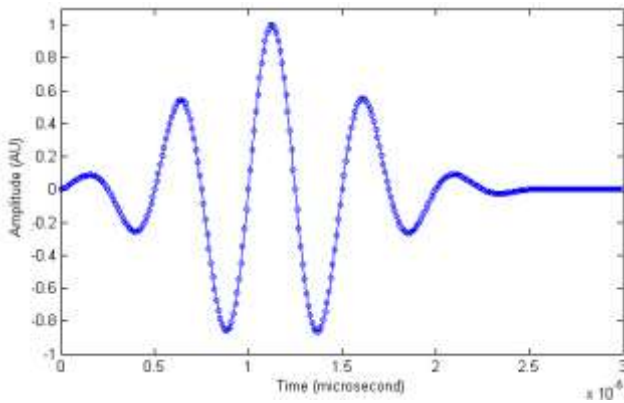


Fig. 2: Sine wave modulated by the Gaussian function at 200 kHz.

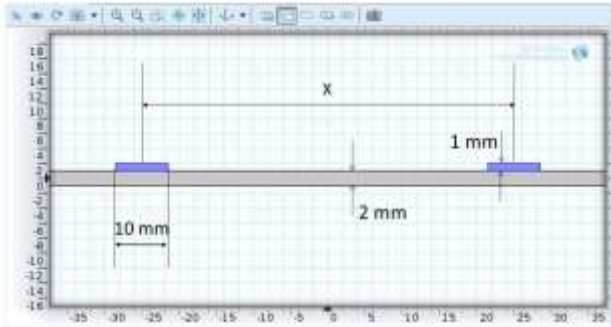


Fig. 3: 2D finite element model in COMSOL interface.

III. FINITE ELEMENTS ANALYSIS

The FE analysis was done by using COMSOL Multiphysics®. The multiphysics interface of this software allows us to link the boundary conditions of one physics with another, such as piezoelectric module which couple the boundary conditions of solid mechanics with electrostatic. A 2D finite element model of the guided waves propagation was created. All the parameters and dimensions were same as used

in the experiments as shown in Fig. 3. ‘Lead Zirconate Titanate (PZT-5A)’ material was selected for transducers and ‘Aluminium alloy 6061’ plate from COMSOL library. The bottom surfaces of both piezoelectric transducers were defined as ‘ground’ boundaries while ‘electric potential’ boundary condition was defined on the top surface of actuator transducer. Rest of the surfaces piezoelectric transducers were defined as ‘zero charge’ boundaries.

The wave propagation simulation was generated, the time duration of the simulation was long enough for the sensor to receive the complete waveform excited by the actuator. The simulation was done on two different condition. The distance between the both transducer was 50 mm in the first simulation and 100 mm in the second simulation. The pattern of the waveform generated at the sensor transducer was exported through post-processing. The data was analyzed by plotting the graph between electric potential amplitude and time as shown in the Fig 5.

IV. RESULTS AND DISCUSSION

The experimental data received at the oscilloscope was exported in the form of CSV file. The data was normalized to compare the wave pattern with the numerical model. The waveform received on the oscilloscope are similar to the waveform generated by numerical model. For better understanding, the data received from the oscilloscope and the data generated by numerical model was normalized and plotted together as shown in Fig 6. For 50 mm distance, the both waveforms are very similar. For 100 mm distance, wavelength and phase angle are closed to each other, however, some deviation in the wave envelopes was observed.

One of the main reasons for the deviation in the experimental and numerical results is the bonding of the transducer with the plate. In FE model, the piezoelectric disc was perfectly combined with the plate, whereas, the actual bonding is not perfect. The wire present between the bottom surface of the piezo disc and the plate is creating air void as shown in the Fig. 1.

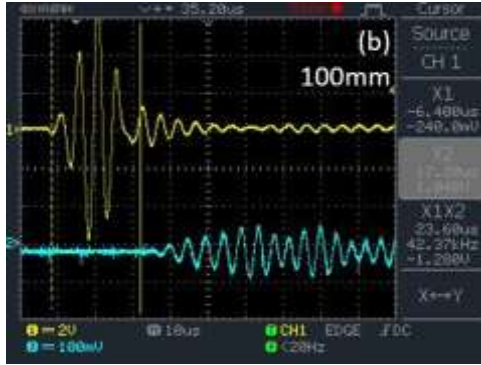
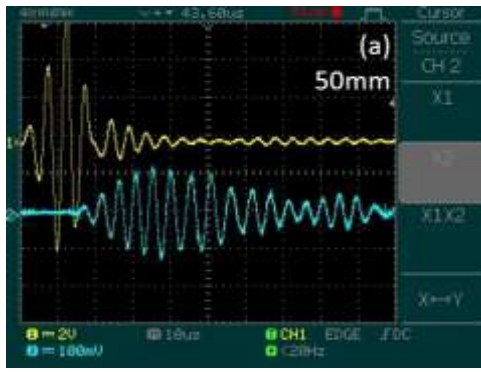


Fig 4: Oscilloscope screen showing exciter signals (above) and signals received (below) at (a) 50 mm distance and (b) 100 mm distance.

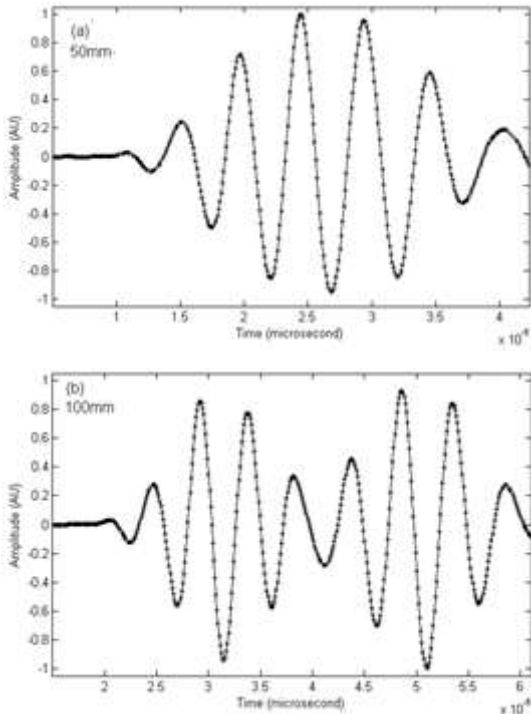


Fig 5: Numerical results of the waveform received at the sensor at (a) 50 mm distance and (b) 100 mm distance.

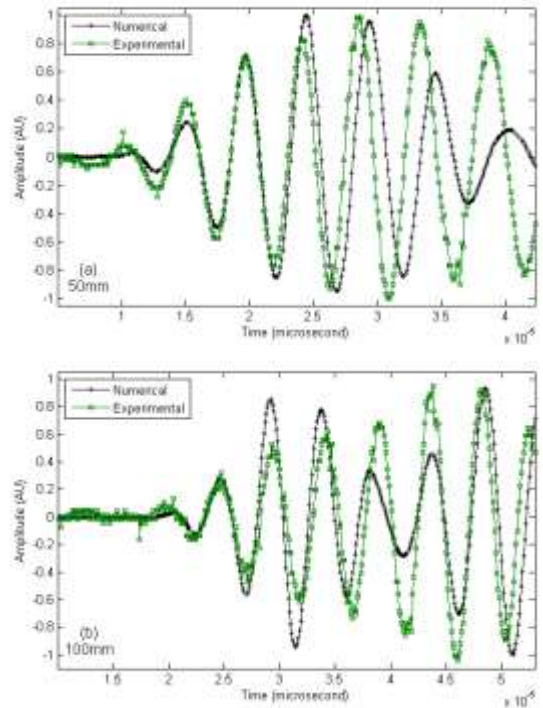


Fig 6: Comparison of the experimental and numerical results for (a) 50mm and (b) 100 mm.

V. CONCLUSION

The comparison of experimental results and finite element model of GWUT was presented in the paper. The experiments was done by using two piezoelectric discs attached on the surface of aluminum plate. A numerical model was made on COMSOL Multiphysics® software. The numerical model was verified and validated with the experimental results and found acceptable to predict the behavior of guided wave propagation at different geometries.

References

- [1] Y.-K. An and H. Sohn, "Integrated impedance and guided wave based damage detection," *Interdiscip. Integr. Asp. Struct. Health Monit.*, vol. 28, pp. 50–62, Apr. 2012.
- [2] J. L. Rose, *Ultrasonic guided waves in solid media*. Cambridge University Press, 2014.
- [3] P. W. Loveday, "Guided wave inspection and monitoring of railway track," *J. Nondestruct. Eval.*, vol. 31, no. 4, pp. 303–309, 2012.
- [4] T. Wandowski, P. H. Malinowski, and W. M. Ostachowicz, "Circular sensing networks for guided waves based structural health monitoring," *Mech. Syst. Signal Process.*, vol. 66–67, pp. 248–267, Jan. 2016.
- [5] V. Giurgiutiu, "Lamb Wave Generation with Piezoelectric Wafer Active Sensors for Structural Health Monitoring," in *SPIE's 10th Annual International Symposium on Smart Structures and Materials and 8th Annual International Symposium on NDE for Health Monitoring and Diagnostics*, 2-6 March 2002, San Diego, CA. paper # 5056-17, 2002.

- [6] B. Lin, V. Giurgiutiu, P. Pollock, B. Xu, and J. Doane, "Durability and survivability of piezoelectric wafer active sensors on metallic structure," *AIAA J.*, vol. 48, no. 3, pp. 635–643, 2010.
- [7] V. Giurgiutiu, *Structural health monitoring with piezoelectrical wafer active sensors*. Academic Press, 2007.
- [8] J. N. J.H. Nieuwenhuis and I. J. O. D.W. Greve, "Generation and detection of guided waves using PZT wafer transducers," *Ultrason. Ferroelectr. Freq. Control IEEE Trans. On*, vol. 52, no. 11, pp. 2103–2111, Nov. 2005.

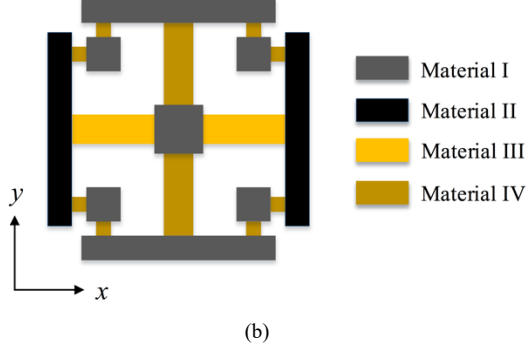


Figure 1. The two-dimensional anisotropic elastic metamaterial, (a) periodic cells, (b) the representative cell.

The representative cell of the proposed model, as shown in Fig. 1(b), is in square shape with length of L . For the current cell, the upper and lower boundaries are built with material I, which is different from material II for the left and right boundaries, and its middle components in rectangle shapes are also built with different materials, III and IV, as depicted in Fig. 1. It should be mentioned that materials I and II have higher mass densities and higher stiffnesses. Materials III and IV have lower mass densities and lower stiffnesses. The five mass resonators in the cell mainly feature translational motions and interact in terms of the displacements and forces with the four boundaries. This proposed anisotropic elastic metamaterial model can exhibit different effective mass densities and stiffnesses along x and y directions.

III. NUMERICAL ANALYSIS

In this section, the proposed anisotropic elastic metamaterial system is analyzed using Finite Element Method (FEM) with the commercial software package COMSOL Multiphysics 5.2a. The wave propagation in this model is numerically evaluated in detail. As a typical example, the effect of the mass of material I on its dynamic behavior is investigated.

A. FEM model

The geometrical parameters used for the numerical analysis of the representative cell are listed in Table 1. Based on the requirements on the masses and stiffnesses of the components, materials have been chosen with properties shown in Table 2. Specially, Steel, Lead and Tungsten are used as the material I for three cases. In the FEM model, quad elements are utilized with element size being around $1.0 \times 10^{-2} L$.

TABLE 1. GEOMETRICAL PARAMETERS

	L	L_1	H_1	LM_1	LM_2
Values (mm)	50.0	40.0	4.0	10.0	7.6
	LK_1	HK_1	LK_2	HK_2	R
Values (mm)	16.0	5.5	4.7	3.0	12.5

TABLE 2. MATERIAL PARAMETERS [12,13]

	Material	Density (Kg/m ³)	Young's modulus (Pa)	Poisson's ratio
I	Steel	7,850	210E9	0.29
I	Lead	11,340	16E9	0.44
I	Tungsten	17,800	3.6E11	0.28
II	Aluminum	2,700	7.0E10	0.33
III	Al-SiC foam	72	7.0E7	0.30
IV	Polyethylene foam	115	8.0E6	0.33

B. Dispersion relation

The dispersion relation of the elastic metamaterial model is determined by conducting modal analysis of the periodic metamaterial system using COMSOL. The attention is paid to the dynamic behaviour of the model along x (ΓX) and y (ΓY) directions. The dispersion curves along ΓX direction and ΓY direction are plotted in black dots for the case with Steel as material I in Fig. 2. It should be mentioned that, for the horizontal axis, along ΓX ($\Gamma \rightarrow X$) direction $K_1 L$ varies from 0 to π with $K_2 L = 0$, and for the vertical axis, along ΓY ($\Gamma \rightarrow Y$) direction $K_2 L$ varies from 0 to π with $K_1 L = 0$, where K_1 and K_2 are the wave numbers in x and y directions, respectively. Along ΓY direction, it can be observed in Fig. 2 that several top branches are under an almost constant frequency, indicating very low energy transmission for these branches. A broad bandgap along ΓY direction is generated, from a low frequency, 845.6 Hz, to infinity. There are also several other bandgaps located in lower frequency ranges along ΓX and ΓY directions. However, their widths are too narrow to be significantly important.

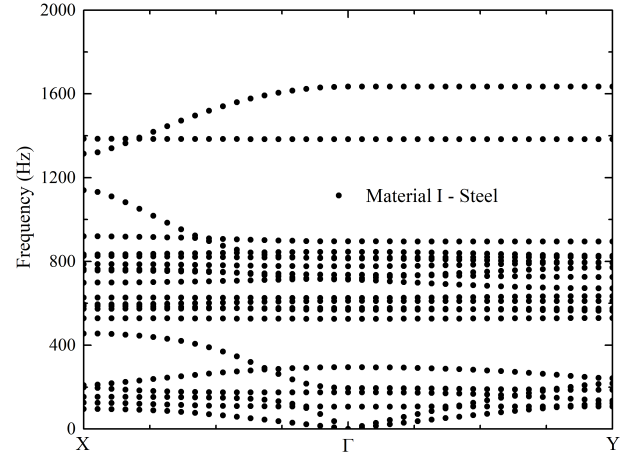


Figure 2. The dispersion relation of the anisotropic elastic metamaterial with Steel as material I.

Fig. 3 shows the dispersion curves with Steel, Lead and Tungsten as the material I, respectively. To clearly demonstrate the broad bandgaps, some insignificant dispersion curves in the lower frequency range are omitted. It can be observed from Fig. 3 that as the density of material I increases, the frequencies of the first top branch with non-zero slope along ΓY direction decrease. The case with tungsten as material I has the widest band gap along ΓY direction among these three configurations, which are from 635.6 Hz to infinity.

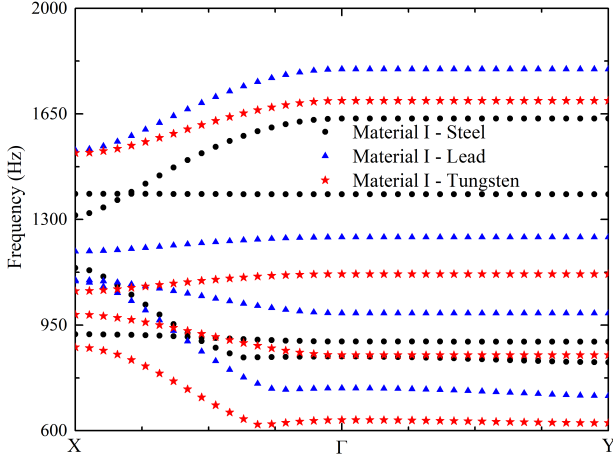


Figure 3. The comparison of the dispersion relation of the anisotropic elastic metamaterials with different material I.

C. Transmission analysis

To investigate the wave filtering ability of the proposed anisotropic elastic metamaterials, two periodic systems, as shown in Fig. 4, have been built for transmission simulations along the two principal directions, x and y . Ten representative cells are embedded in the matrix and external horizontal/vertical harmonic displacements are applied on the left sides to provide longitudinal/transverse (P/S) wave inputs. Perfect matched layers (PML) have been added at the two ends of the models to absorb outgoing waves and periodic boundary conditions have been applied at the upper and lower edges. The output displacements at the data lines are collected to calculate the corresponding transmission ratios.

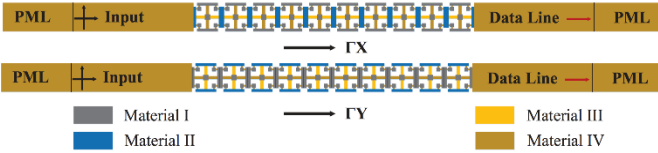


Figure 4. The numerical set-up for transmission analysis along ΓX and ΓY directions.

The transmission ratio of the proposed model under displacement input along ΓX direction has been plotted in Fig. 5 for the case where Tungsten is used as material I. For horizontal displacement input, the frequency ranges with very low transmission ratios match the ranges of the band gaps in the corresponding dispersion relation. For vertical displacement input, it can be observed in Fig. 5 that the displacement output is very small, indicating that transverse waves cannot propagate through the model along ΓX direction at current frequency range. To further illustrate the wave propagation in this model, Fig. 6 shows the displacement distribution of the metamaterial system with elastic wave propagating along ΓX direction based on the numerical simulation set-up depicted in Fig. 4. Here, red color represents high displacement amplitude and blue color indicates very low displacement amplitude. As shown in Fig. 6, longitudinal wave at frequency 300 Hz can propagate through the system, whereas longitudinal wave at frequency 1000 Hz is almost completely blocked. As expected, the transverses at these two frequencies cannot propagate in the system. In this regard,

the proposed anisotropic elastic metamaterial model behaves like a fluid along ΓX direction and this unusual property can be implemented to efficiently filter transverse waves.

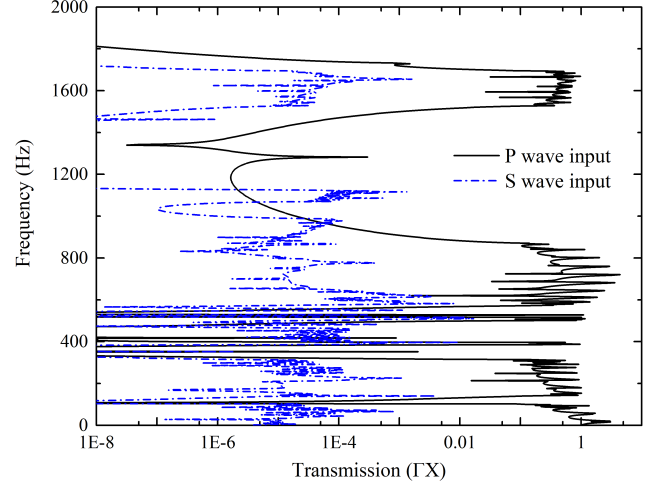


Figure 5. The transmission ratio along ΓX direction.

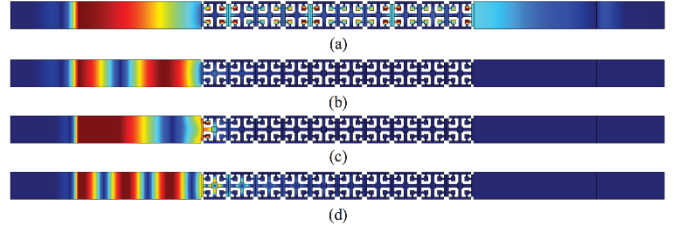


Figure 6. The displacement distribution along ΓX direction at, (a) 300 Hz for P wave input, (b) 1000 Hz for P wave input, (c) 300 Hz for S wave input, and (d) 1000 Hz for S wave input.

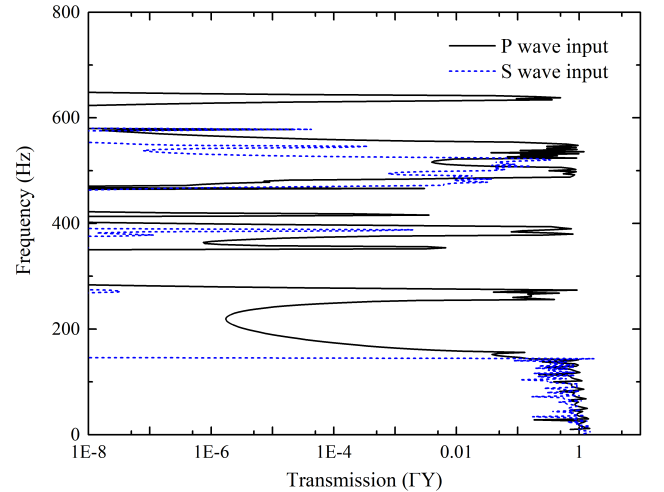


Figure 7. The transmission ratio along ΓY direction.

The transmission ratio of the proposed anisotropic elastic metamaterial model along ΓY direction is plotted in Fig. 7. There are several frequency ranges with very low transmission ratios for horizontal and vertical displacement inputs. These frequency ranges are consistent with the band gaps in the dispersion relation. The displacement distribution of the

metamaterial system for elastic wave propagating along ΓY direction is illustrated in Fig. 8, for exciting frequencies of 100 Hz and 600 Hz, which are located in the passing band and band gap, respectively. As expected, the phenomena of wave propagation and wave filtering are clearly observed. The simulation results indicate that the proposed anisotropic elastic metamaterial models have broad bandgaps for elastic waves propagating along ΓY direction, which range from very low frequencies to infinity, and this property can be efficiently used to attenuate elastic waves.

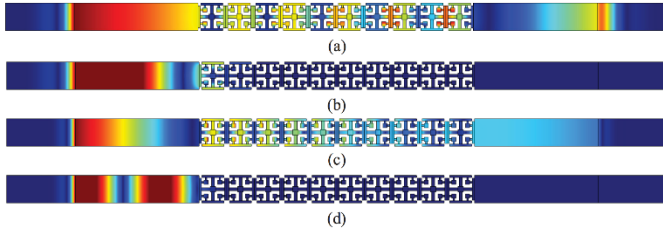


Figure 8. The displacement distribution along ΓY direction at, (a) 100 Hz for P wave input, (b) 600 Hz for P wave input, (c) 100 Hz for S wave input, and (d) 600 Hz for S wave input.

IV. CONCLUSIONS

In this paper, elastic wave propagation in new anisotropic elastic metamaterials is numerically investigated. The dispersion relation of this material is analyzed and the corresponding transmission simulation is conducted. The numerical results clearly demonstrate that broad bandgaps are generated in the low frequency ranges, the widths of which could be extended with heavy material I. Specially, in comparison to the existing anisotropic elastic metamaterials, the proposed material can behave like a solid by supporting the wave propagation of longitudinal and transverse waves, and can also behave like a fluid in which only longitudinal waves can travel. The results presented in this paper could provide a new potential avenue for designing anisotropic elastic metamaterial model for wave filtering.

ACKNOWLEDGMENT

This research is supported by the Natural Science and Engineering Research Council of Canada.

REFERENCES

- [1] S. Cummer, J. Christensen and A. Alu, "Controlling sound with acoustic metamaterials," *Nat. Rev. Mater.*, vol. 1 (16001), March 2016.
- [2] G. Ma and P. Sheng, "Acoustic metamaterials: from local resonance to broad horizons," *Sci. Adv.*, vol. 2, February 2016.

- [3] X. Zhou, X. Liu and G. Hu, "Elastic metamaterials with local resonances: an overview," *Theor. App. Mech. Lett.*, vol. 4, pp. 041001, June 2012.
- [4] Z. Liu, X. Zhang, Y. Mao, Y. Zhu, Z. Yang, C. Chan and P. Sheng, "Locally resonant sonic material," *Science*, vol. 315 (5819), pp. 1734-1736, September 2000.
- [5] Y. Ding, Z. Liu, C. Qiu and J. Shi, "Metamaterial with simultaneously negative bulk modulus and mass density," *Phys. Rev. Lett.*, vol. 99, pp. 093904, August 2007.
- [6] S. Lee and C. Park, Y. Seo, Z. Wang and C. Kim, "Composite acoustic medium with simultaneously negative density and modulus," *Phys. Rev. Lett.*, vol. 104 (5), pp. 054301, February 2010.
- [7] Y. Lai, Y. Wu, P. Sheng, and Z. Zhang, "Hybrid elastic solids," *Nature Mater.*, vol. 10, pp. 620-624, Jun 2011.
- [8] X. Liu, G. Hu, G. Huang and C. Sun, "An elastic metamaterial with simultaneously negative mass density and bulk modulus," *Appl. Phys. Lett.*, vol. 98 (25), pp. 251907, June 2011.
- [9] S. Lee and O. Wright, "Origin of negative density and modulus in acoustic metamaterials," *Phys. Rev. B*, vol. 93 (024302), January 2016.
- [10] X. Wang, "Dynamic behaviour of a metamaterial system with negative mass and modulus," *Int. J. Solids Struct.*, vol. 51(78), pp. 1534-1541, April 2014.
- [11] Z. Li, and X. Wang, "On the dynamic behaviour of a two-dimensional elastic metamaterial system," *Int. J. Solids Struct.*, vol. 7879, pp. 174-181, January 2016.
- [12] Z. Li, H. Hu, and X. Wang, "A new two-dimensional elastic metamaterial system with multiple local resonances," *Int. J. Mech. Sci.*, vol. 149, pp. 273-284, December 2018.
- [13] Z. Li, C. Wang and X. Wang, "Modelling of elastic metamaterials with negative mass and modulus based on translational resonance," *Int. J. Solids Struct.*, vol. 162, pp. 271-284, May 2019.
- [14] N. Kaina, F. Lemoult, M. Fink and G. Lerosey, "Negative refractive index and acoustic superlens from multiple scattering in single negative metamaterials," *Nature*, vol. 525, pp. 77-81, September 2015.
- [15] Y. Liu, X. Su and C. Sun, "Broadband elastic metamaterial with single negativity by mimicking lattice system," *J. Mech. Phys. Solids*, vol. 74, pp. 158-174, January 2015.
- [16] G. Milton and J. Willis, "On modifications of Newton's second law and linear continuum elastodynamics," *P. Roy. Soc. A*, vol. 463, pp. 855-880, January 2007.
- [17] Y. Gu, X. Luo and H. Ma, "Low frequency elastic wave propagation in two dimensional locally resonant phononic crystal with asymmetric resonator," *J. Appl. Phys.*, vol. 105, pp. 044903, February 2009.
- [18] A. Liu, R. Zhu, X. Liu, G. Hu and G. Huang, "Multi-displacement microstructure continuum modeling of anisotropic elastic metamaterial," *Wave motion*, vol. 49 (3), pp. 411-426, January 2012.
- [19] R. Zhu, Y. Chen, Y. Wang, G. Hu and G. Huang, "A single-phase hyperbolic metamaterial with anisotropic mass density," *J. Acoust. Soc. Am.*, vol. 139, pp. 3303-3310, June 2016.
- [20] Z. Li, C. Wang and X. Wang, "Dynamic behaviour of a two-dimensional elastic metamaterial with eccentric local resonators," *The 8th International Conference on Metamaterials, Photonic Crystals and Plasmonics*, July 25-28, 2017, Incheon Korea.

Transportation Systems

MULTIBODY DYNAMICS IN JULIA

BP Minaker PhD PEng

Department of Mechanical, Automotive, & Materials Engineering
 University of Windsor
 Windsor, Ontario, Canada
bminaker@uwindsor.ca

ABSTRACT

Julia is a new high level dynamically typed programming language designed for high-performance numerical analysis and computational science, without the typical need of separate compilation to be fast. It is open source and freely available online. The release candidate for Julia 1.0 (Julia 1.0.0-rc1) was released on 7 August 2018 and the final version a day later. Julia is developed at the Massachusetts Institute of Technology (MIT), and its creators[1] have recently been awarded the James H Wilkinson prize by the Society for Industrial and Applied Mathematics (SIAM) for contributions to numerical software. Julia has a wide variety of toolboxes contributed by the numerical software community; the differential equations toolbox is notable for the wide variety of problems that can be solved and algorithms that have been implemented.

The software ‘Equations of Motion’ (EoM), developed by University of Windsor Vehicle Dynamics and Control Research Group, can be used to generate linear or linearized equations of motion for mechanical systems, and is particularly well suited to vehicle dynamics[2]. EoM is also entirely open source, is freely available online at the hosting site github.com, and has recently been ported to the Julia language. This paper discusses the EoM software, and its implementation in Julia.

When analyzing a multibody system, EoM will read the geometry, inertia, stiffness and constraint information from the input data, and build the necessary stiffness and constraint Jacobian matrices required to find all the preload and constraint forces. Once these are known, the stiffness matrices are updated with the tangent stiffness terms. The kinematic differential equations relating position and velocity are then combined with the Newton-Euler equations of motion, reduced to a minimal coordinate set, and cast in descriptor state space form, before being further reduced to an equivalent standard state space form. From this state matrix form, EoM will perform a linear analysis, i.e., computing the modal analysis, and frequency response.

In EoM, each component of a system is defined as an “item”. The most basic type of item is a “body”, which defines a generic rigid body. The other types of items define either rigid or flexible connectors between the rigid bodies. Available types of items are listed in Table 1.

Table 1: Available types of items in EoM

Type of item	Definition
body	a rigid body
spring	a two-point elastic spring, with linear or torsional stiffness and damping, non-zero free length
link	a two-point massless rigid link
rigid_point	a generic point constraint with a variable number of constraint forces and moments
flex_point	a point spring with translational and/or rotational stiffness and damping
nh_point	a non-holonomic constraint to prevent velocity but not displacement
beam	a beam spring with mass, inertia, bi-directional bending and shear stiffness
load	constant forces or moments applied to the system
actuator	applied force or moment, proportional to an input signal
sensor	used to measure displacement, velocity, or acceleration

The dynamic typing system in Julia allows each EoM item type to be defined as a data structure, greatly simplifying the process of gathering and sorting data during system definition. The paper provides details on the specifics of the implementation, and example systems are analyzed using EoM and Julia.

[1] J. Bezanson, A. Edelman, S. Karpinski and V.B. Shah, “Julia: A Fresh Approach to Numerical Computing”, *SIAM Review*, 59: 65–98, 2017.

[2] B.P. Minaker, “The tangent stiffness matrix in rigid multibody vehicle dynamics”, *Mathematical and Computer Modelling of Dynamical Systems*, 2014

3D Random Road Implementation in Motionview® Full Car Simulation

Zhe Ma

Mechanical, Automotive, Material Engineering
 University of Windsor
 Windsor, Canada
 mal1r@uwindsor.ca

Bruce Minaker

Mechanical, Automotive, Material Engineering
 University of Windsor
 Windsor, Canada
 bminaker@uwindsor.ca

ABSTRACT

Numerical prediction of vehicle performance under different road conditions has proven to be a useful tool for vehicle suspension development. A random road is good description of road roughness for vehicle dynamics simulation in a certain moving direction. A random road with different roughness can be generated by using a sum of a series of sinusoidal waves of random phase angle to form the final random road profile. The waves have an amplitude that decreases with increasing frequency according to a specific pattern based on the desired roughness. The road roughness level K is defined by the PSD value of the road profile at a frequency equal to 0.1 cycles/m, according to the ISO 8606 road roughness classification. Table 1 shows the level K and PSD value at different road class.

Table 1: ISO 8608 Road roughness classification

Road Class	K	PSD at 0.1 cycles/m [10 ⁻⁶ m ³ /cycle]
A-B	3	32
B-C	4	128
C-D	5	512
D-E	6	2048
E-F	7	8192
F-G	8	32768
G-H	9	131072

Here, a class A road is very smooth, and class H is extremely rough, off-road style terrain. Sample random roads generated in this fashion are shown in Figure 1.

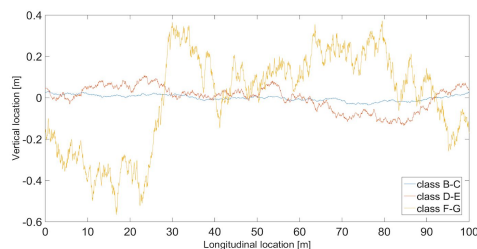


Figure 1: Sample random roads, level K= 4, 6, 8

When considering the complex motion of a full vehicle model during a simulation, a 2D random road does not represent the condition of a real road surface very well. A 3D random road will be a better representative of the real road, especially when the vehicle's moving direction is changing with time. By using OpenCRG® MATLAB® code, a 3D random road surface can be formed by summing two 2D random road profiles in different axes. In this paper, a 120x120 meters random road surface has been generated by combining two level 6 2D random road profile.

The new 3D random road has been implemented in Altair Motionview®, and used for multi-body vehicle dynamic modeling. A full car model was built using the Motionview templates, and its parameters were modified to meet the specific requirements of the project.

A number of different simulations were conducted using this arrangement. To illustrate the effect of the 3D random road, trials along a number of different vehicle paths have been performed, and the tire contact patch location in the z direction for each wheel has been reported. By using these data and some special purpose MATLAB® code, the PSD vs frequency plots are generated. From these PSD results, the roughness level of formed 3D road can be compared against the individual 2D road profile. Results from the 3D random road simulation are given in the paper. As a example, the PSD vs frequency plot of steady steering case of front left wheel is shown in Figure 2.

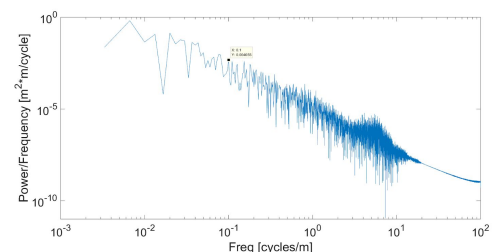


Figure 2: PSD of 3D random road, circular path

Control-Oriented Dynamic Model of a Three-way Catalytic Converter Utilizing Orthogonal Collocation

Amer Keblawi
Systems Design Engineering
University of Waterloo
Waterloo, Canada
akeblawi@uwaterloo.ca

John McPhee, PhD
Systems Design Engineering
University of Waterloo
Waterloo, Canada

ABSTRACT

A control-oriented model of an Otto-engine catalytic converter is an indispensable component for a complete engine emissions and fuel efficiency model. Such a catalytic converter model may be coupled with a mean value engine model to develop engine controllers that include engine emissions constraints. For this model, one dimensional PDEs (Partial Differential Equations) are utilized to obtain solutions of exhaust gas and substrate temperatures and concentrations (1, 3, 4). Orthogonal Collocation is utilized to obtain real time solutions along the one dimensional space. The main contribution in this work is to obtain real time solutions for the aforementioned states taking into consideration both spatial variation and time dynamics for all states in the model.

The considered catalytic converter model consists of a PDE representing the exhaust air temperature, a PDE representing the substrate temperature, a set of PDEs representing gas species concentrations, a set of PDEs representing substrate species concentrations, and an ODE (Ordinary Differential Equation) representing the oxygen storage inside the catalytic converter. In this study, six species are considered resulting in a system consisting of a total of 14 PDEs and an ODE in time. As an extension to the models mentioned (1, 3, 4), spatial variation and time dynamics are considered for all states. In the catalytic converter, chemical reactions occur between the different species; 15 such reactions are considered in this analysis. The rates of these reactions are of a first order dynamic nature following Arrhenius equation. These rates and the rate of species consumptions are state-dependent source terms that are included in the partial differential equation system of the model.

In orthogonal collocation, an approximation function is used to approximate the real solution. The approximation function is constructed such that the value of the function is equal to the real solution at the considered approximation nodes. For this study, the approximation function is constructed from Lagrange polynomials spanning the spatial domain and the nodes are considered to be the roots of the Legendre polynomial spanning the spatial domain. This method is applied over the whole spatial domain of the model transforming the system of PDEs into a system of ODEs. In this study, 5 such polynomials are used to transform the original PDE system into a system of 71 ODE equations. Further details of such an implementation is presented in Hassanpour & McPhee (2) for diesel particulate filter models.

After applying orthogonal collocation, it is noticed that the computer solution time is within 2 seconds for ten thousand seconds of simulation time on a standard desktop computer. This solution is much faster than real time computations in most systems without sacrificing the time dynamics of any state, resulting in a real time, high-fidelity, physics-based model that is suitable for non-linear, multi-objective control analysis and controller design.

ACKNOWLEDGEMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and MapleSoft Inc.

REFERENCES

1. Andrianov, D. I., Brear, M. J., & Manzie, C. (2012). A Physics-Based Integrated Model of a Spark Ignition Engine and a Three-Way Catalyst. *Combustion Science and Technology*, 1269-1301.
2. Hassanpour, S., & McPhee, J. (2018). A control-oriented modular one-dimensional model for wall-flow diesel particulate filters. *International Journal of Engine Research*, 329-346.
3. Pontikakis, G. N., Konstantas, G., & Stamatelos, T. (2004). Three-Way Catalytic Converter Modeling as a Modern Engineering Design Tool. *Journal of Engineering for Gas Turbines and Power*, 906-923.
4. Ramanathan, K., & Sharma, C. S. (2011). Kinetic Parameters Estimation for Three Way Catalyst Modeling. *Industrial & Engineering Chemistry Research*, 9960-9979 .

Comparison of Different Steering Ratios on Steering by Wire Vehicle

Huiyong Zhao

School of Automotive Engineering
Hubei University of Automotive Technology
Shiyan City, China
email: zhohiyong@qq.com

Guangde Zhang

School of Automotive and Traffic Engineering
Wuhan University of Science and Technology
Wuhan City, China
email: 1007513365@qq.com

Baohua Wang

School of Automotive Engineering
Hubei University of Automotive Technology
Shiyan City, China
email: wbbhenz@126.com

Mutaz keldani

Department of Automotive, Mechanical and
Manufacturing Engineering
University of Ontario Institute of Technology
Oshawa, Canada
email: mutaz.keldani@uoit.net

Abstract—To find variable steering ratio calculating method, a steering by wire (SBW) vehicle is modelled and several steering ratio calculating methods are designed and tested with the co-simulation of Carsim and Simulink. Firstly, the SBW vehicle model is established by means of replacing the steering system of Carsim vehicle model with an SBW system model, which is built with components of Simulink/simscape library in Simulink. This SBW system consists of the steering gear ratio calculation model, the electric drive steering feel model, steering gear model and their PID control model. Then, the ideal gear ratio calculating method is discussed and a new algorithm based on PID controller is added to in the steering gear ratio calculation model. Finally, a double lane change and 90-degree road curve conditions are simulated with the co-simulation method. Simulation results indicate that the steering gear ratio calculated with the algorithm based on the PID controller is superior to others in reducing the lateral acceleration and making the gain of yaw rate dividing steering wheel angle near the target value.

Keywords—variable steering gear ratio, algorithm, PID controller, simulation, SBW

I. INTRODUCTION

Steer by wire system (SBW) removes the conventional mechanical linkages between the steering wheel and the front wheel, and controls the vehicle steering according to vehicle status and driver's commands. This system has advantages of both easy spatial layout and improving handling performance of the vehicle [1]. So, there are many researches on the SBW vehicle, which focus on the field of active steering, variable

ratio design, road feel simulating, fault redundancy, and so on [2~7].

As one part of the vehicle's handling performance, driver's road feedback feeling is related to steering system parameters, which includes the gain of yaw rate, steering wheel angle and lateral acceleration of the car. The former is related to the vision of the driver, and the latter is related to the feel of the driver [8]. Those parameters could be adjusted by changing the steering ratio of SBW vehicle. Some researchers proposed several types of variable gear ratio design schemes [9,10]. Shi G.B et al. put up three control programs of ideal steering ratio and compared their influence on the handling and stability performance of the steering by wire vehicle [11]. Zheng H.Y. et al. presents a steady control strategy for the ideal steering ratio and a stability control algorithm to correct the steering angle dynamically [8]. Arslan et al. adopted an energy optimization control method to analyze the handling characteristic of vehicle in the scenarios of double lane change [12]. But their researches are mostly open loop simulation with specific linear vehicle model. Little attention has been paid to closed loop simulation to the author's knowledge.

This paper tries to establish SBW vehicle model with the co-simulation of Simulink and Carsim and put out a variable gear ratio algorithm based on PID controller after comparing the influence of different gear ratio. The remainder of the paper is organized as follows. Section 2 introduces the modelling of the SBW vehicle; Section 3 describes the structure and the model of SBW controller, and designed the steering gear ratio model with ideal gear ratio calculating method and the algorithm based on PID controller; Section 4 presents the

vehicle, $\dot{\phi}$ and a_y are the yaw rate and the lateral acceleration, respectively.

A. Structure of SBW system

As shown in Fig.1, SBW system has two mechanical subsystems and one controller, which are road feel unit, steering actuator unit and SBW controller. The road feel unit measures the steer angle signal and provides road feel torque by road feel motor. Steering actuator unit steers front wheels to a certain angle by steering motor according to the steering angle signal received from the controller. The SBW controller calculates road feel torque and steering angle according to received signals, which also controls the road feel motor and steering motor to provide a certain torque or rotate a certain angle. Received signals are status signals sent from the vehicle controller, which are current signals of road feel motor and steering motor, and the angle and angular speed signals of road feel motor and steering motor.

