SICET 2025

NEXT GENERATION ENGINEERING FOR
DIGITAL AND SUSTAINABLE SOLUTIONS

NATIONAL SCIENCE FOUNDATION

SLIIT

4TH INTERNATIONAL CONFERENCE ON

ENGINEERING AND TECHNOLOGY

CONFERENCE PROCEEDINGS

09th to 12th September,2025 | @ SLIIT- Malabe

FACULTY OF ENGINEERING

25 YEARS OF EXCELLENCE SLIIT UNI
THE KNOWLEDGE UNIVERSITY

www.sicet.sliit.lk



4TH SLIIT INTERNATIONAL CONFERENCE ON ENGINEERING AND TECHNOLOGY (SICET) 2025

"NEXT GENERATION ENGINEERING FOR DIGITAL AND SUSTAINABLE SOLUTIONS"

PROCEEDINGS

This Institution does not, as a body, hold itself responsible for statements made or opinions expressed either in the papers read or the discussions which have occurred at the Meetings. – Chair , SICET 2025

The articles published in the proceedings are based on engineering research and some are of professional interest. All published papers have been refereed in anonymity by at least two subject specialists.

ISSN 2961-5011

9 772961 501001

Published by:
Faculty of Engineering,
Sri Lanka Institute of Information Technology,
SLIIT Malabe Campus, New Kandy Rd,
Malabe 10115,
Sri Lanka
E-mail: sicet@sliit.lk
Website: https://sicet.sliit.lk

SICET 2025 Organizing Committee

Prof. Dushantha Jayakody	Conference Chair
Eng. Anuradha Kumarasinghe	Conference Co-Secretary
Eng. Sudara Withana	Conference Co-Secretary
Prof. Manjula Fernando	Technical Programme Committee Chair
Dr. Lakmini Malasinghe	Technical Programme Committee Co-Chair
Prof. Prasanna Gunawardane	Publication Chair
Ms. Sihara Gallage	Publication Co-Chair
Dr. Lasantha Seneviratne	Finance Chair
Dr. Shavindree Nissanka	Finance Co-Chair
Dr. Chandima Somarathna	Workshops Chair
Eng. Sarath Wickramasuriya	Workshops Co-Chair
Dr. Eranga Wijesinghe	Media, Communication and Publicity Chair
Eng. Hemantha Bandara	Logistics Chair
Mr. Susantha Wanniarachchi	Logistics Co-Chair
Mr. Gayal Rangajeewa	Inauguration and Closing Chair
Dr. Nushara Wedasinghe	Inauguration and Closing Co-Chair
Mr. Pramuditha Coomasaru	Registrations and Invitations Chair
Ms. Anushka Panawenna	Panel Discussions/ Keynotes Chair
Mr. Pramuditha Coomasaru	Excursion Chair
Prof. Priyan Mendis	Track Chair (International)
Prof. Young-Do Choi	Track Chair (International)
Dr. Perumalla Venkata Mallikarjun Rao	Track Chair (International)
Dr. Asela K. Kulatunga	Track Chair (International)
Prof. Murad Khan	Track Chair (International)
Prof. Rui Dinis	Track Chair (International)
Prof. Prayag Tiwari	Track Chair (International)
Prof. Sumedha Rajakaruna	Track Chair (International)
Prof. Sérgio D. Correia	Track Chair (International)
Dr. Mohan Siriwardena	Track Chair (International)
Prof. Alessandro Puiatti	Track Chair (International)
Dr. Tharindu Ponnimbaduge	Track Co-Chair (International)
Prof. Dejan Vukobratovic	Track Co-Chair (International)
Prof. Mashook Khan	Track Co-Chair (International)
Prof. Honnyong Cha	Track Co-Chair (International)
Dr. Sivaram NM	Track Co-Chair (International)
Prof. Mansik Jeon	Track Co-Chair (International)
Prof. Jeehyun Kim	Track Co-Chair (International)
Prof. Shiromi Karunaratna	Track Co-Chair (Local)
Prof. Priyan Dias	Track Co-Chair (Local)
Dr. Priyantha Bandara	Track Co-Chair (Local)
Prof. Anurudda Puswewala	Track Co-Chair (Local)
Prof. Saman Bandara	Track Co-Chair (Local)
Dr. Nandika Miguntanna	Track Co-Chair (Local)
Dr. Nathali de Silva	Track Co-Chair (Local)
Mr. Hanojhan Rajarajam	Track Co-Chair (Local)
Dr. Hansani Weerathunga	Track Co-Chair (Local)
Dr. Ruwan Chnadrasena	Track Co-Chair (Local)
Dr. Kumara Jayawickrama	Track Co-Chair (Local)
Dr. Shavindree Nissanka	Track Co-Chair (Local)
Dr. Lakmini Malasinghe	Track Co-Chair (Local)
Prof. Chulantha Kulasekare	Track Chair (Local) – Industry Connect
Mr. Tilanka Wijesinghe	Track Chair (Local) – Industry Connect

SICET 2025 Conference Board

Prof. Lakshman Ratnayake	Chairman/Chancellor
Prof. Lalith Gamage	President/CEO/Vice-Chancellor
Prof. Nimal Rajapakse	Senior Deputy Vice-Chancellor & Provost
Prof. Rahula Attalage	Pro Vice-Chancellor (Academic)
Prof. Samantha Thelijjagoda	
Prof. Ayantha Gomes	Dean, Faculty of Engineering
Dr. Priyantha Bandara	Outgoing Chair SICET 2024
Dr. S.V.T Janaka Perera	Head, Department of Civil Engineering
Prof. Migara Liyanage	Head, Department of Mechanical Engineering
Dr. Ruwan Chandrasena	Head, Department of Electrical and Electronic Engineering
Dr. Mudith Karunaratne	Head, Department of Materials Engineering
Mr. Ch. QS. Tilanka Wijesinghe	Head, Department of Quantity Surveying
Dr. Nihal Somaratne	Senior Lecturer (Higher Grade)

SICET 2025 Advisory Committee

Prof. Li Ainong	Chinese Academy of Sciences, China
Prof. Onyx Wing-hong Wai	The Hong Kong Polytechnic University, China
Prof. Takeshi Fujino	Saitama University, Japan
Prof. George Mann	Memorial University of Newfoundland, Canada
Prof. Asanga Ratnaweera	University of Peradeniya, Sri Lanka
Dr. Riccardo Taormina	Delft University of Technology, Netherlands
Prof. Takashi Asaeda	Saitama University, Japan
Prof. Marko Beko	Técnico Lisboa - Universidade de Lisboa, Portugal
Prof Stefan Panić	University of Pristina, Serbia

List of Reviewers

Mr. Adeetha Hettiarachchi	Kyungpook National University
Ms.Akarshani Amarasinghe	University of Sri Jayewardenepura
Prof. Asanga Rathnaweera	University of Peradeniya
Prof. Asoka Perera	Sri Lanka Institute of Information Technology
Dr. Buddhi Wijesiri	Queensland University of Technology
Dr. Chamod Hettiarachchi	University of Moratuwa
Dr. Chandana Jayawardana	Open University of Sri Lanka
Dr. Chandana Kulasuriya	Open University Sri Lanka
Prof. Chandana Perera	Sri Lanka Institute of Information Technology
Mr. Charith Sucharitharathna	Sri Lanka Institute of Information Technology
Prof. Chulantha Kulasekere	Sri Lanka Institute of Information Technology
Damayanthi Herath	University of Peradeniya
Dr. Damith Senanayake	ISN Psychology Melbourne
Mr. Deemantha Siriwardana	Sri Lanka Institute of Information Technology
Mr. Deshan Kalupahana	University of Sri Jayewardenepura
Dr. Dhammika Dharmaratne	Sri Lanka Institute of Information Technology
Mr. Dilum Perera	Sri Lanka Institute of Information Technology
Mr. Dimuthu Kodippili Arachchige	University of Illinois Chicago
Mr. Dimuthu Wanasinghe	University of Western Australia
Dr Chandana Perera	University of Ruhuna
Dr. H. R Pasindu	University of Moratuwa
Dr. J.K.D.D.T Jayanetti	Sri Lanka Institute of Information Technology
Dr. Kumara Jayawickrama	Sri Lanka Institute of Information Technology
Porf. Dush Jayakody	Lusofona University
Prof. Dushantha Nalin K. Jayakody	Sri Lanka Institute of Information Technology
Dr. Ehsan Pashajavid	Curtin University
Ms. Fathima Imara Mohamed Nazar	State University of New York at Albany
Mr. Ganegoda Vidanage Charaka Rasanga	CINEC Campus
Dr. Ganga Samarasekara	University of Sri Jayewardenepura
Dr. Gayashika Fernando	Sri Lanka Institute of Information Technology
Prof. George Mann	Memorial University of Newfoundland
Mr. Hanojhan Rajahrajasingh	Sri Lanka Institute of Information Technology
Dr. Hansani Weeratunge	Sri Lanka Institute of Information Technology
Dr. Harinda Fernando	Sri Lanka Institute of Information Technology
Ms. Imasha Pasqual	Sri Lanka Institute of Information Technology
Ms. Iresha Attanayake	Open University Sri Lanka
Dr. Iromi Ranaweera	University of Ruhuna
Mr. Jaliya Wijayaraja Mudiyanselage Mathmaluwe Gedara	Sri Lanka Institute of Information Technology
Ms. Jannat Luna	Kyungpook National University
Dr. Kahala Abhayasinghe	Curtin University
Dr. Kamani Gettapola	Experian Australia Pty Ltd
Mr. Kanishke Gamagedara	American Robotics

Dr. Kapila Dissanayaka	Sri Lanka Institute of Information Technology
Ms. Kavinya Perera	Sri Lanka Institute of Information Technology
Dr. Kudabadu Kumara	University of Ruhuna
Ms.Kumari Wijekoon	Sri Lanka Institute of Information Technology
Dr. Kumudu Amarawardhana	University of New South Wales Canberra
Mr. Lahiru Jayasuriya	Rajarata University of Sri Lanka
Dr. Lahiru Lindamulla	Sri Lanka Institute of Information Technology
Dr. Lakmini Malasinghe	Sri Lanka Institute of Information Technology
Dr. Lasith Yasakethu	Sri Lanka Technological Campus
Mr. Lasitha Weerakoon	University of Maryland
Dr.Madhusha Ishanthi Sudasinghe	University of Colombo
Prof. Manjula Fernando	Sri Lanka Institute of Information Technology
Ms. Manjula Siriwardhana	Open University of Sri Lanka
Mr. Mayura Pavan Manawadu Manawaduge	The University of Melbourne