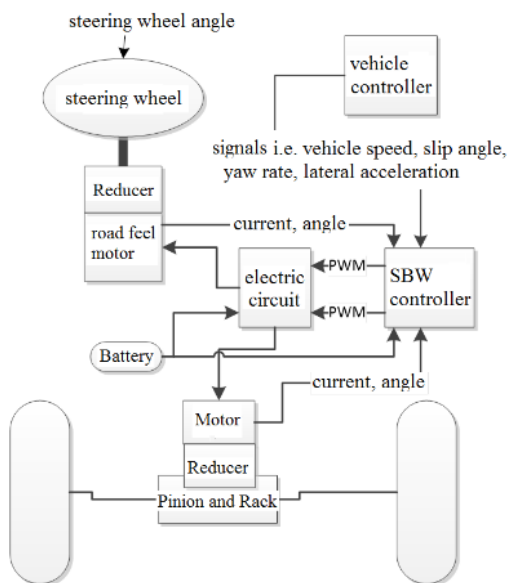


Figure 1. Structure diagram of SBW

The simulation model is composed of three closed loop models, which are driver model loop, road feel loop, steering unit loop, as shown in the Fig.2. Where, the road information, driver model and vehicle are provided by the Carsim, road feel model, controller model and electric driver steering model is built with the Simulink software. In Fig.2, y_{ideal} and y_{real} are denoted as the ideal and real lateral displacement, respectively, δ_{sw} are denoted as the steering wheel angle, U_m and U_{ms} are denoted as the voltage of the road feel motor and the steering motor, respectively. T_{swr} and T_{pin} are denoted as the road feel torque and the resistance torque of the pin, respectively, δ_{str} is the real steering motor angle, v is the longitudinal speed of the

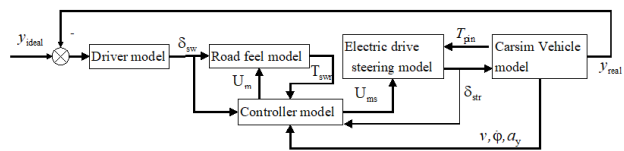


Figure 2. Schematic diagram of SBW vehicle simulation model

B. Road Feel Unit Modelling

As shown in Fig.1, steering wheel is linked to the road feel motor through steering column and reducer to compose the road feel unit. This unit could be built with dynamic equations or Simulink/simscape components in Simulink. As Simulink/simscape model could display the structure during the modeling, its components are used to build the motor and reducer, as shown in Fig.3. The input signals are the angle of the steering wheel and the voltage of the road feel motor, while the output signal is the real output torque of road feel, as shown in the Fig.3.

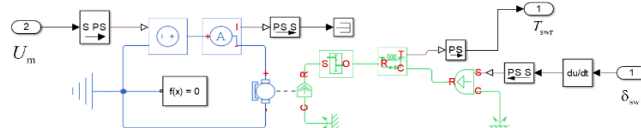


Figure 3. Road feel model based on Simulink/simscape components

In the simulation model, the value of the reducer ratio of road feel motor is set to 10, and parameters of the motor is shown in table 1.

TABLE I. PARAMETERS OF DC MOTOR

Power (W)	armature winding resistance (Ω)	Torque constant	unload current (A)	unload rated rotating speed (rpm)	maximum torque (Nm)	Average Torque (Nm)
600	0.15	0.0625	1.5	2400	5	3.8

C. Steering Actuator Unit Modelling

The Steering actuator unit is composed of steering motor, reducer and pinion and rack steering device. Steering motor drives the pinion rotating after its torque amplified by reducer. In this simulation, damper coefficient, friction and inertia value of the reducer are very small and ignored during modelling.

(1) *Pinion and Rack*

Denote the ratio of pinion and rack as K_g (unit: mm/rad), then the relationship between output angle θ_s (unit: rad) of pinion and the displacement s (unit: mm) of rack is shown in Eq.(1). Define the relationship between the displacement of rack and the angle of front wheel as $g(s)$, and let the movement step of rack be Δ , then i_{WL2P} and i_{WR2P} , the ratio from pinion to the left side and right side of front wheel could be calculated from Eq.(2) and Eq.(3).

$$s = \theta_s \cdot K_g \quad s = \theta_s \cdot K_g \quad (1)$$

$$i_{WL2P} = \frac{1}{\frac{g(s + \Delta) - g(s)}{\Delta} \cdot K_g} \quad (2)$$

$$i_{WR2P} = \frac{1}{\frac{g(-s - \Delta) - g(-s)}{\Delta} \cdot K_g} \quad (3)$$

According to the force flow from pinion to the front wheels, the resistance torque of pinion could be calculated with alignment torques of kingpin acquired from Carsim and gear ratios calculated from Eq.2 and Eq.3. $g(s)$, the relationship between the displacement of rack and the angle of front wheels, is described with tables in Carsim vehicle. Thus, i_{pf} , the gear ratio from the pinion to the front steering axle, could be calculated with Eq.(4).

$$i_{pf} = i_{WL2P} + i_{WR2P} \quad (4)$$

(2) Steering Motor Unit

The steering motor unit, composed of steering motor and reducer, is used to drive the pinion rotating when power on. Its Simulink/simscape model is shown in Fig. 4. Where, the voltage U_{ms} and the torque T_{pin} are input signals, and the angle δ_{str} is the output signal. Here, the ratio of reducer of steering motor is set to 20.

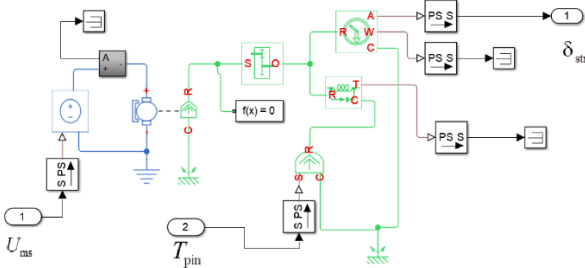


Figure 4. Power driving unit

III. CONTROLLER MODEL

In this section, the schematic diagram of the SBW controller is discussed, the method of road feel calculating is presented, and the steering gear ratio calculated with an algorithm based on the PID controller and the ideal ratio algorithm is put out.

A. Structure of the Controller

The controller model usually is composed of motor control, gear ratio calculation and adjusting and road feel torque calculation. Motor control usually adopts PI control algorithm, whose parameters could be adjusted by PID tool in Simulink. The schematic diagram of the controller is shown as in Fig.5.

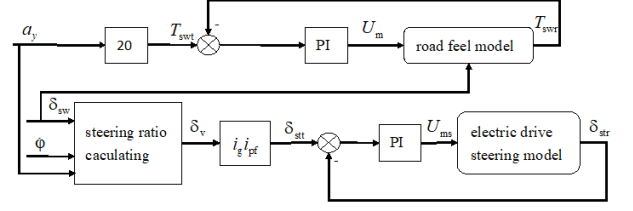


Figure 5. Schematic diagram of the controller

B. Calculation of Road Feel Torque

The road feel torque is what the driver must overcome when steering, which is mainly delivered from tire during vehicle steering. This torque is highly relevant to the lateral acceleration of the vehicle, and usually is expressed with the gradient of steering wheel angle divided lateral acceleration. The range of this gradient changes along with the speed of the vehicle, usually between 15~38 when the speed is lower than 60km/h, and between 31~18 when the speed is in the range of 60~140km/h [13]. Some researchers proposed the idea of comprehensive road fee and deduced the calculation method of road feel torque and strength after analyzing the feeling of the driver base on weber-fechner law of sensation [10].

In this paper, this gradient is set to a fixed value 20Nm/g to simplify the modelling and simulating of the system. Thus, the target road feel torque could be calculated with Eq.(5) [12].

$$T_{swt} = 20\dot{a}_y \quad (5)$$

where, a_y is the lateral acceleration, T_{swt} is the target road feel torque.

C. Variable Steering Gear Ratio Design

(1) Ideal Variable Gear Ratio

According to the theory of the vehicle dynamics [13], the bicycle model of the vehicle is easily deduced, as is shown in Eq. (6).

$$\begin{bmatrix} mv & 0 \\ 0 & J_z \end{bmatrix} \begin{bmatrix} \dot{\beta} \\ \ddot{\phi} \end{bmatrix} = \begin{bmatrix} -C_{av} - C_{ah} & \frac{-C_{av} \cdot l_v + C_{ah} \cdot l_h}{v} - mv \\ -C_{av} \cdot l_v + C_{ah} \cdot l_h & \frac{-C_{av} \cdot l_v^2 - C_{ah} \cdot l_h^2}{v} \end{bmatrix} \begin{bmatrix} \beta \\ \phi \end{bmatrix} + \begin{bmatrix} C_{av} \\ C_{av} \cdot l_v \end{bmatrix} \delta_v \quad (6)$$

The fixed gain of yaw rate dividing steering angle has advantages of preventing the real gear ratio declining and avoiding the increasing of under steering characteristic when the speed exceeds characteristic speed. So, defining the gain of yaw rate dividing steering angle as K_ϕ , its value could be calculated with Eq.(7) [13].

$$K_\phi = \frac{\dot{\phi}}{\delta_{sw}} \quad (7)$$

Thus, the relationship between δ_{sw} and δ_v could be deduced with Eq.(8).

$$\begin{aligned} \delta_v &= \frac{1}{b_1} s + \frac{-(a_{11} + a_{22}) - b_2/b_1}{b_1} + \frac{a_{11}a_{22} - a_{12}a_{21} + (\frac{b_2}{b_1})^2 + (\frac{a_{11} + a_{22}}{b_1})b_2}{b_1 s + b_2} \quad (8) \\ &= \frac{1}{i_{sw}} \end{aligned}$$

where, $b_1, b_2, a_{11}, a_{12}, a_{21}, a_{22}$ are coefficient numbers, whose value could be calculated with equations in the Appendix.

The variable gear ratio calculated with Eq.(8) is changed dynamically with the changing of δ_{sw} and the speed of the vehicle. While its steady value is only related with the velocity of the vehicle.

(2) Algorithm based on PID controller

As is known that most vehicle parameters are difficult to be acquired, a PID control method is a good way to substitute for the ideal gear ratio. The idea of this method is using the error between the calculated parameter K_ψ and the given value, such as 0.2. With the PID control algorithm, the relationship between the error and the value of the variable gear ratio is established.

To extend the range of the gear ratio i_{sw} , the result of the PID algorithm is limited to the range of (-2, 2), and an exponential function is used to amplify this result. Thus, this ratio i_{sw} could be calculated according to Eq.(9) and Eq.(10).

$$x = \begin{cases} 2 \\ K_p(K_\psi - 0.2) + K_i \int (K_\psi - 0.2) dt + K_d \dot{K}_\psi \\ 2 \end{cases} \quad (9)$$

$$i_{sw} = 10^x \quad (10)$$

where, x is the output value of PID controller, K_D is proportionality constant of PID, K_I is integral constant of PID, K_P is differential constant of PID. The value of K_D, K_I and K_P is set as 3, 0.1 and 0.01, respectively.

IV. SIMULATION AND ANALYSIS

In this section, two simulating scenarios are selected to find the influence of different variable gear ratio methods on the vehicle handling performance.

A. Simulation Conditions

To find the influence of the variable gear ratio on the vehicle, two scenarios are selected in this simulation. One is the double lane condition (DLC) at the speed of 80km/h, which is used to test the influence when driver controlling the vehicle at emergency situations. The other is 90-degree road curve condition at the speed of 100km/h, which is used to test gear ratio's effect when driver operating the vehicle to pass through a 90-degree curve road. Road data of the double lane condition and the 90-degree road curve condition are shown in Fig.6 and Fig.7, respectively.

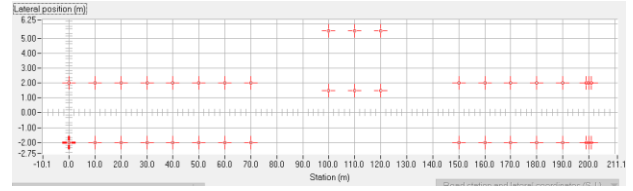


Figure 6. Road data of double lane change

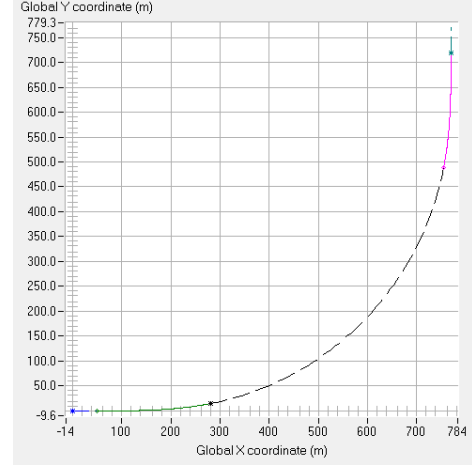


Figure 7. 90-degree road curve

In each scenario, the influence of four types of gear ratio calculating are compared, which are PID calculating, ideal gear ratio dynamically changing, ideal gear ratio steady value and fixed ratio. In the fixed ratio mode, the value of gear ratio from the steering wheel to the pinion is set to 1. The vehicle model selected in Carsim is D-class, Sedan, whose parameters are shown in the table 2.

TABLE II. VEHICLE PARAMETERS

Symbol	Description	Value
m	The weight of vehicle	1690 kg
J_z	Yaw inertia of the vehicle	233.05 kg*m ²
C_{av}	Cornering stiffness of the front tire	1874 Nm/deg
C_{ah}	Cornering stiffness of the rear tire	2624 Nm/deg
l_v	the distance from the mass center of the vehicle to the front axle	1.11 m
l_h	the distance from the mass center of the vehicle to the rear axle	1.67 m

B. Simulation Results

(1) Lateral Displacement Tracking

From Fig.8, it is easy to know that the lateral displacement error is smallest when using fixed ratio than using variable ratio, and biggest when using ratio calculated with the PID algorithm in the DLC simulation. But this error is almost the same when using the dynamic ideal ratio or the steady ideal ratio.

From Fig.9, all curves are almost the same in 90-degree road curve condition, which suggests that the vehicle have the same tracking performance in the operating mode.

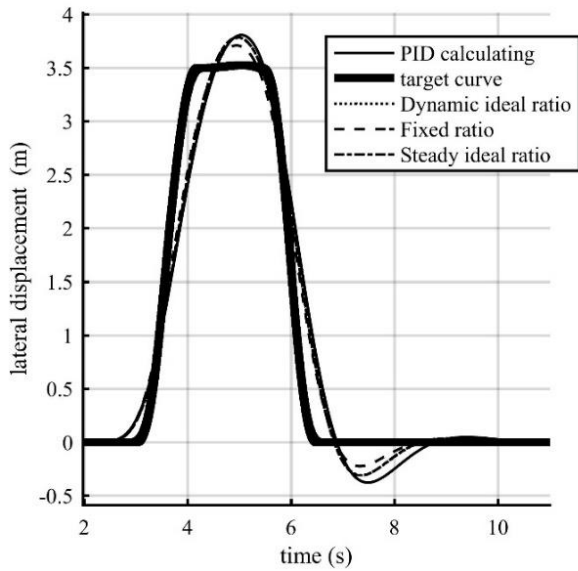


Figure 8. Lateral displacement results in DLC

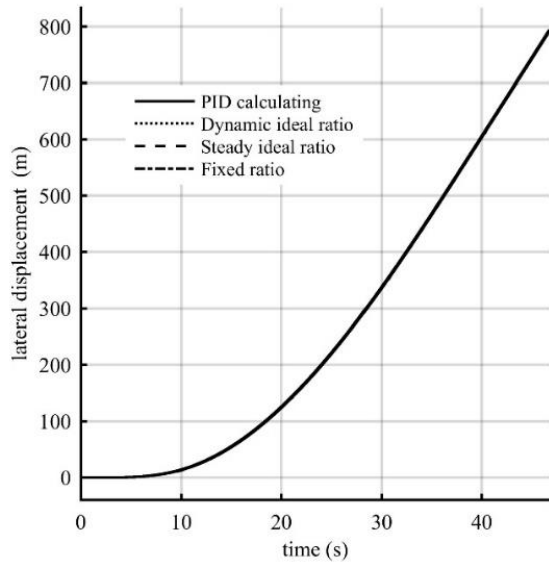


Figure 9. Lateral displacement results in 90-degree road curve condition

(2) Steering Wheel Angle

From Fig.10, the value of steering wheel angle is smallest in fixed ratio mode biggest in PID calculating ratio mode. So, adopting the PID calculating mode to get gear ratio is best at extending the range of steering wheel angle during double lane change control mode, which is also verified with Fig.11.

In the Fig.11, steering wheel angle always changes during 90-degree road curve condition scenario. Its effect on the driver should be discussed in the future research. To improve this, a

better value of K_D , K_I and K_P should be designed in the future, too.

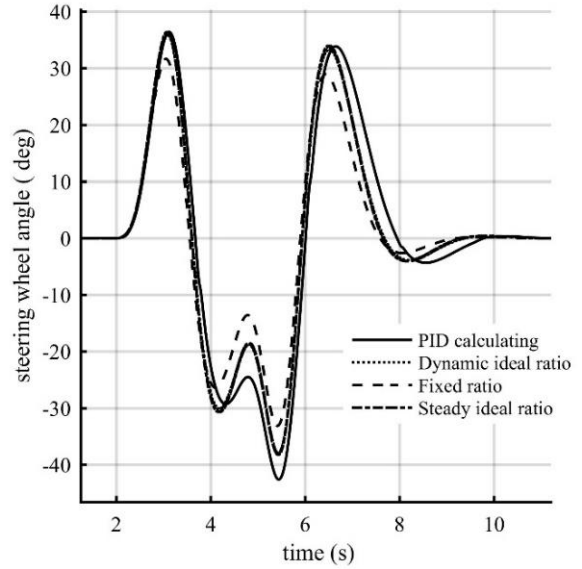


Figure 10. Steering wheel angle results in DLC

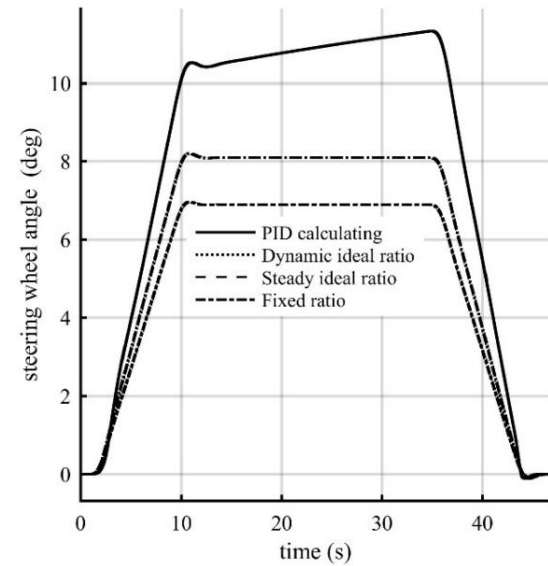


Figure 11. Steering wheel angle results in 90-degree road curve condition

(3) Response of lateral acceleration

Fig.12 suggests the lateral acceleration of the vehicle is smaller when adopting variable gear ratio mode than fixed ratio mode during double lane change control scenario. This acceleration is smallest when acquiring gear ratio with PID calculating mode than others. This could improve the rollover or sideslip performance of the vehicle. However, Fig.13 indicates that lateral acceleration of the vehicle is almost same during 90-degree road curve operation scenario.

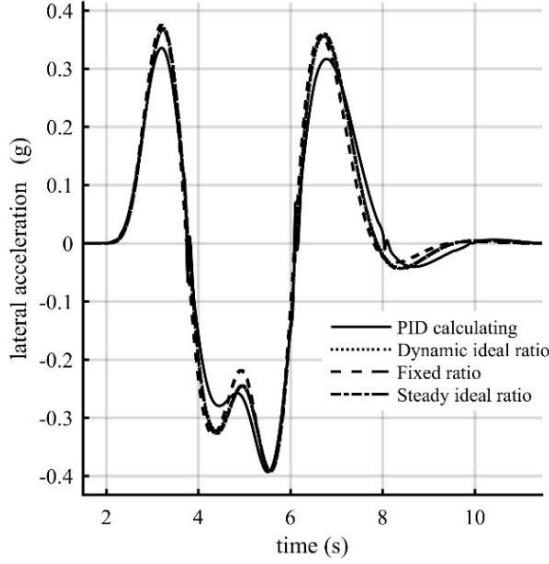


Figure 12. Lateral acceleration results in DLC

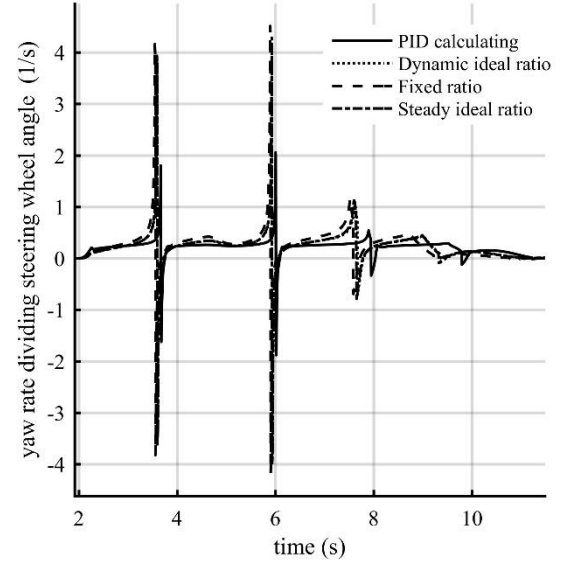


Figure 14. $\dot{\phi}/\delta_{sw}$ in DLC

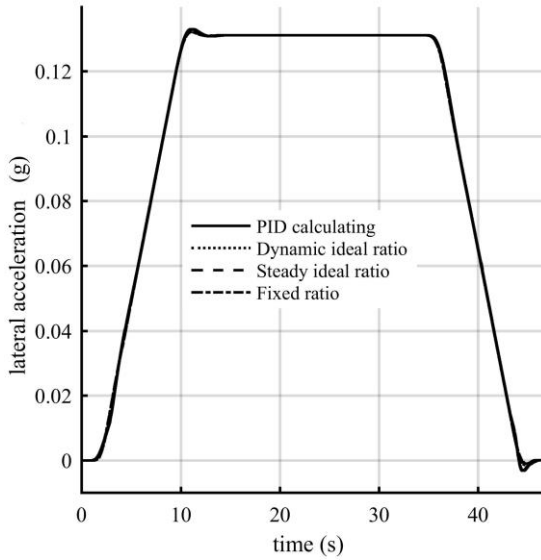


Figure 13. Lateral acceleration results in 90-degree road curve condition

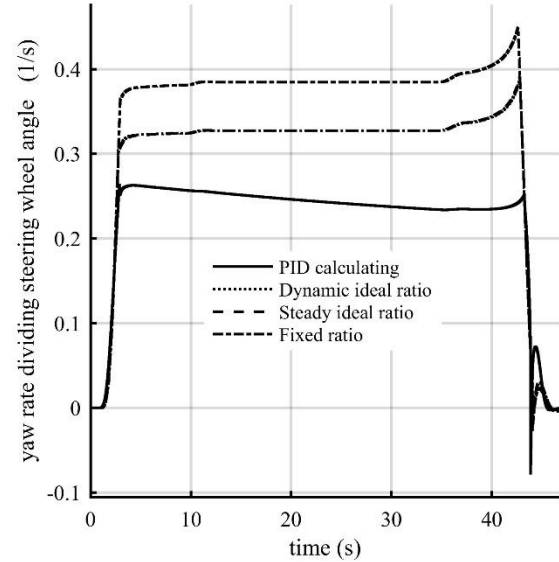


Figure 15. $\dot{\phi}/\delta_{sw}$ in 90-degree road curve condition

(4) Gain of Yaw Rate Dividing Steering Angle

Both Fig.14 and Fig.15 suggest that adopting variable gear ratio could improve the gain of yaw rate dividing the steering wheel angle greatly, and using PID calculating to acquire gear ratio do the best in making this gain near the target value. Fig.15 also suggests the ratio achieved from PID calculating has not been converged to steady state, which indicated the reason in the Fig.11 why the steering wheel angle continually changes in the 90-degree road curve condition.

V. CONCLUSIONS

In this paper, a SBW vehicle is modelled and examined with Simulink and Carsim Co-simulation, where the steering by wire system is built with components of Simulink/Simulink/simscape library to substitute the traditional steering system of Carsim vehicle model. The method of the ideal gear ratio and the ratio got from the PID algorithm are discussed. A double lane change condition and a 90-degree road curve condition are selected in the simulation. Four steering gear ratios are implemented in the simulation to test the influence on the SBW vehicle model.

Through the simulation, the findings derived from the paper are as follows:

1) the fixed steering gear ratio vehicle could have a better lateral displacement tracking, while the variable gear ratio is good at reducing the lateral acceleration under the double lane change condition and good at making the gain of yaw rate dividing steering wheel angle near the target value in both double lane change and 90-degree road curve conditions.

2) the impact of dynamic ideal ratio value on the performance of the vehicle is almost the same as the steady ideal ratio value. And in the vehicle operating mode, such as 90-degree road curve condition, the lateral acceleration of the vehicle is almost the same no matter which method is used for gear ratio calculating.

3) the steering gear ratio acquired from PID algorithm designed in this paper is superior to other methods greatly. Utilizing this method, vehicle's handling performance and lateral stability could be improved greatly.

In the future research, values of PID coefficients need to be optimized to make the gain of yaw rate fast to steady state.

ACKNOWLEDGMENT

The research is mainly supported by National Natural Foundation of China (grant No. 51175155, 51375151), Hubei Key Laboratory of Automotive Power Train and Electronic Control (grant No. ZDK1201502), Startup Foundation for Doctors of Hubei University of Automotive Technology (BK201410) and partly supported Hubei Province Collaborative Innovation Centers for Automotive Part and Assembly Technology, Automobile Conservation Technology (grant No.2015XTZX0403).

REFERENCE

- [1] C.F. Zong, G. Li, et al. Study Progress and Outlook of Chassis Control Technology for X-by-wire Automobile. China Journal of Highway and Transport, 2013, 26(2), pp160-170.
- [2] Tian Chengwei. Research on Fault Tolerant Control Method for Steer-by-wire Automobile. Jilin University Phd thesis, 2010.
- [3] Rodriguez-Angeles, A. and J.A. Garcia-Antonio, Active disturbance rejection control in steering by wire haptic systems. ISA Transactions, 2014. 53(4): p. 939-946.
- [4] S. Luo. Research on Steer-by-wire System Predictive Control Strategy for Active Safety of Vehicle. Jiangsu university Phd thesis, 2010.
- [5] L. He. Research on double motor control method of line control steering system based on FlexRay bus. Jilin University Phd thesis, 2011.
- [6] Yamaguchi, Y. and T. Murakami. Adaptive Control for Virtual Steering Characteristics on Electric Vehicle Using Steer-by-Wire System. Industrial Electronics, IEEE Transactions on, 2009. 56(5): pp. 1585-1594.
- [7] Zong, C. , Xiang, H. , He, L. , & Sha, F. . (2012). Study on control method of dual-motor for steer-by-wire system. International Conference on Consumer Electronics. IEEE.
- [8] Y. H. Zheng, C. F. Zhong, C.W. Tian. Control algorithm for steer-by-wire-system with ideal steering ratio. Journal of Jilin Univerisity (Engineering and Technology Edition) ,2007, 37(6),pp. 1229-1235.
- [9] J.W. Wei, M.X. Wei, W.Z. Zhao. Control law of varied steering ratio based on driver-vehicle-road closed-loop system. Journal of Jiangsu University(Natural Science Edition) ,2011,32(6), pp.652-657.

- [10] H.X. Xie, X. Gao, C.G. Xia. Characteristic Analysis of Steering Wheel Desired Torque Based on Comprehensive Road Feel. Mechanical Science and Technology for Aerospace Engineering, 2012, (7), pp. 1104-1109.
- [11] G.B. Shi, W.Z. Zhao et al. Influence of Variable Steering Ratio for Steer-by-Wire System on Vehicle Handling Stability. Transactions on Beijing Institute of Technology, 2008,28(3),pp. 207-210.
- [12] Arslan, M. S. and N. Fukushima. Energy optimal control design for steer-by-wire systems and hardware-in-the-loop simulation evaluation. Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME, 2015, vol.137(7).
- [13] M. Mitschke, H. Wallentowitz, "Dynamik der Kraftfahrzeuge". Berlin Heidelberg: Springer-Verlag, 2004.

APPENDIX:

$$b_1 = \frac{C_{av} \cdot l_v}{J_z \cdot K_\phi}$$

$$b_2 = \frac{a_{21} C_{av}}{mv} - \frac{C_a l_v}{J_z \cdot K_\phi} a_{11}$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \frac{-C_{av} - C_{ah}}{mv} & \frac{(-C_{av} \cdot l_v + C_{ah} \cdot l_h - mv) K_\phi}{v} \\ \frac{-C_{av} \cdot l_v + C_{ah} \cdot l_h}{J_z \cdot K_\phi} & \frac{-C_{av} \cdot l_v^2 - C_{ah} \cdot l_h^2}{J_z v} \end{bmatrix} \Bigg/ mv$$

Free Vibration Properties of Ring Gears

Jalal Taheri Kahnamouei
Department of Mechanical Engineering
Memorial University
St. John's, NL, Canada
jtk441@mun.ca

Jianming Yang
Department of Mechanical Engineering
Memorial University
St. John's, NL, Canada
jyang@mun.ca

Abstract—Ring gears are used in many applications, planetary gear trains (PGTs) in particular. The ring gear dynamic properties have a significant effect on the overall dynamics of the system with ring gears installed. In this paper, two models are developed to compute the natural frequencies and mode shapes of a ring gear. In model I the ring gear is modeled as a smooth ring, while in the model II is modeled as a smooth ring with attached masses. In both models, curved beam elements are used to model the smooth ring. The results are compared against that from Ansys, showing that model II gives more accurate results.

Keywords- *curve beam element, natural frequency, ring gear*

I. INTRODUCTION

Planetary gear trains (PGT) are widely used in fields such as the aero-related, automation and energy industries. The dynamics of PGTs in operation is critical to safety and performance; thus, many studies have been conducted on the general aspects of PGT dynamics [1,2].

With increasing operational speeds, the dynamic deformation of the ring gear has become an outstanding issue. Thus, the free vibration characteristics of the ring, including natural frequencies, and mode shapes, need to be examined carefully in the design stage. In the open literature, most works modeled the ring gear as a smooth curved beam or as thin-walled circular ring. For example, Mallik et al. [3] analyzed the natural frequencies and mode shapes of the in-plane vibration of thin-wall rings by using a wave technique. Davis et al. [4] considered shear deformation and employed the two-node curved beam finite element model. Yand [5] and Wu [6] applied the finite element method to dynamic responses with the curved beam theory. Petyt et al. [7] carried out analysis on free radial vibration of the ring by using the 3D finite element model with curved beam elements under different boundary conditions. Sabir et al. [8] studied the in-plane vibration of a curved beam with the finite

element method. A real ring gear, in fact, is not smooth; rather, there are gear teeth on the ring. Some researchers have noticed that that a smooth ring with a uniform cross-section might be an over simplification; thus, the results based on this model are questionable in some cases. Tanna et al. [9] compared the modal frequencies and vibration modes of smooth rings against some real ring gears. Chen et al. [10,11] studied on the dynamic responses of elastic ring gear modeled by a curved Timoshenko beam.

To get a quantitative figure of the effect of the gear teeth on the dynamics, this paper investigates the free vibration of a ring gear with two methods. In the first method, the ring gear is modeled as a smooth ring. While the second method assumes that the ring gear has the rim and the teeth, wherein the rim is modeled as a smooth ring and each tooth is represented by a lumped-masses. In both methods, the finite element method is used. The results obtained from the two method are compared against data from a published paper [13]. It indicates that the results from the second method are more accurate than the first method.

II. DYNAMIC MODEL

A ring gear under investigation in this paper is shown in Figure 1. For the purpose of FE modeling in later sections, a global coordinate frame XYZ is placed at the ring's center with the Z-axis perpendicular to the XY plane.

A. Model I

In this model, the ring gear is modeled as a smooth elastic ring with a uniform cross-section. The outer diameter (OD) of the ring is taken as the same with that of the ring gear, while the inner diameter (ID) of the ring is taken as the diameter of the pitch circle as shown in Figure 2(a). To obtain the dynamic properties, the ring is first divided into curved beam elements.

For convenience, the number of elements is chosen to equal the number of teeth. An arbitrary element is taken out, as shown in Figure 2(b), with a local coordinate frame, including the radial, tangential and rotational, is defined. The 6 degrees of freedom of the element include u_x , u_t and u_θ for each node. To obtain the element mass and stiffness matrices, the method in reference [14] is employed, which gives.

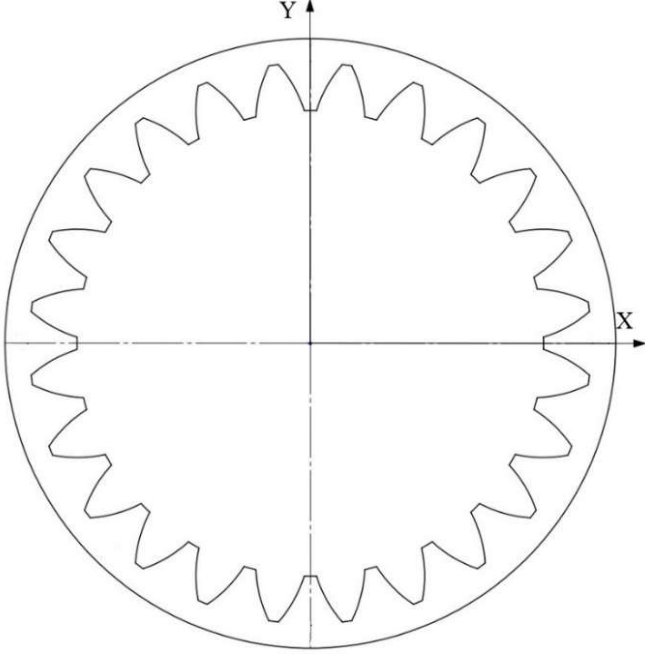


Figure 1: A Ring gear

$$[k] = \frac{EI}{R^3} [V][Q]^{-1} \quad (1)$$

$$[m] = \rho R [Q]^{-1T} \int_{\theta_1}^{\theta_2} ([P]^T [\Lambda] [P] d\theta) [Q]^{-1} \quad (2)$$

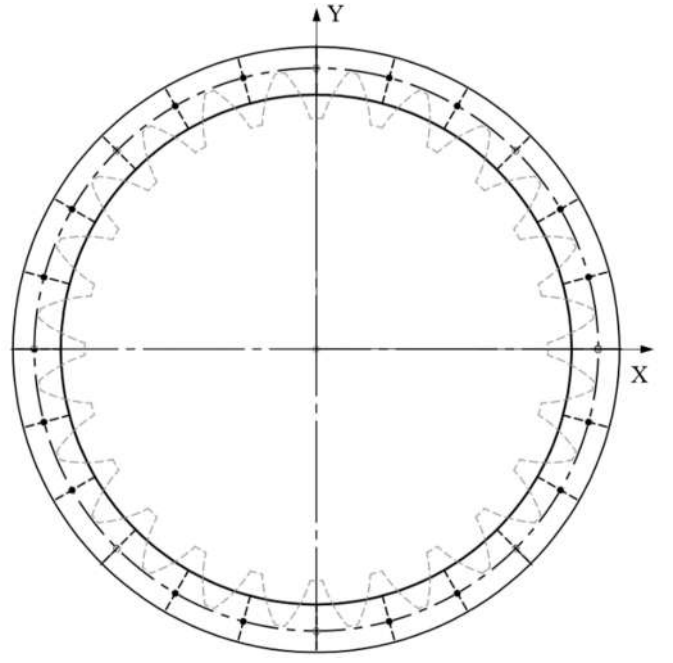
where m and k are the element mass and stiffness matrices, respectively. I is the moment of inertia. E and ρ are the module of elasticity and the mass density of the material, respectively. Λ , Q , V , and P are complex parameters, which are not given here in the purpose of saving space. Readers interested can find them in the Appendix and refer to [14] for details.

Assembling the equations of all elements gives the global equation as:

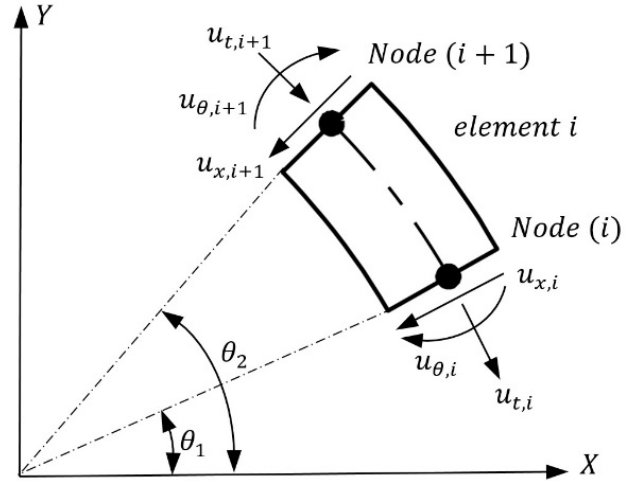
$$[M]\{U\} + [K]\{U\} = 0 \quad (3)$$

where M , K are the global mass and stiffness matrices, respectively, and U is the global displacement vector defined as:

$$\{U\} = \{u_{x,1} \ u_{t,1} \ u_{\theta,1} \ \dots \ u_{x,i} \ u_{t,i} \ u_{\theta,i}\} \quad (4)$$



(a)



(b)

Figure 2: (a). Smooth ring (b). a ring element

B. Model II

In this model, the ring gear is modeled as a smooth ring attached with lumped masses which represent the effect of the teeth. Each tooth is considered as a lumped mass, including mass m_T and moment of inertia around the axis through the center of the tooth J_T as shown in Figure 3(a). The smooth ring is determined with an OD the same with model I, and an ID the diameter of the diameter of the dedendum circle. The mass and stiffness matrices of the smooth ring can be computed in the same way with model I with Eqs. 1 and 2.

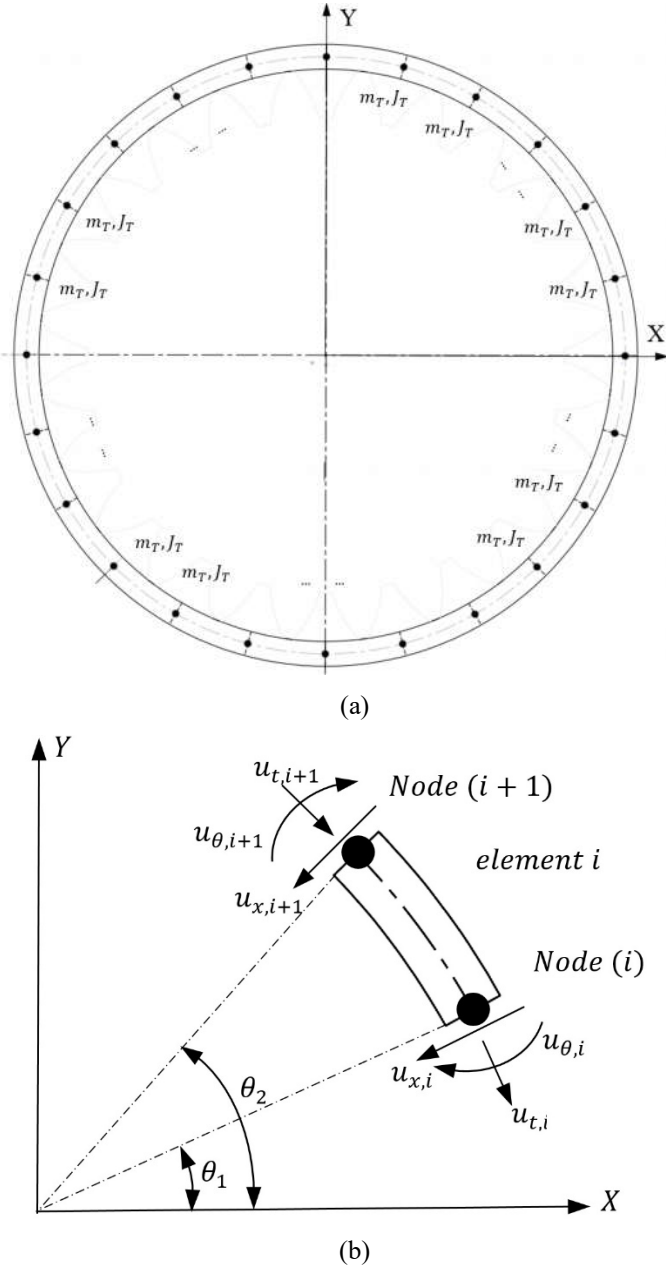


Figure 3: (a). Lumped-mass model (b). Ring element

With the tooth masses and moment inertial considered as lumped masses attached to the FEM nodes as shown in Figure 4. Then the internal forces at the two nodes of an element can be represented as:

$$F_{x,i} = F_{x,i+1} + m_T u_{x,i} \quad (5)$$

$$F_{t,i} = F_{t,i+1} + m_T u_{t,i} \quad (6)$$

$$M_{\theta,i} = M_{\theta,i+1} + J_T u_{\theta,i} \quad (7)$$

where $F_{x,i}$, $F_{t,i}$ and $M_{\theta,i}$ are the internal forces and moments of the elements and are given in Appendix and node number i is equal to $1, 2, 3, \dots, N$. Based on equations 5 to 7, assembling the equation of elements and boundary conditions gives the global equation as:

$$[M + M_T]\{U\} + [K]\{U\} = 0 \quad (8)$$

where M , K are the global mass and stiffness matrices, respectively, and U is the displacement vector. In equation 8, $[M_T]$ is defined as:

$$[M_T] = \text{diag}(m_T^1 \ m_T^2 \ \dots \ m_T^N) \quad (9)$$

where

$$[m_T^i] = \text{diag}(m_T \ m_T \ J_T) \quad i = 1, \dots, N \quad (10)$$

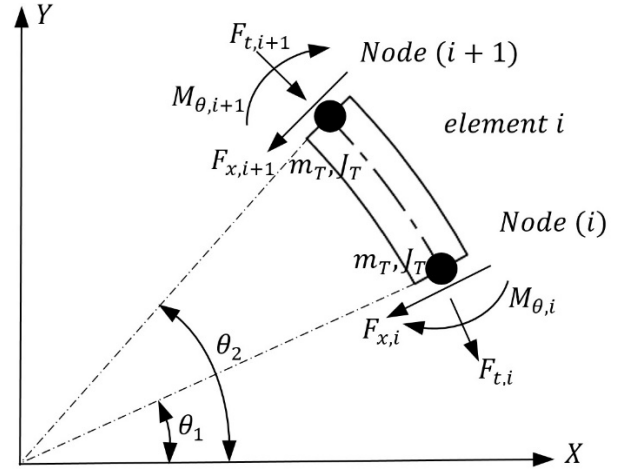


Figure 4:

III. SIMULATION AND RESULTS

In this study, numerical examples have been conducted on a ring gear with both model I and model II. Also for comparison purpose, Ansys is used in the analysis. The results and comparison against Ansys are given Figure 5. The geometry and material properties of the ring gear in simulation are given in Table I.

TABLE I. RING GEAR PROPERTIES

Parameters	Value
Outer radius (mm)	201.14
Root radius (mm)	182.8
m_T (kg)	0.028
I_T (kg.m ²)	554.8×10^{-9}
ρ (kg/m ³)	7858
E (Pa)	205×10^9
Number of teeth	64

For validation of the curved beam element method, we compared the results from model I against a published data [13]. In the comparison, the natural frequencies were transformed into dimensionless. All the dimensions and material properties are taken as the same as Wu [13]. Table II shows that the results from [13] and the curved beam element are matched very well.

TABLE II. DIMENSIONLESS NATURAL FREQUENCIES OF THREE EQUALLY SUPPORTED RING COMPARE WITH THE DATA FROM WU [13]

$\omega \sqrt{\frac{\rho A R^4}{E I_y}}$					
$\omega_{1,2}$	$\omega_{1,2}[13]$	$\omega_{3,4}$	$\omega_{3,4}[13]$	ω_5	$\omega_5[13]$
1.640	1.643	7.529	7.589	10.205	10.465

Although curved beam element is accurate in computation of dynamics of a smooth ring, its applicability to ring gear is till need verified. For the ring gear given in Table I, the natural frequencies are calculated with model I, II and Ansys. All results are given in table III. The relative errors of model I and II against Ansys results are given in table IV and Figure 5 as well.

From table III, it is clear that all the results are pretty close to that from Ansys. However, the accuracies are different.

If Ansys results are taken as the benchmarks, the comparison in table IV clearly shows that model II has much higher accuracy than that of model I. For example, the relative error of model I for the first 6 frequencies ranges from 8.6% to 19.4%. While the model II for the same problem give a relative error in between 1.8-7.6%.

TABLE III. COMPARISON OF THE NATURAL FREQUENCIES OF RING GEAR OBTAINED FROM SMOOTH RING MODEL, LUMPED-MAS MODE, AND ANSYS

	<i>Natural frequencies</i>		
	<i>Ansys (I)</i>	<i>Model I</i>	<i>Model II</i>
1	590	641	601
2	1650	1811	1697
3	3129	3474	3242
4	4977	5631	5216
5	7166	8309	7605
6	9652	11527	10390

TABLE IV. ERRORS BETWEEN THE SMOOTH RING MODEL LUMPED-MASS MODEL AND ANSYS RESULTS

	<i>% Error</i>		
	<i>Ansys vs. model I</i>	<i>Ansys vs. model II</i>	<i>Model I vs. II</i>
1	8.6	1.8	6.6
2	9.7	2.8	6.7
3	11.0	3.6	7.1
4	13.1	4.8	7.9
5	15.9	6.1	9.2
6	19.4	7.6	10.9

The comparison results are also shown in Figure 5.

It is observed that all the natural frequencies from model I are higher than that from model II. This make sense given that in model I the tooth effect is neglected. While in the model II, each tooth is modeled as an attached mass and moment inertial; thus, a more accurate results can be obtained. It is reasonable to deduct that with the decrease of the rim thickness the results from model I would give even higher error, and model II would become a better model for ring gear dynamics.

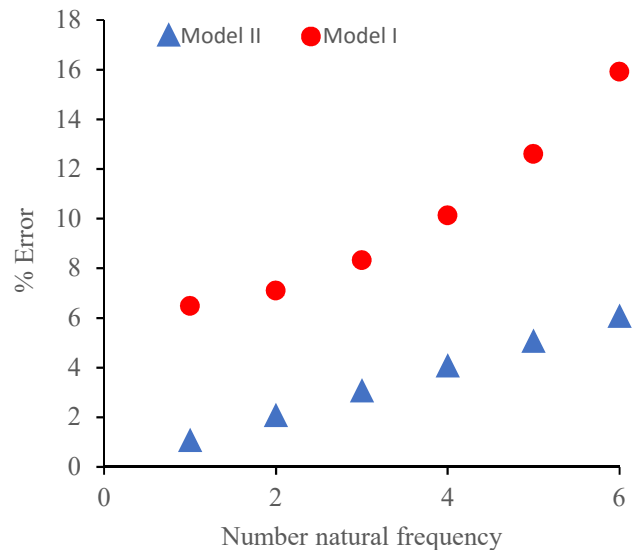


Figure 5: Comparative % errors between the Smooth ring Vs. ANSYS and lumped-mass Vs. ANSYS results

Figure 6 shows the mode shapes of the ring gear for the first six natural frequencies. As expected, mode shapes are less affected by the model methods.

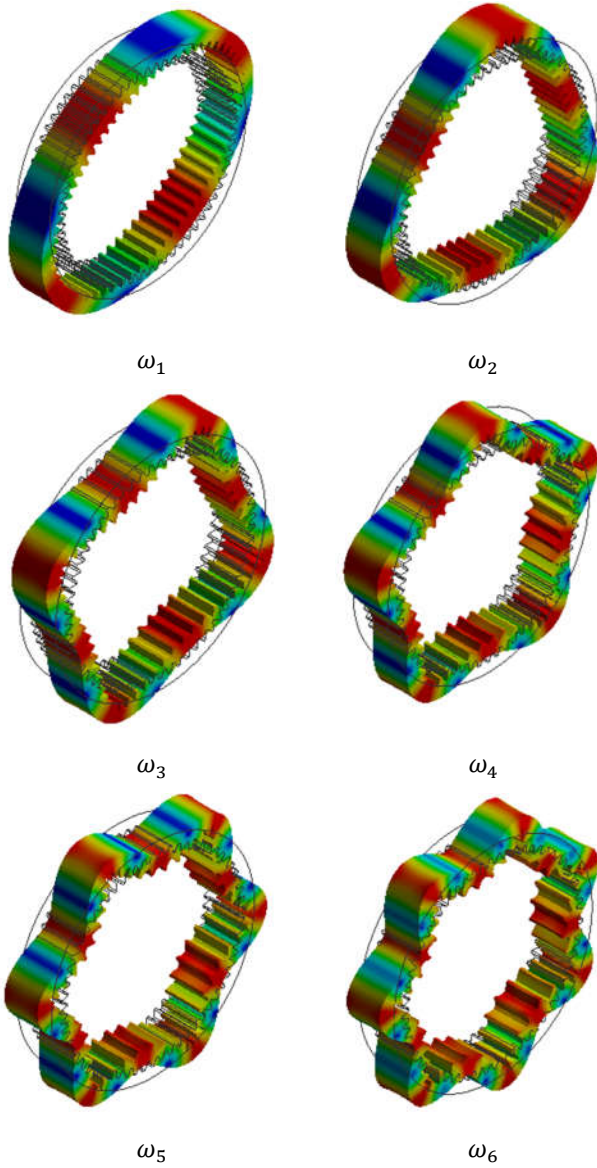


Figure 6: Ring gear's mode shapes

IV. CONCLUSIONS

In this paper, free vibration of ring gears is studied with two models. Model I simplifies a ring gear as a smooth ring with the pitch circle diameter as the ID. While model II treats a ring gear as a smooth ring with the dedendum circle diameter as the ID, and each tooth is modeled as a mass and a moment inertial. For computation, in both models curved beam elements are used. For verification purpose, the results are compared with Ansys results as well. It is indicated that model II is a better description of ring gear dynamics.

REFERENCES

- [1] C.G. Cooley, and R.G. Parker, "A review of planetary and epicyclic gear dynamics and vibrations research," *Applied Mechanics Reviews*, vol. 66(4), p.040804, 2014.
- [2] J. Yang, and L. Dai, "Survey of dynamics of planetary gear trains," *International Journal of Materials and Structural Integrity*, vol. 1(4), pp.302-322, 2008.
- [3] A.K. Mallik, and D.J. Mead, "Free vibration of thin circular rings on periodic radial supports," *Journal of Sound and Vibration*, vol. 54(1), pp.13-27, 1977.
- [4] R. Davis, R.D. Henshell, and G.B. Warburton, "Constant curvature beam finite elements for in-plane vibration," *Journal of Sound and Vibration*, vol. 25(4), pp.561-576, 1972.
- [5] F. Yang, R. Sedaghati, and E. Esmailzadeh, "Free in-plane vibration of general curved beams using finite element method," *Journal of sound and vibration*, vol. 318(4-5), pp.850-867, 2008.
- [6] J.S. Wu, F.T. Lin, and H.J. Shaw, "Free in-plane vibration analysis of a curved beam (arch) with arbitrary various concentrated elements," *Applied Mathematical Modelling*, 37(14-15), pp.7588-7610, 2013.
- [7] M. Petyt, and C.C. Fleischer, "Free vibration of a curved beam," *Journal of Sound and Vibration*, vol. 18(1), pp.17-30, 1971.
- [8] A.B. Sabir, and D.G. Ashwell, "A comparison of curved beam finite elements when used in vibration problems," *Journal of Sound and Vibration*, vol. 18(4), pp.555-563, 1971.
- [9] R.P. Tanna, and T.C. Lim, "Modal frequency deviations in estimating ring gear modes using smooth ring solutions," *Journal of Sound Vibration*, vol. 269, pp.1099-1110, 2004.
- [10] Z. Chen, Y. Shao, and D. Su, "Dynamic simulation of planetary gear set with flexible spur ring gear," *Journal of Sound and Vibration*, vol. 332(26), pp.7191-7204, 2013.
- [11] Z. Chen, and Y. Shao, "Mesh stiffness of an internal spur gear pair with ring gear rim deformation," *Mechanism and Machine Theory*, vol. 69, pp.1-12, 2013.
- [12] X. Wu, and R.G. Parker, "Modal properties of planetary gears with an elastic continuum ring gear," *Journal of Applied Mechanics*, vol. 75(3), p.031014, 2008.
- [13] X. Wu, and R.G. Parker, "Vibration of rings on a general elastic foundation," *Journal of Sound and Vibration*, vol. 295(1-2), pp.194-213, 2006.
- [14] J.S. Wu, and L.K. Chiang, "Dynamic analysis of an arch due to a moving load," *Journal of Sound and Vibration*, vol. 269(3-5), pp.511-534, 2004.

APPENDIX

$$F_x = F'_t$$

$$F_t = \frac{EA}{R} (u_x + u'_t) \quad \frac{M_\theta}{R}$$

$$M_\theta = \frac{EI}{R} (u_x + u''_x)$$

$$[V] = \begin{bmatrix} 0 & 0 & 0 & 2\sin\theta_1 & 0 & 2\cos\theta_1 \\ 0 & 0 & 0 & 2\cos\theta_1 & 1 & 2\sin\theta_1 \\ R & 0 & 0 & 2R\cos\theta_1 & R^{-1} & 2R\sin\theta_1 \\ 0 & 0 & 0 & 2\sin\theta_2 & 0 & 2\cos\theta_2 \\ 0 & 0 & 0 & 2\cos\theta_2 & 1 & 2\sin\theta_2 \\ R & 0 & 0 & 2R\cos\theta_2 & R^{-1} & 2R\sin\theta_2 \end{bmatrix}$$

$$\begin{aligned}
& [Q] \\
& \begin{matrix} 1 & \cos\theta_1 & \sin\theta_1 & \theta_1\sin\theta_1 & 0 & \theta_1\cos\theta_1 \\ C\theta_1 & \sin\theta_1 & \cos\theta_1 & \sin\theta_1 & \theta_1\cos\theta_1 & 1 \\ C\theta_1R^{-1} & 0 & 0 & 2R^{-1}\sin\theta_1 & R^{-1} & 2R^{-1}\cos\theta_1 \end{matrix} \\
= & \begin{matrix} 1 & \cos\theta_2 & \sin\theta_2 & \theta_1\sin\theta_2 & 0 & \theta_2\cos\theta_2 \\ C\theta_2 & \sin\theta_2 & \cos\theta_2 & 2\cos\theta_2 & 1 & \cos\theta_2 + \theta_2\sin\theta_2 \\ C\theta_2R^{-1} & 0 & 0 & 2R\cos\theta_2 & R^{-1} & 2R^{-1}\cos\theta_2 \end{matrix}
\end{aligned}$$

$$\begin{aligned}
& [P] \\
= & \begin{bmatrix} 1 & \cos\theta & \sin\theta & \theta\sin\theta & 0 & \theta\cos\theta \\ C\theta & \sin\theta & \cos\theta & \sin\theta & \theta\cos\theta & 1 \\ C\theta R^{-1} & 0 & 0 & 2R^{-1}\sin\theta & R^{-1} & 2R^{-1}\cos\theta \end{bmatrix}
\end{aligned}$$

$$C = 1 + \left(\frac{I}{AR^2} \right)$$

$$[\Lambda] = \text{diag}(A \ A \ I)$$

Fogging Simulation of Humid Air on the Surface of Windshield in Electric Car Room

Zhou Qinghui, Liu Liyan, Xiao Long, Su Yongyuan, Xi Yifan

School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture Beijing, China
Liyan_liu2019@163.com, zhouqinghui2008@126.com, xiaolongno5@163.com

Abstract—The frost and fog is a common phenomenon in automobile. To study the phase-change process of the humid air on the surface of the windshield, the numerical model was built on the base of CFD when fogging and its simulation was also fulfilled through FLUENT software. Furthermore, the experiments about fogging time and fogging amount were also completed on a fog test bench of windshield. Compared the simulation and experimental data, it is found that the results are consistent.

Keywords—component; wet air; phase change; Fogging process; wettability; condensation

NOMENCLATURE

ρ — Air density, kg/m^3 ;
 u, v, w — The speed of X, Y, Z direction, m/s ;
 P — air pressure, Pa ;
 P_S — water vapor partial pressure, Pa ;
 CP — specific heat, $\text{J/kg}\cdot\text{K}$;
 T — temperature, K ;
 k — heat transfer coefficient, $\text{W}/(\text{m}^2 \cdot \text{K})$;
 ST — Viscous dissipation term;
 CS — Volume concentration of component S ;
 DS — Diffusion coefficient of components, m^2/s ;
 SS — productivity (is 0);
 μ' — Pulsating velocity mean square root;
 μ_{avg} — average velocity m/s ;
 Re_{DH} — Reynolds number in terms of equivalent diameter;
 A — Overflow area of air outlet mm^2 ;
 S — Wetting length m ;
 d_i — Equivalent diameter m ;
 a — Return air length m ;
 b — Width of air outlet m ;
 ϕ — relative humidity.

I. INTRODUCTION

The indoor environment is an important guarantee for the safety and comfort of the driver and passenger in a car. Due to the limitations of insulation performance, the indoor environment is greatly affected by the weather, such as wind, frost, snow and rain. The difference of indoor and outdoor temperature will cause the fog on the surface of the windshield, which seriously affects the visibility of the windshield. The fogging mechanism of the windshield surface can provide an important theoretical basis for defogging and defog of automobile glass. In the previous study, the possibilities of fogging of the windshield was reduced by providing a continuous anti-fog curtain at the windshield, thereby increasing the proportion of return-air utilization[1,2]. The water content and saturated water content were analyzed in the air layer near the automobile glass, and their influence was also studied on the convective heat transfer coefficient by a fog test bench of windshield[3].

With the development of new materials technology, many scholars studies the anti-fogging performance of window glass by using hydrophilic and hydrophobic coatings. The study shows that the super-hydrophilic coating makes the contact angle of the droplets with the glass surface less than 5° , so that the droplets adhere to the glass surface in the form of a liquid film. And the super-hydrophobic coating makes the contact angle of the droplets with the glass surface greater than 150° , causing droplets to adhere to the glass surface in the form of droplets[11]. Due to the special properties of super-hydrophilic and super-hydrophobic coatings, most scholars are working to obtain hydrophobic membranes with larger contact angles and hydrophilic membranes with smaller contact angles[12-15]. However, there are few experimental tests of performance comparison between super-hydrophilic and super-hydrophobic coatings in automobiles. Simultaneously, there are also few visibility studies on fogging.

In this paper, the anti-fog performance is studied on NC319 super-hydrophobic coating and NC309 super-hydrophilic coating, considering the transient changes on fogging and the influence of different humidity.

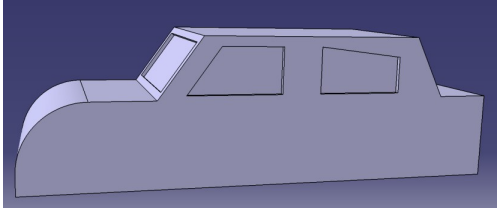
II. NUMERICAL MODEL OF WINDSHIELD ON FOGGING

A. Physical Model

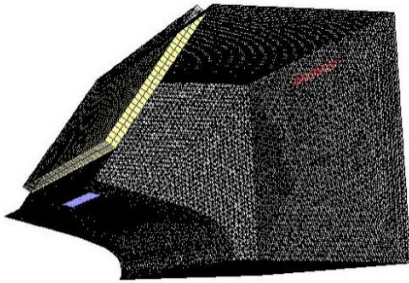
The physical model in the numerical calculation is according to the size of the actual sedan. Its model size is shown in Table I. Because of the fogging process of the car based on the windshield and the anti-fog performance of the glass coating with different wetting properties, the other structures, such as the rear view mirror, door handle, wheel, and front and rear seats inside the car, are ignored in the model, simplified model of real vehicle as shown in Fig. 1(a).

TABLE I. THE SIZE OF PHYSICAL MODEL

Name	Size (m)
Dimensions (length*width*high)	4.792*1.576*1.456
Size of Windshield (length*width*high)	1.376*0.589*0.045
Outlet of wind ((length*width)	0.4*0.04



(a) the shape of model



(b) Fog model

Figure 1. Physical model of the car

Model of automobile fogging as shown in Fig. 1(b), the fogging model has two bodies in total. One body is a windshield, which is divided into a hexahedral mesh, and the other body is indoor humid air, which is divided into a tetrahedral mesh. The air supply port and the return air outlet are divided by a quadrangle. The attribute definition of the selected material for the fogging model is shown in Table II.

TABLE II. THE ATTRIBUTE DEFINITION OF MATERIAL

Parameter	Wet air	Dry air	Liquid water	Glass	Aluminum
Density (kg/m ³)		1.225	998.2	2500	2719
Specific heat (j/kg-k)		1006.43	4182	750	871
Thermal conductivity (w/m-k)	0.0454	0.0242	0.6	1.4	202.4
Viscosity (kg/m-s)	1.72e-05	1.7894e-05	1.003e-03		

B. Control Equation

This paper is based on the VOF (Volume of Fluid) model in the Euler multiphase flow model of FLUENT, which is a surface tracking method under a fixed Euler grid that captures the free interface between one or more mutually incompatible fluids, and then adopt evaporation-condensation model. The surface tension is also taken into account to calculate the instantaneous state of the gas-liquid interface. The two phases are respectively the gas phase and the liquid phase. And there are mass transport and heat exchange in both phases.

Because the flow of air near the surface of the glass is turbulent, the K-ε model is selected. The volume fraction of all phases in each control body is 1 in the VOF model. The volume fraction of the gas phase is recorded as α_G , and the fraction of the liquid phase is α_L , So,

$$\alpha_G + \alpha_L = 1 \quad (1)$$

Volume fraction continuous equation:

$$\frac{\partial \alpha_G}{\partial t} + \nabla \cdot (\alpha_G \vec{v}) = 0 \quad (2)$$

$$\frac{\partial \alpha_L}{\partial t} + \nabla \cdot (\alpha_L \vec{v}) = 0 \quad (3)$$

continuous equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \quad (4)$$

Mass conservation equation:

$$\frac{\partial}{\partial t}(\rho) + \nabla \cdot (\rho \vec{v}) = 0 \quad (5)$$

Momentum conservation equation:

$$\frac{\partial}{\partial t}(\rho \vec{v}) + \nabla \cdot (\rho \vec{v} \vec{v}) = \rho \vec{g} + \nabla \cdot [\mu (\nabla \vec{v} + \nabla \vec{v}^T)] + \vec{\tau}_{VOL} - \nabla p \quad (6)$$

Energy equation:

$$\frac{\partial(\rho T)}{\partial t} + \frac{\partial(\rho u T)}{\partial x} + \frac{\partial(\rho v T)}{\partial y} + \frac{\partial(\rho w T)}{\partial z} = \frac{\partial}{\partial x} \left(\frac{k \partial T}{C_p \partial x} \right) + \frac{\partial}{\partial y} \left(\frac{k \partial T}{C_p \partial y} \right) + \frac{\partial}{\partial z} \left(\frac{k \partial T}{C_p \partial z} \right) + S_T \quad (7)$$

Component transport equation:

$$\begin{aligned} & \frac{\partial(\rho c_s)}{\partial x} + \frac{\partial(\rho u c_s)}{\partial x} + \frac{\partial(\rho v c_s)}{\partial y} + \frac{\partial(\rho w c_s)}{\partial z} \\ &= \frac{\partial}{\partial x} \left(D_s \frac{\partial(\rho c_s)}{\partial x} \right) + \frac{\partial}{\partial y} \left(D_s \frac{\partial(\rho c_s)}{\partial y} \right) + \frac{\partial}{\partial z} \left(D_s \frac{\partial(\rho c_s)}{\partial z} \right) + S_s \end{aligned} \quad (8)$$

Water vapor mass fraction:

$$Y = \frac{0.622 \frac{\phi P_s}{P - P_s}}{1 + 0.622 \frac{\phi P_s}{P - P_s}} \quad (9)$$

C. Equations

This paper simulates an incompressible flow, defining an inlet boundary condition as a constant velocity inlet and an outlet boundary condition as a pressure outlet. The inlet and outlet temperatures are set according to the calculation conditions. Since the car is exposed to the external environment, the outer surface of the glass is set as a convective heat exchange boundary. And the inner surface of the glass is set as coupling thermal boundaries, considering heat exchanges with the external environment through the glass, and simultaneous exchanges with the air layer near the indoor glass surface. In the fogging simulation of super-hydrophilic and super-hydrophobic coating, the contact angle is set in FLUENT to achieve hydrophilic and hydrophobic process.

III. MODEL VERIFICATION AND ANALYSIS

In order to study the fogging of the inner surface of windshield under different environmental conditions, a fog test bench of windshield was built. The schematic diagram of the experimental bench is shown in Fig.2. The red line and the yellow line are respectively used to help determine the different areas of the glass.

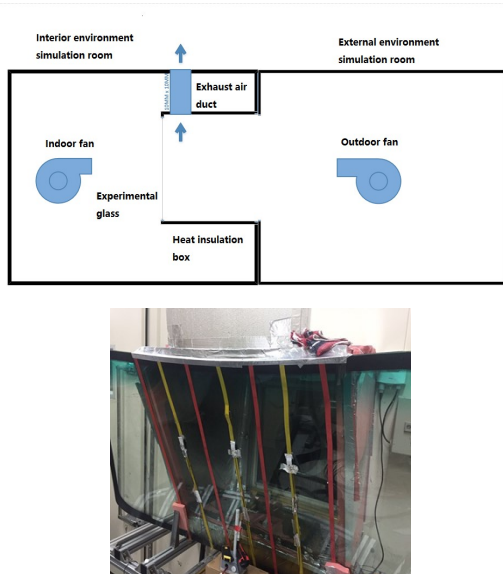


Figure 2. The schematic diagram of the experimental bench

The experimental working conditions are the same as those in the model, as shown in Table III. In this experiment, NC319 super-hydrophobic coated glass, NC309 super-hydrophilic coated glass, and ordinary glass were selected as research objects. The advantages and disadvantages of anti-fog performance were studied in the three kinds of glass under the same working conditions, and the changes of fogging were also observed. The experimental results are shown in Fig.3.

Outdoor temperature is 10 ° C. Wind speed is 4 m / s. Indoor temperature is 20 ° C, and initial relative humidity is 80%.

TABLE III. WORKING CONDITIONS

working condition	Outdoor temperature is 10 ° C, wind speed is 4 m / s, and indoor temperature is 20 ° C					
	40%	50%	60%	70%	80%	90%
Speed(m/s)	humidity					
0	1	2	3	4	5	6
1	7	8	9	10	11	12
2	13	14	15	16	17	18



Figure 3. The fogging on the different coating

(the ordinary glass and super-hydrophobic coated glass begin to fog
ordinary glass and super-hydrophobic coated glass seriously affect the sight)

When the ordinary glass and super-hydrophobic coated glass begin to fog, the super-hydrophilic coated glass does not change from Fig.3. When the indoor relative humidity is increased to 73%, ordinary glass and super-hydrophobic coated glass seriously affect the sight, but the super-hydrophilic coated glass is still not fogged and has good visibility. However, when the indoor humidity is increased to 87% or more, the fog may also occur on the surface of hydrophilic coated glass. Because the contact angle of the hydrophilic coating is less than 90°, the water droplets are distributed in the form of a uniform film on the surface of the glass, which reduces the influence of the droplets on the visibility of the glass. When the hydrophilic coating covers the glass surface, the glass surface will be very smooth. Therefore, it is necessary to arrange a thermocouple for measuring the temperature of the glass in the experiment. However, the super-hydrophobic coating without anti-fog properties will affect the clearness of the glass to a certain extent, because the contact angle of the hydrophobic coating is more

than 150°. Water droplets are not distributed on the surface in the form of a film, but small droplets of water. Therefore, the anti-fog property of the super-hydrophilic coated glass is superior to that of ordinary glass, and the anti-fog property of the ordinary glass is superior to that of the super-hydrophobic coating.

IV. NUMERICAL SIMULATION RESULTS AND COMPARATIVE ANALYSIS

A. Fogging process analysis

The experimental conditions are outdoor temperature 10°C, wind speed 4m/s, indoor temperature 20°C, relative humidity from 40% gradually increased to 90%. According to the experimental results, when the ordinary glass is completely fogged, the relative humidity in the room has reached 73%. Therefore, when the model is simulated and analyzed by fluent, the humidity is controlled to 80% in the ordinary glass. The simulation conditions are consistent with the experimental conditions.

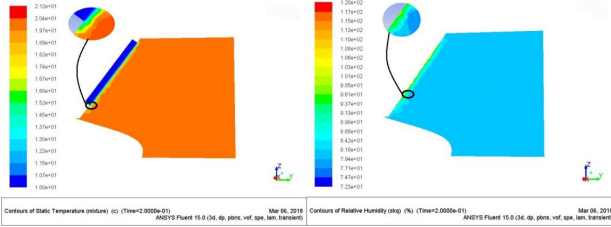


Figure 4. Temperature and humidity distribution diagram on the x section of the compartment

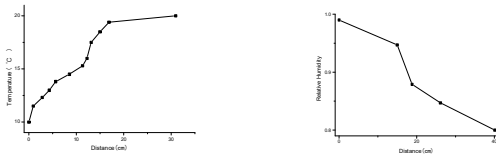


Figure 5. Temperature and humidity gradient distribution in vertical glass direction

From Fig. 4, we can see that, because of the temperature difference between the glass and the interior of the car, the heat and mass transfer is carried out. The temperature of the glass surface rises from 10 °C to 11.5 °C in the period of 0-0.2 s. And the temperature distribution is uniform and lower than the dew point temperature of the surrounding air. There is a significant temperature gradient in the air near the glass. As the distance from the glass increases, the temperature increases, and the temperature in the interior of the vehicle stabilizes at 19 °C. From Fig. 5, we can also see that there is a humidity gradient in the air near the glass in the period of 0-0.2 s. The closer to the glass surface, the higher is the relative humidity. When the humidity near the surface of glass reaches 94%, the relative humidity in the interior of the car is also reduced from the initial 80% and stabilized at 78%. Glass

fogging is an instantaneous process in a short time. Under the temperature difference and humidity difference, water vapor will transfer along the temperature gradient and humidity gradient, which means that the lower gradient will be transmitted from the place where the initial gradient is higher. The water vapor converges on the surface of the glass, the relative humidity on the surface will become higher and higher, and finally the phase change occurs. The water vapor condenses into small water droplets and adheres to the glass surface. In other words, the glass surface is fogged.

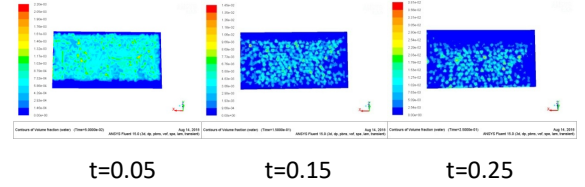


Figure 6. Water volume fraction at different times

Fig.6 shows the volume fraction of water at different times during the fogging process. On the surface of glass, liquid water is generated and uniformly distributed. Due to the randomness of the liquid water molecules, the water volume fraction on the glass surface is different, but the difference is not very large. For example, when $t=0.05s$, the minimum water volume fraction of the glass surface is $2.20E-05$, and most of the volume fraction is at the edge of the glass. The maximum is $2.20e-03$ in the right position of the center of the glass. This is probably because of the combined effects of water adsorption in the glass, self-gravity of water, and randomness of water polymerization. Considering the gradient difference, the water vapor will change its phase, and the water vapor molecules change into liquid water molecules, which can affect the visibility of the glass and the driver's driving safety. During the period from $t=0.05s$ to $t=0.25s$, the coverage area of liquid water on the glass gradually decreases, and its position also gradually moves down. This is because, when heat transfer and mass transfer continuously carried out on the glass surface and in the air layer near the wall, the liquid water molecules randomly polymerize on the surface of glass, and the radius of the droplet increases. When it is increased to a certain radius, the water droplets will slide down the glass. The droplets will be stationary until the next balanced state.

B. Comparison of experimental results and simulation

Because the amount of fog on the glass surface cannot be measured in the experiment, the simulation by software become more important in the study. In order to verify the simulation, the comparison should be completed between the experiment and simulation. The temperature and humidity of the ordinary glass surface detected by sensors on the fog test bench.

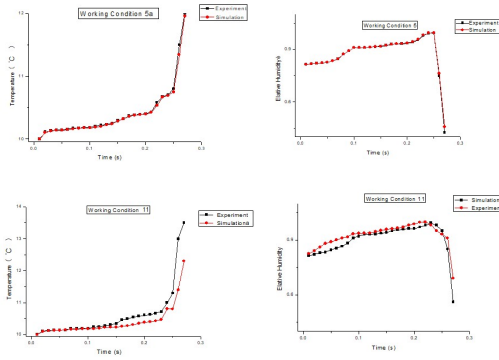


Figure 7. Comparison of experimental results and simulation

Fig.7 is a comparative analysis of the results of simulation and experiment. Within 0-0.3s, the temperature and humidity changes in the simulation are similar to the temperature and humidity changes measured by the fog test bench. And the overall trend of the curve is also consistent. The heat will transmit between the inside and outside of the car because of temperature difference. As the heat transfers, the humidity inside the car also changes, and finally gas-liquid phase changes can occur. At the beginning, the temperature transfer is slower. This is because the heat and mass transfer has just begun. Gradually, when the temperature inside the room drops, the temperature of windshield increases until reaching the dew point temperature. The water vapor phase changes and liquefies into small droplets covering the surface of glass, so that the heat and mass transfer is accelerated, and the temperature and humidity changes rapidly. After reaching a new balance, the heat and mass transfer stops and the fogging stops. After comparing with the experimental results, the simulation model has a high reference in the analysis on the windshield fogging.

V. FURTHER DISCUSSION

Super-hydrophilic coatings and super-hydrophobic coatings can change the chemical composition and microstructure of the glass surface, therefore changing the wetting properties of the glass surface. The super-hydrophilic coating makes the glass surface highly hydrophilized, which cause the contact angle between the glass surface and water is close to 0° . So the water vapor does not easily condense into small water droplets on the glass surface, but forms a transparent water film. However, the superhydrophobic coating makes the glass surface highly hydrophobic, and the contact angle between the glass surface and water is close to 180° , so that the formed water droplets cannot adhere to the glass surface and roll off.

Based on the fogging simulation, the paper changes the contact angle between the glass surface and water, simulates the temperature, humidity and fogging amount, and compares the different fogging properties among ordinary glass, the super-hydrophilic coated glass and super-hydrophobic coated glass at the location $(-0.6701012, 1.532249, -2.070696)$.

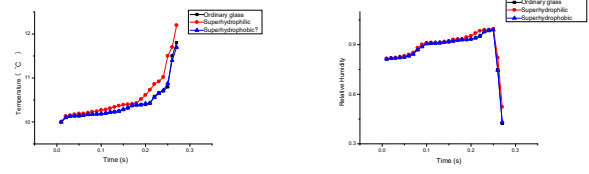


Figure 8. Comparison of temperature and humidity among the different coated glass under the same working conditions

Fig.8 is a comparison of temperature and humidity among the different coated glass under the same working conditions. It can be seen that the temperature of NC309 super-hydrophilic coated glass is slightly higher than that of NC319 super-hydrophobic coated glass and ordinary glass. The temperature of NC319 super-hydrophobic coated glass is not much different from that of ordinary glass. However, the three kind of glass has generally the same trend of changes. According to experiment, the heat transfer coefficient of glass with the super-hydrophilic coating is higher than the ordinary glass, so the temperature is also correspondingly higher. This is because the coating itself changes the wetting characteristics of the glass surface, making the surface of the super-hydrophilic coated glass smoother than ordinary glass, thereby affecting the heat exchange process of the surface of the glass. But super-hydrophobicity has little effect on the heat and mass transfer of windshield, so it can be neglected. In terms of humidity, when the temperature of the surface of the glass changes, the humidity also changes with the temperature. When the glass is fogged, the maximum moisture content by the visibility of the glass is larger than the maximum moisture content by the dew point temperature. Therefore, moisture content has a certain difference between by the visibility of the glass and by the dew point temperature.

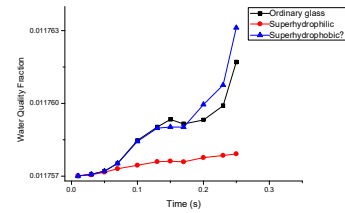


Figure 9. Fogging amount on the surface among the different coated glass

Fig.9 is a comparison of the fogging amount on the surface among the different coated glass at a relative humidity of 80%. Liquid water on the surface can be formed on the super-hydrophilic coated glass, super-hydrophobic coated glass, and ordinary glass. However, the water volume fraction of the super-hydrophilic coated glass increases slowly, and even the increase can be ignored, so the super-hydrophilic coated glass has a good anti-fog performance. This is because the super-hydrophilic coating changes the chemical composition and

microstructure of the glass surface, making the glass surface highly hydrophilized. So the water droplets can adhere to the glass surface in the form of a water film. During the 0-0.1s, the trend of the volume fraction is consistent between the superhydrophobic coating and the ordinary glass. But the super-hydrophilic coating makes the droplets more easily fall off the glass surface because of the smaller water droplets, so super-hydrophobic coating has a certain self-cleaning ability. In term of fogging amount, the superhydrophobic coating makes the glass surface highly hydrophobic. In other word, it is difficult for small water droplets to adhere to the glass surface. The superhydrophobic fogging amount is higher than that of ordinary glass, and the anti-fogging property is poor.

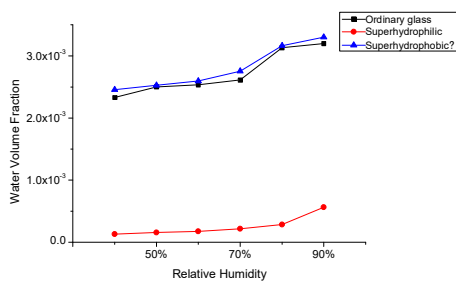


Figure 10. Fogging amount on the surface among the different coated glass

Fig.10 is a comparison of fogging amount on the surface of three glasses at different relative humidity at $t = 0.1$ s. As the relative humidity increases, the fogging amount on the surface also increases. The increase of ordinary glass and super-hydrophobic coated glass are more obvious, but the increase of super-hydrophilic coated glass is not obvious. A significant increase occurred in the super-hydrophilic coated glass when the relative humidity increased to 90%. Therefore, NC309 super hydrophilic coating has good anti-fog properties, but NC319 super hydrophobic coating has poor anti-fog performance.

VI. CONCLUSION

- During the fogging process on the inner surface of windshield, there is a temperature gradient and a humidity gradient in the air layer near the inner surface of the glass. The fogging process is in a short time, and the droplets are evenly distributed on the surface, but there are differences in individual positions.

- The heat transfer coefficient of glass with the super-hydrophilic coating is higher than the ordinary glass. Correspondingly, the inner surface temperature of glass with the super-hydrophilic coating is also higher than the ordinary glass.
- When the glass is fogged, the maximum moisture content by the visibility of the glass is larger than the maximum moisture content by the dew point temperature.
- NC309 super hydrophilic coating has good anti-fog properties. When the indoor humidity exceeds 90%, the affecting sight will only occur. But NC319 super hydrophobic coating has poor anti-fog performance.

REFERENCES

- [1] J.Tlee, S.Kwon, and Y.Lim, "Effect of Air-Conditioning on Driving Range of Electric Vehicle for Various Driving Modes," Asia Pacific Automotive Engineering Conference. 2013.
- [2] F.Qin,Q.Xue,G.M.AVelez,"Experimental investigation on heating performance of heat pump for electric vehicles at -20°C ambient temperature, " Energy Conversion & Management ,pp. 102:39-49,2015.
- [3] J.Liu, H.Zou,G.Zhang, "Experimental Study and Numerical Simulation concerning Fogging Characteristics and Improvement of Return Air Utilization for Electric Vehicles," Applied Thermal Engineering, 2018.
- [4] G.Zhang, H.Zou, F.Qin, " Investigation on an improved heat pump AC system with the view of return air utilization and anti-fogging for electric vehicles," Applied Thermal Engineering.2018.
- [5] T.Waku, N.Watanabe, J.Matsumoto, "An Efficient Numerical Approach for Evaluating the Windshield Defogging Time," SAE 2004 World Congress & Exhibition. 2004.
- [6] Numerical Investigation of Droplets Condensation on a Windshield: Prediction of Fogging Behavior
- [7] X.Y.Yang, Y.Li, S.L.Zhao, "Application of FLUENT software in heat transfer analysis [J]. Industrial heating, "2018.
- [8] S.S.Jeon, S.J.Kim, G.C.Park,"CFD Simulation of Condensing Vapor Bubble using VOF Model," World Academy of Science Engineering & Technology, 2009(60).
- [9] G.Croce, P.D'Agaro, F. DMora," Numerical simulation of glass fogging and defogging[J]. International Journal of Computational Fluid Dynamics, "19(19):437-445,2005.
- [10] J.Wu, F.Jiang, H.Song," Analysis and validation of transient thermal model for automobile cabin[J]. Applied Thermal Engineering," 122,2017.
- [11] X.Fan," Numerical simulation and analysis of wind shield defogging," Jilin University, 2009.
- [12] Q.Xie, C.Xiao, W.Xie," Research on hydrophobic film on glass surface ,Journal of Hubei Engineering University,"20:64-66,2000.
- [13] M.Hikita, K.Tanaka, T.Nakamura," ransparent super-hydrophobic coatings based onfluoroalkylsilane-silica hybrid materials," Journal Surface Science Society of Japan,26(9): 559,2005.
- [14] W.L.Tonar, J.S.Anderson, D.J.Cammenga," Electro-optic device having a self-cleaning hydrophilic coating,,"U.S. Patent 6991339, 2006.
- [15] W.Y.Gan, S.W.Lam, K.Chiang," Novel TiO2 thin film with non-UV activated superwetting and antifogging behaviours. Journal of Materials Chemistry,"17(10): 952-954, 2007.

SDP-based energy management strategy with lithium plating prevention for PHEVs

Siyang Wang

Department of Automotive, Mechanical and
Manufacturing Engineering,
University of Ontario Institute of Technology,
Oshawa, Canada

Xianke Lin

Department of Automotive, Mechanical and
Manufacturing Engineering,
University of Ontario Institute of Technology,
Oshawa, Canada

Abstract—Recent studies in battery health have shown that lithium-ion battery anode potential plays an important role in triggering the problem of lithium plating. The paper develops an energy management strategy (EMS) for plug-in hybrid electric vehicles (PHEVs) that can reduce lithium plating behavior by preventing anode potential from dropping below 0V vs Li/Li^+ . The simulation results show that the proposed strategy can reduce the occurrence of anode potential dropping below zero by 41% while achieving an excellent fuel economy.