Prof. Meththika Vithanage	Faculty of Applied Sciences, University of Sri Jayewardenepura
Dr. Migara Liyanage	Sri Lanka Institute of Information Technology
Mrs. Jeewani Abegoonawardana	General Sir John Kotelawala Defence University
Dr. Murad Khan	Kuwait College Of Science and Technology
Dr. Nadeeka Miguntanna	General Sir John Kotelawala Defence University
Dr. Namali Sirisoma	General Sir John Kotelawala Defence University
Dr. Nandika Miguntanna	Sri Lanka Institue of Information Technology
Mr. Nipun Shantha Kahatapitiya	Kyungpook National University
Prof.Niranga Amarasingha	Sri Lanka Institute of Information Technology
Mr. Nisal Weerasinghe	Sri Lanka Institute of Information Technology
Mr. Nishanthi Gunarathna	Sri Lanka Institute of Information Technology
Dr. Nushara Wedasingha	Sri Lanka Institute of Information Technology
Dr. Nuwan Herath	University Of Peradeniya
Ms. Oshadhi Gunathilake	Sri Lanka Institute of Information Technology
Prof. Pattiyage Ayantha Gomes	Sri Lanka Institute of Information Technology
Dr. P.P.C.R Karunasekara	General Sir John Kotelawala Defence University
Mr. Pramuditha Coomasaru	Sri Lanka Institute of Information Technology
Prof. Priyan Dias	Sri Lanka Institute of Information Technology
Dr. Priyantha Bandara	Sri Lanka Institute of Information Technology
Ms, Priyanwada De Silva	Sri Lanka Institute of Information Technology
Mr. Priyashantha Tennakoon	SLTC Research University
Prof. Kumari Fernando	General Sir John Kotelawala Defence University
Dr. R.P.S. Chandrasena	University Of Ruhuna
Prof. Ramya Kumanayake	General Sir John Kotelawala Defence University
Mr. Rangajeeva Samarathunga	General Sir John Kotelawala Defence University
Ms. Rasanjalee Rathnayake	Informatics Institute of Technology
Ms. Rathnamalie Kotambage	Sri Lanka Institute of Information Technology
Mr. Rathnaweera Bandara	Sri Lanka Institute of Information Technology
Ms. Raveesha Illesinghe	University of Peradeniya
Dr. Renuka Ariyawansha	Sri Lanka Technological Campus
Mr. Rochana Ranasinghe	Sri Lanka Institute of Information Technology
Dr. Ruchire Wijesinghe	Sri Lanka Institute of Information Technology
Mr. Ruminda Wimalasiri	Open University of Sri Lanka
Prof. Ruwan Gopura	University Of Moratuwa
Ms. Ruwanga Amarasingha	Sri Lanka Technology Campus
Prof. Saman Bandara	University Of Moratuwa
Prof. Samantha Wijewardane	University of South Florida
Mr. Sameera Samarasinghe	Energy Queensland
Mr. Samitha Herath	Monash University
Dr. Sandun Konara	Department of Electrical and Information Engineering, University of Ruhuna
Ms. Sanjeevi Chandrasiri	Sri Lanka Institute of Information Technology
Prof. Sarath Hettiarachchi	Sri Lanka Institute of Information Technology
Mr. Shanika Matharage	The University of Manchester
Dr. Shavindree Nissanka	Sri Lanka Institute of Information Technology

Dr. Shayani Mendis	General Sir John Kotelawala Defence University
Ms. Shehani Jayasinghe	Sri Lanka Institute of Information Technology
Prof. Shiromi Karunaratne	Sri Lanka Institute of Information Technology
Mr. SM Abu Saleah	Kyungpook National University
Mr. Sudara Withana	Sri Lanka Institute of Information Technology
Prof. Prasanna Gunawardane	Sri Lanka Institute of Information Technology
Prof. Sumedha Rajakaruna	Curtin University
Dr. Sushovan Chatterjee	Cooch Behar Government Engineering College
Ms. Thanushika Jathunga	Sri Lanka Institute of Information Technology
Prof. Tharaka Samarasinghe	University of Moratuwa
Mr. Tharaniyawarma Kumaralingam	Sri Lanka Institute of Information Technology
Ms. Thilini Amarasooriya	Sri Lanka Institute of Information Technology
Ms.Thilini De Silva	University Of Ruhuna
Dr. Thillaiampalam Sivakumar	University Of Moratuwa
Ms.Thisakya Ransarani	Sri Lanka Institute of Information Technology
Prof. Thumeera Wanasinghe	Memorial University of Newfoundland
Ms. Tinaya Wijethunge	Kyungpook National University
Dr. Udaya Wijenayake	University of Sri Jayewardenepura
Dr. Urvashi Chaudhary	Lusofona University
Dr. Vajira Edirisinghe	Sri Lanka Institute of Information Technology
Ms. Vishalya Sooriarachchi	Lusofona University
Prof.Wahidul Biswas	Curtin University
Prof. Walallawita Wimalsiri	Sri Lanka Institute of Information Technology
Dr. Waranatha Abeygunasekara	University Of Peradeniya
Prof. Young-Do Choi	Mokpo National University

Preface



It is my great pleasure to present the proceedings of the 4th SLIIT International Conference on Engineering & Technology (SICET 2025). SICET continues to grow as a platform where rigorous scholarship meets practical impact—bringing together researchers, students, industry practitioners, and policymakers to co-create solutions for the public good.

This year's programme reflects both scale, international and quality. We received 176 submissions, from which 102 papers were selected through a double-blind review process. To broaden participation and practical skills, we ran 11 pre-conference workshops, 5 of which were international, featuring hands-on tutorials and emerging methods from global partners. Ensuring fairness and depth in evaluation, over 250 reviewers contributed to the double-blind process, with meta-review and TPC decisions taken per track to balance novelty, reproducibility, and societal impact.

SICET's intellectual breadth spans 11 technical tracks—from Next-Generation Communications & Networking, AI & Data-Driven Engineering, and Clean & Renewable Energy to Mechatronics & Robotics, Thermal & Fluid Engineering, Biomedical Engineering & Smart Health Systems, Sustainable Built Environment, Circular Economy, ICT for Societal Development, Next-Generation Transportation, and Construction Dynamics & Digitalisation. Our community is proudly international, with authors and institutions represented from Australia, Bangladesh, Canada, India, Japan, Kosovo, Maldives, Portugal, Russia, South Korea, the United Kingdom, and the United States, alongside Sri Lankan universities and research institutes.

The technical agenda is anchored by two distinguished Keynote Speakers—Prof. Sofiène Affes (INRS, Canada) and Prof. Andrey Koucheryavy (SPbSUT, Russia)—and complemented by two timely panel discussions: *"The Future of Research in Sri Lanka: Vision for 2030"* and *"AI in Engineering: Our Perspective."* Reflecting our commitment to translation and impact.

Our publication pathway ensures visibility and archival value. Proceedings are planned via the ACM International Conference Proceedings Series (ICPS), with selected papers invited to partner journals, including Advances in Science & Technology (AST), Key Engineering Materials (KEM), ARIS² etc.

I am deeply grateful to the Vice Chanceller and the University management, Dean of Engineering, Technical Programme Committee, Publication Committee, Track Chairs, over 250 reviewers, Workshop chairs and Logistics Committee, the Organising Committee, and our many volunteers for their exemplary service from Faculty of Engineering of SLIIT; to our authors for entrusting us with their best work; and to our partners and supporters for enabling this community to thrive. I must thank all our Sponsers who made this event a success. I invite all participants to engage with these proceedings—build on the ideas presented here, challenge assumptions, and collaborate across disciplines and borders to advance knowledge and deliver meaningful impact.

Prof. Dush N. K. Jayakody
General Chair, SICET 2025

Message from the Chancellor



It is with great pleasure that I extend my warmest congratulations to the organizers of the 4th SLIIT International Conference on Engineering & Technology (SICET 2025) and to all the authors whose work is featured in these proceedings. SICET continues to serve as a reputable, multidisciplinary forum that unites researchers, practitioners, and students to exchange ideas that advance engineering and technology for the benefit of society.

This year, the conference received 176 submissions from around the globe, including contributions from the UK, Australia, Canada, Europe (notably Portugal and Cyprus), China, Pakistan, and many others. Additionally, we welcomed representation from nearly all universities and research institutes in Sri Lanka. Following a rigorous peer-review process conducted by multiple independent reviewers, 102 papers were accepted for presentation across various technical tracks, demonstrating our community's commitment to quality and relevance. Selected papers will be published in the conference proceedings and featured in select journals, ensuring global visibility and lasting accessibility for the research presented.

SICET 2025 encompasses eleven technical tracks that cover both core and emerging areas of engineering. I am particularly pleased to highlight the introduction of an Industry Track this year, an initiative aimed at strengthening collaboration between industry and academia, fostering technology transfer, and inspiring solutions with real-world impact. This vital connection between research and practice is crucial for Sri Lanka's innovation ecosystem and for the professional growth of our graduates.

I would like to express my sincere gratitude to the Technical Program Committee, Track Chairs, Reviewers, Publication Partners, Sponsors, and the many dedicated staff and student volunteers whose efforts made this conference possible. To all authors and participants, I wish you fruitful discussions, the establishment of meaningful new partnerships, and continued success in your research endeavors.

With best wishes for a productive and memorable SICET 2025.

Prof. Lakshman L. Ratnayake
Chancellor - SLIIT

Message from the Vice-Chancellor



It is with great pleasure that I welcome you to the 5th International Conference on Engineering and Technology (SICET 2025), organised by the Faculty of Engineering at SLIIT. SICET has established itself as a respected forum where research, practice, and education come together to address important challenges in engineering and technology. This year, the conference brings together participants from around the world, along with a strong representation from Sri Lanka's universities and research institutes. This demonstrates both the diversity of our scholarly networking and the importance of international collaboration in today's world.

As Sri Lanka moves forward in a knowledge-based economy, producing highly capable researchers and graduates is essential. SLIIT is committed to supporting this effort through various research funding opportunities, industry partnerships, and international engagements. I am confident that the discussions at this conference will focus on real challenges and practical solutions that can contribute to a sustainable and progressive future. This year's conference has received more than 170 submissions from international scholars and various universities and research institutes in Sri Lanka. Such participation reflects the confidence of the academic community in SICET as a valuable platform. The technical program that spans over a wide range of areas, from next-generation communications and artificial intelligence to clean energy, smart health systems, and sustainable cities, showcases that these themes closely align with SLIIT's research priorities and its commitment to addressing real needs in society. The inclusion of an Industry Track further reflects our culture of collaboration, where academic research is connected with practical applications to create a meaningful impact for communities and industry alike.

The publication opportunities offered by SICET through international partners and SLIIT's own journals ensure that the outcomes of this conference will reach beyond these sessions and support future research and development. We are also privileged to host keynote speakers Prof. Sofiène Affes from the Institut national de la recherche scientifique (INRS), Canada, and Prof. Andrey Koucheryavy from The Bonch-Bruevich Saint-Petersburg State University of Telecommunications (SPbSUT), Russia, whose insights will guide and inspire the discussions.

The success of this conference comes from the contributions of its participants, and their commitment gives SICET its true value. I would also convey my appreciation to the organising committee and the Faculty of Engineering for the effort and dedication that made this event possible. All participants are encouraged to engage actively, share knowledge, and build meaningful collaborations. I am confident that SICET 2025 will mark another important step in advancing engineering, technology, and related fields.

Professor Lalith Gamage
Vice Chancellor/ Chief Executive Officer, SLIIT

Message from the Senior Deputy Vice-Chancellor & Provost



SICET 2025 is both a showcase of frontier research and a classroom without walls. This edition brings our community together in a hybrid format, welcoming participants on site and online. The spirit of SICET is collaborative and future-facing—uniting scholars, students, industry, and policymakers to translate ideas into solutions that benefit society.

The Technical Programme reflects academic rigor and translational relevance. From 176 submissions, 102 papers were selected through a double-blind review process—with 35 international paper submissions, sessions balanced for novelty, innovation, and societal impact.

The audience is proudly global, with participation from institutions in Sri Lanka, Australia, Bangladesh, Canada, India, Japan, Kosovo, the Maldives, Portugal, Russia, South Korea, the United Kingdom, and the United States—a testament to SICET's growing international reach. This diverse participation underscores the global impact of our work. Two distinguished Keynote Speakers anchor the agenda: Prof. Sofiène Affes (INRS, Canada) and Prof. Andrey Koucheryavy (SPbSUT, Russia). In addition, two timely panel discussions—*"The Future of Research in Sri Lanka: Vision for 2030"* and *"AI in Engineering: Our Perspective"*—bring together national leaders from the NSF, NRC, the Ministry of Science & Technology, senior academics, and advisors to examine policy, talent pipelines, and innovation pathways.

The conference features a unique Industry Connect stream that highlights deployable solutions and university–industry partnerships with tangible outcomes. These outcomes include pipelines for internships, joint R&D, and sponsored capstones, demonstrating the real-world impact of our work. The publication pathway underscores quality and visibility.

I extend my sincere appreciation to the authors, reviewers, Track Chairs, the Technical Programme Committee, the organizing team, sponsors, and the Faculty of Engineering of SLIIT. Their invaluable contributions have been instrumental in making SICET 2025 a resounding success. I encourage participants to mentor generously, ask bold questions, and convert the insights gained this week into sustained collaborations that yield global benefits.

Prof. Nimal Rajapakse
Senior Deputy Vice-Chancellor & Provost, SLIIT

Message from the Dean, Faculty of Engineering



It is a pleasure to extend my warmest greetings to all participants of SICET 2025 and to present this message on behalf of the Faculty of Engineering (FoE), SLIIT. SICET has become a signature achievement of our university—an arena where high-quality engineering research is shared, challenged, and transformed into solutions that benefit society.

This year's conference again reflects both breadth and rigour. The Technical Programme spans 11 tracks and a rich slate of 11 pre-conference workshops, 5 of which were delivered by international experts. Upholding academic integrity, the double-blind review was supported by over 250 reviewers, ensuring careful assessment and constructive feedback for every submission. The resulting programme showcases outstanding work across communications and networking, AI and data-driven engineering, energy and sustainability, mechatronics and robotics, biomedical systems, the built environment, and more.

On behalf of the Faculty of Engineering, I wish to record my sincere appreciation to our leadership for their continuous guidance and encouragement: the Chancellor, the Vice-Chancellor, and the Senior Deputy Vice-Chancellor & Provost. I also thank our General Chair, Prof. D. N. K. Jayakody and the Organising Committee, Track Chairs, the Technical Programme Committee, and the many reviewers, workshop leads, volunteers, and staff whose dedication has made SICET 2025 possible. We are grateful as well to our distinguished keynote speakers and to our partners from academia, industry, and government for their collaboration.

I trust that you will find the conference, its presentations, and these proceedings valuable to your current and future research and development work. I encourage you to use SICET to exchange ideas, build new collaborations, and strengthen the bridges between universities, research institutes, industry, and society.

Wishing you a stimulating and rewarding SICET experience.

Prof. Ayantha Gomes
Dean, Faculty of Engineering, SLIIT

Acknowledgement

We are pleased to host the 4th SLIIT International Conference on Engineering and Technology (SICET 2025) in Colombo and Malabe, Sri Lanka. This multidisciplinary forum provides a valuable platform for scholarly exchange, welcoming contributions across diverse research areas including Sustainable Built Environment, Sustainable Consumption and Production, ICT for Societal Development, Transportation Systems, Mechanics, Robotics and Intelligent Systems, Thermal and Fluid Engineering, Energy and Renewable Energy Systems, Hydraulics and Water Resources, Communication and Networking, Electrical Power Systems, Machine Dynamics and Design, Contract and Risk Management, and Industry Connect technical activities.

SICET aims to unite researchers, academics, industry professionals, and policymakers in a collaborative environment to share insights, disseminate findings, and explore the latest advances in engineering and technology. To maintain the highest standards of quality, all submissions have undergone a rigorous double-blind peer review process conducted by distinguished local and international experts. Only papers that exemplified originality, innovation, and scholarly excellence were accepted, thereby ensuring the conference's credibility and impact.

Further, we would like to express our heartfelt appreciation to all who played a vital role in the success of this event. We are grateful to the authors and presenters for their insightful contributions, the reviewers for their meticulous evaluations, the session chair and co-chair for their valuable contributions within the sessions, and the organising committee for their unwavering commitment and hard work. We also give our fullest gratitude to the sponsors for their valuable contributions and support.

We are excited about the potential for further collaboration and knowledge creation stemming from the research presented here. We encourage your ongoing exploration in cutting-edge research areas, and we eagerly anticipate your continued engagement and contributions to advance engineering and technology through SICET 2026.

On behalf of the Organizing Committee,

Prof. Prasanna Gunawardane
Publication Chair, SICET 2025

Table of Contents

List of Reviewers **5**

Preface **9**

Message from the Chancellor **10**

Message from the Vice-Chancellor **11**

Message from the Senior Deputy Vice-Chancellor & Provost **12**

Message from the Dean, Faculty of Engineering **13**

Acknowledgement **14**

Table of Contents **15**

Keynote 01:Optimization Methods in Computational Intelligence for Joint Wireless Channel Parameter Estimation over Advanced Radio Interface Technologies **18**

Keynote 02: Triple Play: AI, MV, and HTC **19**

A Numerical Investigation of the Potential of Dimpled Surface Configurations to Improve Aerodynamic and Aeroacoustic Performance of Airfoils **1**

Evaluation of the Knowledge Base in Agriculture and Food to Reduce and Prevent Chronic Kidney Disease of Unknown Etiology (CKDu) **11**

A Numerical Investigation of Valve Timing and Intake Pressure Effects on Performance and Emissions in a Hydrogen Port Fuel Injection Engine **22**

The Influence Of Project Managers' Decision-Making Styles On Schedule Variance In Building Construction Projects **32**

High-Speed Path Tracking of a Small Mobile Robot Using PD and ADRC Controllers with Experimental Validation **42**

Development of an AI-Based Model with Low Computational Complexity for Accurate Load Demand Forecasting **52**

Design, Simulation, and Optimization of a Hybrid Hydrogen Fuel Cell–Battery Energy System for Sustainable Electric Vehicle Applications Using MATLAB Simulink **63**

Enhancing Healthcare Predictive Models Through Privacy- Preserving Synthetic Data Generation **71**

A Cost-Effective Battery Retrofit for Non-Hybrid Grid-Tied PV Systems to Reduce Solar Energy Loss During Grid Outages **78**

Preserving Built Heritage in Sri Lanka through Digital Twin Technology: Opportunities and Implications **85**

FreshSight: An Accessibility-Focused Approach to Produce Freshness and Shelf-Life Detection for Food Safety and Waste Reduction **94**

Towards Safer Elderly Care: A Convolutional Neural Network Solution for Fall Detection ... **104**

Development of an AI-Based Model with Low Computational Complexity for Accurate Wind Energy Forecasting **112**

Enhancement of Quality Management Through Lean in Sri Lankan Construction Industry .. 122

Multifactorial Drivers of Chronic Kidney Disease of Unknown Etiology (CKDu): A Review of Cadmium Exposure, Ultraviolet B B Radiation, and the Potential Role of Vitamin D Toxicity 132

Developing A Web-Based Augmented Reality Tool For Promoting Sustainable Fashion Consumption 142

Piezoelectric Energy Harvesting in Interactive Environments: A Case Study Using a Playpen-Based Prototype 154

Development Of An Ai-Based Model With Low Computational Complexity For Accurate Solar Energy Forecasting 162

Li-ion Battery Cooling - A Computational Study of Different Phase Change Material Configurations 172

Public Sector Role as a Key Stakeholder Towards the Circular Economy in the Built Environment in Sri Lanka 184

A Comparative Study on TiO_2/Graphite–PEG and Graphite/Carbon Fibre-Paraffin Shape Stabilized Phase Change Materials for Thermal Energy Storage Applications 198