Keywords—component; lithium-ion battery; PHEV; EMS; lithium plating; SDP; anode potential

I. INTRODUCTION

The market for plug-in hybrid electric vehicles (PHEVs) market is continuously growing in recent years. As the most widely used electric power source for PHEVs, lithium-ion batteries (LIBs) is considered as one of the most promising technologies for their high energy and power density, great performance in cycling and light weight [1, 2]. However, regardless of the advantages of LIBs, LIBs suffer from several limitations related to safety and battery degradation problems [3]. Lithium plating caused by metallic lithium deposition at the graphite anode becomes one of the dominant factors leading to battery aging and safety issues. The metallic lithium that deposits at the anode side can cause capacity fade and even trigger an internal short circuit by the generated dendrites [4–6]. It is therefore important to find the cause of lithium plating in order to prevent it from occurring.

Recently, many experiments have been conducted to investigate the factors that cause lithium plating behavior. Ecker *et al.* [7] revealed that batteries with lower anode potentials are more likely to have lithium plating, according to their experiments performed on commercial batteries at low temperatures. A LIB model developed by Yang *et al.* [8] suggested that lithium plating would happen when the local anode potential drops below 0V vs Li/Li^+ . Tippmann *et al.* [9] proposed a pseudo-2D electrochemical model to discover the factors that result in battery degradation and a qualitative correlation is found between the degradation in the simulation and the negative local anode potential below 0V vs Li/Li^+ . From the aforementioned studies, it is found that anode potential

in the battery acts as the main factor in provoking lithium plating phenomenon. When the anode potentials drop below 0V vs Li/Li^+ , lithium plating is prone to occur. Therefore, to prevent metallic lithium depositing at anode side graphite particles, anode working potentials that are below 0V vs Li/Li^+ must be avoided.

The present paper proposes an SDP-based control strategy that introduces lithium plating reduction into the PHEV energy management problem. In this multi-objective control strategy, during optimizing energy consumption, the occurrence of anode potential below zero is minimized.

The remainder of the paper is organized as follows. In Section II, the model of the PHEV and the lithium-ion battery is presented. Section III introduces the problem formulation and the proposed strategy. Section IV shows the simulation results. Finally, Section V concludes with a summary.

II. SYSTEM MODELING

A. Modeling of plug-in hybrid electric vehicle (PHEV)

The studied PHEV with power-split configuration is shown in Fig. 1. This architecture includes planetary a gear device, an internal combustion engine (ICE), a motor/generator 1 (M/G1), inverter, an electric motor/generator 2 (M/G2) and battery pack. Based on the design of hybrid electric vehicles (HEVs), PHEVs can be supplied by both the mechanical power from the internal combustion engine (ICE) and electric power from the battery pack. In addition, a larger battery pack capacity enables PHEV to have better electric energy sustainability over a long driving distance. Among different configurations, power-split configuration has the best performance as well as high flexibility in energy management for that it integrates the advantages of series hybrid and parallel hybrid while avoiding their drawbacks [10].

To calculate the power demand of PHEVs, two power paths are considered, the mechanical power path and electric power path.

In the mechanical power path, the power demand from wheels can be represented by,

$$P_{req} = (F_f + F_w + F_j) \cdot v \quad (1)$$

$$\begin{cases} F_f = mg\mu \\ F_w = \frac{1}{2}AC_d\rho_{air}v^2 \\ F_j = \delta m\dot{v} \end{cases} \quad (2)$$

where the required power P_{req} includes rolling resistance F_f , aerodynamic resistance F_w and acceleration/deceleration resistance F_j . In Eq. (2), m and μ are the vehicle mass, and friction coefficient in rolling resistance F_f . In aerodynamic resistance F_w , A , C_d , ρ_{air} and v are the vehicle frontal area, aerodynamic coefficient, air density and vehicle velocity, respectively. δ denotes the conversion ratio of vehicle rolling mass in acceleration/deceleration resistance F_j .

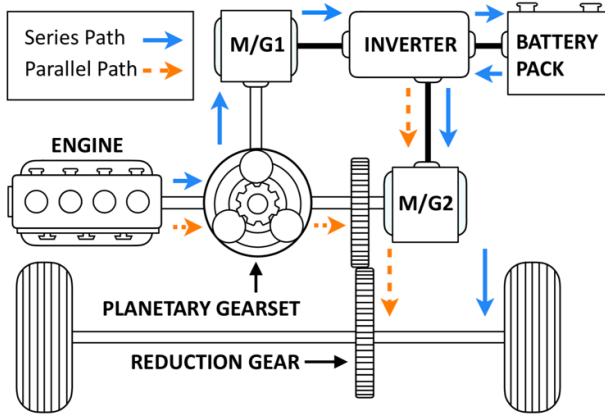


Fig. 1 Power-split PHEV configuration

In order to reduce the computational cost of the experiment, a simplified battery equivalent circuit model is employed in simulating PHEVs. It has an open voltage circuit with an internal resistance [11]. The electric power of the battery pack is thus modeled as,

$$P_{batt} = V_{oc}I_{batt} - I_{batt}^2 R_{batt} \quad (3)$$

where V_{oc} , I_{batt} and R_{batt} are the open circuit voltage, battery current and battery internal resistance, respectively.

The state of charge (SOC) is an index to measure the level of the remaining capacity over the maximum capacity. The rate of change of SOC is written as,

$$\dot{SOC} = -\frac{I_{batt}}{Q_{batt}} \quad (4)$$

where Q_{batt} represents the maximum battery capacity.

The resulting equation of the rate of SOC is derived by combining the Eq. (3) and Eq. (4),

$$\dot{SOC} = -\frac{V_{oc} - \sqrt{V_{oc}^2 - 4P_{batt}R_{batt}}}{2Q_{batt}R_{batt}} \quad (5)$$

Since the electric power is exchanged between battery pack, M/G1 and M/G2. The governing equation for the electric power path is written as,

$$\begin{cases} P_{batt} = P_{MG1} + P_{MG2} \\ P_{MG1} = T_{MG1} \cdot \omega_{MG1} \cdot \eta_{MG1}^k \cdot \eta_{i1}^k \\ P_{MG2} = T_{MG2} \cdot \omega_{MG2} \cdot \eta_{MG2}^k \cdot \eta_{i2}^k \end{cases} \quad (6)$$

where P_{MG1} is the power of M/G1 and P_{MG2} is the power of M/G2. T_{MG1} and T_{MG2} are the torque of M/G1 and M/G2, respectively. ω_{MG1} and ω_{MG2} are the rotational speed of M/G1 and M/G2, respectively. η_{MG1}^k and η_{MG2}^k are the power transfer efficiency of M/G1 and M/G2, respectively. η_{i1}^k and η_{i2}^k are the corresponding transfer efficiencies of the inverters. k denotes the direction of charging/discharging. When $k = 1$, the battery is being charged, and when $k = -1$, the battery is being discharged.

B. Lithium-ion battery

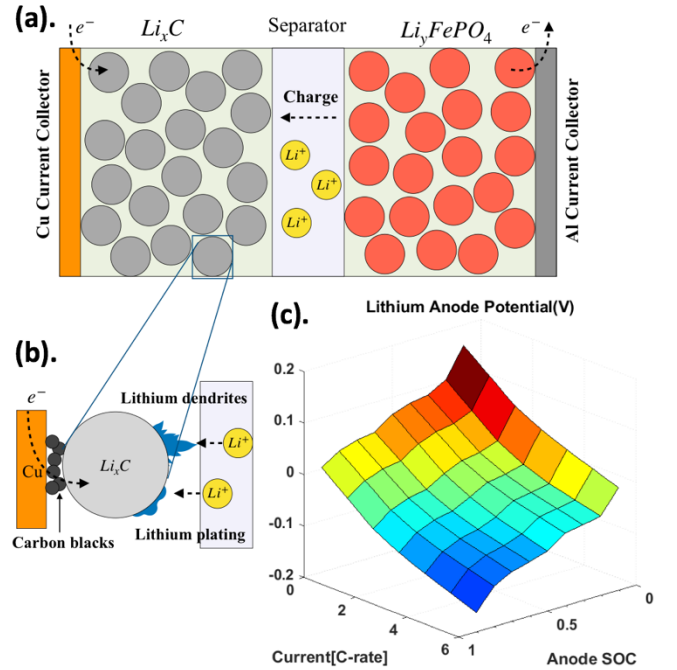


Fig. 2 Physics-based battery model and anode potential map at different SOC and C rates.

The pseudo-2D electrochemical battery model is used in this study. The details of the physics-based model can be found in reference [12–14]. The anode potentials at different SOC and different charging current are extracted from the battery model, as shown in Fig. 2 c. The anode potential map is used in the SDP optimization process.

III. SDP-BASED ENERGY MANAGEMENT STRATEGY WITH LITHIUM PLATING REDUCTION

A. Problem formulation

In this study, the control problem is formulated into a multi-objective optimal problem. The objectives are to optimize the fuel and energy consumption and reduce the occurrence of anode potentials that drop below 0V vs. Li/Li^+ .

The optimization problem is written as

$$\min: J^g = \lim_{N \rightarrow \infty} E \left[\sum_{k=0}^N C(x_k, u_k) \right] \quad (7)$$

$$\text{state variables: } x_k = [SOC, v, a] \quad (8)$$

$$\text{control variables: } u_k = [\omega_e, T_{M/G2}] \quad (9)$$

$$\text{subject to: } x \in X, u \in U(x) \quad (10)$$

where J^g is the objective function. The state variables are battery state of charge (SOC), vehicle speed v and vehicle acceleration a . The control variables are engine speed ω_e and torque of M/G2, $T_{M/G2}$. They are subjected to the physical constraints as written in Eq (10).

The cost function $C(x_k, u_k)$ to be minimized is defined as,

$$C(x_k, u_k) = k_1 \alpha_f W_f + W_e + k_2 (-P_A) \quad (11)$$

In the cost function, W_f is the instantaneous fuel cost in g/s and α_f converts it to the unit MJ/s. W_e is the instantaneous electric consumption in MJ/s. P_A denotes the anode potential value at each time step. The aim is to reduce the occurrence of anode potentials that drop below 0V vs. Li/Li^+ , therefore the negative sign is applied to fit it into this minimization problem.

The weighting factors k_1 is to make a tradeoff between fuel consumption and electric energy consumption, which is given by,

$$k_1 = \frac{\text{Price of Gasoline per MJ in USD}}{\text{Price of Electricity per MJ in USD}} \quad (12)$$

We use $k_1 = 0.75$ according to the market price for gasoline that is \$2.854 per gallon [15] and electricity price that is \$0.1042 per kWh [16] as of Aug. 2018.

k_2 serves as a weighting factor in determining the importance of lithium plating objective in the cost function.

B. Stochastic dynamic programming (SDP)

The optimization problem is solved by using stochastic dynamic programming (SDP). In order to derive the optimal policy, the Bellman optimal equation is utilized to perform the policy evaluation,

$$J_g^*(n) = c(x, u) + \gamma \sum_{n'}^{end} p_{nn'} J_g^*(n') \quad (13)$$

where $J_g^*(n)$ is the optimal cost at state n and $J_g^*(n')$ is the optimal cost at the next state n' . $c(x, u)$ is the instantaneous cost at each time instance. $p_{nn'}$ is the transition probability between state n and n' . Discount factor γ ($\gamma \in (0, 1)$) represents the weight applied on the future expected optimal cost.

The optimal policy can be derived once the optimal cost of each state is computed, as follows

$$g = \arg \min_{g \in G} J_g^*(n) \quad (14)$$

IV. RESULTS

In the simulation, the vehicle is running on a selected driving cycle with 65% as initial SOC, 0 m/s as initial velocity and 0 m/s² as initial acceleration.

Table 1 presents the results between conventional SDP and the proposed SDP-based energy management strategy with lithium plating reduction. The weighting factor k_2 is set to be 5.

It can be seen from Table 1 that with the proposed strategy applied, the fuel and energy management of PHEVs achieve nearly as has little increase when compared with the ones of conventional SDP due to the added objective of lithium plating reduction in the optimization function. However, the percentage of occurrence of negative anode potentials over the entire trip drop from 16.7% to 9.8% which equals to 41% reduction. In this way, fuel and energy consumption can be ensured while the lithium plating can be reduced by reducing the negative anode-side working potentials.

Table 1 Simulation results (SOC correction applied)

Algorithm	SDP	SDP _{Li}
Fuel cons. (g)	1581.6	1593.1
Energy cons. (MJ)	58.3	58.6
Negative anode potential (%)	16.7%	9.8%

Fig. 3 shows the anode working potentials between the two strategies. It is clearly shown that with the proposed strategy employed, it ensures less anode working potentials working points to occur below 0V vs. Li/Li^+ to reduce lithium plating.

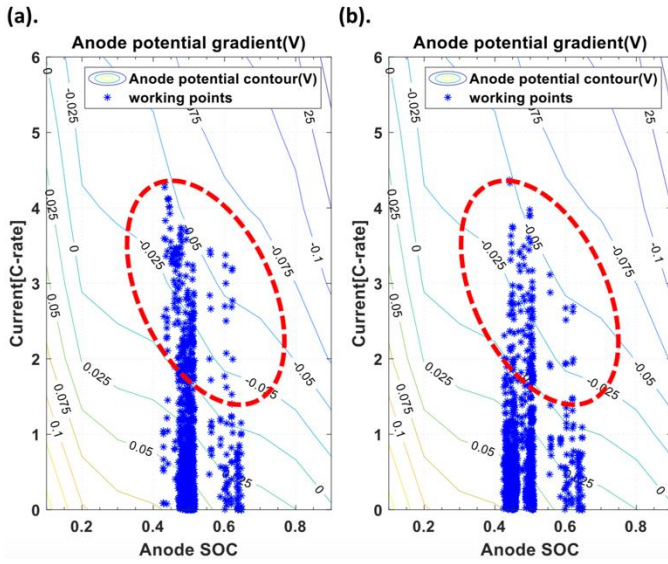


Fig. 3 (a) anode potential distribution by conventional SDP (b) anode potential distribution by SDP with lithium plating

Therefore, the proposed SDP-based lithium plating reduction strategy results in nearly the same fuel and energy consumption but much less battery anode potential working points below 0V vs. Li/Li^+ .

V. CONCLUSION

The paper presents an SDP-based PHEV energy management strategy that integrates the lithium plating reduction method. In this multi-objective strategy, weighting factors are applied to make a tradeoff between the goals of energy management and lithium plating reduction. Simulation results indicate that with the proper parameters selected, the proposed strategy can achieve nearly good energy consumption reduction as the conventional method, however, less battery anode working points tend to occur in the negative area which helps to improve the battery health by reducing the lithium plating behavior.

ACKNOWLEDGMENT

This work was supported by the NSERC Discovery Program (RGPIN-2018-05471).

REFERENCES

- [1] Q. Liu *et al.*, "Understanding undesirable anode lithium plating issues in lithium-ion batteries," *RSC Adv.*, vol. 6, no. 91, pp. 88683–88700, 2016.
- [2] S. Ma *et al.*, "Temperature effect and thermal impact in lithium-ion batteries: A review," *Progress in Natural Science: Materials International*, vol. 28, no. 6, pp. 653–666, 2018.

- [3] P. Arora, "Mathematical Modeling of the Lithium Deposition Overcharge Reaction in Lithium-Ion Batteries Using Carbon-Based Negative Electrodes," *J. Electrochem. Soc.*, vol. 146, no. 10, p. 3543, 1999.
- [4] N. Harting, N. Wolff, and U. Krewer, "Identification of Lithium Plating in Lithium-Ion Batteries using Nonlinear Frequency Response Analysis (NFRA)," *Electrochimica Acta*, vol. 281, pp. 378–385, 2018.
- [5] X. Xu and N. Chen, "A state-space-based prognostics model for lithium-ion battery degradation," *Reliability Engineering & System Safety*, vol. 159, pp. 47–57, 2017.
- [6] H. Ge *et al.*, "Investigating Lithium Plating in Lithium-Ion Batteries at Low Temperatures Using Electrochemical Model with NMR Assisted Parameterization," *J. Electrochem. Soc.*, vol. 164, no. 6, A1050-A1060, 2017.
- [7] M. Ecker, P. Shafiei Sabet, and D. U. Sauer, "Influence of operational condition on lithium plating for commercial lithium-ion batteries – Electrochemical experiments and post-mortem-analysis," *Applied Energy*, vol. 206, pp. 934–946, 2017.
- [8] X.-G. Yang, Y. Leng, G. Zhang, S. Ge, and C.-Y. Wang, "Modeling of lithium plating induced aging of lithium-ion batteries: Transition from linear to nonlinear aging," *Journal of Power Sources*, vol. 360, pp. 28–40, 2017.
- [9] S. Tippmann, D. Walper, L. Balboa, B. Spier, and W. G. Bessler, "Low-temperature charging of lithium-ion cells part I: Electrochemical modeling and experimental investigation of degradation behavior," *Journal of Power Sources*, vol. 252, pp. 305–316, 2014.
- [10] S. J. Moura, D. S. Callaway, H. K. Fathy, and J. L. Stein, "Tradeoffs between battery energy capacity and stochastic optimal power management in plug-in hybrid electric vehicles," *Journal of Power Sources*, vol. 195, no. 9, pp. 2979–2988, 2010.
- [11] S. J. Moura, *Techniques for Battery Health Conscious Power Management via Electrochemical Modeling and Optimal Control (en_US)*. Available: https://deepblue.lib.umich.edu/bitstream/2027.42/86511/1/sjmoura_1.pdf.
- [12] J. Newman and W. Tiedemann, "Porous-electrode theory with battery applications," *AIChE J.*, vol. 21, no. 1, pp. 25–41, 1975.
- [13] X. Lin *et al.*, "A Comprehensive Capacity Fade Model and Analysis for Li-Ion Batteries," *J. Electrochem. Soc.*, vol. 160, no. 10, A1701-A1710, 2013.
- [14] X. Lin, X. Hao, Z. Liu, and W. Jia, "Health conscious fast charging of Li-ion batteries via a single particle model with aging mechanisms," *Journal of Power Sources*, vol. 400, pp. 305–316, 2018.
- [15] *National Gas Price*. [Online] Available: <https://gasprices.aaa.com/>. Accessed on: Dec. 18 2018.
- [16] *State Electricity Profiles - Energy Information Administration*. [Online] Available: <https://www.eia.gov/electricity/state/>. Accessed on: Dec. 18 2018.

A FAULT-TOLERANT STEERING CONTROL ALGORITHM USING MPC FOR DRIVING SAFETY OF ALL-TERRAIN CRANES

Kwang-seok Oh

Department of Mechanical Engineering
Hankyong National University
Anseong-si, Korea
oks@hknu.ac.kr

Jaho Seo

Department of Automotive, Mechanical and
Manufacturing Engineering
University of Ontario Institute of Technology
Oshawa, Canada
Jaho.Seo@uoit.ca

Abstract—This paper proposes a fault-tolerant steering control algorithm for driving safety of all-terrain cranes. For this, the model predictive control algorithm (MPC) is proposed with physical constraints such as an equality constraint, which was used for the tolerant steering control. In the study, it is assumed that the steering axles are locked when the sensor or actuator fault of a steering system is detected. The performance evaluation of the developed control algorithm was conducted using a vehicle model with virtual fault conditions under a path tracking scenario in Matlab/Simulink. Simulation results show that the proposed fault-tolerant control algorithm enables an all-terrain crane to track the desired path although a fault occurs in the steering system.

Keywords- *model predictive control; fault tolerant; all-terrain crane; steering control; driving safety*

I. INTRODUCTION

Recently, autonomous driving technologies in several industries such as automotive and logistics have been developed for improving safety and efficiency. In these industrial sectors, autonomous driving systems have adopted various sensors and actuators for the purpose of control and environmental monitoring. If there is a fault in sensors and actuators, this may result in fatal accidents. To improve the functional safety of autonomous driving systems, diverse approaches have been introduced as follows.

Kim [1] developed a sensor and actuator fault diagnostic algorithm using the Kalman filter-based observer with a combination of multiple sensors. Oh [2] proposed a predictive fault detection algorithm for longitudinal autonomous driving by applying the sliding mode observer with prediction algorithms. Kim [3] designed a fault detection and isolation algorithm for current and position sensors for individual in-wheel motors of electric vehicles using a parity equation. Garoudja [4] proposed a statistical approach-based fault detection/diagnostic algorithm of a photovoltaic system. Li [5] developed a fault reconstruction algorithm of a linear system using the adaptive sliding mode method. Loureiro [6] provided an integrated fault diagnosis and fault-tolerant control

algorithm for health monitoring of MIMO intelligent autonomous vehicles. Through a review on the previous studies, it is observed that fault detection and isolation algorithms for tolerant control have contributed to the functional safety in the considered systems.

In this study, a fault-tolerant steering control algorithm with the MPC was designed and evaluated for driving safety of all-terrain cranes using a crane planar model. For this control design, a planar crane model was derived based on vehicle dynamics. Steering angles of a crane were computed by applying the error dynamics model and MPC algorithm where an equality constraint was applied for fault tolerant control. As a steering fault condition adopted in this study, we assumed that steering axles in fault are fixed.

The rest of the paper is organized as follows. Section 2 describes a fault tolerant steering control algorithm of all-terrain crane. Section 3 describes an evaluation of the designed controller's performance. Finally, concluding remarks are provided in Section 4.

II. FAULT-TOLERANT STEERING CONTROL ALGORITHM

A. Conventional Steering System of All-Terrain Crane

A steering control algorithm for all-terrain cranes was designed based on the conventional steering system. The conventional crane's steering system considered in this study consists of ten wheels (5 axles). The 1st and 2nd axles are connected mechanically based on the Ackerman steering rule. Five steering modes (road steering, all-wheel steering, crab steering, reduced swing out, and independent real-axle steering) are available in the crane's conventional steering system by fixing axles selectively according to operational tasks [7]. Among five modes, the road steering mode was considered in this study where axles are locked depending on the crane's driving speed (Fig. 1). The proposed MPC based steering algorithm was designed to compute optimal steering angles of 1st, 3rd, 4th, and 5th axles.

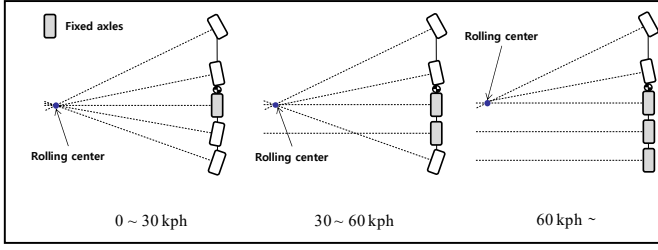


Figure 1. Conventional steering system of crane

A detailed explanation on the designed MPC-based steering control system is provided in the next section.

B. MPC-based Fault-Tolerant Steering Control System

In order to design the MPC based steering control algorithm, error dynamics of the crane was derived from the crane planar model as follows.

$$\dot{\vec{e}} = A\vec{e} + B\vec{\delta}. \quad (1)$$

where \vec{e} represents error state vector, $[e_1 \ \dot{e}_1 \ e_2 \ \dot{e}_2]^T$. e_1 and e_2 represent lateral error and yaw angle error, respectively. Matrices A and B in equation (1) are as follows. And elements of matrices are defined in the Appendix.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & 1 \\ 0 & a_{42} & a_{43} & a_{44} \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ 0 & 0 & 0 & 0 & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \end{bmatrix}. \quad (2)$$

Based on the error dynamics, a cost function for the MPC is defined as follows.

$$\min_{\vec{\delta}} J = \vec{y}^T Q \vec{y} + \vec{\delta}^T R \vec{\delta}. \quad (3)$$

$$\vec{y}_p = [\vec{y}(k+1) \ \vec{y}(k+2) \ \dots \ \vec{y}(k+N_p)]^T. \quad (4)$$

$$\vec{\delta}_p = [\vec{\delta}(k+1) \ \vec{\delta}(k+2) \ \dots \ \vec{\delta}(k+N_p)]^T. \quad (5)$$

$$Q = \text{diag}[\bar{Q}(k+1) \ \dots \ \bar{Q}(k+N_p)]^T. \quad (6)$$

$$R = \text{diag}[\bar{R}(k+1) \ \dots \ \bar{R}(k+N_p)]^T. \quad (7)$$

where \vec{y}_p is the predictive output matrix, $\vec{\delta}$ is the predicted input vector that can be calculated by a Quadratic Programming (QP) algorithm to minimize the cost function J . The matrices Q and R in the cost function represent the weighting matrices. N_p stands for a prediction step of the MPC. Based on N -step prediction, a predictive output can be derived as follows.

$$\vec{y}_p = M\vec{e}(k) + H\vec{\delta}_p + F\vec{\delta}(k). \quad (8)$$

where $\vec{e}(k)$ and $\vec{\delta}(k)$ represent error state and steering angle at the current step, respectively. And the matrices M , H , and F are defined as follows.

$$M = [A \ A^2 \ \dots \ A^N]^T. \quad (9)$$

$$H = \begin{bmatrix} 0 & 0 & \dots & 0 \\ B & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-2}B & A^{N-3}B & \dots & 0 \end{bmatrix}. \quad (10)$$

$$F = [B \ AB \ \dots \ A^{N-1}B]^T. \quad (11)$$

By rearranging equation (3) with (8), the cost function defined in equation (3) can be expressed as a quadratic form.

$$J = \frac{1}{2} \vec{\delta}_p^T (H^T Q H + R) \vec{\delta}_p + (\vec{e}^T(k) M^T Q \vec{e}(k) + \vec{\delta}^T(k) F^T Q H) \vec{\delta}_p. \quad (12)$$

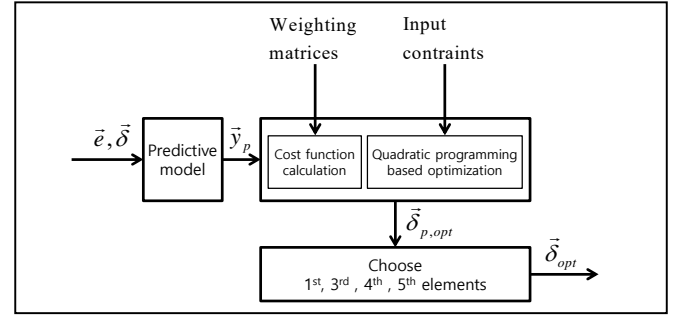


Figure 2. Model schematics of the detailed MPC

In the figure, the optimal steering angles $\vec{\delta}_{p,opt}$ that minimizes the defined cost function can be computed by quadratic programming with constraints such as boundary, equality, and inequality constraints. In this study, physical constraints such as steering wheel angle limits, change rate limits, and mechanical constraints between 1st and 2nd axles were applied.

In order to design a fault-tolerant steering controller, an equality constraint was applied to the MPC algorithm. It is assumed that steering axles of the crane are fixed when faults of steering angle sensors or actuators are detected in the axles. Based on this assumption, the equality constraint was formulated so that the MPC algorithm can compute optimal steering angles by considering the steering angles fixed after fault detection. Equation (13) presents an example of equality conditions applied to the MPC based on the fault condition described in Fig. 3.

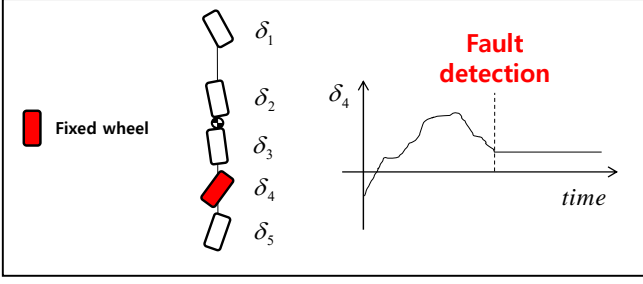


Figure 3. Fault condition of steering system for tolerant control

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & -R & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_{1,k+1} \\ \delta_{2,k+1} \\ \delta_{3,k+1} \\ \delta_{4,k+1} \\ \delta_{5,k+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \delta_{4,k} \\ 0 \end{bmatrix}. \quad (13)$$

where R represents a steering angle ratio between 1st and 2nd axles. When the 4th steering angle is in fault, the optimal steering angles of 1st, 3rd, and 5th axes can be calculated by applying the equality constraint in (13) to the MPC algorithm. A general form of the equality constraint for fault-tolerant steering control can be expressed in a matrix form as follows.

$$\begin{bmatrix} A_{eq,11} & 0 & 0 & 0 & 0 \\ 1 & -R & 0 & 0 & 0 \\ 0 & 0 & A_{eq,33} & 0 & 0 \\ 0 & 0 & 0 & A_{eq,44} & 0 \\ 0 & 0 & 0 & 0 & A_{eq,55} \end{bmatrix} \begin{bmatrix} \delta_{1,k+1} \\ \delta_{2,k+1} \\ \delta_{3,k+1} \\ \delta_{4,k+1} \\ \delta_{5,k+1} \end{bmatrix} = \begin{bmatrix} \delta_{1,k} \\ 0 \\ \delta_{3,k} \\ \delta_{4,k} \\ \delta_{5,k} \end{bmatrix}. \quad (14)$$

A_{eq} in equation (14) has a value of either 0 or 1 depending on fault condition. Specifically, $A_{eq,11}$ has the value of 1 if the 1st axle is in fault. The next section describes the performance evaluation of the designed fault-tolerant steering control algorithm.

III. PERFORMANCE EVALUATION

For the performance evaluation, a planar crane model constructed on Matlab/Simulink environment was used and a curved-path tracking scenario was considered. The following sub-sections present the curved-path scenario, fault condition, and evaluation results.

A. Scenario

Figure 4 shows the curved path considered for an evaluation. The MPC-based fault-tolerant steering control algorithm was designed so that the crane can track the curved path using the information of lateral and yaw angle errors despite the steering system's faults arising from sensors or actuators.

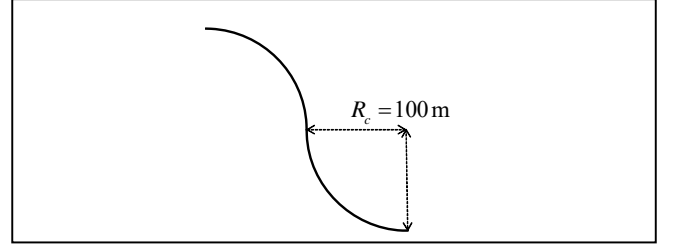


Figure 4. Curved path scenario

A fault signal was applied to the steering system of the crane suddenly during the path tracking.

B. Fault Condition

Figure 5 describes a fault condition applied for the performance evaluation of the designed fault-tolerant steering control algorithm.

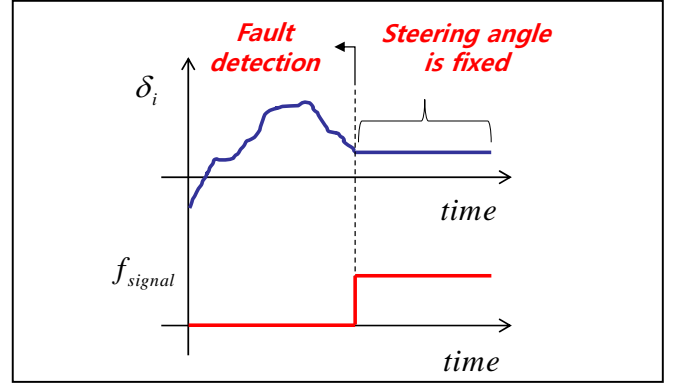


Figure 5. Fault condition applied to the steering system

As shown in the figure, the steering axle in fault is assumed to be fixed when sensor or actuator faults are detected by the fault diagnostic system. f_{signal} represents an applied fault signal in the steering system.

C. Evaluation Results

Figure 6 shows a model schematic to describe the structure of the developed fault-tolerant steering control algorithm.

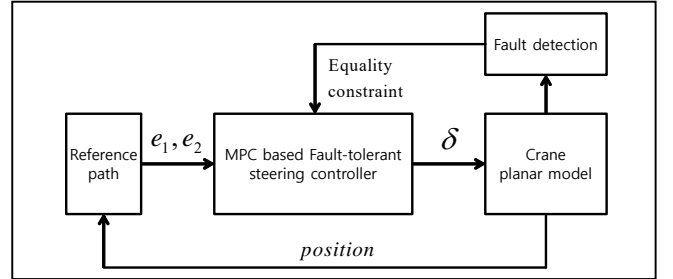


Figure 6. Model schematics for performance evaluation

Figs. 7 - 13 show evaluation results with the assumption that the 1st and 4th axles are in fault at 15 sec, and thus steering

wheel angles are fixed at those axes after 15 sec under the curved-path scenario. Since the 1st and 2nd axes are connected mechanically, the steering wheel angle at the 2nd axle becomes constant after 15 sec according to the status of 1st steering angle (i.e., fixed). And the performance evaluation was conducted under a constant speed condition (45 kph) of the crane. In Fig. 8, it is observed that the MPC computes the optimal 3rd and 5th wheel angles in the fault-tolerant control based on the defined equality constraint of the MPC. This equality constraint enables the optimal steering wheel angles in normal (at 3rd and 5th axes) to be computed in the considered scenario by considering the status of the axes in fault (1st and 2nd axes). Especially, the steering angle of the 5th axle in Fig. 8 was computed reasonably without large oscillation compared to the case of without tolerant control in Fig. 7 since the equality constraint was applied to the MPC based on a consideration of fault conditions. A comparative study between without- and with tolerant control in Fig. 11 (lateral error), 12 (yaw angle error), and 13 (trajectory) also confirms that the fault-tolerant steering control achieves a reduction of error rates and accurate tracking accuracy through calculation of the optimal steering wheel angles at the 3rd and 5th axes.

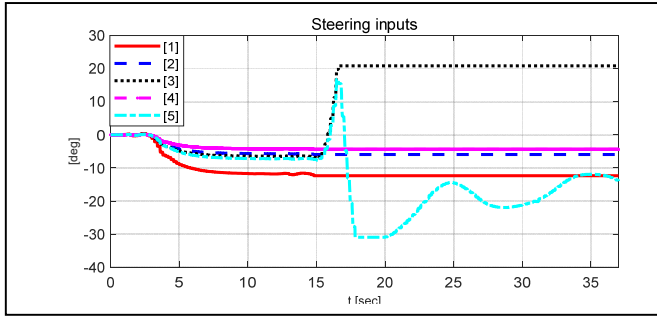


Figure 7. Steering inputs w/o tolerant control: curved path, velocity 45 kph

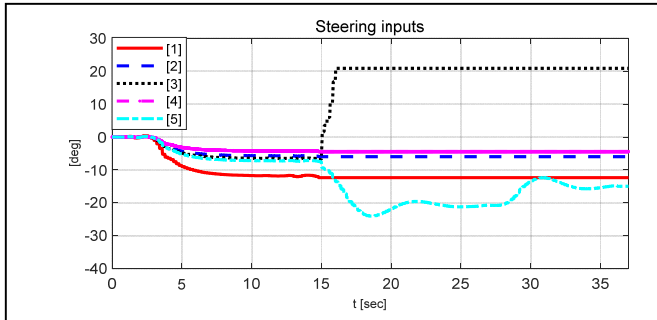


Figure 8. Steering inputs w/ tolerant control: curved path, velocity 45 kph

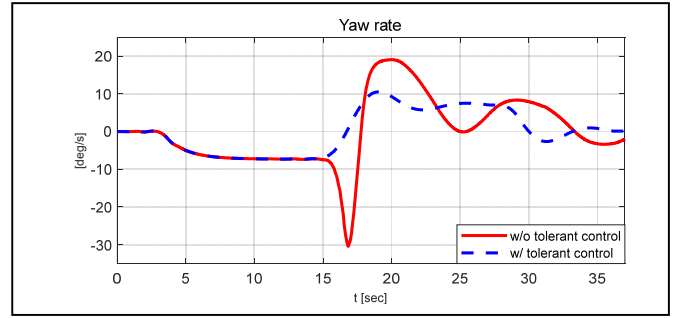


Figure 9. Yaw rate: curved path, velocity 45 kph

In case of yaw rate (Fig. 9) and lateral velocity (Fig. 10), the maximum values with tolerant control in each case are smaller than ones without tolerant control. Because fault conditions were not considered by the equality constraint for the MPC in the steering system without tolerant control, its lateral and yaw angle errors become relatively larger after 15 sec. As it can be seen in figures, the MPC based fault-tolerant steering control algorithm is able to compute optimal steering angles despite occurrence of a fault.

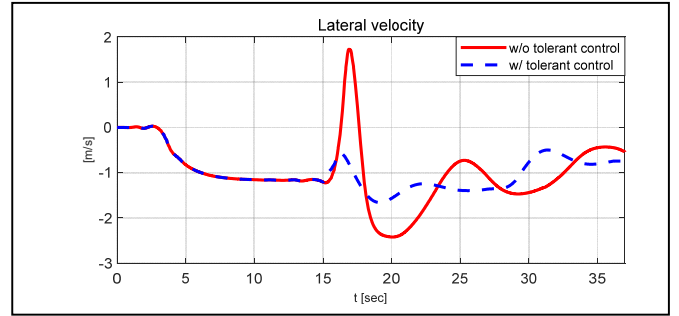


Figure 10. Lateral velocity: curved path, velocity 45 kph

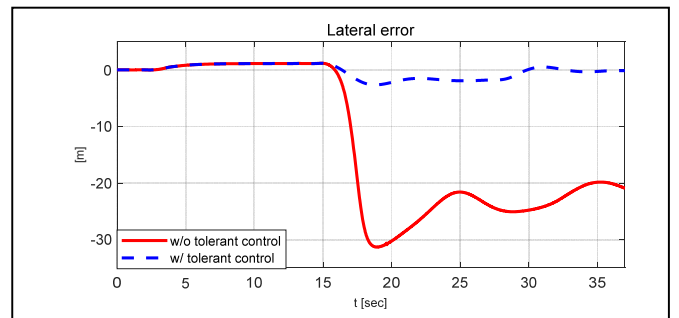


Figure 11. Lateral error: curved path, velocity 45 kph

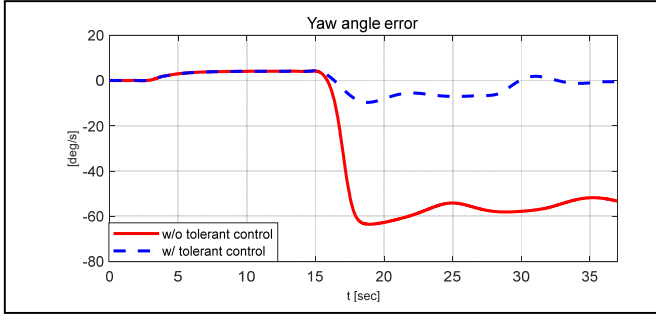


Figure 12. Yaw angle error: curved path, velocity 45 kph

Fig. 14 shows the fault condition applied for the performance evaluation. In addition, the equality constraint was changed based on the equation (14).

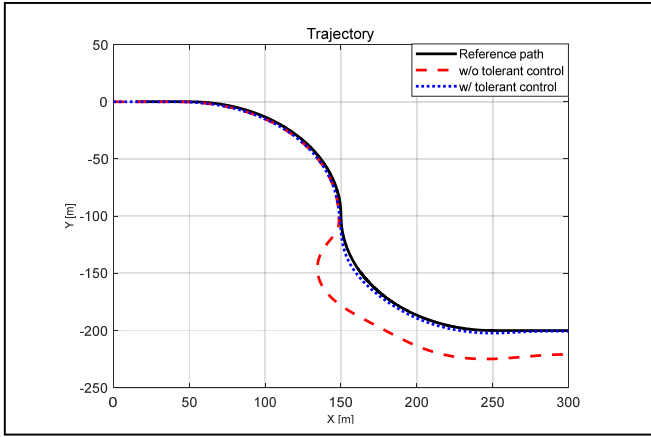


Figure 13. Trajectory: curved path, velocity 45 kph

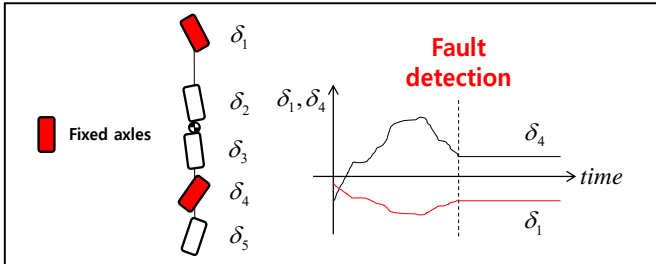


Figure 14. Fault condition applied for the performance evaluation

The maximum values of yaw rate and lateral velocity in case of without tolerant control are larger than ones in case of with tolerant control as seen in Figs. 9-10. The lateral error in Fig. 11 was maintained at a small constant level (max 1.2 m, min -2.7 m) even though a fault occurs. The absolute value of lateral error in case of without tolerant control is about 25 meter. Figure 12 shows that the tolerant control contributes to reduction of the maximum yaw angle error (10 deg with control vs 60 deg without control). From Fig. 13, one can note

that the crane with fault-tolerant steering control algorithm shows better tracking performance to follow the reference path.

IV. CONCLUSION

This study proposed the MPC-based fault-tolerant steering control algorithm for driving safety of all-terrain cranes. To cope with faults in the steering system, The MPC algorithm was designed based on the assumption that steering axles in fault are fixed when the steering angle sensor or actuator fault is detected. The equality constraint in the MPC algorithm was applied to enable the steering wheel in fault to maintain the same steering angle as one at the previous step. To conduct the performance evaluation of the developed control algorithm, a curved path scenario was considered using a planar crane model constructed in Matlab/Simulink. Evaluation results show that the fault-tolerant steering control algorithm has the ability to compute optimal steering angles, and thus to accurately track a desired path although a fault occurs in the steering system. Development of fault detection and reasonable countermeasure algorithms that can improve the performance of tolerant control is considered as a future work. Finally, it is expected that the fault-tolerant steering control algorithm developed in this study can be extensively used to design the fail-safe algorithm in an actual crane's steering system for improving the functional safety.