A Numerical Investigation of Valve Timing and Intake Pressure Effects on Performance and Emissions in a Hydrogen Port Fuel Injection Engine 205

Diabetic Retinopathy Screening Using Image Processing 216

Penetratingz the Defenses: An Investigation into the Achilles' Heel of HTTP and SSH 224

Construction Dynamics And Digitalization 231

Feature Analysis of Blood Spatter Patterns with Image Processing 237

Adoption Of A Circular Economy Framework In The Design Phase Of The Refurbishment Project 247

Analysis And Optimization Of OLSR And AODV Routing Protocols For Highly Mobile Autonomous Aerial Vehicle Networks: Experimental Performance Evaluation In Various Application Scenarios 259

Analysis of Thermal Performance of Shell and Tube Heat Exchangers: A Correlation and CFD Based Approach 281

Circular Economy Practices in Road Rehabilitation and Development in Sri Lanka 291

Fog Computing Power For Telepresence Suit As User Terminal For Metaverse 305

Highly Efficient 3D Object Transmission System for HTC Services in 6G Networks 316

Improving Post-Harvest Rice Drying Efficiency through a Low-Cost Halogen Dryer Design for Rural Communities 325

***Macaranga peltata* Leaf Extract Mediated Green Synthesis of Iron Nanoparticles and Their Application in Organic Dye Removal 337**

Predicting Cognitive Test Performance from New Onset Behavioral and Personality Changes in Adults over 50 using Post-Selection Boosted Random Forest Classifier 349

Solar Hotspot Detection Using VHDL-Simulated Fixed-Point SVM: A Methodology Toward FPGA Realization Solar Hotspot Detection via FPGA-SVM 357

Synergistic Charge Dynamics and Light Harvesting in TiO_2/MgO Composites for Efficiency Enhancement in CdS Quantum Dot-Sensitized Solar Cells 374

Enhancing Modern Education Through an AI-Integrated Learning Management and Support System (LMSS) 385

Development of Low-cost Slipper by using NR/EVA Blend with Recycling Materials for reducing Environment Pollution in Footwear Industry 389

Predictive Modeling for Personalized Cancer Therapy Using Reinforcement Learning..... 399

A Data-Driven Approach to Predicting Ischemic Heart Disease Risk in Monaragala: Integrating Lifestyle and Symptom Factors with Machine..... 411

Keynote 01:Optimization Methods in Computational Intelligence for Joint Wireless Channel Parameter Estimation over Advanced Radio Interface Technologies

Prof. Sofiene Affes
Wireless Communications Group
Institut national de la recherche scientifique (INRS), Canada

ABSTRACT

Advanced Radio Interface Technologies (RITs) combine broadband signalling—hence multi-carrier operation and richly multipath propagation—with multi-antenna transceivers. In these regimes, joint estimation of channel parameters (angles of arrival/departure, delays, Doppler/frequency offsets, gains/phases, etc.) becomes a central yet challenging inference problem. Objective or cost functions are often nonconvex, multimodal, and simulator-defined, with scarce gradients, tight pilot budgets, and low signal-to-noise ratios (SNRs). Therefore, among computational intelligence (CI) categories that encompass 1) neural networks and 2) fuzzy systems, the third or 3) population-based and bioinspired optimization (PBO/BO) methods – such as particle swarm optimization (PSO), differential evolution (DE), genetic algorithms (GA), grey wolf optimizer (GWO), and related swarms – have gained traction as global search engines that either directly minimize maximum-likelihood (ML) or mean-square error (MSE) criteria or act as robust initializers for hybrid pipelines. In this talk, first we integrate a disciplinary taxonomy relating artificial intelligence (AI), optimization, and Monte Carlo inference to place CI and PBO/BO within a broader computational context worth contemplating. Then we survey the current state of the art on CI optimization for wireless channel parameter estimation and analyze the strengths and weaknesses of each CI subcategory versus the others and against conventional estimation methods. We synthesize algorithmic patterns, objectives, accuracy, convergence/complexity trends, and empirical findings, etc., over advanced RITs, and we discuss most recent progress and open challenges.

BIOGRAPHY



Sofiene Affes received the Engineering Diploma in Telecommunications and the Ph.D. Degree (Hons.) in Signal Processing from Telecom Paris-Tech (ENST), Paris, France, in 1992 and 1995, respectively. He has been since with INRS, Montreal, QC, Canada, as a Research Associate until 1997, an Assistant Professor until 2000, an Associate Professor until 2009, and from then on as a Full Professor. From, 2013 till 2019, he was the Director of PERWADE, a unique 4M$ research-training program on wireless in Canada involving 27 academic and industrial partners from 8 universities and 10 industrial organizations. Dr. Affes has been twice the recipient of a Discovery Accelerator Supplement Award from NSERC, from 2008 to 2011 and from 2013 to 2016. From 2003 to 2013, he held a Canada Research Chair in Wireless Communications. And from 2017 to 2022, he was the recipient of a Cyrille-Duquet Research Chair in Telecommunications. In 2008 and 2015, he received the VTC and ICUWB Chair Recognition Award and Certificate from IEEE VTS and MTT-S, respectively, for exemplary contributions as General Co-Chair to the success of both conferences held at their Fall 2006 and 2015 editions in Montreal. In 2017, he served as General Chair of IEEE PIMRC also held in Montreal. His team received Best Paper Awards at IEEE Globecom, IEEE ICASSP, and IEEE VTC. He has been or is currently an Editorial Board Member for the IEEE Transactions on Communications, Signal Processing, and Wireless Communications, for the Wiley Journal on Wireless Communications & Mobile Computing, the MDPI Sensors and Sci Journals, and the Electronics and Signal Processing Journal.

Keynote 02: Triple Play: AI, MV, and HTC

Prof A. Koucheryavy
Chaired Professor, Telecommunication Network and Data Transmission Department
The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, Russia

ABSTRACT

At the beginning of the 21st century, with the convergence of communication networks, the term Triple Play, which included voice, video, and data services, became a popular term for the provision of services by next-generation networks. At that time, this approach meant the unity of the convergent network and the universality of the services provided to the user.Now, about a quarter of a century after that time, the development of communication networks is taking place in the direction of creating various kinds of universes based on artificial intelligence technologies. Even urban universes have appeared as a development of smart cities. At the same time, the design Artificial Intelligence (AI) plus Metaverse (MV) was considered quite popular for some time. At the global standardization symposium in New Delhi in October 2024 during WTSA-24, it was proposed to supplement this dual with a holographic type of communications (HTC). This was necessary because there is a fairly widespread opinion that society will become holographic by 2030. This proposal was made by the ITU-T Focus Group Leader Richard Lee in 2018 at the first meeting of the focus group. Considering the great potential of holographic universes and the development of technologies in the field of providing holographic telepresence services, it seems useful to use the triune design of AI + MV + HTC at present. The report further discusses the development of artificial intelligence in the field of communication networks, the transformation of metauniverses into holographic network universes (HolNetVerse), work in the field of creating various terminals for holographic interactions and implementing telepresence services, the creation of holographic universities, the creation of holographic cities using the example of St. Petersburg, the development of a telepresence suit with parametric feedback, the capabilities of such telepresence suits together with a network universe for remote rehabilitation of patients using the example of recovery for children with upper limb injuries. All these applications show how effective the use of the new Triple Play AI + MV + HTC can be.

BIOGRAPHY



After graduating from Leningrad University of Telecommunications in 1974, A. Koucheryavy joined Telecommunication Research Institute LONIIS, where he worked till October 2003 (from 1986 to 2003 as the First Deputy Director). Dr. A. Koucheryavy holds Professor position at the Bonch-Bruevich St. Petersburg State University of Telecommunications (SUT) since 1998. There, in 2011 he became a Chaired Professor in “Telecommunication Networks and data transmission” department. Dr. A.Koucheryavy was an advisor of the Central Science Research Telecommunication Institute (ZNIIS) from 2003 to 2010. Co-founder of the International Teletraffic Seminar (1993, 1995, 1998, 2002); founder of the model network for digital networks at LONIIS (1997); co-founder of the model network for packet networks at ZNIIS (2004); co-founder of the Internet of Things Laboratory (2012) and Quality of Experience and IPTV Laboratory (2014) at SUT. Chair of the Scientific school on teletraffic theory in LONIIS (1990 – 2003); Founder and scientific school chair “Internet of Things and self-organizing networks” in SUT (2010 up to now); Steering committee member of IEEE technically co-sponsored series of conferences ICACT, NEW2AN and ICFNDS, Chair of ICACT 2020 International Steering Committee. SG11 ITU-T vice-chairman 2005 – 2008, 2009 – 2012. WP3/WP4 SG11 chairman 2006 – 2012, WP4 SG11 vice-chairman 2015-2016, Chairman of SG11 from 2016 up to March 2022. Co-founder of International Testing Center for new telecommunications technologies at ZNIIS under ITU-D competence. Host and technical program committee member of the “Kaleidoscope 2014” at SUT. Founder of the model network for telepresence services in SUT (2021). Honorary member of Popov’s society (2002).

SLIIT International Conference on Engineering and Technology
Malabe, Sri Lanka, September 2025

A Numerical Investigation of the Potential of Dimpled Surface Configurations to Improve Aerodynamic and Aeroacoustic Performance of Airfoils

Nisal Nelaka Fernando[1, a *], Indrajith Nissanka[2, b] and Nalaka Samaraweera[3, c]

[1,2,3]Department of Mechanical Engineering, Faculty of Engineering, University of Moratuwa, Moratuwa 10400, Sri Lanka

[a]fernandoinn.24@uom.lk, [b]nissankai@uom.lk, [c]nalakas@uom.lk

KEYWORDS: *Dimpled Airfoil; Aerodynamics, Aeroacoustics, Cfd, Naca 0012, Detached Eddy Simulations (Des), Passive Flow Control*

ABSTRACT

This study investigated the potential of dimpled surface configurations to enhance the aerodynamic and aeroacoustic performance of airfoils. Computational Fluid Dynamics (CFD) simulations were carried out on a NACA 0012 airfoil featuring surface dimples, under flow conditions relevant to low-speed aerodynamic applications such as unmanned aerial vehicles (UAVs), light aircraft, and small-scale wind turbines. The simulations were conducted at a Reynolds number of 700,000 and a Mach number of 0.21, representing typical subsonic operating conditions. Two angle of attack, 5° and 10°, were examined to represent attached flow and near-stall behavior, respectively. Aerodynamic performance was evaluated through lift and drag coefficients, while aeroacoustic characteristics were analyzed using Overall Sound Pressure Level (OASPL) with directivity plots and frequency spectrum analysis based on the Ffowcs Williams–Hawkings (FW-H) acoustic analogy. Key findings indicate that the dimpled configuration enhances flow behavior by increasing lift and reducing drag at a 10° Angle of Attack (AoA), primarily through delayed separation and modified stall onset characteristics. Aeroacoustic analysis showed a noise reduction of 2–7 dB at various receiver positions at a 10° AoA, with reductions varying by observer angle and frequency, confirming the directional sensitivity of noise emissions. These insights contribute to the understanding of passive flow control mechanisms and their dual impact on aerodynamic performance and noise reduction in airfoil design

INTRODUCTION

Aeroacoustic noise generated by turbulent fluid flow and unsteady aerodynamic forces, has become a growing concern in many low-speed aerodynamic applications such as small-scale unmanned aerial vehicles (UAVs), wind turbines, and industrial fans. This form of noise not only impacts human health and comfort but also affects the surrounding environment by interfering with wildlife communication and behavior [13]. Regulatory bodies such as the Federal Aviation Administration (FAA) and the Environmental Protection Agency (EPA) have responded by implementing strict noise limitations often below 65 dB for residential zoning near airfields [9] and 52 dB for wind energy installations, depending on land use and frequency [6]. As a result, the demand for passive noise reduction technologies has increased significantly.

Simultaneously, maintaining or improving aerodynamic performance remains essential, as any noise reduction strategy that leads to increased drag or reduced lift may offset its practical value. An increase in drag leads to higher energy consumption, while reduced lift can compromise efficiency or system capability. Therefore, there is a strong need for passive flow control techniques that can address both aerodynamic and acoustic concerns without introducing complex structural changes.

Dimples, which are shallow, concave surface features, have emerged as a promising passive flow control strategy in airfoils. By promoting localized turbulence, dimples help re-energize the boundary layer, potentially delaying flow separation and enhancing lift during low-speed operation. Unlike traditional vortex generators, which often involve protruding elements that can increase drag at higher speeds and complicate the surface structure [11,12], dimples provide a flush configuration that is easier to manufacture and maintain. However, the turbulent structures induced by dimples can also influence the aeroacoustic characteristics of the surface, either positively [18] or negatively, depending on their configuration and flow conditions [19]

While many studies have explored the aerodynamic effects of dimples mainly on flat plates at low Reynolds numbers [7,14,16], very few have investigated their combined aerodynamic and aeroacoustic performance on airfoils at moderate Reynolds numbers suitable for small UAVs, Wind Turbines and small Aircraft, Ananthan et al. (2022) [2] investigated a shallow dimple configuration at zero angle of attack and reported a drag increase of 1.9% with no detailed acoustic analysis. To the best of the author's knowledge, no comprehensive study has been conducted to examine how dimple geometry influences both lift–drag characteristics and noise generation across multiple angles of attack at Reynolds numbers relevant to real-world applications.

This study aims to investigate the influence of dimpled surface configurations on both aerodynamic and aeroacoustic performance of airfoils, focusing on low-speed applications such as UAVs, small wind turbines, and light aircraft operating at a Reynolds number of approximately 700,000. The objectives are to evaluate the lift–drag characteristics, analyze the noise reduction potential, and correlate flow structures with changes in aerodynamic and aeroacoustic performance. To achieve these, Computational Fluid Dynamics (CFD) simulations were conducted on a selected airfoil profile with dimple arrangements and different angles of attack. The results provide insights into the dual impact of dimpled surfaces on performance and noise, supporting their feasibility as a passive flow control method for broader low-speed aerodynamic applications.

METHODOLOGY

Geometry and Dimple Configuration

The NACA 0012 airfoil was selected for this investigation due to the extensive availability of high-quality experimental aerodynamic and aeroacoustic data at Reynolds numbers around 700,000, its symmetrical geometry, and its widespread use as a benchmark in CFD-based aerodynamic and aeroacoustic studies. The airfoil had a chord length of 150 mm and a span of 450 mm (3c), forming a quasi-2D configuration to isolate mid-span aerodynamic and aeroacoustic behavior while minimizing computational cost. This approach aligns with prior studies, which identify turbulent boundary layer trailing-edge noise and separation-induced stall noise at the mid-span as dominant contributors for low-speed airfoils (Figure 1(a)).

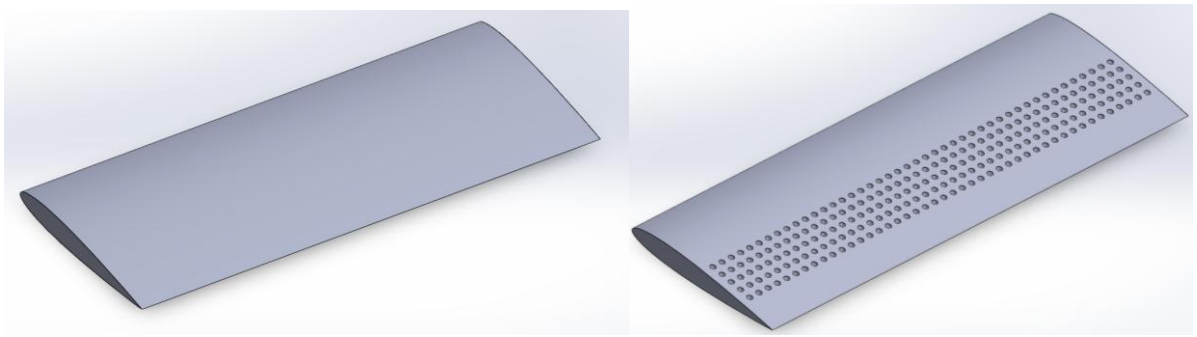
A semi-spherical dimple configuration was introduced on the suction side of the airfoil as shown in the Figure 1(b). Each dimple had a diameter of 5 mm, equivalent to 3.3% of the chord, which falls within the range (2%–7%) found effective in prior aerodynamic studies [4,5]. The dimples were arranged starting at 53% of the chord and continued up to 80%, distributed in five longitudinal rows to influence boundary layer behavior beyond mid-chord an area prone to early flow separation at higher angle of attack.

The use of semi-spherical dimples was preferred over hexagonal or rectangular geometries because of their balance between drag performance and noise reduction. While hexagonal dimples offered slightly better drag reduction (~1.2%), semi-spherical ones consistently demonstrated superior aeroacoustic behavior, particularly in low to moderate Reynolds number regimes [3,17]. Moreover, rectangular dimple patterns were found to generate more turbulent structures and less stable flow patterns [8], making them less suitable for passive noise control strategies.

This configuration aimed to improve aerodynamic performance by maintaining attached flow and simultaneously reducing turbulent noise without requiring significant structural modification making it practical for aerospace and wind energy applications where minimal geometric alteration is preferred.

Flow and Boundary Conditions

The simulations were conducted at a freestream Mach number of 0.21, corresponding to a chord-based Reynolds number of 700,000, calculated using a standard speed of sound of 340 m/s. This Reynolds number was selected as it falls within the typical mid-span operating range of UAVs, small wind turbines, and light aircraft. Two angles of attack were considered: 5° to represent attached-flow conditions, and 10° to represent pre-stall conditions. These choices align with aerodynamic performance regimes commonly studied for such low-speed applications



(a) (b)

Figure 1. NACA 0012 Computer Aided Model (a)Clean Wing (b)Dimpled Wing

Computational Domain and Meshing

A quasi-2D computational domain was used, extending 6c upstream and 12c downstream of the airfoil. A polyhedral mesh was adopted to accurately capture surface curvature, with local refinement applied in the dimple regions. Near-wall prism layers were included to maintain y+<1 around the dimples, ensuring adequate resolution of the boundary layer. Each 5 mm dimple was resolved with at least 20 elements across its diameter to capture the relevant flow features. A mesh sensitivity study was conducted to establish the optimum mesh density for both aerodynamic and aeroacoustic predictions. As shown in Figure 2, no significant differences in Cl and Cd were observed beyond 1.4 million elements. Therefore, a final mesh with approximately 1.4 million polyhedral elements was used, as depicted in Figure 3.

A velocity inlet was defined with a magnitude of 71.3 m/s to match the desired Mach number. The outlet was set as a pressure outlet with ambient pressure conditions. The airfoil surface was treated as a no-slip wall, while the remaining domain boundaries were specified as symmetry planes to reduce computational cost and simulate a quasi-2D flow environment.