REFERENCES

- [1] Y. Jeong, Kim, J. Yoon, H. Chong, B. Ko, and K. Yi, "Vehicle sensor and actuator fault detection algorithm for automated vehicles," Intelligent Vehicles Symposium (IV), IEEE, pp. 927-932, 2015.
- [2] K. Oh, S. Park, J. Lee, and K. Yi, "Functional perspective-based probabilistic fault detection and diagnostic algorithm for autonomous vehicle using longitudinal kinematic model," Microsystem Technologies, vol. 24, no. 8, pp. 1-11, 2018.
- [3] Y. Kim, N. Jeon, and H. Lee, "Model based fault detection and isolation for driving motors of a ground vehicle," Sensors and Transducers, vol. 199, no. 4, pp. 67-72, 2016.
- [4] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, "A statistical-based approach for fault detection and diagnosis in a photovoltaic system," in Systems and Control (ICSC), 2017 6th International Conference, pp. 75-80, 2017.
- [5] H. Li, Y. Yang, Y. Wei, and F. Dai, "Observer-based fault reconstruction for linear systems using adaptive sliding mode method," Journal of Robotics, Networking and Artificial Life, vol. 3, no. 4, pp. 236-239, 2017.
- [6] R. Loureiro, S. Benmoussa, Y. Touati, R. Merzouki, and B. O. Bouamama, "Integration of fault diagnosis and fault-tolerant control for health monitoring of a class of MIMO intelligent autonomous vehicles," IEEE Transactions on Vehicular Technology, vol. 63, no. 1, pp. 30-39, 2014.
- [7] K. Oh, J. Seo, J. Kim, K. Yi, "MPC-based approach to optimized steering for minimum turning radius and efficient steering of multi-axle crane," International Journal of Control, Automation and Systems, vol. 15, no. 4, pp. 1799-1813, 2017.

APPENDIX

$$a_{22} = -\frac{2C_1}{mv} - \frac{2C_2}{mv} - \frac{2C_3}{mv} - \frac{2C_4}{mv} - \frac{2C_5}{mv}$$

$$\begin{aligned}
a_{23} &= \frac{2C_1}{m} + \frac{2C_2}{m} + \frac{2C_3}{m} + \frac{2C_4}{m} + \frac{2C_5}{mv} \\
a_{24} &= -\frac{2C_1L_1}{mv} - \frac{2C_2L_2}{mv} + \frac{2C_3L_3}{mv} + \frac{2C_4L_4}{mv} + \frac{2C_5L_5}{mv} \\
a_{42} &= -\frac{2L_4C_1}{Iv} - \frac{2L_2C_2}{Iv} + \frac{2L_3C_3}{Iv} + \frac{2L_4C_4}{Iv} + \frac{2L_5C_5}{Iv} \\
a_{43} &= \frac{2L_4C_1}{I} + \frac{2L_2C_2}{I} - \frac{2L_3C_3}{I} - \frac{2L_4C_4}{I} - \frac{2L_5C_5}{I}
\end{aligned}$$

$$\begin{aligned}
a_{44} &= -\frac{2L_1^2C_1}{Iv} - \frac{2L_2^2C_2}{Iv} - \frac{2L_3^2C_3}{Iv} - \frac{2L_4^2C_4}{Iv} - \frac{2L_5^2C_5}{Iv} \\
b_{21} &= \frac{2C_1}{m}, \quad b_{22} = \frac{2C_2}{m}, \quad b_{23} = \frac{2C_3}{m}, \quad b_{24} = \frac{2C_4}{m}, \quad b_{25} = \frac{2C_5}{m} \\
b_{41} &= \frac{2L_4C_1}{I}, \quad b_{42} = \frac{2L_2C_2}{m}, \quad b_{43} = \frac{2L_3C_3}{m}, \quad b_{44} = \frac{2L_4C_4}{m}, \\
b_{45} &= \frac{2L_5C_5}{m}
\end{aligned}$$

Grey Wolf Optimizer for Automotive Brake-By-Wire (MRB) System Design

Adel Younis

Assistant Professor, Dept. of Mechanical Engineering
Australian College of Kuwait
Kuwait, Kuwait
Email: a.younis@ack.edu.kw
(Corresponding Author)

Zuomin Dong

Professor, Dept. of Mechanical Engineering
University of Victoria
Victoria, BC, Canada
zdong@uvic.ca

Fadi Al Khatib

Assistant Professor, Dept. of Mechanical Engineering
Australian College of Kuwait
Kuwait, Kuwait
Email: f.alkhatib@ack.edu.kw

Abstract— Advanced optimization methods have been continuously introduced and improved over the past decades to improve search efficiency and robustness in solving various and complex optimization problems. In recent years, considerable efforts have been carried out to introduce new nature-inspired optimization algorithms. These advanced stochastic global optimization search methods imitate the biological phenomena in nature to further improve the search efficiency and robustness of the global optimization methods. In this work, the performance of Grey Wolf Optimizer (GWO) is demonstrated through a highly nonlinear design problem – the optimal design of automotive brake-by-wire (MRB) system design. Unlike the conventional brakes, an MRB employs the interaction between a magnetorheological fluid and an applied magnetic field to generate the retarding braking torque. The objective of employing the GWO algorithm was to maximize the braking torque (retarding torque) and minimize the weight of the brake structure. The optimized design parameters illustrated GWO's capability to converge to acceptable and accurate results with less computational time compared to conventional global optimization methods.

Keywords—component; swarm intelligence; gray wolf optimizer; global optimization; brake-by-wire

I. INTRODUCTION (HEADING 1)

This Many optimization algorithms have been introduced over the past years. New ideas and new algorithms evolve to improve the performance of such algorithms. Engineering design problems are highly complex in nature. Therefore,

deriving the gradient is very expensive; requires intensive computational efforts and resources. For that reason, non-gradient optimization algorithms, stochastic and heuristic optimization algorithms [1, 2], became attractive. These algorithms have become popular and have found a lot of interest because of their superior performance and their capability in extracting information on the search space without experiencing any computational complexities.

Despite the fact that many optimization algorithms available today to handle complex engineering problems for different applications, the need for efficient, effective and robust algorithms maintains pace in algorithm evolution and presents a true challenge. Thus, swarm intelligence optimization algorithms recently started to evolve to handle such challenges. Most of swarm intelligence-based algorithms use multi-agents, inspired by the collective behavior of social creatures, like bees, fireflies and ants, as well as from other animal societies like herds of birds, school of fish or pack of wolves. The classical particle swarm optimization (PSO) [3] uses the behavior of fish and birds, while firefly algorithm (FFA) [4] uses the flashing behavior of fireflies. Cuckoo search (CS) [5] is based on the brooding parasitism of some cuckoo species, while bat algorithm uses the echolocation of foraging bats. Ant colony optimization (ACO) [6] uses the chemical language of social ants while the class of bee algorithms are all based on the foraging behavior of honeybees. Swarm intelligence-based algorithms are so popular and widely used nowadays because of its ability of sharing information among multiple agents, so that self-organization, co-evolution and learning during iterations help contribute largely in increasing efficiency of such algorithms.

The other reason is that multiple agents can be parallelized easily to deal with large-scale optimization. This has made such optimization algorithms suitable for highly nonlinear and complex design optimization problems involving expensive analysis and simulation processes such as finite element analysis (FEA) and computational fluid dynamics (CFD).

In this paper, a highly nonlinear and complex real-life engineering design problem is selected to be optimized using swarm intelligence based, the GWO algorithm. The main objective of this work is to determine the optimum design variables for MRB system so that the braking torque can be maximized and the weight of the structure of MRB system can be minimized.

A basic configuration of MRB was introduced by Park et al. [7]. As shown in Fig. 1, in this configuration, a rotating disk (3) is enclosed by a static casing (5), and the gap (7) between the disk and casing is filled with the MR fluid (a typical MR fluid is a type of functional fluid which has a suspension of magnetic particles in inert carrier liquids). A coil winding (6) is embedded on the perimeter of the casing and when electrical current is applied to it, magnetic fields are generated, and the MR fluid in the gap becomes solid-like instantaneously. The shear friction between the rotating disk and the solidified MR fluid provides the controllable retarding torque and the magnitude of the shear friction is directly related to the applied magnetic flux density on the fluid.

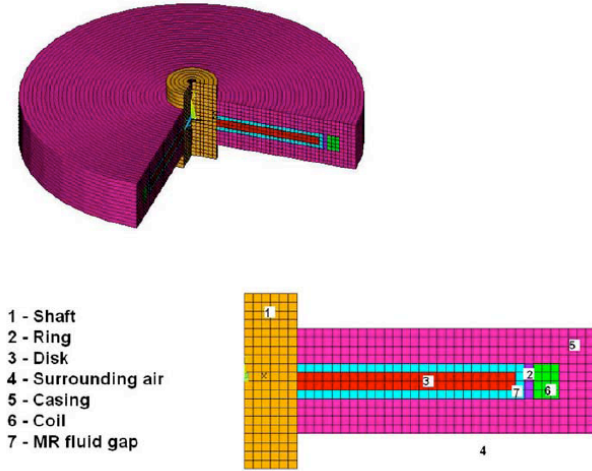


Figure 1: Basic MRB Design (Park et. al. [8])

II. MRB MODEL

First, The MRB design optimization is based on the model for calculating the braking torque resulting from different MRB configurations [9]. The idealized characteristics of the MR fluid can be described effectively by using the Bingham plastic model [10, 11, 12, 13]. According to this model, the total shear stress τ is:

$$\tau = \tau_H \operatorname{sgn}(\dot{\gamma}) + \mu_p \dot{\gamma} \quad (1)$$

where τ_H is the yield stress due to the applied magnetic field, μ_p is the no-field plastic viscosity of the fluid and $\dot{\gamma}$ is the shear rate. The braking torque for the geometry shown in Fig. 2 can be defined as follows:

$$T_b = \int_A \tau r dA = 2\pi N \int_j^z (\tau_H \operatorname{sgn}(\dot{\gamma}) + \mu_p \dot{\gamma}) r^2 dr \quad (2)$$

Where A is the working surface area (the domain where the fluid is activated by applied magnetic field intensity), z and j are the other inner and outer radii of the disk, N is the number of disks used in the enclosure and r is the radial distance from the center of the disk.

Assuming the MR fluid gap in fig. 2 to be very small (e.g., ~ 1 mm), the shear rate can be obtained by:

$$\dot{\gamma} = \frac{rw}{h} \quad (3)$$

with the additional assumption of linear fluid velocity distribution across the gap and no slip conditions. In Equation (3), w is the angular velocity of the disk and h is the thickness of the MR fluid gap. In addition, the yield stress, τ_H can be approximated in terms of the magnetic field intensity applied specifically onto the MR fluid, H_{MRF} , and the MR fluid dependent constant parameters, k and β , i.e.

$$\tau_H = k H_{MRF}^\beta \quad (4)$$

By substituting Eq. (3) and (4), the braking torque equation in Eq. (2) can be rewritten as

$$T_b = 2\pi N \int_j^z \left(k H_{MRF}^\beta \operatorname{sgn}(\dot{\gamma}) + \mu_p \frac{rw}{h} \right) r^2 dr \quad (5)$$

Then, Equation (5) can be divided into the following two parts after the integration:

$$T_H = \frac{2\pi}{3} N k H_{MRF}^\beta (r_z^3 - r_j^3) \quad (6)$$

$$T_\mu = \frac{\pi}{2h} N \mu_p (r_z^4 - r_j^4) w \quad (7)$$

where T_H is the torque generated due to the applied magnetic field and T_μ is the torque generated due to the viscosity of the fluid. Finally, the total braking torque is $T_b = T_\mu + T_H$. From a design point of view, the parameters that can be varied to increase the braking torque generation capacity are: the number of disks (i.e. N), the dimensions and configuration of the magnetic circuit (i.e. r_z , r_j) in addition other structural design parameters shown in Fig. 3, H_{MRF} that is directly related to the applied current density in the electromagnet and materials used in the magnetic circuit.

In order to solve for the braking torque, the magnetic field distribution over the MRB domain has to be calculated. Therefore, a nonlinear finite element model (FEM) for the

MRB is created and, using this model, the braking torque generation for different configurations can be calculated. For the current global optimization algorithm, a one disk MRB configuration, shown in fig. 3, is adapted as the benchmark design.

III. GREY WOLF OPTIMIZER

The Grey Wolf Optimizer (GWO) is a recently proposed method inspired by grey wolves in wildlife [14]. The GWO simulates the leadership hierarchy and hunting strategy of grey wolves in nature. GWO like other nature-inspired algorithm such as GA, PSO, and ACO population based method. There are four different types of groups of grey wolves, which are alpha α , beta β , delta δ and omega ω . Those groups can be used to simulate the leadership hierarchy in the pack. The alpha for example, is commonly responsible for making decisions on hunting, sleeping place, and so on and their decisions must be obeyed by the rest of pack. Due to its dominant role, alpha is placed at the top level of pack. Beta is the second level in the gray wolves family that supports the alpha in decision-making or other pack initiatives. The beta can be either male or female, and it is maybe the best candidate to replace the alpha in any case. The beta acts as counsellor to alpha and applies its orders; while at the same time guide the lower-level groups. Delta is the third level of the team and has to respect alpha and beta, but it dominates the omega. The duty of delta is to protect and guarantee the safety of the pack. Omega is considered to be the lowest ranked among the grey wolf family and play the role of scapegoat. Omega wolves always have to submit to all the other groups (α, β, δ). They are the last wolves that are allowed to eat and their duty is to take care of the newborn. The first three groups (α, β, δ) are considered to lead other group of wolves ω to find where the promising search space exists. During the search on design space, all members update their positions according to the positions of α, β , and δ . In order to complete implementing GWO three steps of hunting including searching for prey, encircling prey, and attacking prey need to be achieved. In the GWO method, through the optimization process, the first three best solutions reached are assumed to be α, β , and δ are likely to be the location of the optimal solution. Through the optimization process, the first three best solutions reached are assumed to be a, p , and S respectively. Meanwhile, the rest of the team members are considered as ω and have to relocation with the respect to α, β , and δ . The hunting is guided by the alpha; however, beta and delta can be also in hunting too. In the algorithm, alpha is the fittest candidate; beta and delta have better knowledge about the potential location of prey. Therefore, the first three best solutions are stored while the rest of search agents including omegas are obliged to update their position based on the position of the best search agent. In the GWO n considered to be the wolf population, where k indicates the current iteration, A and C are random parameters $A = (1, 0)$, $C = (1, 1)$, x_p represent the position vector of the prey, x is the position of grey wolf, a is random value to update position.

Pseudo code of GWO

- 1: Define the objective function

$$f(x), x \in (x_1, \dots, x_n)^T$$
- 2: Initialize the grey wolf population x_i ($i = 1, 2, \dots, n$)
- 3: Initialize a, A , and C (coefficient) vectors
- 4: For each agent calculate the fitness rank them by order $x_\alpha =$ the best search agent, $x_\beta =$ the second best search agent and $x_\delta =$ the third best search agent.
- 5: **While** ($k < \text{Max of iterations}$) or ($f_{\min} > \text{tolerance}$)
- 6: **For** each search agent
- 7: Update the position of the current search agent
- 8: **End for**
- 9: Update a, A , and C
- 10: Calculate the fitness for all search agents
- 11: Update x_α, x_β , and x_δ .
- 12: Update iteration number $k = k+1$
- 13: **End While**
- 14: The final output are x_i^* and f^*

IV. FORMULATION OF THE OPTIMIZATION PROBLEM

The cross-section of the selected one disk configuration of the MRB is shown in Fig. 2. As a next step, the chosen design configuration was optimized for higher braking torque and lower weight. In setting up such an optimization problem for the MRB, a cost function was defined by including the braking torque and weight as functions of the dimensional parameters of the magnetic circuit shown in Fig. 3.

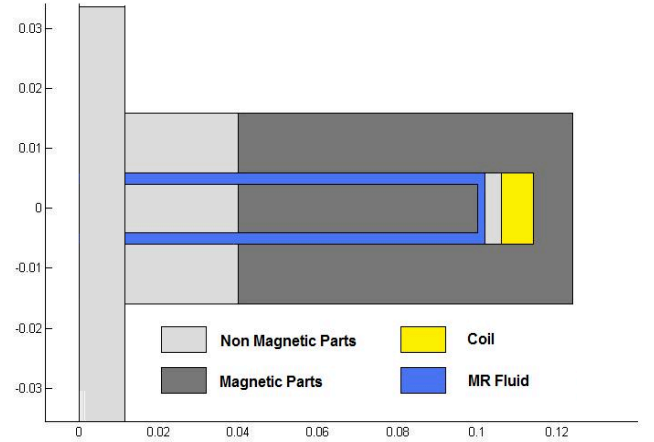


Figure 2. Chosen MRB Based on the Design Criteria (Younis et al. [9])

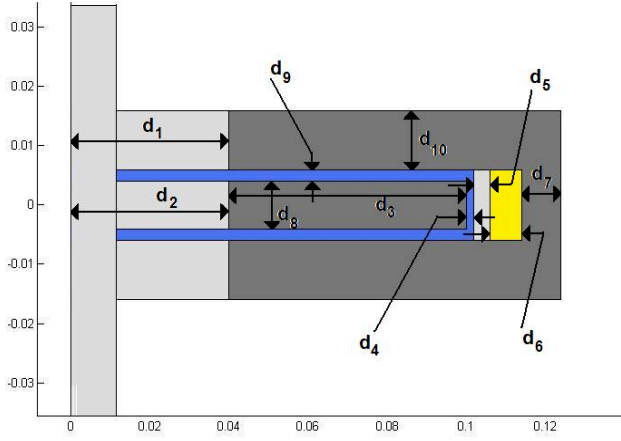


Figure 3. Dimensional Parameters Related to Magnetic Circuit Design (Park *et al.* [8] and Younis *et al.* [9])

The objective function of the MRB optimization problem is defined as

$$\text{Minimize } f(d) = k_w \frac{W}{W_{ref}} - k_t \frac{T_H}{T_{ref}} \quad (1)$$

where;

$$k_w + k_t = 1 \quad (2)$$

Subject to:

$$W < 150 \text{ N, and } d_{brake} < 240 \text{ mm} \quad (3)$$

W_{ref} was obtained considering the overall system weight of the CHB. Moreover, since the braking torque generated by the proposed MRB configuration is comparably less than that of the CHB, T_{ref} was selected to be 20 Nm. This reference torque value was selected by checking a number of random MRB designs, which satisfied the constraints.

As the constraints for the optimization problem, the weight of the actuator was set to be smaller than the weight of the CHB, i.e. $W < 150 \text{ N}$. In addition to this, since the brake should fit into a wheel, the diameter of the MRB is set to be smaller than the inner diameter of the wheel. In this study, the brake is optimized for the standard 13" wheel whose inner diameter is 240 mm, i.e. $d_{brake} < 240 \text{ mm}$.

V. OPTIMIZATION PROBLEM RESULTS AND DISCUSSION

For a given MRB design configuration, COMSOL Multiphysics™, a commercial FE software package, was used to calculate the magnetic flux density within MRB. MATLAB was then used to solve for the weight and the braking torque using the given brake configuration and the FEA result. GWO parameters used for this particular application (MRB) are shown in Table I.

Table I: GWO parameters

Population size	40
α	0.8346
β	0.1128
δ	0.1521
ω	1
ϵ	1

The global optimization program has been run multiple times and the best optimum result is accepted and used as the measure for search efficiency, as shown in Table II. A PC with Intel (R) Core (TM) i7-7500 CPU@ 2.7GHz processor and 16 GB memory is used for the design optimization task.

TABLE II: OPTIMUM DESIGN VARIABLES (YOUNIS ET AL. [9])

Parameters	Optimum Value (mm)	LB-UB (mm)
d_1	50	5-50
d_2	50	5-50
d_3	30	30-120
d_4	1.1	1-3
d_5	5	5-10
d_6	22	5-30
d_7	9.9	5-20
d_8	20	1-20
d_9	1	1-3
d_{10}	15.2	5-20

The obtained optimal design parameters from the GWO algorithm are shown in Tables II and III. One of the very important considerations of MRB system with many rotating parts is the weight. The optimal design using GWO leads to a minimum weight of 13.5 kg. In this application, the GWO algorithm led to more accurate optimum. The obtained minimum weight is reported in Table III.

TABLE III. OPTIMIZED MRP PROTOTYPE SPECIFICATIONS

Main Specifications of MRP	Optimization Algorithm Used
	GWO
Diameter (mm)	240.00
Coil Wire Size	AWD 21 (dia. 0.77mm)
MR Fluid Used	MRF-132DG
Maximum Current Applied (A)	1.80
Maximum Current Density Applied (A/m ²)	2.54E+6
Magnetic Materials Used	Steel 1018
Weight (kg)	15.4
Amount of MR Fluid Used (m ³)	5.44E-05
Maximum Braking Torque (N.m)	34.6
Weight (Kg)	13.5

The other important consideration is the maximum torque generated during braking. The amount of the generated braking torque obtained by GWO is 34.9 Nm as reported in Table III.

The relationship between the applied current and the generated torque is shown in Fig. 4. It can be observed from Fig. 4 that the relationship is proportional. when the applied current increases the generated torque increases.

VI. CONCLUSION

In this paper, a swarm intelligence global optimizer, Grey Wolf Optimizer (GWO) was used for determining the optimum design parameters of the brake-by-wire (MRB) design system. The optimal design parameters of the MRB system were identified with maximum braking torque generated at the rotating disks and minimum brake system weight. GWO yielded results that were promising in terms of computational time, accuracy and robustness. GWO capabilities and efficiency as a swarm intelligence global optimization algorithm for complex engineering applications requiring intensive numerical computation were demonstrated.

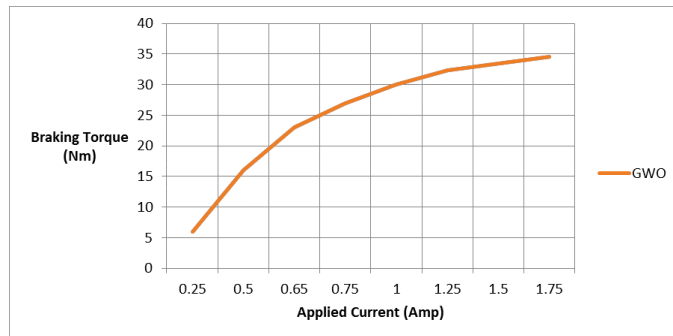


Figure 4: Effect of increasing the applied current on braking torque generation

REFERENCES

- [1] P. Gill, E. Murray, and M. Wright, eds., 1981, Practical Optimization Academic Press, New York.
- [2] B. Baritumpa, and E. Hendrix, "On The Investigation of Stochastic Global Optimization Algorithms," *Journal of Global Optimization*, 31(4), pp. 567-578, 2005.
- [3] J. Kennedy, and R. Eberhart, "Particle Swarm Optimization," *Proceedings of IEEE International Conference on Neural Networks*, pp. 1942-1948, 1995.
- [4] X. S. Yang, "Firefly algorithm, Levy flights and global optimization," in *Research and Development in Intelligent Systems XXVI*, Springer, London, UK, 2010; 209-218.
- [5] Yang XS, Deb S. Engineering Optimisation by Cuckoo Search, *Int. J. Mathematical Modelling and Numerical Optimisation* 2010; 4: 330-343.
- [6] M. Dorigo, V. Maniezzo, and A. Colomi, "Ant System: Optimization by A Colony of Cooperating Agents," *IEEE Transactions On Systems, Man, and Cybernetics-Part B Cybernetics*, 26(1), pp. 29-41, 1996.
- [7] E. Park, D. Stoikov, D. Falcao, and A. S., "A Performance Evaluation of an Automotive Magnetorheological Brake Design with a Sliding Mode Controller," *Mechatronics*, 16, pp. 405-416, 2006.
- [8] K. Karakoc, E. Park, and A. Suleman, "Design Consideration for an Automotive Magnetorheological Brake," *Mechatronics*, 18, pp. 434-447, 2008.
- [9] Younis A., Karakoc K., Dong Z., Park E., Suleman A., 2011, "Application of SEUMRE global optimization algorithm in automotive magnetorheological brake design", *Struct Multidisc Optim* (2011) 44:761-772, 2011
- [10] R. Phillips, 1969, "Engineering Applications of Fluids with a Variable Yield Stress," University of California, Berkeley, CA
- [11] J. An, and DS. Kwon, "Modeling of a Magnetorheological Actuator Including Magnetic Hysteresis," *Journal of Intelligent Material Systems and Structures*, 14(9), pp. 541-550, 2003.
- [12] W. Kordonsky, "Elements and Devices Based on Magnetorheological Effect," *Journal of Intelligent Material Systems and Structures*, 4(1), pp. 65-69, 1993.
- [13] K. Weiss, J. Carlson, and D. Nixon, "Viscoelastic Properties of Magneto- and Electrorheological Fluids," *Journal of Intelligent Material Systems and Structures*, 5(11), pp. 772-775, 1994
- [14] Mirjalili S. How affective is the Gray Wolf Optimizer in training multi-layer perceptrons. *ApplIntell* DOI 2015 10.1007/s10489-014-0645.

A REVIEW OF ACTIVE AERODYNAMIC CONTROL FOR INCREASING SAFETY OF HIGH-SPEED ROAD VEHICLES

Mohammed Hammad and Yuping He

Department of Automotive, Mechanical and Manufacturing Engineering
University of Ontario Institute of Technology, Oshawa, Ontario, Canada

Abstract—Under high speed and high lateral acceleration (i.e., high g) cornering maneuvers, the lateral tire-road forces are saturated, and the vehicle will lose its lateral stability. In order to increase the safety of road vehicles, various active safety systems have been proposed and developed. These active safety systems include active front/rear steering, torque vectoring, differential braking, active camber control, active roll moment control, etc. Essentially, all these active safety systems are manipulating the longitudinal/lateral forces and aligning torques at the contact patches of tire/road to improve the yaw and roll stability of road vehicles. With a given road condition and vehicle payload, the longitudinal/lateral forces and aligning torques due to the interfaces of tire/road are limited. A lot of severe road vehicle accidents frequently occur under cornering maneuvers at high speeds. Aerodynamic drag and downforce or lift increase significantly with vehicle speed. It is indicated that active aerodynamic control is effective for enhancing the lateral stability of high-speed vehicles. However, little attention has been paid to this research topic. This paper is intended to review the investigations conducted in the last two decades. It is seen that interesting vehicle aerodynamic control studies have been conducted, and numerous open problems need to be addressed.

Keywords – *Aerodynamic control; active safety systems; lateral stability; high-speed road vehicles; CFD.*

I. INTRODUCTION

Worldwide around 1.25 million people are killed per year in road vehicle accidents [1]. To increase vehicle safety, various active safety systems have been commercialized [2]. The past two decades have witnessed a dramatic development of active safety systems. In 2002, Toyota launched lane keeping assistance systems [3]. In 2007, the U.S. Government established FMVSS126, a vehicle safety standard that requires all passenger cars sold in North America to include an electronic stability control (ESC) system starting in 2012 [4]. In 2014, the Society of Automotive Engineers (SAE) first defined the levels of automated driving systems for road vehicles [5]. To date, various active safety systems have been proposed and developed to increase the safety of road vehicles. These active safety systems include active front/rear steering, torque vectoring, differential braking, active camber control, active roll moment control, etc. [6-10]. Essentially, these active safety systems

mainly manipulate the tire/road forces to enhance the lateral stability of vehicles [11]. However, these active safety systems are hindered by the saturation of tire/road forces at high lateral accelerations and on icy slippery roads [12]. Huge number of severe road vehicle accidents frequently occur under cornering maneuvers at high speeds. Aerodynamic drag and downforce or lift increase significantly with vehicle speed [13]. It is indicated that active aerodynamic control is effective for enhancing the lateral stability of high-speed vehicles [11]. However, little attention has been paid to this research topic.

Road vehicle aerodynamics concerns the effects arising due to motion of the vehicle through, or relative to, the air [14]. Drag and lift are two widely used terms for signifying aerodynamic forces in horizontal and vertical directions, respectively. Drag reduction of cars to improve longitudinal performance and fuel economy is a well-known research area [15, 16]. When it comes to high speed vehicles, efforts to increase negative lift or downforce and thus the road holding capability of the vehicle are well documented [17]. Some work on identifying and stabilizing the dynamic states of a vehicle under crosswinds is also done, especially on bigger vehicles, such as buses and articulated heavy vehicles [18]. Recently, attempts have been made in accurate and robust mathematical modelling of the aerodynamic forces and coefficients to fit into full scale simulations [19-20].

Active aerodynamic control is to manipulate moving surfaces (bodies) to generate varying aerodynamic forces and moments for enhancing vehicle dynamics and increasing vehicle safety. It is important to note that all the vehicle dynamic control systems discussed so far rely on forces developed at the tire-road surface and would cease to perform adequately under the saturation of such forces. Active aerodynamic control systems take into account the forces/torques due to aerodynamics for vehicle dynamics control. The research work in this area, although relatively less in quantity, has spanned two decades and has seen an exponential growth and has a promising future. Therefore, it is worthwhile to review all studies have been conducted so far.

The remainder of the paper is organized as follows. Section II reviews the effects of active aerodynamic control systems on the longitudinal, vertical, and lateral dynamics of high-speed vehicles. Section III introduces various vehicle models and wing configurations for active aerodynamic control systems. Section IV outlines the fundamental law governing the relationship of

drag or lift with the key system parameters. Section V overviews the main control strategies used in active aerodynamic control systems. Section VI briefly introduces the CFD-based aerodynamic data used for active aerodynamic control. Finally, conclusions are drawn in Section VII.

II. EFFECTS OF ACTIVE AERODYNAMIC CONTROL ON VEHICLE DYNAMICS

The effects due to aerodynamics pose impacts on the longitudinal, lateral, and vertical dynamics of road vehicles. It is well-known that aerodynamic drag significantly affects the fuel economy and the longitudinal dynamics of road vehicles traveling at high speeds [21]. It is demonstrated that the coefficient of drag of a truck may be reduced by 35% by controlling the boundary layer separation by the momentum injection method using a rotating cylinder [22]. An active flow control method was proposed to reduce aerodynamic drag of a van [23]. An actively translating rear diffuser device was proposed to reduce the aerodynamic drag experienced by passenger cars [24]. Due to the stringent government standard/regulation requirements on improving the fuel economy and decreasing greenhouse gas emissions, research on active airflow control gains increasing attention.

Various attempts have been made to enhance vertical vehicle dynamic and improve ride quality of road vehicles. Doniselli et al. examined the aerodynamic effects on ride quality of a vehicle running at high speed on randomly profiled roads [25]. Savkoor proposed an aerodynamic control device for reducing pitch and heave of truck cabins [26]. Meijaard et al. investigated the potential for vehicle ride quality improvement using both suspension and aerodynamic actuators [27]. It is shown that combining the active suspension and aerodynamic actuators can achieve a considerable improvement of the ride comfort with a small deterioration of the road holding. Active aerodynamic surfaces were explored to improve vehicle road-holding without affecting ride quality [28]. On the other hand, active aerodynamic surfaces were proposed to improve the ride comfort, and the results showed that the proposed method shows improvements of the order of 30% in ride quality without negative effects on road-holding at high speeds [29]. It is well-known that there is a trade-off between ride quality and road holding capability. The aforementioned studies indicate that active aerodynamic control provide an effective way to relieve the trade-off.

Active aerodynamic control strategies have also been studied to improve the lateral dynamics of road vehicles, i.e., handling performance and lateral stability. It was shown that aerodynamically generated direct yaw and roll moments can significantly improve the handling and lateral stability of road vehicles [30]. An active aerodynamic system was examined to improve the handling and safety of a small-size race car in a line-change maneuver and driving on wet roads [31]. Active inverted wings were proposed to increase the downward forces on the tires for enhancing the road holding, thereby improving handling and stability of race cars [32]. To improve the handling and safety of high-speed vehicles, normal force distribution was controlled by manipulating the angle of attack of the front and rear spoilers, and the performance of the active aerodynamic control system was analyzed through numerical simulation

based on a nonlinear vehicle model [33]. A physical prototype of an actively controlled rear wing was fabricated, and the essential design parameters, functionality and features of the active wing were introduced [11].

For the purpose of examining the paradox due to the co-existence of drag and lift of spoilers, CFD simulation was conducted, and the effects of the aerodynamic spoilers' effects on the lateral and longitudinal dynamics were analyzed [34]. To investigate the downforce and drag, as well as their relation, CFD simulations were performed for a car with an active rear split spoiler at different spoiler angle of attack and a different speeds [35]. The simulation results were compared with the experimental data derived from a wind tunnel on the physical car and the spoiler prototype. It is demonstrated that a proposed active split rear wing can improve the lateral stability of a race car over tight cornering maneuver, and will not degrade the longitudinal dynamics of the vehicle [36].

III. VEHICLE MODELS AND WING CONFIGURATIONS

The objectives of active aerodynamic control systems are to improve the longitudinal, vertical, and lateral dynamics of high-speed road vehicles. To date, majority of studies on active aerodynamic control are based on numerical simulations. To this end, various vehicle models with different complexity and fidelity have been generated and developed. In addition, to design and develop controllers for active aerodynamic control systems, simplified linear or nonlinear vehicle models have been widely used. It is, hence, very important that an appropriate vehicle model is built or chosen for application. The equations of the dynamic model has to include the aerodynamic forces as actuators in order to bring them to use in the controller design. Hence, alongside choosing a dynamic model, selection of the number of wings and their positioning is also an important consideration.

Savkoor and Chou [30] chose a four degrees of freedom (4-DOF) bicycle model, setting the longitudinal velocity to be a kinematic constraint, to analyze vehicle handling under aerodynamic influence. Their work proposed a few configurations of the aerodynamic surfaces, such as spoilers and rudder devices as shown in Fig 1. Savkoor [26] designed an in-plane 13-DOF truck model to implement an active aerodynamic control system, as shown in Fig 2, to counteract the heave and pitch of a truck cabin.

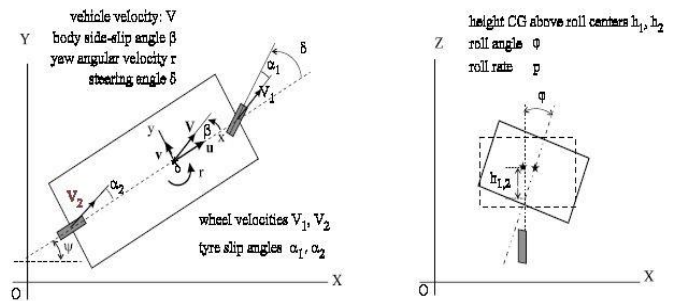


Figure 1. 4-DOF bicycle model [30].

The paper examined and compared three different locations of the wing on top of the cabin to get optimum results. Meijaard, et al. [27] used a 4-DOF half car model, as shown in Fig 3, that incorporates bounce and pitch motions of the car and bounce of two un-sprung masses in their study of vehicle ride improvement. Quarter car models, as shown in Fig 4, were used by Corno et al. in [28, 29] to conduct simulations on improvement of road-holding and ride control under the effect of active aerodynamic control. In the first of the two publications, the wing was located at the center of the tires and in the latter, on the sprung mass. Hammad and He [36] selected a bicycle model with two rear wings to work on improving the lateral stability of high speed vehicles.

Linear models as such for estimating handling behavior have the tire forces estimated as a linear product of cornering stiffness and slip angle.

Application of aerodynamics becomes meaningful at higher speeds and tight cornering maneuvers where the vehicle operates beyond the linear dynamic range and exhibits nonlinear characteristics. Such models in these cases do not represent a real world vehicle as all the states are function of vehicle forward speed and hence the application of active aerodynamic is not entirely accurate. Similarly, quarter car models or half-car

models also does not truly represent the dynamics of a vehicle at high speeds.

Models with greater complexity and accuracy were used in the works by Diba et al. [31, 32, 37], who developed an 8-DOF nonlinear model, as shown in Fig. 5, to study handling behavior under the action of aerodynamic forces. In these works, four wings have been deployed at four corners of the vehicle to control downforce on each tire separately. Ahangarnejad and Melzi [33] further worked on handling of a race car using a commercially sourced 14-DOF model, as shown in Fig. 6. They deployed two active controlled wings at the front and rear of the vehicle body.

Few of the studies have been solely based on CAD models and CFD simulations which do not involve any dynamic modelling. These studies, most of which concern drag reduction have served the purpose of proposing ideal design and placement of the aerodynamic surfaces on the vehicle body. He et al. [11, 34, 35] conducted extensive simulations involving different software packages to build and validate a wing model that ideally serves the purpose in active aerodynamic control systems so far that handling is concerned.

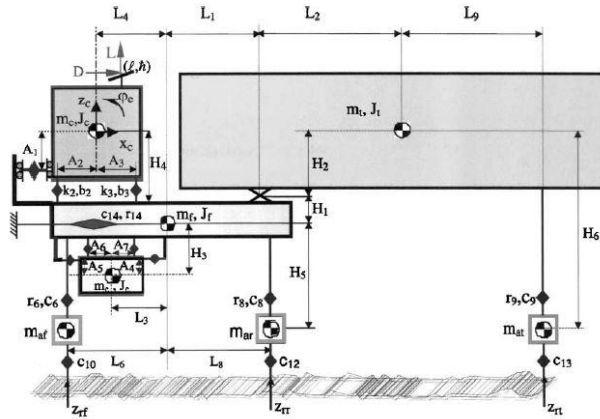


Figure 2. 13-DOF in-plane truck model [26].

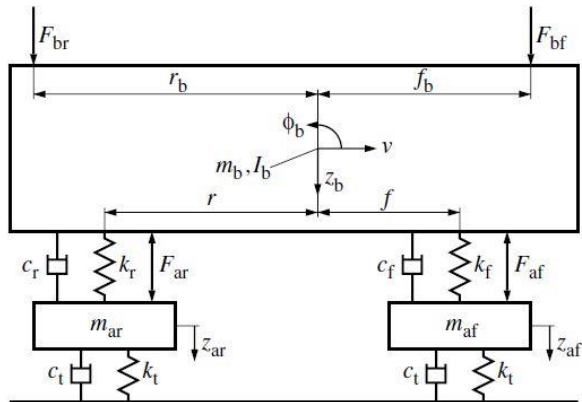


Figure 3. 4 DOF half car model [27].

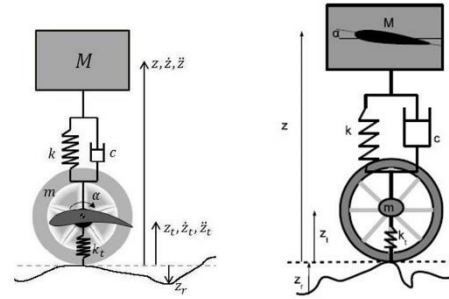


Figure 4. Quarter car models [28, 29].

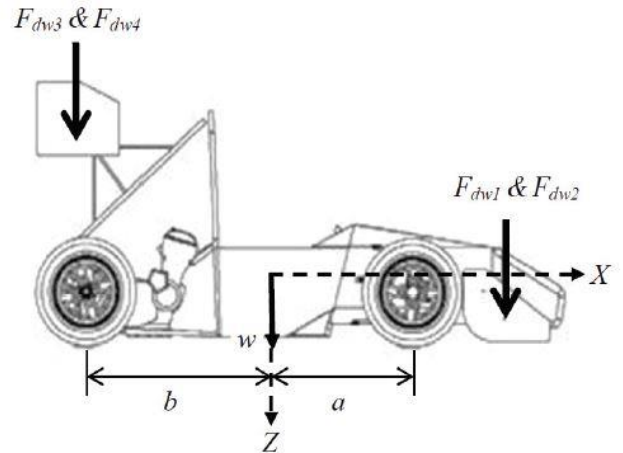


Figure 5. 8 DOF full car model [32]

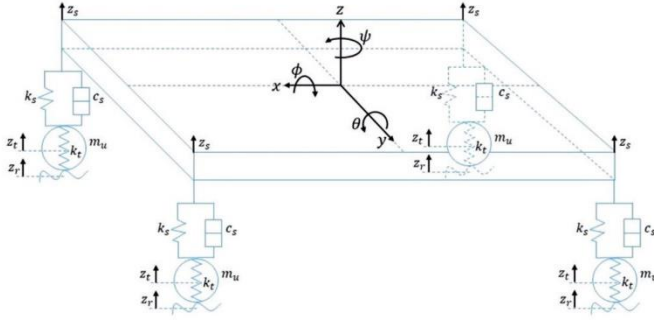


Figure 6. 14-DOF full car model [33].

Other studies have presented CAD vehicle models with aerodynamic surfaces to study the reduction in drag forces [22, 23, 24].

IV. AERODYNAMIC FORCES

Proper modelling of aerodynamic forces is another necessary step to develop an ACTIVE AERODYNAMIC CONTROL system. The equations of drag and lift are commonly found and given by

$$Drag = C_D \rho A \frac{V^2}{2} \quad (1)$$

$$Lift = C_L \rho A \frac{V^2}{2} \quad (2)$$

where C_D and C_L are aerodynamic coefficients, ρ is the air density in kg/m^3 , A is Area of the aerodynamic surface in m^2 and V is the relative velocity of wind and the aerodynamic surface, in the above equations. In case of moving surfaces the angle with which the surfaces move is a factor of the aerodynamic forces. The forces in this case would be given by

$$Drag = C_D \rho A \frac{V^2}{2} \alpha \quad (3)$$

$$Lift = C_L \rho A \frac{V^2}{2} \alpha \quad (4)$$

where α is the angle of attack. Moreover, most importantly, the aerodynamic coefficients are not constant as well and they are functions of the angle of attack as well. This fact was not realized in [30]. Later developments, however, took this aerodynamic characteristic into account in [28, 29] but the implementation in control system design is not presented. Recent papers have acknowledged this fully and also presented the usage methodology [31, 32, 33, 37].