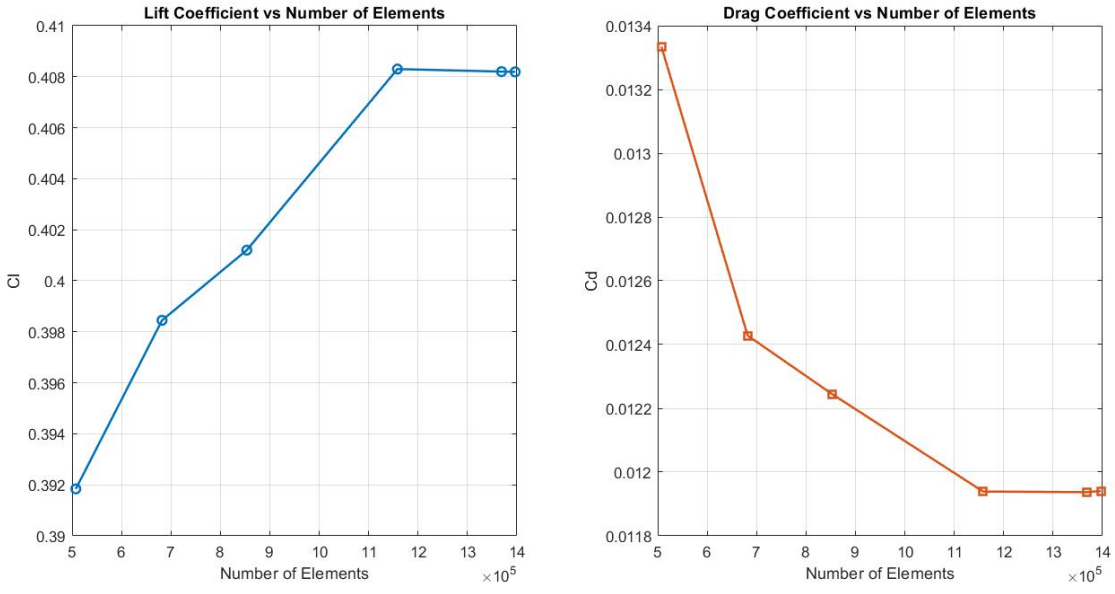


Figure 2. Mesh Independence Study

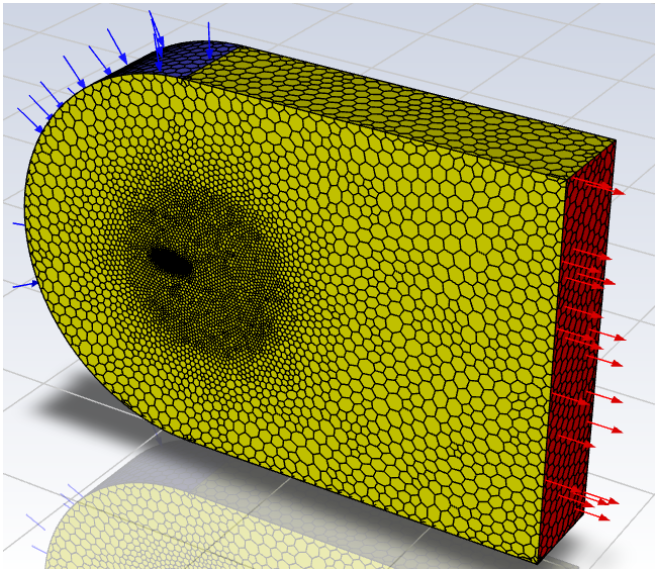


Figure 3. Computational Mesh with Polygonal Elements

Turbulence Model and Acoustic Model

A hybrid turbulence modeling approach was adopted using the Delayed Detached Eddy Simulation (DDES) model based on the SST k-ω formulation, offering improved accuracy in capturing unsteady flow structures critical for aeroacoustic.

For the aeroacoustic analysis, the FW-H equation was employed to predict sound propagation in the far field. The FW-H formulation is an extension of Lighthill's acoustic analogy and is well-suited for compressible, turbulent flows involving solid boundaries. It allows for the calculation of noise generated by both volume sources (turbulence) and surface sources (airfoil motion and pressure fluctuations) [10]. This method is widely followed for aeroacoustic studies due to its ability to decouple the sound generation mechanisms from the propagation effects, enabling efficient post-processing of CFD results to obtain far-field acoustic metrics.

In this study, the permeable FW-H formulation was used, allowing integration over a virtual surface surrounding the airfoil. This approach is particularly advantageous when simulating unsteady flows using hybrid turbulence models such as DDES, as it captures transient pressure fluctuations that are directly linked to acoustic emission mechanisms.

Aerodynamic and Acoustic Validation

To assess the accuracy of the CFD model, the aerodynamic results were validated against experimental data from Sheldahl and Klimas (1981) [15], who tested the NACA 0012 airfoil at a Reynolds number of 700,000. The comparison was carried out for lift and drag coefficients across varying angle of attack. The simulated Cl and Cd values showed strong agreement with the experimental trends, particularly in the linear region. A slight deviation observed around an angle of attack of 10°, which is near the stall onset, is attributed to the simulation approach used at this condition. At higher angle of attack, the flow becomes unsteady due to the formation of vortex bubbles, which causes fluctuations in Cl and Cd. To obtain stable, time-averaged aerodynamic coefficients for comparison with experimental data, steady-state RANS (Reynolds-Averaged Navier–Stokes) simulations were used at these angles. In contrast, Detached Eddy Simulation (DES) was employed in the rest of the study to capture unsteady flow behavior and perform aeroacoustic analysis. This difference in modeling approach may have contributed to the localized discrepancy seen in the aerodynamic validation.

To validate the acoustic model, the computed SPL spectrum at Re = 700,000 for the NACA 0012 airfoil was compared with experimental data from Al Tlua et al. (2021) at Re = 500,000. Despite differences in Reynolds number, chord length, and receiver location, the comparison focuses on overall spectral trends rather than absolute SPL values (Figure 5). Both datasets show a similar broadband decay in SPL with increasing frequency and lack tonal peaks, indicating the simulation captures key noise-generation mechanisms. Prior studies [20,21] confirm that broadband trailing-edge noise spectra remain consistent across moderate Reynolds number changes under low-Mach conditions. Thus, the qualitative agreement supports the DES–FW-H approach as a reliable method for

predicting airfoil noise in such regimes.

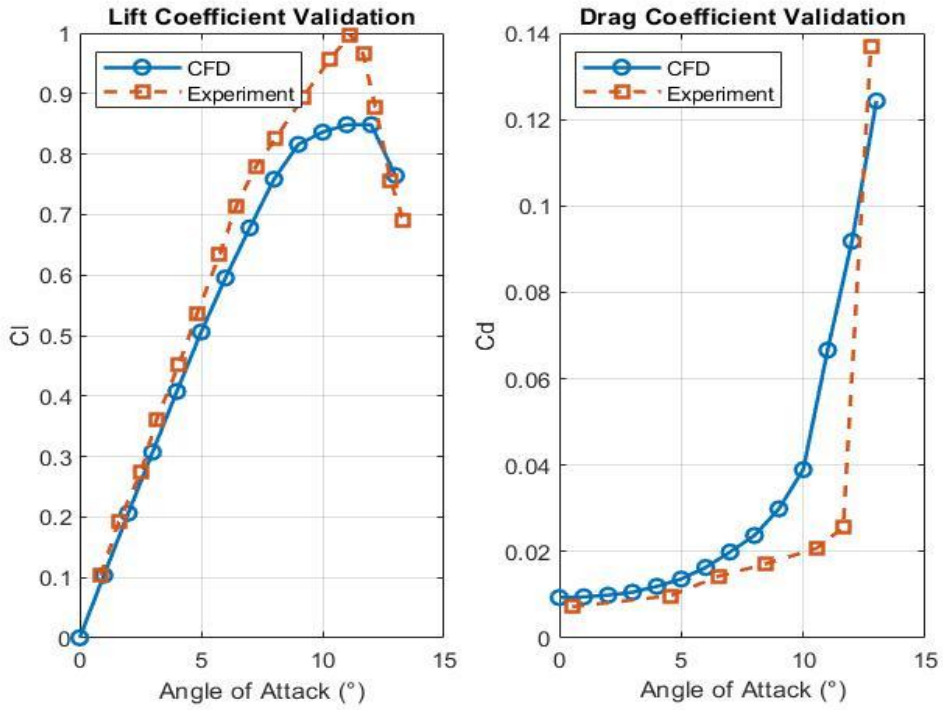


Figure 4. Aerodynamic Validation Cl vs AoA and Cd vs AoA curves

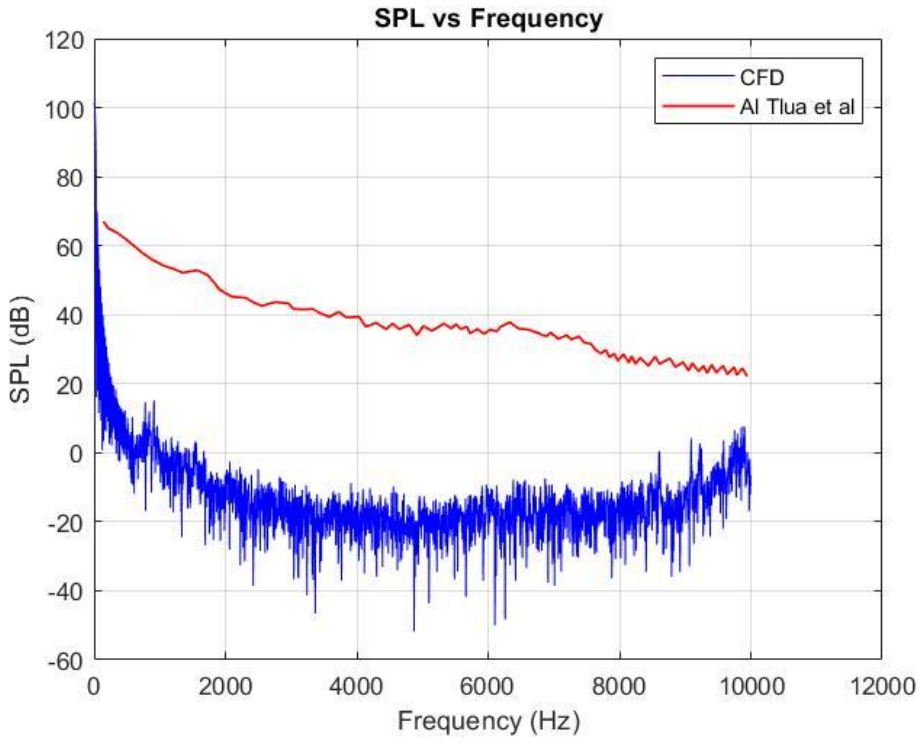


Figure 5. Acoustic Validation: SPL vs Frequency of CFD and Experimental Study

RESULTS AND DISCUSSION

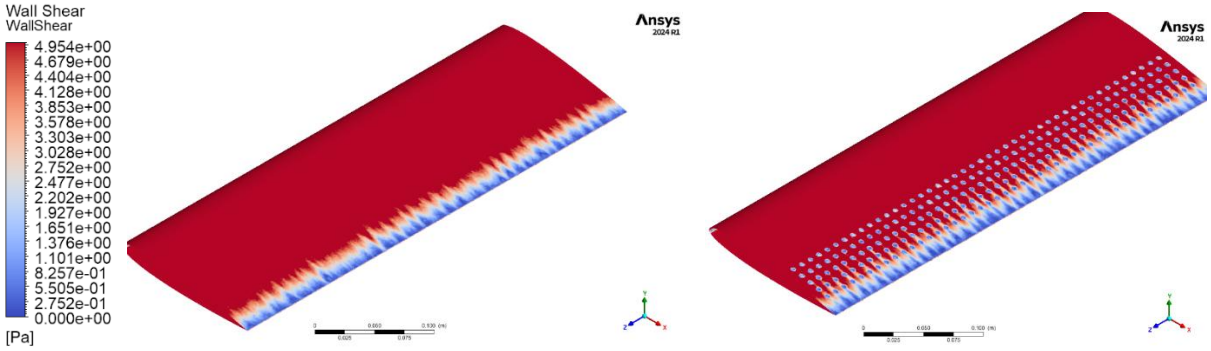
Aerodynamic Performance

For aerodynamic performance evaluation, lift coefficient (Cl), drag coefficient (Cd), and lift-to-drag ratio (Cl/Cd) were extracted for both the clean and dimpled NACA 0012 airfoils at angle of attack 5° and 10°. Velocity and pressure contours were also generated, however no significant differences between the configurations were observed. Therefore, wall shear stress plots and Q-criterion iso-surfaces were presented to visualize the near-wall

behavior and vortex structures in the simulated scenario.