V. CONTROL STRATEGIES

The control system design and strategy is perhaps the crux in development of an active aerodynamic control system. Its accurate development based on the dynamics of the vehicle and the desired objective is an essential feature of an active aerodynamic control system. This chapter presents a brief overview of the control methods and strategies used over the years.

When the idea first prefaced, the control strategy was merely proposed but not actually implemented. In [30] two different strategies, namely full state feedback and open loop feedback

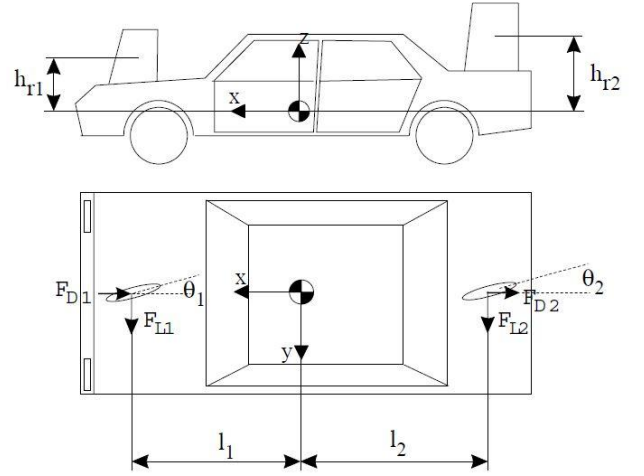


Figure 7. Aerodynamic devices (rudders) mounted on a vehicle [30]

		Transferred Load (N)								
Velocity (m/s)		1	2	3	4	5	6	7	8	9
	4	0	5.0	13.1	20.4	29.2	42.2	42.2	42.2	42.2
	5	0	0	4.2	9.8	14.6	19.2	24.3	30.6	42.2
	6	0	0	0	2.5	7.1	10.7	13.9	17.1	20.4
	7	0	0	0	0	0.2	4.6	7.6	10.2	12.5
	8	0	0	0	0	0	0	2.1	5.0	7.1
	9	0	0	0	0	0	0	0	0	0
		Angle of Attack (deg)								

Figure 8. Example of a lookup table used in an ACTIVE AERODYNAMIC CONTROL system [29]

with pre-filter and a combination of both were suggested to calculate the demand on the roll and yaw actuators. This demand would be fulfilled by trying various configurations of rear split spoiler and/or rudder devices at front and/or rear of the vehicle as shown in Fig. 7. An extension and reinforcement of the same concept can be found in [29], the control system consists of two layers where in the first layer system states are estimated and based on them the required forces are calculated using a lookup table to stabilize the system. Another lookup table, as shown in Fig. 8, is used to estimate the attack angles of the two wings, four in case of [20, 25], so as to generate the required forces. The second part of the control system consists of a PID controller to move the actuators so as to rotate the wings by the given angle of attack.

In [30] an LQR logic was used to minimize a certain defined performance index according to the required objectives. In [28], the control laws are discussed but the control system is not fully described. In [37], a control strategy is proposed which leads the way to [31, 32]. A description of a physical prototype of the split wings, as shown in Fig. 9, and its hardware is given in [11] without presenting the control algorithm. In [38], a more robust control method in the form of H_∞ was used to control the designed performance indices. The LQR method was used again in [36] to test the impact of active aerodynamic control on handling behavior.



Figure 9. Physical prototype of active aerodynamic control system mounted on a vehicle [11].

As implied in previous discussion, accurate development of active aerodynamic control systems need development of a full vehicle model which is nonlinear in nature. The LQR and H_∞ control strategies, which although explain the benefits of active aerodynamics, are not applicable to nonlinear systems. The control systems built for nonlinear models are a combination of pre-determined action and PID control for actuators. Despite such control systems being easy to implement and understand, are tedious to build and, notably, not robust. There is a scope to improve and sharpen the control system technique and make it more robust.

VI. ACQUISITION OF AERODYNAMIC DATA

Aerodynamic data means the CFD simulations or wind tunnel tests to estimate the aerodynamic coefficients of the concerned aerodynamic surface at varying angle of attacks. This is an important set of data that is fitted into the dynamic model to complete a holistic active aerodynamic control system. This subsection in very brief overviews the steps taken by researchers in this regard. It also includes testing of different aerodynamic surfaces and estimation of best location.

It is needless to say that researchers who only focused on CAD models and CFD models have abundant aerodynamic data to approximate the optimum location of actuators. Some are a further step ahead, determining the best airfoil shapes for active aerodynamic control application, as mentioned previously. A good amount of CFD simulations are presented in [26] to estimate the best location for mounting the aerodynamic actuators. While only CFD results are presented in every other work, [33] contains the data that is obtained from wind tunnel tests.

A scope of improvement is found in this regard as all the work presenting study on different wing shapes and sizes do not go on to design the control systems and all the works that includes design of control systems is done based on the results of a single wing design. Here arises a need to couple both the works for more fruitful results.

VII. CONCLUSIONS

The past two decades have witnessed the increasing research efforts on active aerodynamic control systems. The published research results indicate that for high-speed road vehicles, active aerodynamic control systems can significantly improve the longitudinal, vertical, and lateral dynamics. In addition to the streamlined body design, active aerodynamic surfaces can further reduce the aerodynamic drag to improve the acceleration performance of the vehicle and fuel economy. It is well-known that in road vehicle suspension design, there is a trade-off between ride quality and tire road holding capability. Active aerodynamic control provides an effective way to relieve the trade-off by adaptively varying the downforce of tires. It is shown that coordinating the active aerodynamic and active suspension actuators can greatly improve ride quality without degrading road holding capability. Moreover, by means of increasing the downforce of tires and adaptively adjusting axle load distributions, vehicle handling and lateral stability can be enhanced. Some research efforts have been made on examining the effects of active aerodynamic control on drags and downforces.

Various vehicle models with different complexity and fidelity were developed for exploring active aerodynamic control systems. For the purpose of numerical simulation, the vehicle models have been derived with high complexity and fidelity, while to design aerodynamic controllers, the objective-oriented vehicle models with less complexity and fidelity have been derived. A variety of control strategies have been proposed for active aerodynamic control, including PID, LQR, etc. Recently, robust control techniques, such as H_∞ , have been used to control the active spoilers. These control strategies have been implemented with varying degrees of success. Interesting CFD simulations have been conducted to investigate the effects of varying the attack angles of spoilers and vehicle forward speeds on aerodynamic drags and downforces. The past studies on active aerodynamic control paved the road for further development of techniques in this field.

In the past studies on active aerodynamic control, few attempts have been made on exploring the coordination of active aerodynamic control with active steering or differential braking or torque vectoring. It is expected that the coordinated aerodynamic control may greatly improve the stability and safety performance of existing active safety systems. Research efforts are required on the co-simulation by combining models of CFD, controllers, and multi-body vehicle systems. With the effective co-simulation, the effects of aerodynamics and control strategies on the vehicle dynamics can be extensively and intensively explored. Very few physical prototypes of active aerodynamic control system have been fabricated, and limited wind tunnel testing data for validating aerodynamic control were reported. To advance active aerodynamic control techniques, further investigations into the fabrication of physical prototype and wind tunnel testing are needed.

ACKNOWLEDGEMENTS

Financial support of this research by the Natural Science and Engineering Research Council of Canada [grant no. DDG-2017-00030] is gratefully acknowledged.

REFERENCES

- [1] World Health Organization. Global status report on road safety 2015. [cited 2018 Oct.13]. http://www.who.int/violence_injury_prevention/roadsafety/status/2015/en/.
- [2] Van Zanten A. Bosch ESP systems: 5 years of experience. 2000; SAE Technical Paper 2000-01-1633.
- [3] Katourakis D, Alireaei M, de Winter JCF, et al. Shared control for road departure prevention. IEEE International Conference on Systems, Man, and Cybernetic, 9-12 October 2011, Anchorage, AK, USA.
- [4] Office of Regulatory Analysis and Evaluation (National Center for Statistics and Analysis), Final regulatory impact analysis: FMVSS No. 126 – Electronic stability control systems. National Highway Traffic Safety Administration, 2007, U.S. Department of Transportation, 169 pages.
- [5] SAE International. Surface vehicle recommended practice J3016: Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Society of Automotive Engineering; 2014.
- [6] A. Higuchi and Y. Saitoh, "Optimal control of four wheel steering vehicle," *Vehicle System Dynamics*, vol. 22, no. 5-6, pp. 397-410, 1993.
- [7] Y. Shibahata, K. Shimada, and T. Tomari, "Improvement of vehicle maneuverability by direct yaw moment control," *Vehicle System Dynamics*, vol. 22, no. 5-6, pp. 465-481, 1993.
- [8] M. Abe, N. Ohkubo, and Y. Kano, "A direct yaw moment control for improving limit performance of vehicle handling-comparison and cooperation with 4WS," *Vehicle System Dynamics*, vol. 25, no. S1, pp. 3-23, 1996.
- [9] M. Nagai, Y. Hirano, and S. Yamanaka, "Integrated control of active rear wheel steering and direct yaw moment control," *Vehicle System Dynamics*, vol. 27, no. 5-6, pp. 357-370, 1997.
- [10] M. Yamashita, K. Fujimori, K. Hayakawa, and H. Kimura, "Application of H ∞ control to active suspension systems," *Automatica*, vol. 30, no. 11, pp. 1717-1729, 1994.
- [11] Y. He, "Design of an Actively Controlled Aerodynamic Wing to Increase High-Speed Vehicle Safety," SAE Technical Paper 0148-7191, 2013.
- [12] D. T. Ayyagari and Y. He, "Aerodynamic analysis of an active rear split spoiler for improving lateral stability of high-speed vehicles," *International Journal of Vehicle Systems Modelling and Testing*, vol. 12, no. 3-4, pp. 217-239, 2017.
- [13] Milliken, W.F., Milliken, D.L., *Race Car Vehicle Dynamics*, SAE International, 1994.
- [14] F. Wild. (2017). What Is Aerodynamics? Available: <https://www.nasa.gov/audience/forstudents/5-8/features/nasa-knows/what-is-aerodynamics-58.html>
- [15] J.-F. Beaudoin and J.-L. Aider, "Drag and lift reduction of a 3D bluff body using flaps," *Experiments in fluids*, vol. 44, no. 4, p. 491, 2008.
- [16] G. Fourié, L. Keirsbulck, L. Labraga, and P. Gilliéron, "Bluff-body drag reduction using a deflector," *Experiments in Fluids*, vol. 50, no. 2, pp. 385-395, 2011.
- [17] X. Zhang, W. Toet, and J. Zeriha, "Ground effect aerodynamics of race cars," *Applied Mechanics Reviews*, vol. 59, no. 1, pp. 33-49, 2006.
- [18] C. Baker, F. Cheli, A. Orellano, N. Paradot, C. Proppe, and D. Rocchi, "Cross-wind effects on road and rail vehicles," *Vehicle system dynamics*, vol. 47, no. 8, pp. 983-1022, 2009.
- [19] J. A. Mohrfeld-Halterman and M. Uddin, "High fidelity quasi steady-state aerodynamic model effects on race vehicle performance predictions using multi-body simulation," (in English), *Vehicle System Dynamics*, vol. 54, no. 7, pp. 963-981, Jul 2016.
- [20] J. A. Mohrfeld-Halterman and M. Uddin, "High Fidelity Quasi Steady State Aerodynamic Model Development and Effects on Race Vehicle Performance Predictions," (in English), *Sae International Journal of Passenger Cars-Mechanical Systems*, vol. 9, no. 2, pp. 603-611, Jun 2016.
- [21] J. Y. Wong, *Theory of ground vehicles*. Fourth Edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 2008.
- [22] S. N. Singh, L. Rai, P. Puri, and A. Bhatnagar, "Effect of moving surface on the aerodynamic drag of road vehicles," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 219, no. 2, pp. 127-134, 2005.
- [23] B. Harinaldi, R. Tarakka, and S. P. Simanungkalit, "Computational analysis of active flow control to reduce aerodynamics drag on a van model," *International Journal of Mechanical & Mechatronics Engineering IJMME-IJENS*, vol. 11, no. 3, 2011.
- [24] S.-O. Kang et al., "Actively translating a rear diffuser device for the aerodynamic drag reduction of a passenger car," *International Journal of Automotive Technology*, vol. 13, no. 4, pp. 583-592, 2012.
- [25] C. Doniselli, G. Mastinu, and M. Gobbi, "Aerodynamic effects on ride comfort and road holding of automobiles," *Vehicle System Dynamics*, vol. 25, no. S1, pp. 99-125, 1996.
- [26] A. Savkoor, "Design of actively controlled aerodynamic devices for reducing pitch and heave of truck cabins," *JSAE Review*, vol. 22, no. 4, pp. 421-434, Oct 2001.
- [27] J. P. Meijaard, A. R. Savkoor, and G. Lodewijks, "Potential for vehicle ride improvement using both suspension and aerodynamic actuators," 2005, vol. 1, pp. 385-390: IEEE.
- [28] M. Corno, S. Bottelli, G. Panzani, M. Tanelli, C. Spelta, and S. M. Savaresi, "Improving high speed road-holding using actively controlled aerodynamic surfaces," 2013, pp. 1493-1498: IEEE.
- [29] M. Corno, S. Bottelli, M. Tanelli, C. Spelta, and S. M. Savaresi, "Active control of aerodynamic surfaces for ride control in sport vehicles," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 7553-7558, 2014.
- [30] A. R. Savkoor and C. T. Chou, "Application of aerodynamic actuators to improve vehicle handling," *Vehicle System Dynamics*, vol. 32, no. 4-5, pp. 345-374, Nov 1999.
- [31] F. Diba, A. Barari, and E. Esmailzadeh, "Active Aerodynamic System to Improve the Safety and Handling of Race Cars in Lane Change and Wet Road Maneuvers," 2012, pp. 417-423: American Society of Mechanical Engineers.
- [32] F. Diba, A. Barari, and E. Esmailzadeh, "Handling and safety enhancement of race cars using active aerodynamic systems," *Vehicle System Dynamics*, vol. 52, no. 9, pp. 1171-1190, Sep 2014.
- [33] A. Hosseinian Ahangarnejad and S. Melzi, "Numerical analysis of the influence of an actively controlled spoiler on the handling of a sports car," *Journal of Vibration and Control*, p. 1077546318754683, 2018.
- [34] T. Sikder, S. Kapoor, and Y. He, "Optimizing Dynamic Performance of High-Speed Road Vehicles Using Aerodynamic Aids," 2016, pp. V007T09A060-V007T09A060: American Society of Mechanical Engineers.
- [35] D. T. Ayyagari and Y. He, "Aerodynamic analysis of an active rear split spoiler for improving lateral stability of high-speed vehicles," *International Journal of Vehicle Systems Modelling and Testing*, vol. 12, no. 3-4, pp. 217-239, 2017.
- [36] M. Hammad, K. Qureshi, and Y. He, "Safety and Lateral Dynamics Improvement of a Race Car Using Active Rear Wing Control," SAE Technical Paper 0148-7191, 2019.
- [37] A. Barari, F. Diba, and E. Esmailzadeh, "Down Force Control of the Low Velocity Racing Car Using Active Aerodynamic Inverse Wings," 2011, pp. 739-745: American Society of Mechanical Engineers.
- [38] Y. Wu and Z. Chen, "Improving Road Holding and Ride Comfort of Vehicle Using Dual Active Aerodynamic Surfaces," 2018, pp. 1-5: IEEE.

DESIGN ANALYSIS OF AN ACTIVE TRAILER STEERING SYSTEM FOR MULTI-TRAILER ARTICULATED HEAVY VEHICLES

Mutaz Keldani

University of Ontario Institute of Technology
Department of Automotive Engineering
Oshawa, Canada
Mutaz.keldani@uoit.net

Yuping He

University of Ontario Institute of Technology
Department of Automotive Engineering
Oshawa, Canada
Yuping.he@uoit.ca

Abstract— Multi-trailer articulated heavy vehicles (MTAHV) usage is increasing around the world due to their positive impact on the environment, gas consumption and cost-effectiveness. Despite the advantages of the MTAHV, they exhibit low lateral stability and poor maneuverability. To address this problem, a 4 degrees of freedom (DOF) yaw-plane MTAHV model is derived and used to design an active trailer steering (ATS) system to improve both the lateral stability and maneuverability of the combination. An active trailer steering system based on a linear quadratic regulator (LQR) controller is used in this work to control the steering of wheels on trailer axles. Also, this work presents a sensitivity analysis to study the effect of several parameters on the stability of the combination, e.g. trailer's payload and the center of gravity (CG) location. The proposed control method is applied to one type of the MTAHV known as B-Train Double. Numerical simulations are performed using TruckSim software to demonstrate the effectiveness of the ATS system.

Keywords- *B-train double; active trailer steering; ATS; LQR; lateral acceleration; RWA ratio; eigenvalues; stability; yaw-plane; 4-DOF model*

I. INTRODUCTION

B-train double is one type of long combinational vehicles (LCVs). A B-train double has a tractor hauling two full-length semitrailers, and the total length of the large vehicle can be as long as 40 meters [1]. A B-train double can transport double the goods at a lower cost compared to a typical tractor-semitrailer combination [1]. Furthermore, the LCV reduces the greenhouse gas emission by approximately one-third, making it one of the most environmentally friendly transportation tools.

Although several benefits are demonstrated, LCVs exhibit unique unstable motions at high speeds due to their length and high CG [2]. The unstable motion modes, such as Jack-knifing, trailer swaying, and rollover, occur at high speeds. Whereas, at low speeds, LCVs face poor path-following off-tracking (PFOT), which results in the poor maneuverability [2].

Trailer sway is known as the side-to-side movement about the hitch caused by several factors, such as crosswind [3]. This

unique motion mode occurs due to the low yaw-damping ratio of the rearmost trailer [4]. The damping ratio is influenced by the forward speed of the combination, where the combination demonstrates a lack of damping ratio at higher speeds [5]. Severe trailer sway could lead to the dynamic phenomenon known as jack-knifing, where the articulated angle reaches a high value. Furthermore, sudden braking while negotiating sharp turns could lead the combination to lose lateral stability and cause it to jack-knife.

Various active safety systems have been proposed and investigated to overcome and reduce the dangers associated with LCVs. These active safety systems include ATS, torque vectoring (TV), active differential braking system (ADBS) [6] and active roll control [7]. Simulation results of the ADBS indicate an enhancement in the lateral stability and maneuverability of the combination [8]. However, this could lead to undesired forward speed reduction and excessive wear of brakes, tires and damage to the roads [9].

The proposed ATS system is designed based on the performance measures, i.e., the lateral acceleration and the rearward amplification (RWA) of the combination. RWA is defined as the ratio of the peak lateral acceleration of the rearmost trailer to that of the lead unit. This work aims to maintain an RWA ratio of 1 [5] to ensure better lateral stability. Articulated vehicles tend to have a low rollover threshold in terms of lateral acceleration, which is well below 0.5g [10]. In other words, controlling the lateral acceleration could result in a reduced RWA ratio where the vehicle's tendency to rollover reduces well below the threshold value [11].

The ATS controller designed in this work is based on the LQR technique. The LQR-based controller is designed and fine-tuned for improving the previously mentioned performance measures. Two test procedures are used to validate the performance of the LQR-based ATS controller: 1) an open-loop single lane change maneuver, and 2) a closed-loop single lane change maneuver [12]. The LQR-based ATS controller is designed for high speeds as high as 100 km/h. In addition, the LQR-based ATS controller is designed under specified operation conditions, e.g. constant trailer payload and forward speed. The numerical simulation results show an improvement

in the lateral trajectory and the lateral acceleration of the combination.

A sensitivity analysis is conducted to predict and study the dynamic stability of the vehicle system. This procedure is used to estimate the stability boundaries of the system, which is also known as the critical speed. To conduct the sensitivity analysis, several typical system parameters are selected, which include trailer's payload and CG location.

The rest of the paper is organized as follows. The second section presents the design and validation of the 4 DOF yaw-plane B-train model. The third section presents results and discussion. Finally, conclusions are drawn in the fourth section.

II. VEHICLE SYSTEM MODELING

Vehicle dynamics are considered to be a highly non-linear system where they do not exhibit a proportional relationship between the input and the output. Moreover, it is nearly impossible to design and derive all the non-linear dynamics in a model, making deriving a non-linear model a challenging task. In addition, a non-linear system reduces the computation speed due to the complexity of the model. To address this problem, this section presents the modelling and validating of a 4-DOF yaw-plane model of a B-train double. The derived model is then validated with a highly non-linear model using TruckSim software package, to ensure that the linear model has a similar response to that of the non-linear.

A. 4-DOF Yaw-plane Model of a B-train double

Fig. 1, shows the configurations of a B-train double, which consists of a tractor hauling two semi-trailers connected by a fifth-wheel coupling. The tractor consists of a front axle and two rear axles, while both trailers consist of three axles each.

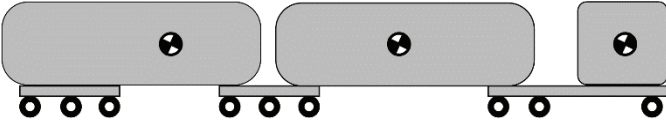


Figure 1. Configuration of a B-train double.

Since this paper focuses on the lateral dynamics of the B-train double, a yaw-plane model is derived using the schematic outline shown in Fig. 2. As shown in Fig. 2, a single track model is used where a single wheel represents each axle.

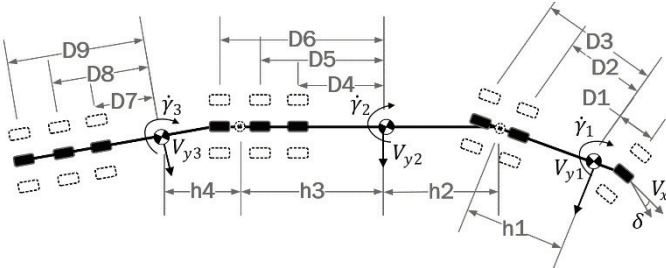


Figure 2. Schematic representation of the 4-DOF yaw-plane model for B-train double.

To simplify and linearize the yaw-plane model, the following assumptions are made: 1) Forward speed is constant for all the units; as a result, V_x is equal to V_{x1}, V_{x2}, V_{x3} . 2) The articulation

angles between the units are insignificant and neglected. 3) The pitch and roll motions are neglected. 4) The units move at a constant speed without any braking force or aerodynamic forces. 5) The tire model used is linear, and the slip angles ($\alpha_{1 \rightarrow 9}$) produced due to the road-tire friction are in the linear region where, $0^\circ \leq \alpha \leq 10^\circ$.

The 4-DOF model is developed using four independent motions, the tractor's lateral velocity (V_{y1}) and yaw rate ($\dot{\gamma}_1$), first trailer's and second trailer's yaw rates ($\dot{\gamma}_2, \dot{\gamma}_3$) respectively. The governing equations of motion for the B-train double are listed below.

$$m_1 * (\dot{V}_{y1} + V_x * \dot{\gamma}_1) = F_{y1} + F_{y2} + F_{y3} - F_{h1} \quad (1)$$

$$I_1 * \ddot{\gamma}_1 = D_1 * F_{y1} - D_2 * F_{y2} - D_3 * F_{y3} + F_{h1} * h_1 \quad (2)$$

$$m_2 * (\dot{V}_{y2} + V_x * \dot{\gamma}_2) = F_{y4} + F_{y5} + F_{y6} - F_{h2} \quad (3)$$

$$I_2 \ddot{\gamma}_2 = D_4 F_{y4} - D_5 F_{y5} - D_6 * F_{y6} + F_{h1} * h_2 + F_{h2} h_3 \quad (4)$$

$$m_3 * (\dot{V}_{y3} + V_x * \dot{\gamma}_3) = F_{y7} + F_{y8} + F_{y9} + F_{h2} \quad (5)$$

$$I_3 \ddot{\gamma}_3 = -1 * (D_7 F_{y7} + D_8 F_{y8} + D_9 F_{y9}) + F_{h2} * h_4 \quad (6)$$

where equations 1,2 represent the lateral velocity and the yaw rate of the tractor respectively. Equations 3,4 represent lateral velocity and the yaw rate of the first trailer and equations 5,6 represent lateral velocity and the yaw rate of the second trailer.

$F_{y1 \rightarrow 9}$ is the lateral forces due to the slip angles produced between the tires and road surface and can be expressed as follow.

$$F_{yi} = -C_i * \alpha_i, i = 1 \rightarrow 9 \quad (7)$$

Due to the fifth-wheel coupling, the lateral motion of the articulation points is exposed to the following kinematic constraints:

$$\dot{V}_{y2} + \dot{\gamma}_2 * h_2 = \dot{V}_{y1} - \dot{\gamma}_1 * h_1 + V_x * \dot{\gamma}_1 - V_{x2} * \dot{\gamma}_2 \quad (8)$$

$$\dot{V}_{y3} + \dot{\gamma}_3 * h_4 = \dot{V}_{y2} - \dot{\gamma}_2 * h_3 + V_x * \dot{\gamma}_2 - V_{x3} * \dot{\gamma}_3 \quad (9)$$

where equation 7 refers to the first fifth-wheel located on the first trailer and equation 8 represents the second fifth-wheel located on the second trailer.

By eliminating the reaction forces (F_{h1}, F_{h2}) at the fifth-wheels, the B-train double can be expressed in the state space representation as follow:

$$\dot{x} = Ax(t) + Bu(t) \quad (10)$$

where, x is the state variables expressed as,

$$x = [V_{y1} \dot{\gamma}_1 V_{y2} \dot{\gamma}_2 V_{y3} \dot{\gamma}_3]^T \quad (11)$$

and u is the input matrix $u = [\delta]^T$.

B. Non-linear TruckSim Model of a B-train double

As mentioned before, developing a 4-DOF linear model simplifies the process of designing and controlling the B-train

double lateral dynamics. However, some of the assumptions made could affect the accuracy of the model. For example, in the linear model, both the roll and pitch angles are neglected while B-train exhibit load transfer at high speeds under transient maneuvers. To investigate this issue, TruckSim software package is used to validate the 4-DOF yaw-plane model. TruckSim consists of several vehicle models including B-train double trailer. The software package has been developed and validated using real-life experimental data using different test procedures [2].

C. 4-DOF Yaw-Plane Model Validation

Since the main aim of this paper is the lateral dynamics of the B-train double, the lateral acceleration and the yaw rate are used to validate the 4-DOF model with the highly non-linear TruckSim model. In this process, both the TruckSim model and the derived model perform an open-loop single lane change (SLC) maneuver at a speed of 88 km/h as specified by SAE J2179 [13]. The test maneuver can be seen in Fig. 3, where a sine wave input is sent to the front axle of the tractor with a frequency of 0.4 Hz and an amplitude of 0.0175 rad \approx 1.002676 degrees. This maneuver was utilized by many researchers to validate their linear models [12, 13, 14, 15].

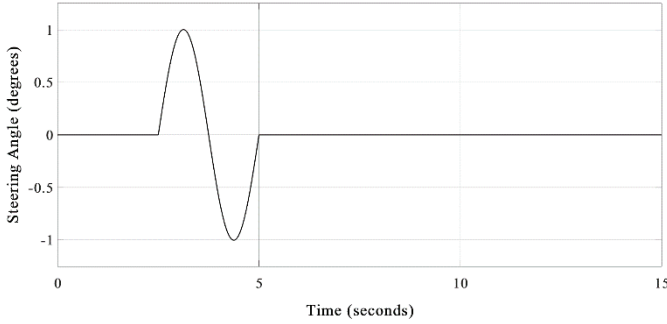


Figure 3. Time history of steering wheel input for the SLC maneuver.

Once both models perform the SLC at the specified conditions, the lateral acceleration and yaw rate graphs of both models get compared to validate the 4-DOF model. Fig. 4 and Fig. 5, show the time history of the lateral acceleration and the yaw rate of the B-train double respectively.

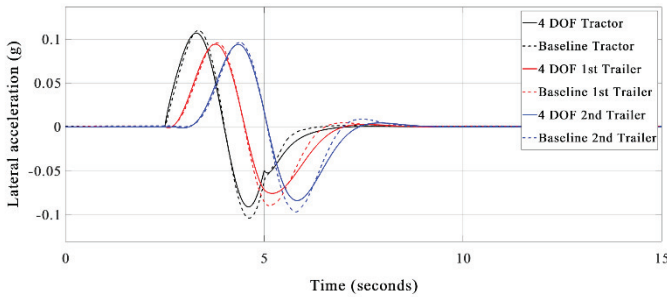


Figure 4. Time history of lateral acceleration of the 4-DOF and TruckSim models.

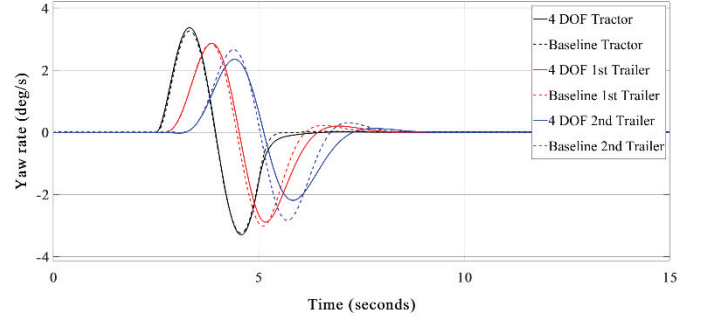


Figure 5. Time history of yaw rate of the 4-DOF and TruckSim models.

Fig. 4, illustrates the lateral acceleration curves for both the 4-DOF model and the TruckSim in response to the SLC maneuver. As expected, the derived model behaved similarly to that of TruckSim. It can also be noted, that the tractor exhibited the highest lateral acceleration peak followed by the first trailer and second trailer correspondingly. As mentioned earlier, the yaw-plane model was developed based on a few assumptions that simplify the model, yet could have an effect on the accuracy of the model. The TruckSim model demonstrates higher lateral acceleration peaks due to the load transfer under the SLC maneuver, unlike the 4-DOF yaw-plane model, since the roll angles are neglected. Moreover, the TruckSim model has a higher overshoot and takes longer to stabilize, due to the non-linearity of the model. Overall, the models are in good agreement from the lateral acceleration standpoint.

The yaw rate curves of the 4-DOF and TruckSim models demonstrate a similar response to that of the lateral acceleration curves shown in Fig. 4. However, in the yaw-rate curves, it can be noted that the tractor is at a better agreement than the first and second trailers. The second trailer of the TruckSim baseline model exhibits higher peaks at the end of the maneuver. The results of the SLC maneuver were as expected and proves the validity of the 4DOF model. Since the 4-DOF model showed similar lateral dynamics to that of the TruckSim, it can further be used in the implementation of the LQR controller.

D. LQR Controller Design

As mentioned previously, the linear model was derived for several reasons such as; simplicity of the design, computation efficiency and linearity of the system allows feasibility for numerous control strategies. The proposed ATS system is designed using the LQR control technique. The state space representation shown in equation (10) is used to implement the LQR controller, where the gain K controls all the state variables mentioned in equations (11). The primary objective of using the LQR is to reduce the value of the cost function J , shown in the equation (11) below. [4]

$$J = \int_0^{\infty} x^T Q x + u^T R u + 2x^T N u) dt \quad (12)$$

The variables x and u were defined earlier as the state variables and input to the system respectively, and Matrices Q and R are defined as the weighting matrices [4]. The system state feedback law in LQR can be defined as, $u = -Kx(t)$, where K

is the gain matrix. The weighting matrices are tuned by the user to define the effect of the controller on each of the system states. The LQR updates the system matrix A and can be expressed as $A_{controller} = A - BK$ [4]. The trial and error method of tuning is used in this work, to obtain the desired performance.

To validate the implemented LQR and its impact on the stability of the system, the co-simulation environment is performed using MATLAB/Simulink and TruckSim. A closed loop SLC maneuver with the same conditions mentioned previously in the model validation sub-secession is executed. The lateral dynamic response of the B-train double of the 4-DOF model is then compared to that of the TruckSim model. Fig. 6, shows the MATLAB/Simulink co-simulation with TruckSim environment.

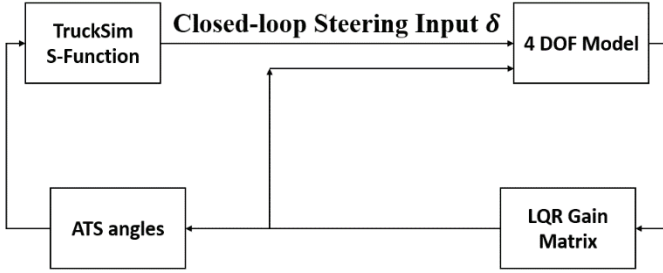


Figure 6. Block diagram of a LQR controller.

As shown in Fig. 6, a driver model in TruckSim software package is used to perform the closed loop SLC maneuver and provides the steering input to the 4-DOF model. The ATS angles then get generated and fed to TruckSim in a closed-loop feedback control system to achieve the desired performance.

E. Sensitivity Analysis

To design an effective control strategy, the critical speed of the B-train double has to be known first. The critical speed is also defined as the stable boundaries of a system [3]. Once the combination exceeds the critical speed, the vehicle units lose their lateral stability. To design an efficient and effective ATS, the LQR must operate beyond the critical speed is reached.

In this section, the eigenvalue analysis of the 4-DOF model is studied to estimate the unstable motion modes [15]. As discussed earlier, the 4-DOF model is written in the state space representation where A resembles the system matrix. To perform the eigenvalue analysis, the system matrix A is utilized to compute the complex eigenvalues that take the following form.

$$S_{1,2} = R_e \pm j\omega_d \quad (13)$$

where R_e and ω_d represent the real and imaginary part of the eigenvalues respectively. The real and imaginary parts of the eigenvalues are then used to calculate the damping ratio. The damping ratio ζ is expressed as the following equation.

$$\zeta = \frac{-R_e}{\sqrt{R_e^2 + \omega_d^2}} \quad (14)$$

The damping ratio is expressed as a function of vehicle forward speed to identify the critical speed of the vehicle. The damping ratio takes a minimum value of -1 and a maximum value of 1. The motion modes of the vehicle units are considered to be stable when the damping ratio is within the positive region ($0 < \xi \leq 1$). The closer the damping ratio to value zero, the closer the vehicle is to lose stability. Moreover, the vehicle motion modes are considered unstable, once the damping ratio is within the negative region ($-1 \leq \xi < 0$). [15]

As stated, to synthesize an optimal controller, it is essential to determine the critical vehicle parameters and their impact on its dynamic performance. Since MTAHVs are used to transport goods, they are subjected to different payloads. Changing the payload of the trailers influence the location of the CG. To address this issue, the conducted eigenvalues analysis consists of the following variations: 1) Different payloads 2) Different CG locations.

One of the advantages of LCVs, is the large available cargo capacity, allowing manufacturers to transport goods for less cost [16]. However, the LCVs are subjected to a maximum of 26000 Kg hauling capacity per trailer [1]. In this sub-section, the stability of the system is evaluated under three different payloads; an empty payload of 0 Kg, 10000 Kg payload and a full payload of 26000 Kg.

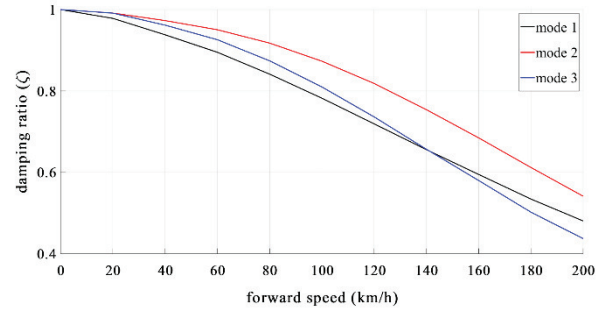


Figure 7. Damping ratio with respect to the vehicle forward speed with empty payload.

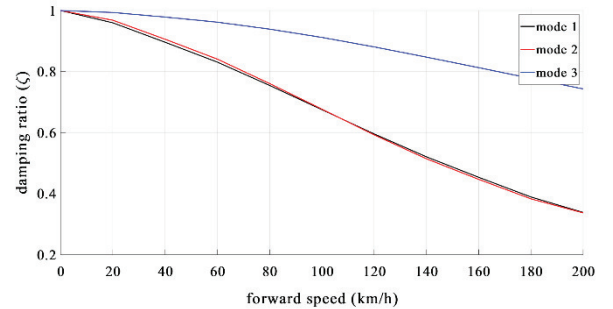


Figure 8. Damping ratio with respect to the vehicle forward speed with 10000Kg payload.

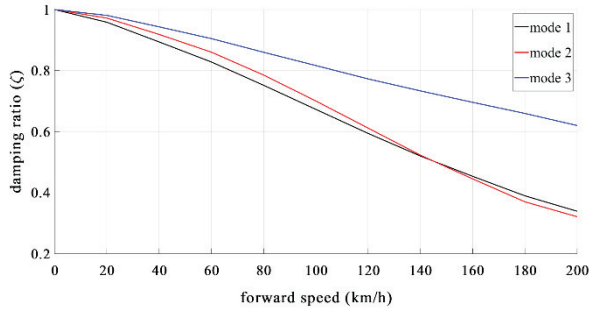


Figure 9. Damping ratio with respect to the vehicle forward speed with full payload.

Fig. 7-9, show the damping ratio with respect to the vehicle forward speed for the 0 Kg, 10000 Kg and 26000 Kg respectively. It can be noted from the figures, that the damping ratio for the three motion modes decreases with an increase in vehicle forward speed. Furthermore, the full payload case scored the least damping ratio when compared to that of the 0 Kg and 10000 Kg payloads. Moreover, all figures demonstrate that the vehicle never experiences instability. However, in the 4-DOF model, the roll angle of the vehicle units is neglected, which could influence the damping ratio at high speeds. Above all, the damping ratio of all three modes is within the positive region, the critical speed is undefined and assumed to be beyond $200 \frac{km}{h}$.

It was concluded after performing the first eigenvalue analysis, that the payload affected the damping ratio, though the system maintained stability. For the second eigenvalue analysis, the effect of the CG location is studied. For this procedure, the payload was 26000 Kg, to further investigate the stability of the worst case scenario. Fig. 10, is a schematic configuration of the trailers, where three different CG locations are labelled. The distances 1, 2, 3 are measured with respect to the hitch point of the trailer. The CG locations 1, 2, 3 takes each of the values 1.385m, 3.385m and 6.385m accordingly.

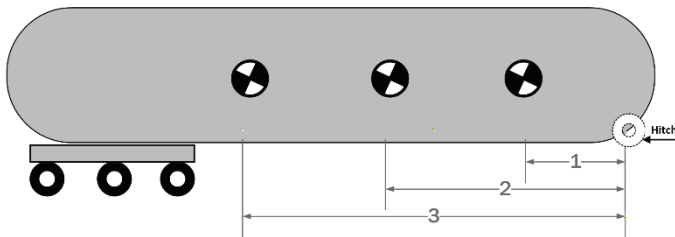


Figure 10. Different CG configurations

Fig. 11-13, show the damping ratio versus the vehicle forward speed for CG locations 1, 2, 3 respectively.

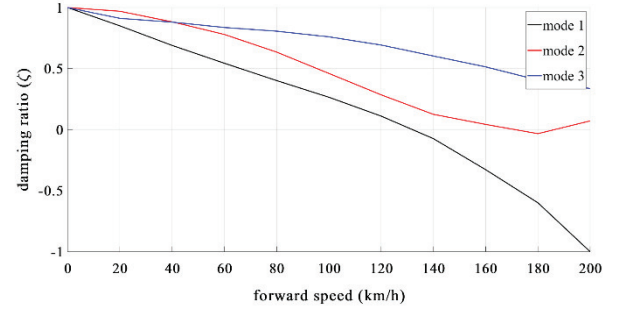


Figure 11. Damping ratio with respect to the vehicle forward speed at CG location 1.

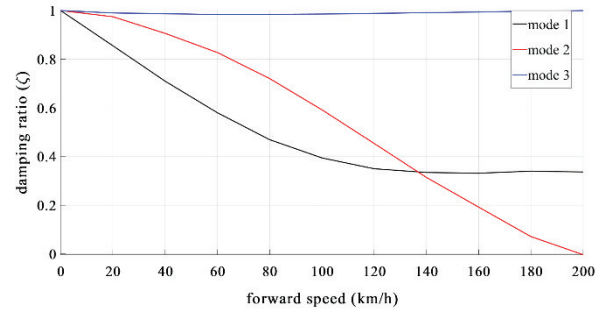


Figure 12. Damping ratio with respect to the vehicle forward speed at CG location 2.

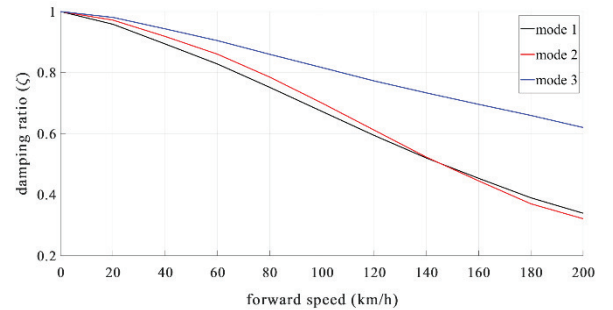


Figure 13. Damping ratio with respect to the vehicle forward speed at CG location 3.

It can be noted from the figures when the CG was located on point 1, one of the modes had a damping ratio in the negative region, which indicates instability. The critical speed can be defined using Fig. 11, to be $132 \frac{km}{h}$. The conducted analysis show that the vehicle is most stable when the CG is at location 3.

III. RESULTS AND DISCUSSION

In this section, the dynamic response of the baseline vehicle and the LQR based ATS vehicle are compared. A closed-loop SLC shown in Fig. 3, is used as an input for both the TruckSim and 4-DOF yaw-plane model. The tests were conducted at a constant speed of 100 km/h to test the effectiveness of the proposed LQR based ATS system. The performance measure is the lateral acceleration peaks and the RWA ratio to improve the

lateral stability and maneuverability of the combination. The primary objective of the ATS system is to reduce the lateral acceleration of the vehicle units while maintaining an RWA ratio value close to 1. Fig. 14-15, demonstrate the lateral acceleration and yaw rate response of the 4-DOF baseline model and the 4-DOF ATS model respectively.

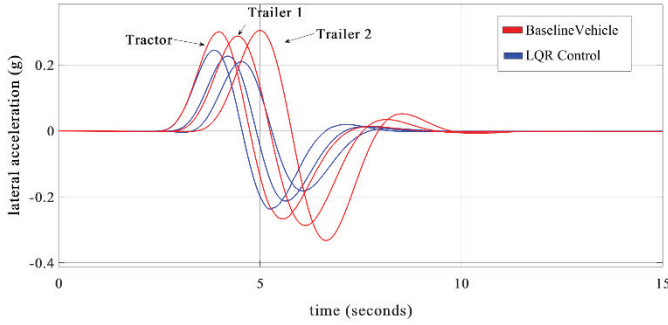


Figure 14. Time history of lateral acceleration for the 4-DOF model with and without LQR controller.

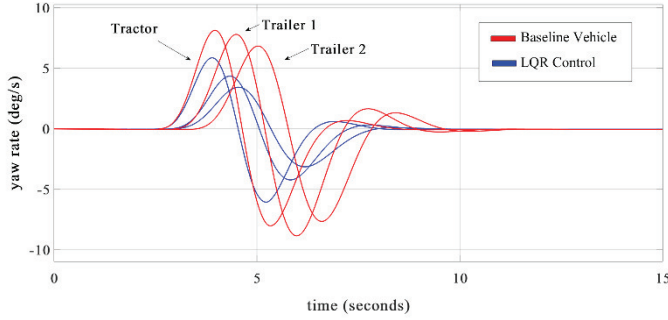


Figure 15. Time history of yaw rate for the 4-DOF model with and without LQR controller.

Few things can be noted from Fig. 14, the LQR succeeded in reducing the lateral acceleration of all the vehicle units. In addition, the baseline model scored an RWA ratio of 1.1663 while the ATS model scored an RWA ratio of 0.9929. The reduction in the RWA ensures an improved performance in lateral dynamics of the B-train double. Moreover, it can be concluded that the LQR controller reduced the overshoot in lateral acceleration curve of the 2nd trailer.

To validate the LQR controller performance, the same analysis was conducted using the TruckSim Baseline model and compared the LQR based ATS model. The conducted results can be seen in Fig. 16-17.

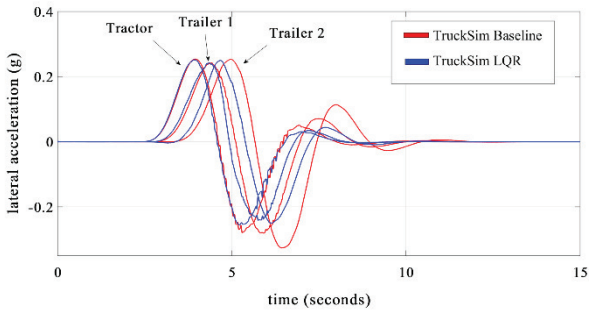


Figure 16. Time history of lateral acceleration for the TruckSim model with and without LQR controller.

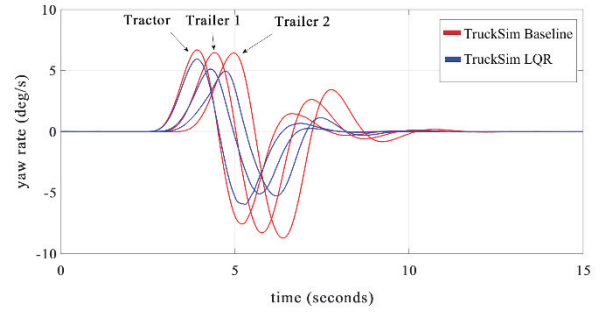


Figure 17. Time history of yaw rate for the TruckSim model with and without LQR controller.

The TruckSim baseline and ATS models demonstrated a similar response to that of the 4-DOF model.

Lastly, to investigate both models Fig. 18-19, show a comparison of the lateral acceleration and yaw rate response of the TruckSim and the 4-DOF models. It can be concluded that the LQR had a positive impact on the lateral stability and both models are in good agreement.

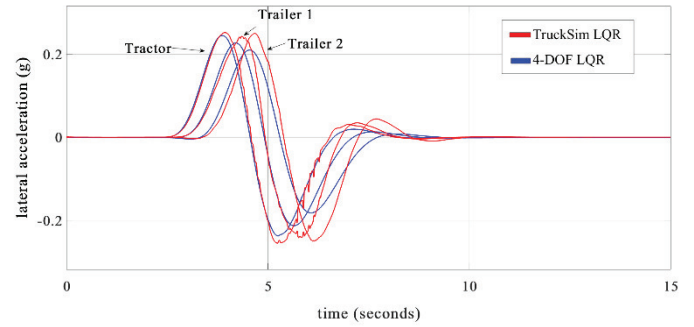


Figure 18. Time history of lateral acceleration for the TruckSim and 4-DOF models with LQR controller.

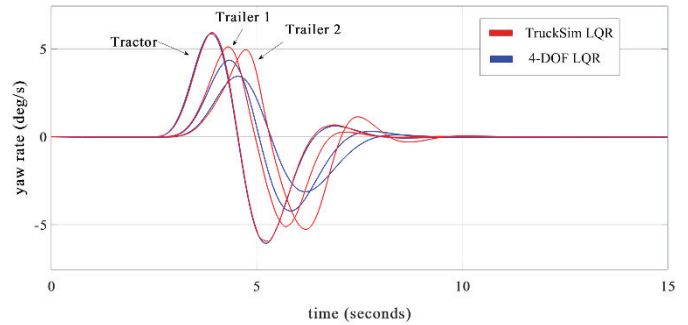


Figure 19. Time history of yaw rate for the TruckSim and 4-DOF models with LQR controller.

IV. CONCLUSIONS

This paper presents the design analysis of an active trailer steering (ATS) system for multi-trailer articulated heavy vehicles (MTAHVs). An ATS controller is designed using the LQR technique. A 4-DOF yaw-plane model is derived and validated using a non-linear TruckSim model. The LQR-based ATS controller is designed to improve the lateral stability of the B-train double under a single lane change maneuver at a speed of 100 km/h. The co-simulation is implemented by integrating

the LQR-based ATS controller designed in MATLAB and the nonlinear B-train model developed in TruckSim. An eigenvalue analysis is carried out to identify the stability boundaries of the system. The linear stability analysis of the MTAHV indicates that the location of the CG has a significant impact on the damping ratio of the system. Lastly, the simulation results suggest that the LQR-based ATS controller can suppress the lateral accelerations while maintaining an RWA ratio close to 1.

REFERENCES

- [1] G. o. Ontario, "The Ontario Ministry of Transportation Home Page," 2009. [Online]. Available: <http://www.mto.gov.on.ca/english/trucks/long-combination-vehicles.shtml>. [Accessed 15 2 2019].
- [2] Md Manjurul Islam, Yuping He, Shenjin Zhu and QiushiWang, "A comparative study of multi-trailer articulated heavy-vehicle models," *Institution of Mechanical Engineering*, vol. 229(9) 1200–1228, 2015.
- [3] Jeannette Montufar, Jonathan Regehr, Garreth Rempel and Robyn V. McGregor, "Long Combination Vehicle (LCV) Safety Performance in Alberta," *Alberta INFRASTRUCTURE AND TRANSPORTATION*, Alberta, 2007.
- [4] Mutaz Keldani, Khizar Qureshi, Yuping He and Ramiro L, "Design Optimization of a Robust Active Trailer Steering System for Car-Trailer Combinations," in *SAE International*, Cobo Center, Detroit Michigan, 2019.
- [5] P. Fancher and C. Winkler, "Directional performance issues in evaluation and design of articulated heavy vehicles," *Vehicle System Dynamics*, vol. 45(7–8), pp. 607–647, 2007.
- [6] S. Zhu, "Coordinated control of active safety systems for multi-trailer articulated heavy vehicles," *Diss. University of Ontario Institute of Technology*, 2016.
- [7] D.J.M. Sampson and D. Cebon, "An Investigation of Roll Control System Design for Articulated Heavy Vehicles," in *4th International Symposium on Advanced Vehicle Control*, Nagoya, 1998.
- [8] C. MacAdam and M. Hagan, "A Simple Differential Brake Control Algorithm for Attenuating Rearward Amplification in Doubles and Triples Combination Vehicles," *Vehicle Syste Dynamics*, vol. 37:sup1, pp. 234–245, 2016.
- [9] Sogol Kharrazi, Mathias Lidberg and Jonas Fredriksson, "A generic controller for improving lateral performance of heavy vehicle combinations," *Institution of MECHANICAL ENGINEERS*, vol. 227(5), pp. 619–642, 2012.
- [10] C. Winkler, "Rollover of Heavy Commercial Vehciles," 2000.
- [11] Rob Di Cristoforo, Chris Blanksby and Anthony Germanchev, "Performance-Based Design of an Innovative Truck-Trailer Configuration," *Journal of the Transportation Research Board*, vol. 1966, p. 110–117, 2006.
- [12] T. Sikder, "Design of Active Trailer Steering Systems for Long Combination Vehicles using Robust Control Techniques," *Diss. University of Ontario Institute of Technology*, 2017.
- [13] Qiushi Wang and Yuping He, "A study on single lane-change manoeuvres for determining rearward amplification of multitrailer articulated heavy vehicles with active trailer steering systems," *Vehicle System Dynamics*, vol. 54:1, pp. 102–123, 2016.
- [14] S. Kapoor, "Fault Tolerant Control of Active Trailer Steering Systems for Multi-Trailer Articulated Heavy Vehicles," *Diss. University of Ontario Institute of Technology*, 2016.
- [15] M. M. Islam, "Parallel Design Optimization of Multi-Trailer Articulated Heavy Vehicles with Active Safety Systems," *Diss. University of Ontario Institute of Technology*, 2013.
- [16] N. Philipp, "Longer combination vehicles (LCV) for Asia and the Pacific region: some economic implications," *United Nations Publications*, 2007.

A Comparison of Hydraulic and Electric Trailer Steering Actuators for Articulated Heavy Vehicles

K Sai Abhilash Patnaik,

Department of Automotive, Mechanical and
Manufacturing Engineering
University of Ontario Institute of Technology
Oshawa, Canada
kondavalasasaiahilash.patnaik@uoit.net

Yuping He

Department of Automotive, Mechanical and
Manufacturing Engineering
University of Ontario Institute of Technology
Oshawa, Canada

Abstract— This paper presents a comparison of an electro-hydraulic actuator to an electric actuator to be applied in an active trailer steering system of articulated heavy vehicles. We used Mathworks software to develop mathematical models for both the electro-hydraulic and electric actuators. The numerical simulation results based on the model for each of the actuators were compared and evaluated. Simulation results demonstrate that the dynamic responses of the electro-hydraulic actuator are faster than those of the electric actuator, and the former is more effective for realistic applications.

Keywords- Articulated heavy vehicles, active trailer steering, electric actuator, electro-hydraulic actuator, mathematical model of actuator, step steering input, sine wave steering input, numerical simulation.

I. INTRODUCTION

Transportation plays an important role in human life on the earth. The logistics industry is vital for a country's economy. This industry is facing a great financial crisis with the rising fuel prices and concerns for global warming. To relieve the energy security crisis and global warming problem, multi-trailer articulated heavy vehicles (MTAHV) may provide an economically feasible solution. Compared with a single trailer articulated heavy vehicle, a MTAHV with double trailers can reduce fuel consumption and greenhouse gas emissions by 33% [1]. Despite the above benefits of MTAHVs, the large vehicles exhibit poor low-speed maneuverability and low high-speed lateral stability. The low high-speed lateral stability of MTAHVs may lead to unstable motion modes, including trailer sway, jackknifing and rollover, which are the main causes for severe highway traffic accidents [2]. It is shown that active trailer steering systems may improve the low-speed maneuverability and enhance the high-speed lateral stability of MTAHVs [3]. The past decade has witnessed the development of active trailer steering systems for MTAHVs.

An active trailer steering system mainly consists of electric control unit (ECU), distributed sensor network, power packs, and a number of steering actuators. Design and development of reliable and cost-effective steering actuators is critical for an active trailer steering system. In the initial design and development state of the steering actuators, mathematical modelling and numerical simulation is the most cost-effective way to explore the performance of these critical subsystems.

In this paper, we will select suitable actuators for the active trailer steering systems of MTAHVs using benchmark investigation of simulation results for different types of actuators. To this end, two different types of actuators, i.e., electro-hydraulic and electric actuators, are briefly reviewed. Then, based on the physical characteristics of these actuators, their mathematical models are derived. Finally, benchmark investigation can be conducted using the numerical simulations based on the two models of the actuators.

II. ELECTRO-HYDRAULIC AND ELECTRIC ACTUATORS

In this section, two types of actuators, namely, electro-hydraulic and electric, are briefly introduced.

A. Electro-Hydraulic Actuators

Electro-hydraulic actuators replace hydraulic systems operated solely with electrical power. Electro-hydraulic actuators eliminate the need for separate hydraulic pumps and tubing, simplifying system architectures and improving safety and reliability. The technology originally was developed for the aerospace industry but has since expanded into many other industries where hydraulic power is commonly used.

Figure 1 shows the configuration of an electro-hydraulic system [4]. This is an energy source, which has hydraulic fluid stored as a storage and when it receives a control signal, it converts its energy into mechanical motion. It is one of the major concepts, upon which the electro hydraulic system may be used for the steering actuators of the active trailer steering systems of MTAHVs. Figure 2 illustrates a sample electro-hydraulic actuator [5].

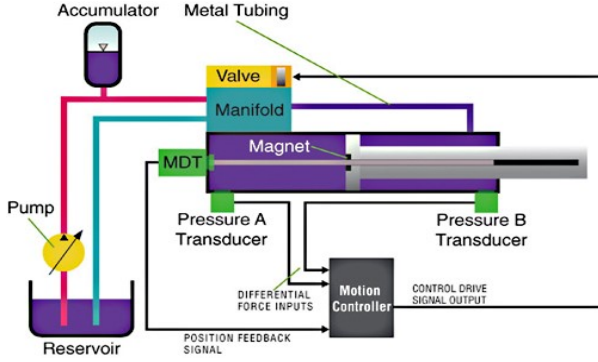


Figure 1: General diagram of an electro-hydraulic system [4].

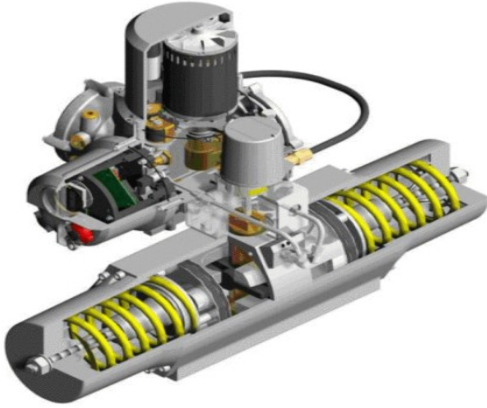


Figure 2: A sample electro-hydraulic actuator [5].

A. Electric Actuators

Figure 3 shows the structure of an electric actuator [6]. The actuator works with the use of a DC motor, which provides the actual movement. When signal input is provided to these actuators, the motor starts rotating in the required direction i.e., clockwise or anticlockwise depending upon the signal input, which in turn moves a lead screw running up and down the cylinder, with a limit switch to control from going beyond required length. This lead screw is used to control the gears, which turn the cylinder. The resulting motion of the cylinder is a linear motion. Figure 4 illustrates a commercially available electric actuator [7].

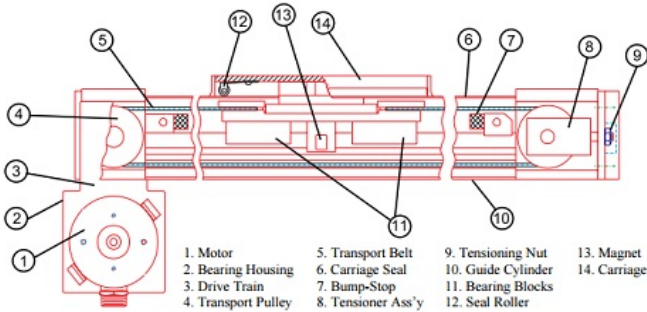


Figure 3: The structure of an electric actuator [6].

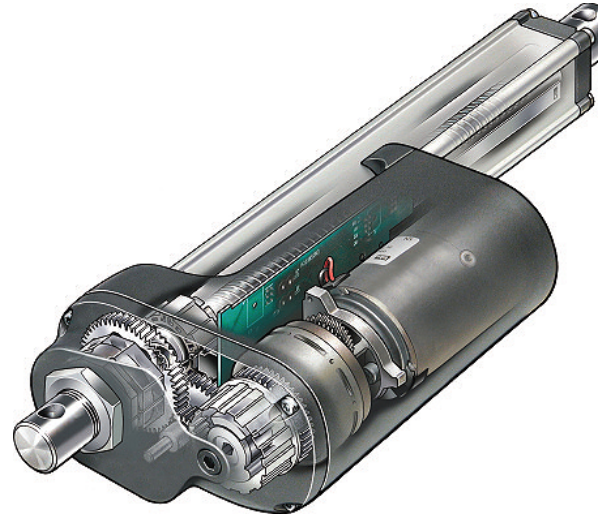


Figure 4: A commercially available electric actuator [7].

III. MATHEMATICAL MODELS OF ACTUATORS

In this section, the mathematical models are derived to represent the aforementioned actuators, namely, electro-hydraulic and electric

A. Electro- Hydraulic Actuator Model

The developed model has a four-way valve with controlled linear actuator configuration, in which there is a two-stage solenoid actuated electro-hydraulic valve. Due to the use of the valve based controlled systems, numerous benefits are achieved, such as low complexity, high flexibility and accurate control. Moreover, these are mainly used in applications such as industrial robots, vehicles, aerospace industry, etc.

To derive the electro-hydraulic actuator model, the governing equations of motion are determined, and the system is then linearized. Moreover, the system design parameters, such as the required piston diameter, nominal supply pressure, valve coefficients, and pump displacement, are determined based on the steady-state forms of the equations of motion. Finally, the linear model is validated using the non-linear Simscape hydraulic model [8, 9].

Figure 5 shows the schematic diagram of the electro-hydraulic system. The governing equation of motion with the payload is expressed as [10],

$$my'' + cy' + ky = \eta_{af} (A_a P_a - A_b P_b) - F - F_0 \quad (1)$$

where m is the mass, k the spring stiffness, c the damping coefficient of the payload, y the position of the payload, A_a and A_b are the piston areas, and P_a and P_b the fluid pressures on the A and B sides of the actuator, respectively, η_{af} is the force efficiency of the actuator, F is the disturbance force, and F_0 is the nominal spring and bias load that is applied to actuator when y is zero.

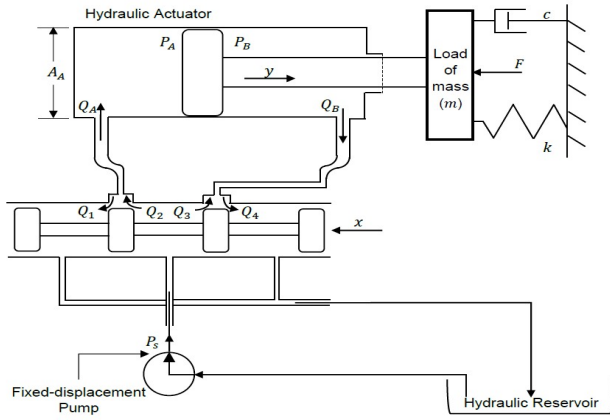


Figure 5: Schematic diagram for the electro-hydraulic system.

It is imperative to examine the model's accuracy and fidelity. Hence, the system model is validated using the nonlinear Simscape software package. Moreover, it allows direct integration of physical component models using block diagrams. Simscape library consists of physical component models, such as electric motors, rectifiers, hydraulic actuators, etc. Figure 6 illustrate the block diagram of the Simscape electro-hydraulic actuator model.

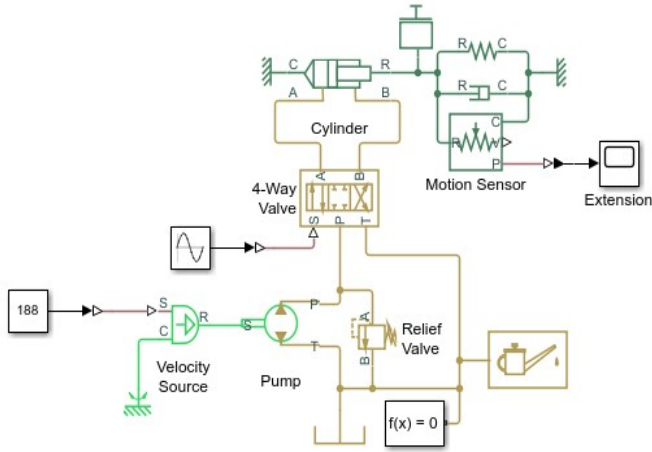


Figure 6: Block diagram of Simscape electro-hydraulic actuator model.

To validate the linear electro-hydraulic actuator model, the actuator's position response is evaluated, and the results are compared with the Simscape model. The model is validated using step inputs. The system is first analyzed using a step input to evaluate the system's rise time, to ensure that model can track the Simscape model. The magnitude of input current in the case is 20 mA.

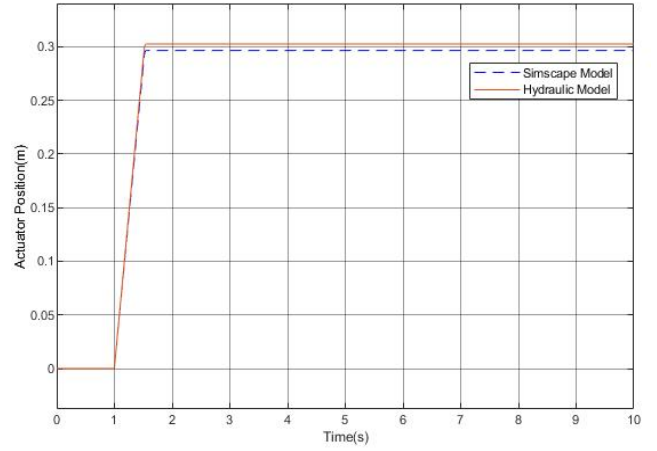


Figure 7: Time history of the electro-hydraulic actuator position for the step input.

Figure 7 show the time history of actuator position for the step input. It is shown that both models have similar rise times. Moreover, the model can accurately track the Simscape model. The results help establish the validity of the linear model.

B. Electric Actuator Model

An electric actuator is an electro-mechanical system, which consists of electric motor, planetary gearhead and ball-screw. The ball-screw is a mechanical element having very low coefficients of friction. Therefore, it is used in this linear electric actuator. In addition, these elements are used to execute linear movements with great performance because they provide a transmission with relatively high rigidity and reduction coefficient that allows having enough force after conversion of rotational to linear movement. The motor is an electric component that transducers voltage in torque and angular displacement [11].

Figure 8 shows the schematic representation of the electric model. Figure 9 illustrates the free body diagram for the linear electric actuator. Based on the components and notation shown in Figures 8 and 9, the mathematical model of the electric actuator is derived.

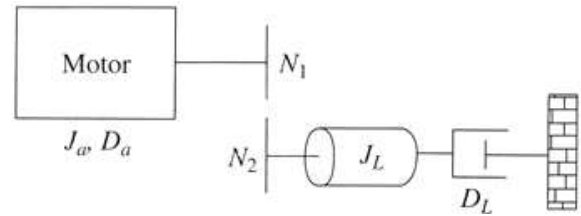


Figure 8: DC Motor that moves a rotational mechanical payload.

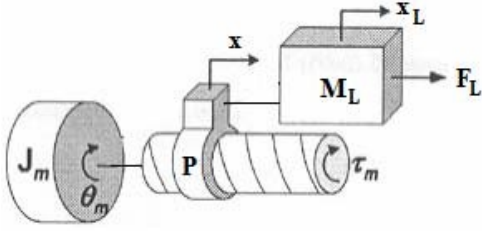


Figure 9: Free body diagram for the linear electrical actuator.

Based on Newton's second law, the governing equations of motion for the model shown in Figure 8 can be derived. Then, using Laplace's transform, we have:

$$\frac{X(s)}{E_a(s)} = \frac{k_t}{P[(J_m s^2 + D_m s)(R_a + L_a s) + k_b k_t s]} \quad (2)$$

where X is the linear advance of the screw, and E_a is the voltage of armature, k_b is the constant of back EMF, k_t is the constant of torque of motor, J_m is the Equivalent inertia in armature, D_m is the equivalent viscous damping in armature, R_a is the resistance of armature, and L_a is the inductance of armature. The governing equations of motion of the electric actuator model can be written in the state-space form as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{D_m}{J_m} & \frac{P k_t}{J_m} \\ 0 & -\frac{K_b}{P L_a} & -\frac{R_a}{L_a} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L_a} \end{bmatrix} u \quad (3)$$

$$Y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (4)$$

where x is the state variable vector, and Y is the output vector. To conduct numerical simulation, the electric actuator is implemented as shown in Figure 10.

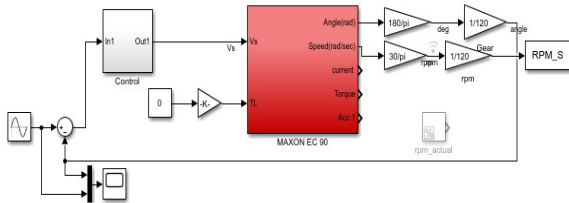


Figure 10: Block diagram of the Simscape electric actuator model.

To validate the electric actuator model, the results are compared with the Simscape model. The system is first analyzed using a sinewave input to examine if the model can track the Simscape model. The magnitude of input current in the case is 20 mA. The simulation result is shown in Figure 11.

IV. RESULTS AND DISCUSSION

In order to compare the performance of the electric actuator model with that of the electro-hydraulic actuator model, the position of the electro-hydraulic actuator model is also simulated under the same sinewave input used for validated the electric actuator model. The time history of the electro-hydraulic actuator position response to the sinewave input is shown in Figure 12.

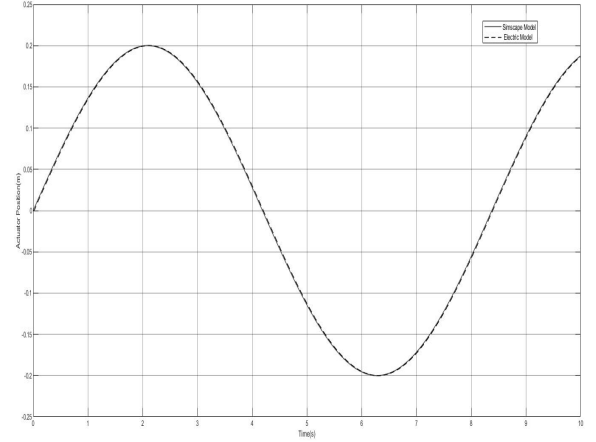


Figure 11: Time history of the electric actuator position response for the sinewave input.

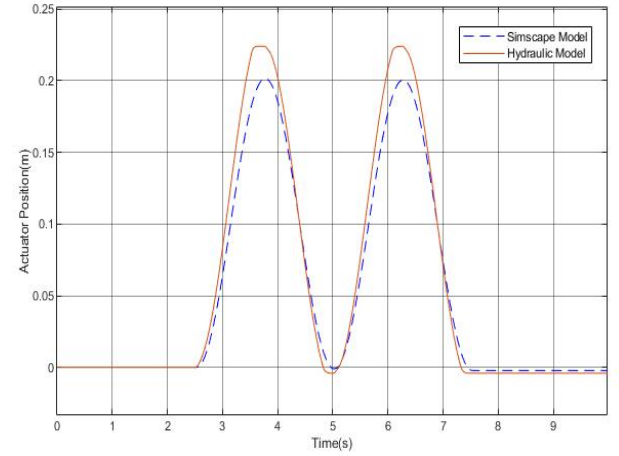


Figure 12: Time history of the electro-hydraulic actuator position response for the Sinewave input.

As we can observe from Figure 12, the sinewave takes approximately 2.5 seconds for a cycle in the case of electro-hydraulic system. As seen in Figure 11, in the case of the electric model, the sine wave takes approximately 4.2 seconds for a cycle. In case of the electric actuator a full sinewave is obtained due to the change in polarity of voltage which causes the motor to run clockwise or anti clockwise, due to which we can observe the full sinewave in motion.

V. CONCLUSIONS

Investigation based on the numerical simulation indicates that the electro-hydraulic model's response time is shorter than that of an electric actuator model. Electro-hydraulic systems are considered more compact and is efficient compared to electric actuator due to less gear involvement which helps in achieving smoother result. Moreover, for a heavy-duty application like MTAHV's, the size of an electric actuator is quite large compared to electro-hydraulic actuator which results in bulkiness of actuation system. The constant usage of high current in electric actuator for heavy duty applications could also lead to an overheating issue which could lead to a fire hazard. Since the response time is the most critical factor for the active trailer steering system of MTAHV's and to address the problem of maneuverability, trailer sway, rollover and jackknifing, we can conclude that the model developed for the electro-hydraulic actuator is more effective than the model developed for the electric actuator. This is a preliminary study, based on the mathematical models for the two actuators hence it should be further evaluated and validated considering the realistic operating conditions of MTAHVs.

REFERENCES

- [1] Yuping He, Md. Manjurul Islam, Shenjin Zhu, and Thomas Hu (2017), "A Design Synthesis Framework for Directional Performance Optimization for Multi-Trailer Articulated Heavy Vehicles with Trailer Lateral Dynamic Control Systems", Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, Vol. 231, No. 8, pp. 1096-1125.
- [2] Qiushi Wang, Yuping He (2015), "Design and Validation of Active Trailer Steering Systems for Multi-Trailer Articulated Heavy Vehicles Considering Driver-Vehicle-Controller Interactions", International Journal of Vehicle Performance, Vol.2, No.1, pp.58-84.
- [3] Xuejun Ding, Steve Mikaric, and Yuping He (2013), "Design of an Active Safety System for Multi-Trailer Articulated Heavy Vehicles Using Real-Time Simulations", Proc IMechE Part D: J Automobile Engineering, Vol. 227, No. 5, pp.643-655.
- [4] < <https://www.picswe.com/pics/electro-hydraulic-servo-99.html> >
- [5] < <http://www.kidskunst.info/linked/mea-inc-industry-leader-in-hydraulic-actuator-systems-6d6561.htm> >
- [6] < <https://engineering.stackexchange.com/questions/2409/how-does-a-rodless-electrical-linear-actuator-work> >
- [7] < <https://www.therobotreport.com/electric-linear-actuators-the-mechatronic-choice/> >
- [8] Kapoor, S., Sikder, T. and He, Y., 2017, July. Fault tolerant control of active trailer steering systems for multi-trailer articulated heavy vehicles. In *Dynamics of Vehicles on Roads and Tracks Vol 1: Proceedings of the 25th International Symposium on Dynamics of Vehicles on Roads and Tracks (IAVSD 2017), 14-18 August 2017, Rockhampton, Queensland, Australia* (p. 471). CRC Press.I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [9] Kim, M.S. and Chung, S.C., 2005. A systematic approach to design high-performance feed drive systems. *International Journal of Machine Tools and Manufacture*, 45(12-13), pp.1421-1435.
- [10] Manring, N., 2005. Hydraulic control systems. New York: Wiley.
- [11] Ruiz-Rojas, E.D., Vázquez-González, J.L., Alejos-Palomares, R., Escudero-Urbe, A.Z. and Mendoza-Vázquez, J.R., 2008, March. Mathematical model of a linear electric actuator with prosthesis applications. In 18th International Conference on Electronics, Communications and Computers (conielecomp 2008) (pp. 182-186). IEEE.

Wind Energy and Engineering

WIND LOADS ON LOW-RISE BUILDINGS IN TORNADO-LIKE VORTICES

Jubayer Chowdhury	Peter J. Vickery	Horia Hangan	Sudhan S. Banik
WindEEE Research Institute Western University London, ON, Canada cjubaye@uwo.ca	Applied Research Associates, Inc. Raleigh, NC, USA pvickery@ara.com	WindEEE Research Institute Western University London, ON, Canada hmhangan@uwo.ca	Applied Research Associates, Inc. Raleigh, NC, USA sbanik@ara.com

ABSTRACT

Tornadoes are the most disastrous, violent and lethal wind events in the world. North America in particular sees more tornadoes, by far, than any other places on the earth [1]. In the United States, more than 4,000 fatalities and US\$20 trillion in damages were reported between 1950-1994 [1]. Canada, distant second to US, also sees on an average 100 tornadoes every year [2]. Based on the degrees of damages, intensity of tornadoes are rated in Enhanced Fujita (EF) scale and ranges from EF0-EF5, where EF0 is the weakest and EF5 is the strongest of tornadoes [4]. Frequency of occurrence of weaker tornadoes (rated EF2 and lower) are considerably higher than that of stronger tornadoes (rated EF3 and higher). Between 1994 and 2012, approximately 92% of the recorded tornadoes in the United States were rated EF2 and lower [5]. This significantly higher probability of occurrence for weaker tornadoes is encouraging as this implies that designing buildings and structures to withstand up-to EF2 rated tornado could considerably reduce the property damages and economic losses due to tornadoes, especially in the tornado prone regions.

Studies have been conducted to evaluate wind loads on building due to tornadic flows. Several buildings with different sizes and shapes have been tested in the tornado simulator at the Iowa State University. Sarkar et al. [6] evaluated tornadic wind loads on a 1:500 geometrically scaled tall building and found that tornadoes with strength F2 or higher would exceed the minimum design load set by ASCE 7-02 by a factor of 1.8 or higher. A 1:100 scaled gable roof building with roof slope of 35° was tested by Haan et al. [8] and it was found that the peak external uplift force coefficients exceed the minimum design wind loads in ASCE 7-05 by factors between 1.8 and 3.2. Roueche et al. [9] utilized the pressure data from Haan et al. [8] and observed that for roof-to-wall connections, peak shear forces were 1.8 times stronger in tornadoes compared to ASCE 7-10 building standard. These studies investigated the peak wind loads on the building as a whole and lack information regarding wind loads on different zones on the building surfaces, which is of importance for designing components and claddings as outlined in the ASCE 7-16 [10] building code.

In the present study, external pressure coefficients are measured on a gable roof building for different tornado strengths (EF1 and EF2 rated tornadoes), two different building orientations (0° and 45°) and two different tornado translation paths in the WindEEE Dome. External pressure coefficients are calculated for different zones on the building surfaces as identified in the ASCE 7-16 provision and compared with the recommended design values for straight-line ABL winds.

REFERENCES

- [1] J. J. Bohonos and D. E. Hogan, "The medical impact of tornadoes in north America," *J. Emerg. Med.*, vol. 17, no. 1, pp. 67–73, Jan. 1999.
- [2] National Centers for Environmental Information, "U.S. Tornado Climatology," National Oceanic and Atmospheric Administration.
- [3] M. K. Tippett, C. Lepore, and J. E. Cohen, "More tornadoes in the most extreme U.S. tornado outbreaks," *Science*, vol. 354, no. 6318, pp. 1419–1423, Dec. 2016.
- [4] J. McDonald and K. Mehta, "A recommendation for an Enhanced Fujita scale," Wind Science and Engineering Centre, Texas Tech University, Lubbock, Texas, 79409–1023, Jun. 2004.
- [5] J. B. Elsner, T. H. Jagger, H. M. Widen, and D. R. Chavas, "Daily tornado frequency distributions in the United States," *Environ. Res. Lett.*, vol. 9, no. 2, p. 024018, 2014.
- [6] P. Sarkar, F. Haan, V. Balaramudu, and A. Sengupta, "Laboratory Simulation of Tornado and Microburst to Assess Wind Loads on Buildings," *Struct. Congr. 2006*, 2006.
- [7] A. Sengupta, F. L. Haan, P. P. Sarkar, and V. Balaramudu, "Transient loads on buildings in microburst and tornado winds," *J. Wind Eng. Ind. Aerodyn.*, vol. 96, no. 10, pp. 2173–2187, Oct. 2008.
- [8] F. L. Haan, V. K. Balaramudu, and P. P. Sarkar, "Tornado-Induced Wind Loads on a Low-Rise Building," *J. Struct. Eng.*, vol. 136, no. 1, pp. 106–116, Jan. 2010.
- [9] D. B. Roueche, D. O. Prevatt, F. L. Haan, and P. L. Datin, "An estimate of tornado loads on a wood-frame building using database-assisted design methodology," *J. Wind Eng. Ind. Aerodyn.*, vol. 138, pp. 27–35, Mar. 2015.
- [10] ASCE/SEI 7-16, Minimum design loads and associated criteria for buildings and other structures. Reston, Virginia: American Society of Civil Engineers, 2017.

TORNADO-INDUCED INTERNAL AND EXTERNAL PRESSURES ON A LOW-RISE BUILDING WITH MULTIPLE OPENINGS

Aya Kassab
WindEEE Research Institute
Western University
London, ON, Canada
akassab@uwo.ca

Peter J. Vickery
Applied Research Associates Inc.
Raleigh, NC, USA
pvickery@ara.com

Jubayer Chowdhury
WindEEE Research Institute
Western University
London, ON, Canada
cjubaye@uwo.ca

Sudhan S. Banik
Applied Research Associates Inc.
Raleigh, NC, USA
sbanik@ara.com

Horia Hangan
WindEEE Research Institute
Western University
London, ON, Canada
hmhangan@uwo.ca

ABSTRACT

Tornado-induced internal pressures in buildings due to flexibility of structures, openings or sudden window breakage attributed to wind-borne debris are one of the substantial causes of the extensive destruction to properties and large number of fatalities [1]. The low-rise buildings, which account for most of the residential buildings (\approx three quarters) in North America [2], are the most severely affected structures from tornado hits. The net wind load on buildings, which is the joint effect of internal and external pressures, during tornadic events exceeds the allowable design wind loads and results in partial or total collapse of the building. Albeit the standard building codes, e.g., ASCE 7-16 [3] or NBCC 2015 [4] have contemporized wind maps with adding tornado wind speeds in tornado-prone regions, the pressure coefficients values are based only on atmospheric boundary layer (ABL) flowfields. One of the causes of this demerit is ascribed to the insufficient reliable analyses of pressure dynamics, particularly internal pressures in tornado simulators. Thus, a need for a comprehensive study of the tornado-induced pressures, internal and external, on low-rise buildings is indispensable. In the current paper, experimental measurements of external and internal pressures on a low-rise building model were carried out in the state-of-the-art large-scale tornado simulator, the WindEEE Dome [5] at Western University to examine the effect of different opening configurations on the net wind loading. Moreover, the effect of different building orientations is analyzed. Further, a comparison between internal and external pressure coefficients from this study and those for ABL flow from the NIST database [6] will be performed.