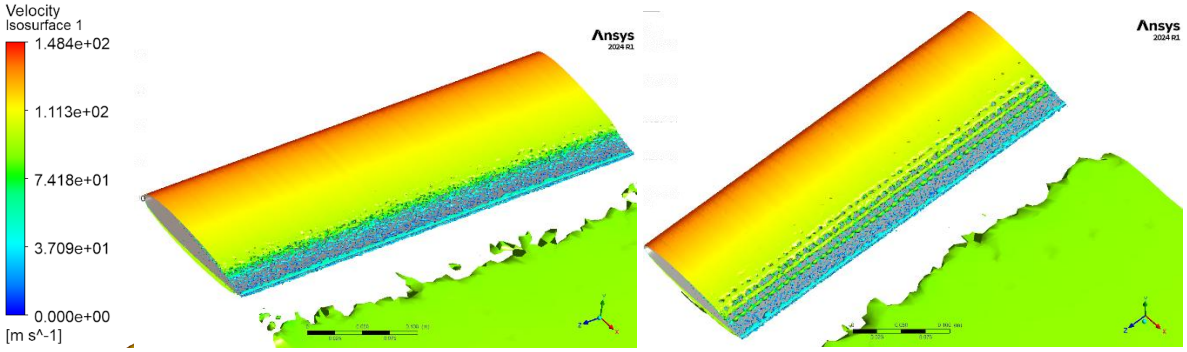
Table 1. Cl, Cd, Cl/Cd vs AoA of Clean and Dimpled Wings

	AoA (degrees)	**Cl**	**Cd**	**Cl/Cd**
Clean Wing	5	0.49	0.0174	28.13
Dimpled Wing	5	0.48	0.0178	26.79
Clean Wing	10	0.84	0.0489	17.07
Dimpled Wing	10	0.86	0.0415	20.74



(a) (b)

Figure 6. Wall Shear at AoA $\mathbf{5}^0$ (a) Clean Wing (b) Dimpled Wing



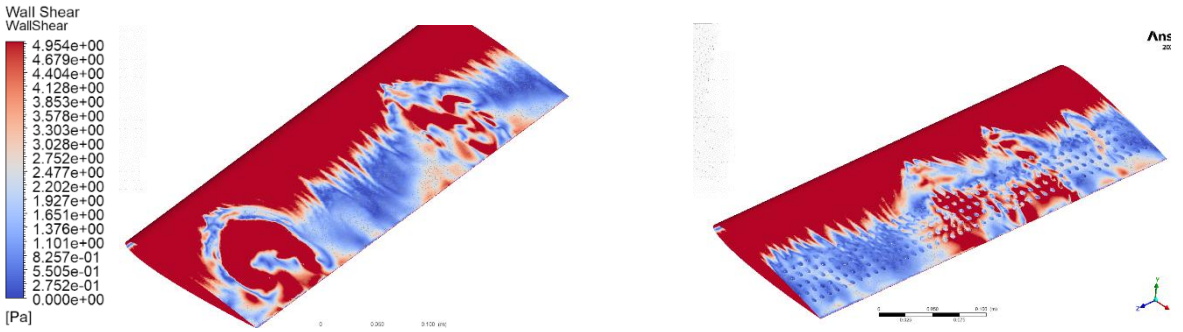
(a) (b)

Figure 7. Q Criterion Iso-surface at AoA $\mathbf{5}^0$ (a) Clean Wing (b) Dimpled Wing

At an angle of attack of 5°, the clean wing produced a lift coefficient (Cl) of 0.49 and a drag coefficient (Cd) of 0.0175, resulting in a lift-to-drag ratio (Cl/Cd) of 28.00. In comparison, the dimpled wing generated a slightly lower Cl of 0.48, reflecting a 2.04% reduction in lift, and a marginally higher Cd of 0.0179, indicating a 2.29% increase in drag. These changes led to a reduction in aerodynamic efficiency, with the Cl/Cd ratio decreasing by 4.21% to a value of 26.82.

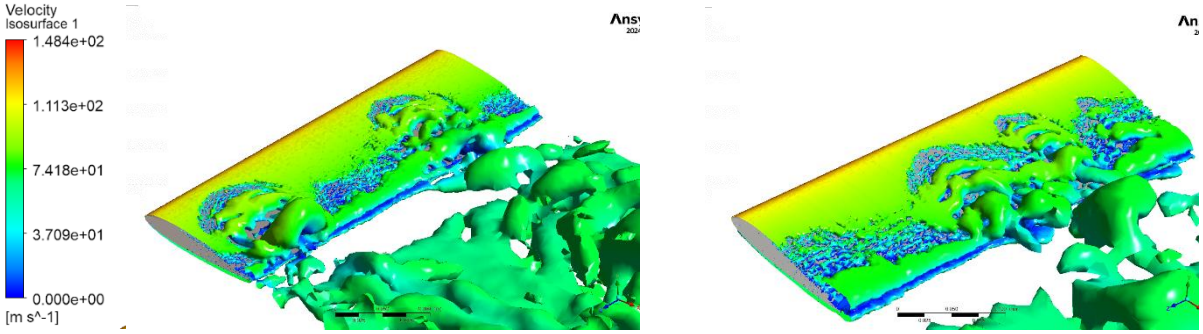
This reduction in aerodynamic efficiency for the dimpled wing at low AoA can be explained by the changes in near wall flow field conditions. Wall shear stress (Figure 6 (a) and (b)) and Q-criterion plots (Figure 7 (a) and (b)) revealed that the introduction of dimples triggered early transition of the boundary layer, turning the otherwise laminar attached flow into a turbulent flow. This artificially introduced turbulence increases skin

friction drag and disrupts the smooth pressure recovery over the airfoil surface. Consequently, although the flow remains attached, the aerodynamic performance deteriorates due to premature degradation of the boundary layer's favorable characteristics, even at a relatively low angle of attack.



(a) (b)

Figure 8. Wall Shear at AoA $\mathbf{10}^0$ (a) Clean Wing (b) Dimpled Wing



(a) (b)

Figure 9. Q Criterion Iso-surface at AoA $\mathbf{10}^0$ (a) Clean Wing (b) Dimpled Wing

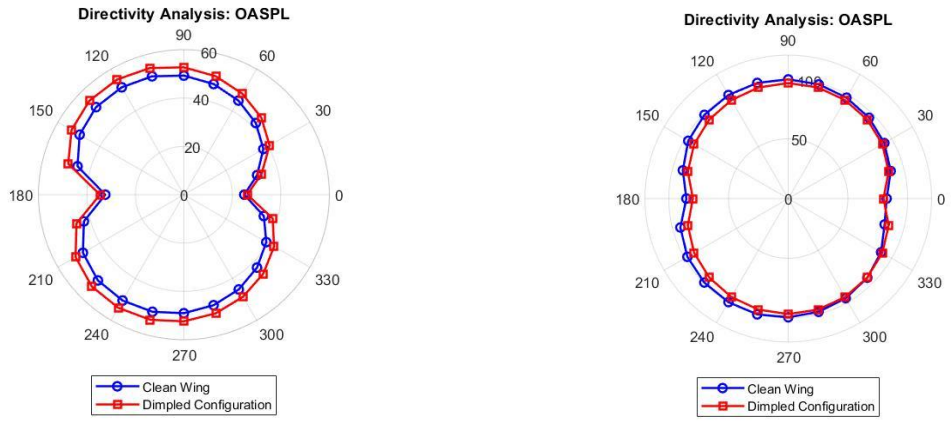
At an angle of attack of 10°, the clean wing produced a lift coefficient (Cl) of 0.84 and a drag coefficient (Cd) of 0.0489, resulting in a lift-to-drag ratio (Cl/Cd) of 17.17. In contrast, the dimpled configuration achieved a higher Cl of 0.86, reflecting a 2.38% increase in lift, and a lower Cd of 0.0415, indicating a 15.14% reduction in drag. These improvements enhanced the aerodynamic efficiency, increasing the Cl/Cd ratio by 20.66% to a value of 20.72.

This performance gain can be attributed to the effect of dimples in energizing the boundary layer. The flow field visualizations using Q-criterion (Figure 9 (a) and (b)) revealed increased vortex activity in the turbulent boundary layer over the dimpled surface. These streamwise vortices introduced by the dimples enhanced momentum exchange, allowing the flow to remain attached for a longer portion of the chord. As a result, both lift enhancement and drag reduction were observed at this higher angle of attack.