REFERENCES

1. E. D. Kuligowski, et al., Preliminary reconnaissance of the May 2013, Newcastle-Moore tornado in Oklahoma., NIST SP-1164, National Institute of Standards and Technology, Gaithersburg, MD (2013).
2. Statistics Canada, National Household Survey (NHS), Condominium dwellings in Canada (2011)
3. ASCE/SEI (ASCE/Structural Engineering Institute), Minimum design loads for buildings and other structures. ASCE/SEI 7-16, Reston, VA (2016)
4. NBCC, National Building Code of Canada 2015-National Research Council Canada (2016)
5. H. Hangan, Current and Future Directions for Wind Research at Western: A New Quantum Leap in Wind Research through the Wind Engineering, Energy and Environment (WindEEE) Dome, J. JAWWE, 35 (4) [No 125] (2010) 277-281.
6. T. C. E. Ho, D. Surry & M. Nywening, NIST/TTU Cooperative agreement—Windstorm mitigation initiative: Further experiments on generic low buildings. BLWT-SS21, Boundary Layer Wind Tunnel Laboratory, Univ. of Western Ontario, London, Ont., Canada (2003)

A LARGE EDDY SIMULATION OF ATMOSPHERIC BOUNDARY LAYER FOR VERTICAL AXIS WIND TURBINE APPLICATION

Maryam Zabarjad Shiraz

Department of Mechanical and Industrial Engineering
Concordia University
Montreal, Canada
m_zabarj@encs.concordia.ca

Marius Paraschivoiu

Department of Mechanical and industrial Engineering
Concordia University
Montreal, Canada
marius.paraschivoiu@concordia.ca

ABSTRACT

Wind energy plays an important role in the energy communities as it is clean and very accessible. A revived interest in Vertical Axis Wind Turbines (VAWTs) has recently arisen for urban locations because of their insensitivity to the wind direction, lower noise emission and structural integrity. Indeed, roof mounted turbines in an urban area is beneficial as the consumption need is close to the production location. The Darrieus rotor of VAWTs has potential to be used in urban areas because of high efficiency, low visual impact and adequately noiseless features. Nevertheless, in contrast to the mechanical and structural simplicity, the Darrieus rotor has complex aerodynamic behavior due to periodically changing of aerodynamic forces, relative velocities and the angle of attack during every rotation cycle. The accurate study of VAWTs aerodynamic behavior and performance is required especially in urban areas as the effect of the buildings can be significant. CFD models are robust tools and can be very useful to simulate the VAWT's performance and flow around the turbine. However, the full CFD models are computationally expensive specifically for roof mounted turbines. Therefore, hybrid models are recommended as they solve the Navier Stokes equations for the flow while modeling effects of the turbine rotor on the flow[1]. The main advantage of the hybrid method is that the viscous airfoil boundary layer is not resolved and therefore the number of elements can be significantly reduced. In this work, 3D Navier Stokes equations are solved for the flow and turbine effects are introduced by means of the momentum source terms in the annular volume swept by the blades. In urban areas, it is demonstrated that the Atmospheric Boundary Layer (ABL) turbulence has a significant effect on the wind turbines performance due to the change of wind speed, angle of attack and blade loads. In this presentation, the wind speed profile is applied at the inlet boundary condition to consider the ABL effect. A LES framework has been employed to study the ABL flow and wind turbine wakes. The LES model is based on a Sub-Grid Scale (SGS) to model small scale motions. The accuracy of LES depends on the ability of the SGS to capture the sub-grid scale turbulent fluxes. Over the last few decades, many methods have been proposed as SGS models, from the simple Smagorinsky model to the more complicated dynamic Smagorinsky model. One of the main challenges in implementation of the eddy viscosity model in LES is to specification of model coefficients. Anisotropy turbulence in ABL flows especially with strong mean shear near to the surface makes the model coefficients determination more challenging. Therefore, one of the LES limitations is applying SGS model in the near-wall region. Hence, a very high-resolution grid is necessary to accurately capture all the turbulent scales, and this is the reason that LES is still computationally expensive. In this model, the refine mesh is not required around the turbine blades as turbine blades effects are considered as a source momentum term therefore the computational time is significantly reduced. However, one of the LES challenges is applying turbulence at the inlet that has a significant impact on the flow dynamics. For this purpose, the turbulent stochastic fluctuations need to be generated in the grid scale quantities. The created fluctuations should satisfy the Navier-Stokes equations which allow specifying the spectral and stochastically varying properties such as intensity and length scales. A Fourier technique is used in this model to generate the turbulent fluctuations that can be added to the mean velocity profile to create an instantaneous flow profile. The model developed is a robust and accurate tool to evaluate the flow around the wind turbines and to evaluate their performance. As this method does not need mesh refinement around turbine's blade, it is not computationally expensive and can be used to study rood mounted VAWTs in urban areas. The first part of the presentation will describe a validation case including a comparison with experimental data [2]. The second part will analyze the performance of a roof-mounted turbine on a single building.

Keywords: Vertical axis wind turbines, Large Eddy Simulation, CFD, Atmospheric boundary conditions.

[1] M. Torresi, B. Fortunato, and S. M. Camporeale. An efficient 3D CFD model for the analysis of the flow field around Darrieus rotors. ASME Turbo Expo 2013: Turbine Technical Conference and Exposition, 8, 2013.

[2] Sidney F. Johnston Jr., Editor, "Proceedings of the Vertical Axis Wind Turbine (VAWT) Design Technology Seminar for Industry", SAND 80-0984, Sandia Laboratory, Albuquerque, 1980.

EVALUATING STRUCTURAL DESIGN LOADS UNDER THUNDERSTORM WINDS

Haitham Aboshosha*

Tarek Ghazal*

Moustafa Aboutabikh*

*Department of Civil Engineering
Ryerson University
350 Victoria St.
Toronto, ON M5B 2K3

INTRODUCTION

Due to climate change, the number extreme events and their consequent losses have been increasing. For example, there was an estimated billions of dollars in damage from hurricane Igor and Matthew that made landfall in 2011 and 2016 over coastal regions stretching from South Carolina in United States to Nova Scotia in Canada. The inland is also affected by non-synoptic high intensity local storms such as tornadoes and downbursts that represent most of insured property losses due to natural catastrophes. The recent tornado that hit Ottawa and Gatineau region in September 2018 damaged hundreds of homes and many hydro poles/towers leaving more than 450k customers with no power. In Canada, the average annual loss due to thunderstorm activity (tornados and downbursts) is in the order of \$200 million. Despite that, most of the design codes (including the NBCC) provide clauses to evaluate wind loads for synoptic wind and not for thunderstorm wind which has very different characteristics. Wind load evaluation for structures utilizing wind tunnels is very well-established, however load evaluation for thunderstorms is still immature.

OBJECTIVE

The current study aims at developing a methodology for wind load evaluation of structures under thunderstorm winds inspired by Alan Davenport's wind loading chain. According to Alan, wind load evaluation consists of the following 5 steps (1) climate analysis, (2) terrain effect, (3) aerodynamic effect, (4) structural dynamics (if any), and (5) evaluation of equivalent static loads and responses. The novelty in this current methodology is related to steps (1) thunderstorm climate analysis and (3) thunderstorm aerodynamic modeling.

METHODOLOGY

Thunderstorms climate analysis is conducted by combining (i) historical meteorology data with (ii) thunderstorm downdraft model to extract important thunderstorm characteristics and (iii) Mont Carlo simulation to generate thunderstorm records following to the statistical distributions of obtained characteristics. The analysis was conducted for various records in Canada and the US and the results showed a strong influence of the climate change (i.e. number and intensity of storms). Thunderstorm aerodynamic simulation is conducted by coupling a system that generates atmospheric boundary layer for different terrain conditions with a gust-generating system. The systems can generate both the transient nature of thunderstorm outflows and its spatial variation.

OUTCOME

The developed method was implemented for the wind tunnel facility at Ryerson University and currently is utilized to study thunderstorm-induced loads for various structures. It worth to mention that the developed method is generic and can be employed with other wind tunnels capable of generating thunderstorm gust fronts. This will ultimately achieve more resilient (and economic) structures to thunderstorm winds.

EXPERIMENTAL SIMULATIONS OF LARGE-SCALE TORNADO-LIKE VORTICES

Arash Ashrafi
WindEEE Research Institute
Western University
London, Ontario, Canada
aashra22@uwo.ca

Horia Hangan
WindEEE Research Institute
Western University
London, Ontario, Canada
hmhangan@uwo.ca

ABSTRACT

Tornadoes are a serious threat to vast regions of North America. An average of 900 tornadoes, according to Wind Hazard Reduction Coalition statistics, occur annually in the US and cause around 80 deaths, 1500 injuries, and US\$850 million worth of damage [1]. Currently, structures are mostly designed for straight line hurricane-like wind speeds which can be considered as two-dimensional stationary flow. In contrast, tornado flow structure is three dimensional and highly non-stationary posing design challenge for structures [2]. Therefore, study of tornadoes and their interaction with engineered buildings is of significant importance in order to mitigate the damage level.

In this regard, several tornado simulators (Ward, Purdue, TTU Vortex Simulator II, TTU VorTECH, ISU tornado/microburst simulator, and WindEEE) have emerged with different capabilities to better understand the tornado flow field and its effect on wind loading of structures [3]. The novel, state-of-the-art wind testing facility, the WindEEE Dome, employs a combination of 54 out of 106 fans to simulate tornado-like vortices. There are six fans in a plenum above the test chamber create the suction and eight fans on each of the six peripheral walls provide inflow. At the WindEEE Dome, tornadoes can be simulated in two modes, mode A and B. In mode A, only six fans in the upper plenum are powered and louver vanes changing the inflow direction which consequently providing a range of swirl ratio. In mode B, the peripheral fans are also powered to provide stronger inflow. Using the peripheral fans in mode B, gives the ability to produce relatively larger tornado-like vortices compared to the mode A by increasing the inflow momentum and consequently increasing the radius of tornado.

One of the essential requirements for any laboratory testing is to follow the laws of similitude. Proper scaling method for tornadic flow was introduced and analyzed based on geometric, kinematic, and dynamic similitudes for mode A tornadoes at the WindEEE Dome. The geometric scale for the mode A tornado-like vortices ranged from 1:300 to 1:200 with velocity scales of 1:3 to 1:2 [4]. However, this range of geometric scales poses some technical challenges for detail investigation of relatively smaller structures; for example, aeroelastic testing of different structural elements of transmission towers, or light poles.

In this study, the same scaling procedure for mode A [4] was used for mode B tornadoes. Comparing with the tornado event recorded near Russell, Kansas, on 25 May 2012 [5], a range of length scales from 1:70 to 1:50 and velocity scales from 1:8 to 1:6 corresponding to EF0 to EF2-rated tornadoes is obtained. Velocity distributions from these two modes (A and B) tornadoes are also compared. Results show that the larger simulated tornadoes (mode B) have approximately twice the radius, half the maximum tangential velocity and half of the maximum height compared to former mode (mode A). This translates in increased geometric scales (up to 4x), decreased velocity scales (down to 2x) and increased swirl ratios for mode B tornadoes.

REFERENCES

- [1] Z. Yang, P. Sarkar, and H. Hu, "An experimental study of a high-rise building model in tornado-like winds," *J. Fluids Struct.*, vol. 27, no. 4, pp. 471–486, 2011.
- [2] H. Hangan, D. Romanic, and C. Jubayer, "Three-dimensional, non-stationary and non-Gaussian (3D-NS-NG) wind fields and their implications to wind-structure interaction problems" *J. Fluids Struct.*, no. xxxx, pp. 1–27, 2019.
- [3] H. Hangan, M. Refan, J. Chowdhary, D. Romanic, and D. Paarvu, "Novel Techniques in Wind Engineering," *J. Wind Eng. Ind. Aerodyn.*, vol. 171, no. June, pp. 12–33, 2017.
- [4] M. Refan and H. Hangan, "Near surface experimental exploration of tornado vortices," *J. Wind Eng. Ind. Aerodyn.*, vol. 175, pp. 120–135, 2018.
- [5] K. Kosiba and J. Wurman, "The Three-Dimensional Structure and Evolution of a Tornado Boundary Layer," *Weather Forecast.*, no. 1960, pp. 1552–1561, 2013.

Calibration of icing severity measurement on a wind turbine nacelle using image analysis of ice accretions

Patrice Roberge, Jean Lemay, Jean Ruel, André Bégin-Drolet*

Laval University Engineering Department
Laval University
Quebec City G1V 0A6, Canada

*Corresponding author. E-mail address: andre.begin-drolet@gmc.ulaval.ca (A. Bégin-Drolet)

Abstract—As countries invest massively on the construction of new wind farms to increase their renewable energy production, multiple challenges still need to be addressed to increase the efficiency of their existing wind farms. A major source of energy loss for wind energy in cold climates is ice accretion on the wind turbine blades. Reliable ice detectors are required to operate proactively the thermal based ice protection systems equipped on wind turbines. This paper presents the icing severity calibration of a Meteorological Conditions Monitoring Station (MCMS), a novel ice detector developed at Laval University. The calibration is made by comparing the ice thickness accreted on a reference cylinder (visual measurement) with the modeled ice thickness based on the icing severity measurement provided by the MCMS. The reference ice thickness was assessed by numerically processing photos of the reference cylinder. The calibration process showed a good agreement between observed and modeled ice thickness for most of the 18 different icing events observed.

Icing; wind turbines; ice detection; ice accretion; image processing; physical modeling.

I. INTRODUCTION

The increasing demand for renewable energy leads Canada to invest massively in wind energy and become a global leader. As a matter of fact, the production of wind energy increased by 15% yearly over the last five years (CanWEA, 2018). An increase of power production can be achieved either by building new wind farms or by improving the efficiency of existing wind farms. Since ice accretion on wind turbines is one of the most important causes of energy loss for wind power production in cold climate, it is crucial to limit these effects. Lacroix (2013) estimated that the loss in energy caused by icing can reach up to 15% of the yearly production. First, ice accumulation on the blade modifies its profile, thus yielding poor aerodynamic performances and causing directly a lower power output (Hudecz, 2013). Secondly, ice accretes usually in a nonuniform manner on each of the blades resulting in additional vibrations and stresses that may reduce the effective lifetime of a wind turbine (Frohboese, P., & Anders, 2007). The last main impact of icing on wind turbines is the risk of injury caused by ice

projectiles generated by the turbine (LeBlanc *et al.*, 2007). The main wind turbine manufacturers have equipped their turbines with ice protection systems (IPS) to mitigate these effects. Most of them consists of heating systems for the blades. Such systems require a significant amount of power to operate (~100 kW) and present a large delay between the activation of the IPS and its effectiveness. Presently, ice detection on wind turbines is mainly done by comparing the power output of a turbine with an expected power based on the wind speed. The problem with this method is that it is limited by the incubation time which is the period between the beginning of the accretion and the moment where it starts affecting the production. Having a reliable ice detector would allow wind turbine operators to proactively activate the IPS and maximize power production. A reliable ice detector would also allow the modelling of the physical phenomena. Such modeling would be greatly useful for the prediction of icing related losses, thus providing crucial information for the site assessment process. Presently there are no commercial ice detectors that are able to detect precisely the beginning, the end and the severity of an icing event (Wickman, 2013). Moreover, there are currently no standards on ice detection and icing event characterization. IEA task 19 is currently working on this issue (Lehtomaki, 2019) and would greatly benefit from a standardized calibration method based on a physical approach. This paper describes the calibration method of a novel approach to ice detection using visual data as a reference.

II. AREA DESCRIPTIONS, METHODS AND MATERIAL STUDIED

The experimental setup consists of a Meteorological Conditions Monitoring Station (MCMS) and a camera (Axis P5635-E Mk II) installed on the nacelle of a wind turbine. The wind turbine is located in an undisclosed wind farm situated in eastern Canada subject to harsh icing events. The MCMS is an instrument developed at Laval University composed of two patented thermal probes (Bégin-Drolet *et al.*, 2017), an ultrasonic anemometer, a humidity sensor, an external temperature sensor, a barometric pressure transducer and a solar radiation sensor. With an onboard microprocessor, the data from all the sensors are combined to generate signals which are related to icing using physical models. The heat loss of the heated

probes is compared to the theoretical heat loss based on the outside temperature and wind speed. A measured heat loss exceeding the modeled heat loss is interpreted as water particles present in the atmosphere impacting the surface of the probe that are extracting extra heat from it. This additional heat can be linked to the droplet impingement flux on the surface of the probe which will be called icing severity for the rest of this paper. The link between icing severity and the measurement of additional heat presented in this paper is made by comparing the MCMS measurements to visual data of a reference cylinder also present on the nacelle of the wind turbine. For this purpose, the ice thickness on the reference cylinder was modeled using the MCMS measurements and then compared to the observed thickness. Using a camera as the only ice detection method is not viable since the daylight hours are limited especially in the winter. The accretion of ice on the camera can also be a problem.

Icing events occur in two different conditions: in the presence of freezing rain or by a cloud at the structure height (in-cloud icing). The freezing rain droplets, because of their large size will solidify in a dense, translucent and almost clear accretion called glaze ice (ISO, 2017). The water droplets in the clouds are significantly smaller than the ones of the freezing rain. Thus, the ice type formed will depend on ambient conditions (*i.e.* wind speed and temperature). For high wind speed and high temperature, the water particles are able to get arranged in a dense manner before freezing, therefore forming glaze ice (ISO, 2017). Opposite to the last case, for low wind speeds and low temperatures, water particles freeze almost instantly on contact generating an opaque, less dense accretion called rime ice (ISO, 2017). For median conditions, mixed ice can form.

Moreover, the icing rate depends not only on the area of the surface but on three efficiencies: the collision efficiency (α_1), the sticking efficiency (α_2) and the accretion efficiency (α_3) (ISO, 2017). The collision efficiency (α_1) is the ratio of the particles that actually impact the surface to the total number of particles having an initial path to the surface. This efficiency is strongly dependent on droplet size and wind speed since smaller droplets will tend to follow the streamlines and avoid the surface. The sticking efficiency (α_2) is linked to the particles bouncing on the surface. For water droplets, the sticking efficiency is essentially equal to 1 but for dry snowflakes it is practically null. Finally, the accretion efficiency (α_3) takes into account the flow of water drops when the available heat absorption capacity of the surface is not enough to freeze the entirety of the impinging water particles.

III. ICE THICKNESS MODEL

The ice thickness model consists of two parts: the accretion model accounting for the increase of mass due to the water particles in the air and the sublimation model simulating the decrease of mass.

A. Accretion Model

The first step to model ice accretion is to assess the icing severity from the extra heat extraction of the probes. Considering that each droplet impacting the surface of the probe is heated from the outside temperature to the surface temperature of the probe (the probe is maintained at a fixed temperature) and the

droplets are traveling at the same speed and direction as the wind, it is possible to define the icing severity (ζ) as per Eq. 1. In this equation ΔQ is the extra heat measured by the probe, H is the probe height (see Fig. 2), D is the probe diameter, c_p is the heat capacity of water, T_s is the surface temperature of the probe, T_∞ is the outside temperature and C_0 is an experimental constant that will be optimized in this process. The nature of C_0 is not known with certainty but it probably arises from the fact that water droplets do not lose all of their additional heat before flowing downwards. Physically, C_0 would depend on other parameters (*e.g.* wind speed, temperature and droplet size), but it is hypothesized that the impact of these parameters on its value is negligible.

$$\zeta = C_0 \frac{\Delta Q}{H D c_p (T_s - T_\infty)} \quad (1)$$

Two parameters are needed to model a 2-D slice of ice accretion on a cylinder: the frontal thickness (E) and the lateral thickness (e) as presented in Fig. 1. The side thickness influences the water gathering area of the cylinder. The linear mass accretion rate (\dot{m}) is defined by Eq. 2.

$$\dot{m} = \zeta (D + 2e) \quad (2)$$

Based on the model for rime ice accretion in ISO-12494, with the total mass accreted (m) it is possible to estimate the value of the frontal and lateral thickness according to Eq. 3 and Eq. 4 where ρ is defined as the ice density and is estimated by interpolating between the ice transition curves (wind speed and temperature dependence for in-cloud icing) presented in ISO-12494.

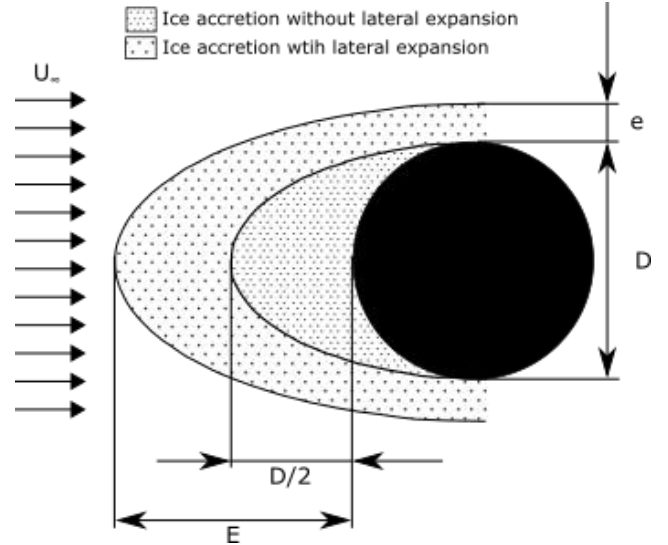


Figure 1: Schematic representation of the dimensions used in the accretion model (adapted from ISO-12494) (ISO, 2017).

$$\begin{aligned} E &= \frac{4 \times 10^6 m}{\rho \pi D} & \text{For } E \leq \frac{D}{2} \\ E &= \frac{D}{2} + 8e & \text{For } E > \frac{D}{2} \end{aligned} \quad (3)$$

$$e = 0 \quad \text{For } E \leq \frac{D}{2}$$

$$e = \frac{1}{32} \left(-10D + \sqrt{68D^2 + \frac{M}{\rho} (8.149 \times 10^7)} \right) \quad \text{For } E > \frac{D}{2} \quad (4)$$

From these equations it is possible to model ice accretion on a cylinder of a similar cylinder as the heated probe.

B. Sublimation Model

The linear sublimation rate can be approximated with the rate equation for convective mass transfer presented in Eq. 5, where h_m represents the convective mass transfer coefficient, ρ_s is the density of water in the air layer just above the surface of the probe and ρ_∞ is the water density in ambient air. In this equation H_c is the half-circumference of an ellipse and can be calculated by the Ramanujan approximation presented in Eq. 6, where a and b are the major and minor axis defined respectively by $\frac{D}{2} + e$ and $\frac{D}{2} - e$.

$$\dot{m} = -h_m H_c (\rho_s - \rho_\infty) \quad (5)$$

$$H_c = \frac{\pi}{2} \left[3(a+b) - \sqrt{10ab - 3(a^2 + b^2)} \right] \quad (6)$$

The convective mass transfer coefficient can be approximated using the convective heat transfer coefficient (h), the binary diffusion coefficient of water in air (D_{AB}), the Lewis number (Le) and thermal conductivity of air (k) as presented in Eq. 7.

$$h_m = \frac{D_{AB} Le^{1/3}}{k} h \quad (7)$$

The density can be derived from the partial pressure of water vapor (P) and the ideal gas formula in which R is the individual gas constant of water vapor and T is the temperature (in °C) as presented in Eq. 8.

$$\rho = \frac{P}{R(T + 273.15)} \quad (8)$$

The vapor pressure of water of air in Eq. 9 (in Pa) can be estimated using the saturation vapor pressure over water equation given by the World Meteorological Organization (2008) and the relative humidity (RH).

$$P_\infty = RH \cdot 611.2 \cdot e^{\frac{17.62 T_\infty}{T_\infty + 243.12}} \quad (9)$$

The vapor pressure on the layer of air immediately over the water surface can be approximated (in Pa) by the saturation pressure over ice given by the World Meteorological Organization (2008) as presented at Eq.10 assuming that the ice surface is at the ambient temperature.

$$P_s = 611.2 \cdot e^{\frac{22.46 T_\infty}{T_\infty + 272.62}} \quad (10)$$

By adding the contribution of the sublimation model to the linear mass rate of the accretion model, the mass accreted can be derived, thus the ice frontal thickness can be computed.

IV. NUMERICAL TREATMENT OF THE IMAGES

A camera placed at a certain distance on a line passing through the MCMS and normal to the wind direction was programmed to take a photo of the sensor every 15 minutes. Ice can accrete on the reference cylinder holding the thermal probe facing the camera since it is not heated. With the position of the camera, every photo allows a measurement of the frontal thickness of ice. A similar approach was used by Jolin *et al.* (2018) to correlate the accretion of ice on the nacelle and the blade. Even though the diameter of the calibration cylinder is slightly larger than the diameter of the probe, it is reasonable to assume that the icing efficiencies variation is negligible. The images were processed using the *MATLAB* software. First of all, the image was stabilized using known edges in order to reduce the effects of the vibration of the camera on the measurements. Fig. 2 presents the MCMS and the measurement zone of the ice thickness. This zone was chosen because of its proximity with the probe and the absence of background elements.

The ice thickness was assessed by finding the high contrast caused by the edge of the accretion. The measurement zone was averaged along its vertical dimension to get a one dimensional variation and to attenuate the image noise. The position of the edge in pixels was converted to mm using the pixel width of the thermal probe width and its real diameter.

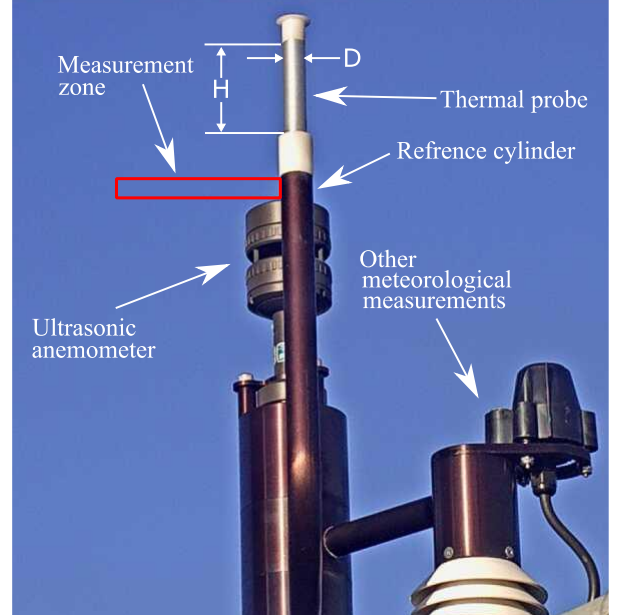


Figure 2: Presentation of the experimental setup consisting of a MCMS on a wind turbine nacelle and the thickness measurement zone for the image processing.

V. RESULTS

The period from the October 2018 26th to January 6th 2019 was analyzed and divided in 18 separate events as presented in Fig. 3. Events of ice shedding due to the shear forces caused by the wind or the accretion weight were manually implemented by forcing the ice thickness to be equal to the observed thickness. Such events occurred infrequently (*i.e.* 7 times over the analyzed period) and are very hard to predict. Fig. 3

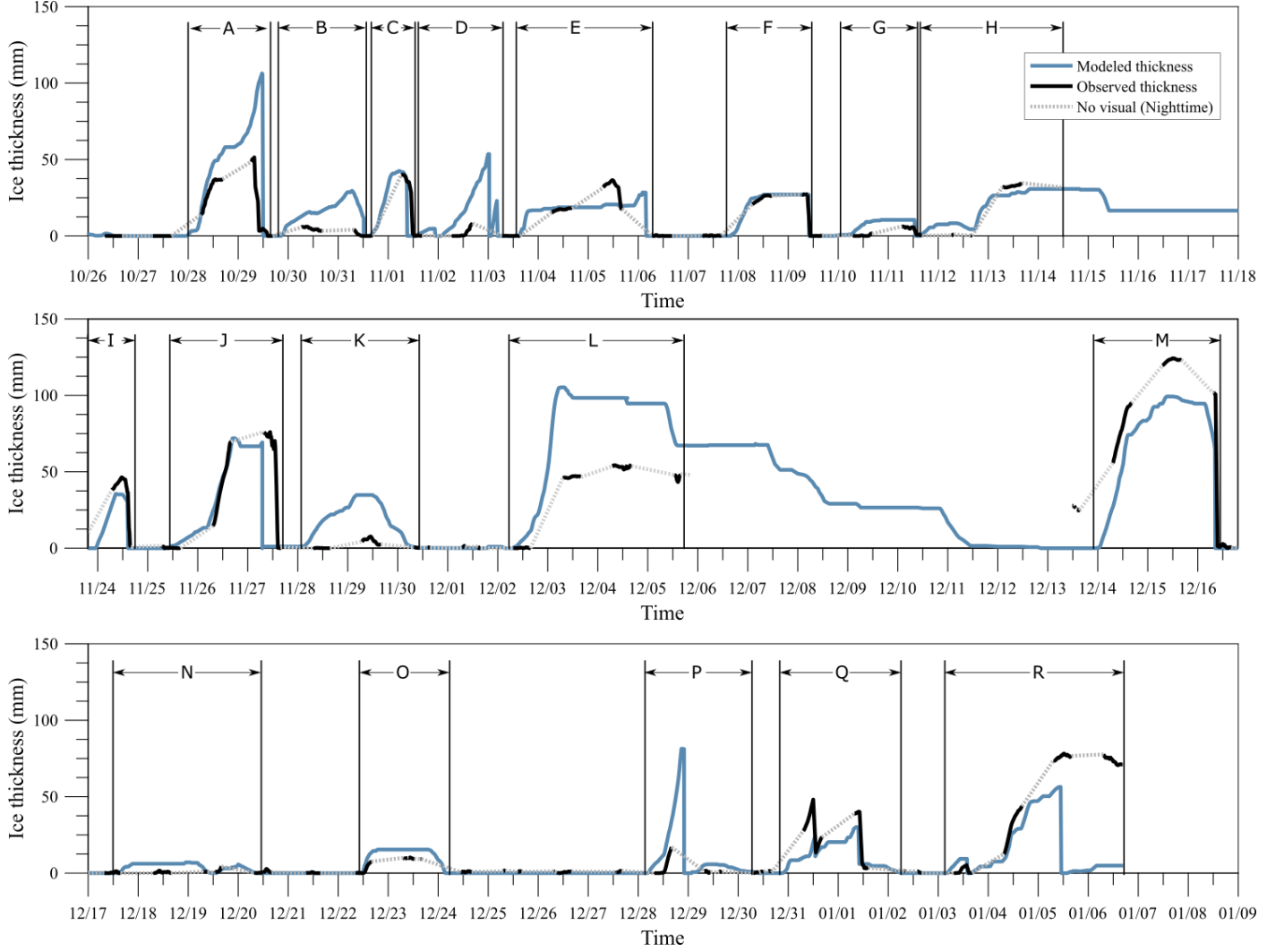


Figure 3: Comparison between the modeled and observed ice frontal thickness from the October 26th 2018 to the January 6th 2019.

presents, in general, a good agreement between the modeled and observed ice frontal thickness. The dashed gray lines represents the tendency interpolated when no visual data was available. Each event was then analyzed individually to get the duration and both observed and modeled characteristic ice accretion rate (IAR) and simultaneous final ice frontal thickness (E_f). An event starts when the first icing conditions are detected and ends when all accreted ice is shed or sublimated. The boundaries of those discretized events are presented on Fig. 3.

The most important accretion rate averaged over an hour and the maximal ice thickness originating from the observed and modeled signals when visual data is available are defined respectively as the characteristic accretion rate and the simultaneous final ice frontal thickness. Table 1 presents a summary of the characteristics of each of the icing events in addition to comments made while analyzing the images. The average difference between the observed and measured characteristic accretion rates is -0.17 mm/h and the standard

deviation is 1 mm/h. The average difference between the observed and measured simultaneous final ice frontal thickness is 0.2 mm and the standard deviation is 24.2 mm. The standard deviations are strongly influenced by few events where no accretion is observed while icing conditions are clearly observable on the camera. In the D, G, K and P events at some time in the beginning of the event icing conditions were easily identifiable on the camera image while no accretions were visible on the reference cylinder. For example, on November 28th, the images presents a very low visibility and visible particles in the air. Small accretions that did not cover the whole surface were visible on the structure. The pattern made by those accretions on the surface changed from image to image suggesting that the particles were colliding with the surface, sticking to it but did not have a sufficient bounding force to resist to the shear forces caused by the wind. Those accretions were probably large particles of wet snow that were probably removed from the surface by wind erosion. In all of those cases, it acted like an incubation period since accretions have always

Event	Duration h	IAR mm/h		E_f mm		Comments
		Obs.	Mod.	Obs.	Mod.	
A	38	6	6	106	51	Ice shed earlier than modeled
B	39	0	1	15	4	Probable ice shed overnight
C	21	0	0	38	42	Accumulation happened overnight
D	41	4	3	8	27	Icing conditions observed but no accretions visible
E	61	2	1	21	36	Probe slightly obstructed
F	38	1	1	27	27	None
G	35	0	1	5	10	Icing conditions observed but no accretions visible
H	69	1	1	32	27	None
I	17	0	2	45	35	None
J	53	5	4	60	60	Data outage at the end
K	57	0	2	6	60	Icing conditions observed but no accretions visible
L	81	0	0	53	98	None
M	60	6	6	124	99	None
N	70	1	1	1	6	Probable ice shed during the night
O	41	4	5	10	15	None
P	54	5	6	17	31	Icing conditions observed but no accretions visible
Q	57	6	4	39	30	None
R	85	6	6	73	56	Ice did not shed when expected

Table 1: Description of the duration, the modeled and observed characteristic ice accretion rate (IAR) and the modeled and observed simultaneous final ice frontal thickness (E_f) of the 18 icing events that occurred from October 26th 2018 to January 6th 2019. Obs. = observed; Mod. = modeled.

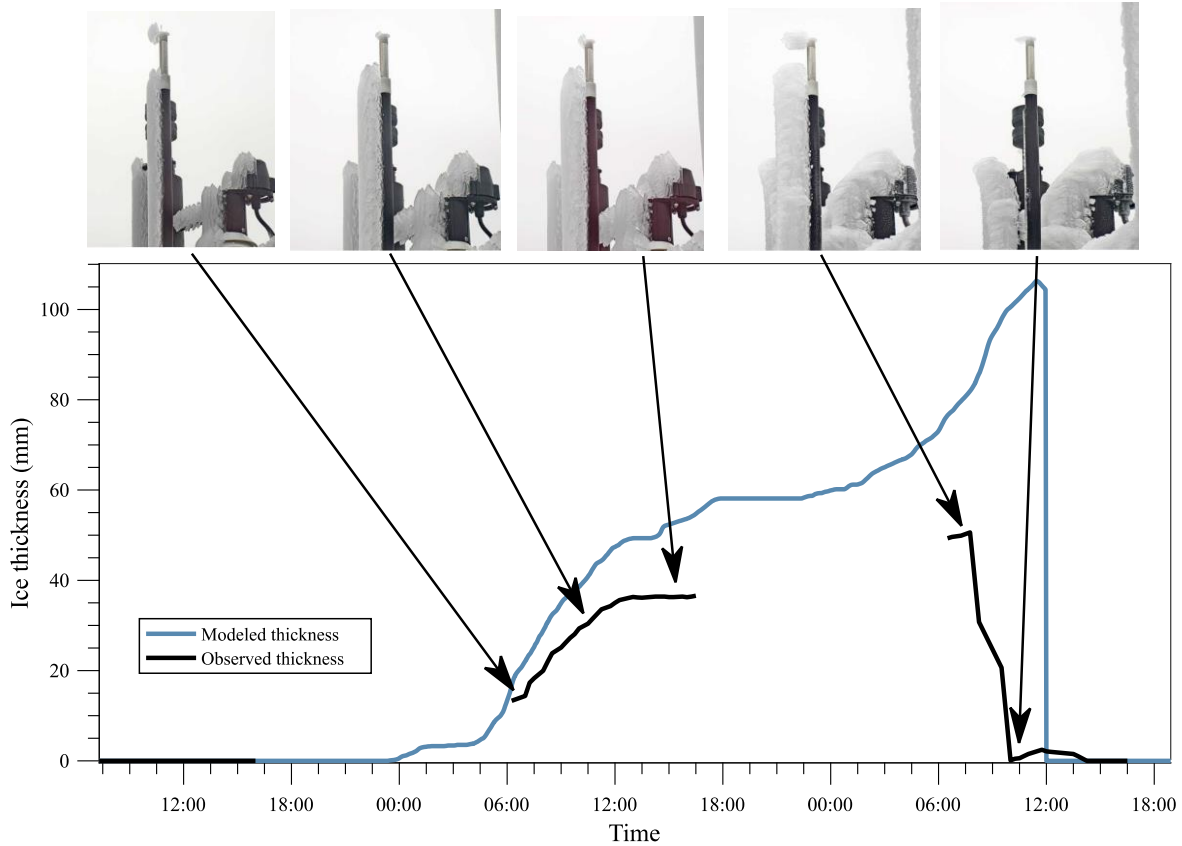


Figure 4: Description of the observed and modeled frontal ice thickness of event A occurring from October 27th 7:30 to October 29th 19:00 in addition to the images taken at 5 different moments along the event. The periods where the observed thickness signal is missing correspond to the moments where the visibility on the camera images was not high enough to be able to decipher the ice thickness (i.e. nighttime).

been observed shortly after those periods. During events B and N, ice probably shed from the measurement cylinder during the night since the ice thickness on the neighboring structures seems to increase therefore resulting in a strong overestimation of the simultaneous frontal ice thickness. The ice shed was not manually implemented since both the accumulation and the ice shed happened during the night.

Fig. 4 presents in further details event A along with the corresponding images. This event is a severe freezing rain event. The figure begins at 7:30 on October 27th and ends at 19:00 on October 29th 2018. The accretion rate on the morning of October 28th is slightly overestimated and the same plateau is observed during the afternoon since no icing conditions were detected. On October 29th the ice shed before the end of the meteorological event probably due to the high mass of the already accreted ice. After ice is shed, the observed accretion continues at a slower rate than the modeled one since it is also influenced by the lateral thickness (e). The end of the period where the ice accretion rate is positive on the visual observations matches really well with the measurements since both thicknesses stop increasing around the same time (around 12:00). The signal corresponding to the observed ice thickness has been smoothed to reduce the effects of the measurement noise.

Even though most of the events in this period are correctly modeled, some of the events are either severely overestimated or underestimated. A first explanation for these discrepancies is that the accretion model assumes an in-cloud rime ice type. The wet growth from in-cloud glaze or freezing rain induces phenomena that can influence the ice accretion form, such as the runback effect and the production of icicles. The agreement of the accretion rate of the beginning of event A, a severe freezing rain event shows that this explanation is probably not the main one. Another potential explanation is that the presence of ice on the top cap and the base of the probe as visible on Fig. 4 can influence the fluid flow around the cylinder and therefore extract more or less heat. Also, the ice buildup on the base did obstruct, at times, the heated surface of the probe, thus reducing the effective area where the liquid water droplets can impact the heated surface. To reduce these unwanted effects, the probes have been redesigned to prevent ice accretion on the top cap and on the base. The top cap must be large enough to prevent 3-dimensionnal effects that can trigger a regime transition of the flow in the range of wind speeds observed. The selected design is an ellipsoidal form where the transition of the radius of the probe is made smoothly. The core of the cap is made of aluminum and the thickness of its insulating part has been chosen so that the surface temperature of the top cap is over 0°C while the metal surface is maintained at 35°C for harsh conditions (*i.e.* wind speed of 20 m/s and an outside temperature of -5°C). The base also has an aluminum core and its thickness has been as well adjusted to prevent ice from forming. The determination of the thickness of the insulators has been done using a finite element method with the *Comsol* software. The outer surface convection coefficient has been estimated

constant at a value corresponding to an average Nusselt number observed around a cylinder for the conditions previously described. Even though during icing events extra heat will be extracted from the probe via the liquid water droplets, it is assumed that once the icing event has ended, the probe will be able to shed ice from its cap and base. In the numerical simulation, it was possible to observe that the minimal temperature over the outside surface is 0.29°C therefore preventing ice accretion. These simulation results have been thereafter confirmed by wind tunnel measurements.

It is possible to observe on Fig. 3 that there is a camera outage between December 6th and December 12th. During this period, no accumulations were measured by the MCMS. The ice accretions visible on the images suggests that the accretion on December 5th and December 13th are the same and the modeled sublimation rate is overestimated. Since the observed accretion was a dense glaze, it is anticipated that the decrease rate of ice would be low due to its high density. The ambient conditions observed during the camera outage were favorable to low density rime ice. The rate at which the frontal thickness varies is derived from Eq. 3 and 4 for both accretion and sublimation. The ice density used in those calculations is evaluated at the current time step conditions (*i.e.* wind speed and temperature). This calculation is only accurate for ice accretion, since the sublimation process involves ice that may have accreted in different conditions, therefore with a different density. For the sublimation model, the ice density value should be based on the density evaluated during ice accretion. This modification will be added to the next version of the model.

VI. CONCLUSION

In this paper, a physical model of ice accretion on a cylinder based on the measurement of icing severity, wind speed and ambient temperature has been presented. This model has been calibrated using visual data retrieved by a camera placed orthogonally to the flow on top of a wind turbine nacelle. The test period took place from October 26th 2018 to December 6th 2019. The analysis showed that the calibrated model was able to replicate the observed ice thickness precisely in most of the events. These results therefore confirm the accuracy of the icing severity measurement calibration of the MCMS. It was noticed that the buildup of ice on the top cap and the base could have influenced the measurements. Ice free edges have been designed to eradicate this unwanted impact. These new probes will soon be installed on wind turbines.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge the Engineering Research Council of Canada (NSERC) (RGPIN-2015-04564 and BESC-M) for the financial support they provided for this research as well as our undisclosed industrial partner who provided us access to a wind turbine.

REFERENCES

- [1] Bégin-Drolet, A., Ruel, J., & Lemay, J. (2017). U.S. Patent No. 9,846,261. Washington, DC: U.S. Patent and Trademark Office.
- [2] Canadian Wind Energy Association. (2018). Wind Energy – A Climate-Fighting Crowd Favourite. URL: <https://canwea.ca/blog/2018/03/09/wind-energy-a-climate-fighting-crowd-favourite/>.
- [3] Frohboese, P., & Anders, A. (2007). Effects of icing on wind turbine fatigue loads. In *Journal of Physics: Conference Series* (Vol. 75, No. 1, p. 012061). IOP Publishing.
- [4] Hudecz, A. (2013). Icing Problem of wind turbine blades in Cold Climate, PhD Thesis, Technical University of Denmark. ISBN: 978-87-92896-80-3.
- [5] ISO (2017). Atmospheric Icing of Structures. (ISO 12494:2017 (E) 2000).
- [6] Jolin, N., Bolduc, D., Swytink-Binnema, N., Rosso, G., & Godreau, C. (2019). Wind turbine blade ice accretion: A correlation with nacelle ice accretion. *Cold Regions Science and Technology*, 157, 235-241.
- [7] Lacroix, A. (2013). Atmospheric Icing Effects on Wind Energy Production in Canada, Winterwind 2013, Ostersund, Seeden, February 12-13.
- [8] LeBlanc, M. P., Morgan, C. A., Bossanyi, E. A., & Garrad, A. D. (2007). Recommendations for risk assessments of ice throw and blade failure in Ontario. Garrard Hassan Canada.
- [9] Lehtomaki, V. & Cattin, R. (2019). IEA Wind Task 19 – Key results from 2016-2018 Future plans 2019-202, Winterwind 2019, Umea, Sweden, February 5-6.
- [10] Wickman, H. (2013). Evaluation of field tests of different ice measurement methods for wind power: focusing on their usability for wind farm site assessment and finding production losses. M.S. thesis, Upsala University, Sweden.
- [11] World Meteorological Organization. (2008). Guide to Meteorological Instruments and Methods of Observation, WMO- No. 8, seventh edition. URL: <https://www.weather.gov/media/epz/mesonet/CWOP-WMO8.pdf>