Aerocoustic Performance

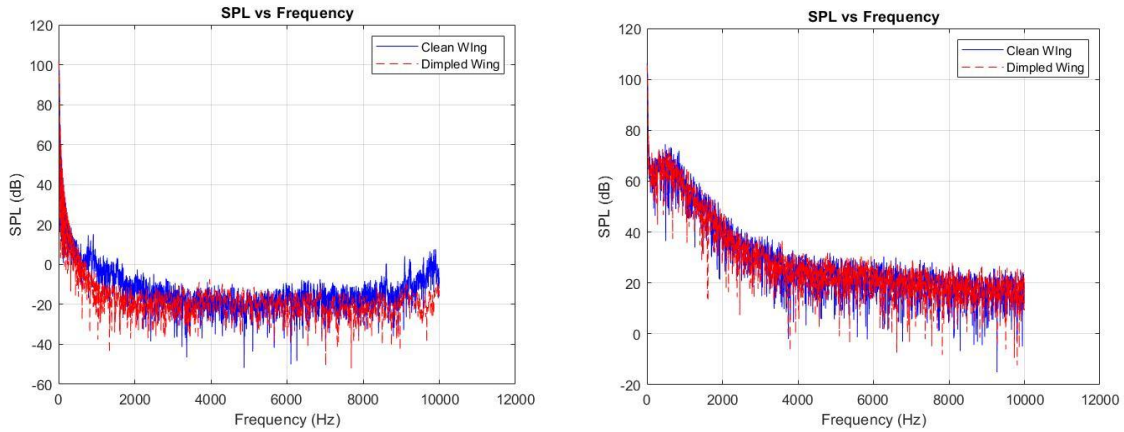
To evaluate the acoustic performance, a directivity analysis was conducted using a circular array of receivers positioned at a radius of 0.5 m from the trailing edge, located on the mid-span plane (half-chord spanwise location). Receivers were placed at 15° intervals around the full 360° circle for both 5° and 10° angle of attack. The Overall Sound Pressure Level (OASPL) was recorded at each receiver location and presented in the form of polar plots to visualize the noise directivity pattern for each case.

In addition to the directivity analysis, spectral analysis was performed at a representative receiver location. Sound Pressure Level (SPL) versus frequency plots were generated for both the clean and dimpled wing configurations at AoA = 5° and AoA = 10°, allowing comparison of the tonal and broadband noise characteristics resulting from the surface modifications.



(a) (b)

Figure 10. Directivity Analysis for (a) AoA 5° (b) AoA 10°



(a) (b)

Figure 11. Frequency Spectral Analysis (a) AoA 5° (b) AoA 10°

At an angle of attack of 5°, the dimpled wing configuration exhibited a slight increase in OASPL compared to the clean wing, with differences of approximately 1–2 dB observed around the receiver circle as shown in Figure 10(a). This increase is attributed to the additional turbulence and small-scale vortical structures introduced by the dimples, as evident from the flow field.

Spectral analysis revealed that the rise in SPL was primarily shifted to the 4000Hz-7000hz frequency range as illustrated in Figure 11(a). These high-frequency components had very low amplitudes and are expected to dissipate quickly in the far field, contributing minimally to perceived noise levels. The low-frequency region of the spectrum, which has a greater impact on far-field acoustic propagation, remained largely unaffected. In fact, certain frequencies like 1000Hz- 3000Hz showed slight reductions in SPL, indicating that the dimpled surface did not significantly degrade the overall acoustic behavior at this low angle of attack.

At an angle of attack of 10°, the dimpled configuration demonstrated a clear improvement in acoustic performance. The OASPL directivity plots showed a reduction of approximately 2–6 dB across most receiver positions compared to the clean wing (Figure 10(b)). This indicates a significant decrease in overall radiated noise due to the presence of dimples.

From a flow physics perspective, this reduction can be attributed to the influence of dimples on the boundary layer behavior under near-stall conditions. The dimples controlled streamwise vortices and enhance turbulent mixing within the boundary layer, which delays separation and stabilizes the flow over the trailing edge region. As a result, the intensity and irregularity of large-scale turbulent structures which are major sources of broadband noise are reduced, leading to lower sound emission.

In the frequency spectrum, no major shift in the trend or dominant frequency components was observed.

However, a general reduction in SPL magnitude across the entire spectrum was noted for the dimpled configuration as can be seen from Figure 11(b). This suggests that while the noise generation mechanisms remained similar, the overall acoustic energy radiated by the airfoil was reduced due to the modified and more organized flow field enabled by the surface dimples.

CONCLUSION

This study investigated the potential of dimpled surface configurations to enhance the aerodynamic and aeroacoustic performance of airfoils under low-speed flow conditions relevant to applications such as UAVs and small-scale wind turbines. Using high-fidelity CFD simulations with a DES turbulence model and the Ffowcs Williams–Hawkings acoustic analogy, the NACA 0012 airfoil was evaluated in both clean and dimpled configurations at angle of attack 5° and 10°.

Aerodynamically, the dimpled surface showed a minor performance penalty at 5° due to early boundary layer transition, increasing drag and reducing lift slightly. However, at 10°, the dimples delayed separation by energizing the boundary layer through enhanced vortex activity, resulting in both lift improvement and drag reduction. This led to a higher aerodynamic efficiency (Cl/Cd) compared to the clean configuration.

From an aeroacoustic perspective, the dimpled surface led to a slight increase in high-frequency noise at 5°, but with low SPL magnitudes that are expected to decay quickly in the far field. At 10°, a substantial reduction in OASPL ranging from 2–6 dB was observed in the directivity pattern, along with a general decrease in SPL across the spectrum. This reduction in radiated noise is attributed to the dimples' ability to suppress large-scale turbulent structures near the trailing edge under near-stall conditions.

Overall, the results demonstrate that properly configured dimples can offer dual benefits: improved aerodynamic performance at higher angle of attack and notable noise reduction, without requiring major structural changes. These findings highlight the potential of passive dimpled surface treatments as an effective design strategy for low-noise, high-lift airfoil applications.

ACKNOWLEDGEMENT

This project was funded by Senate Research Committee (SRC) grant of University of Moratuwa, Sri Lanka (Grant no: SRC/LT/2021/26)

REFERENCES

[1] B. Al Tlua, J. Rocha, Experimental investigation of NACA-0012 airfoil instability noise with sawtooth trailing edges, Wind Eng. (2021). https://doi.org/10.1177/0309524x211060548

[2] S. Ananthan, J. Sitaraman, N. Hariharan, Effect of shallow dimples on the aerodynamic performance of a symmetric airfoil, Aerospace 9 (2022) 1–15.

[3] J. Chen, L. Wu, T. Zhang, Aeroacoustic performance of various surface dimple configurations, AIAA J. 57 (2019) 4560–4572. https://doi.org/10.2514/1.J058070

[4] H. Choi, W.-P. Jeon, J. Kim, Control of flow over a bluff body, Annu. Rev. Fluid Mech. 40 (2008) 113–139. https://doi.org/10.1146/annurev.fluid.39.050905.110215

[5] Y. Choi, K. Kwon, D. Lee, Aerodynamic performance of airfoils with surface dimples at low Reynolds numbers, AIAA J. 52 (2014) 1212–1223. https://doi.org/10.2514/1.J052571

[6] Environmental Protection Agency, Guidelines for wind energy noise, 2020.

[7] B. Etter, Numerical investigation of dimpled surfaces for drag reduction, Master's thesis, University of Illinois, 2007.

[8] C. L. Etter, Comparative study of spherical and non-spherical dimples for flow control, Experiments in Fluids 61 (2020). https://doi.org/10.1007/s00348-020-02994-3

[9] Federal Aviation Administration, Noise exposure map guidelines, 2023.

[10] J. E. Ffowcs Williams, D. L. Hawkings, Sound generation by turbulence and surfaces in arbitrary motion, Philos. Trans.

R. Soc. Lond. A 264 (1969) 321–342. https://doi.org/10.1098/rsta.1969.0031

[11] M. Fouatih, K. Khellil, F. Chabane, Numerical investigation of vortex generators in enhancing airfoil performance, Renew. Energy 92 (2016) 335–346.

[12] M. S. Genç, U. Kaynak, Y. S. Uslu, Aerodynamic performance of passive flow control devices on airfoils, Aerospace Sci. Technol. 98 (2020) 105679.

[13] C. R. Kight, J. P. Swaddle, How and why environmental noise impacts animals: An integrative, mechanistic review, Ecol. Lett. 14 (2011) 1052–1061.

[14] S. Panda, H. V. Warrior, Flow control using shallow dimples at low Reynolds number, J. Fluids Eng. 143 (2021) 031202.

[15] R. E. Sheldahl, P. C. Klimas, Aerodynamic characteristics of seven symmetrical airfoil sections through 180-degree angle of attack for use in aerodynamic analysis of vertical axis wind turbines, 1981. https://doi.org/10.2172/6548367

[16] P. R. Spalart, S. Deck, M. L. Shur, K. D. Squires, M. K. Strelets, A. Travin, A new version of detached-eddy simulation, resistant to ambiguous grid densities, Theor. Comput. Fluid Dyn. 20 (2019) 181–195.

[17] L. Van Campenhout, J. Lefebvre, M. Brossard, Experimental investigation of dimple geometries on airfoil aerodynamics, J. Fluids Eng. 143 (2021). https://doi.org/10.1115/1.4048536

[18] S. Kumar, S. K. Singh, S. Jha, K. Baskaran, K. Srinivasan, S. Narayanan, On the reductions of airfoil broadband noise through circular dimples, Appl. Acoust. 217 (2024) 109819. https://doi.org/10.1016/j.apacoust.2023.109819

[19] C. L. Perry, Mitigation of Aeroacoustic Noise of a Fixed Wing Using Passive Flow Control, Georgia Southern Commons, 2023. https://digitalcommons.georgiasouthern.edu/etd/2632

[20] D. Moreau, L. Brooks, C. Doolan, Experimental investigation of flat plate self-noise reduction usingtrailing edge serrations, 2012. https://www.icas.org/icas_archive/ICAS2012/PAPERS/221.PDF

[21] T. Brooks, D. Stuart, M. Marcolini, NASA Reference Publication 1218 Airfoil Self-Noise and Prediction, 1989. https://ntrs.nasa.gov/api/citations/19890016302/downloads/19890016302.pdf

Evaluation of the Knowledge Base in Agriculture and Food to Reduce and Prevent Chronic Kidney Disease of Unknown Etiology (CKDu)

R.T.K. Ariyawansha[1a*], B.F.A. Basnayake[2b], D.A.N. Dharmasena[2c], Ashoka Gamage[3d]

[1]Department of Environmental Technology, Faculty of Technology, Sri Lanka Technology Campus, Meepe, Padukka 10500, Sri Lanka
[2]Department of Agricultural Engineering, Faculty of Agriculture, University of Peradeniya, Peradeniya 20400, Sri Lanka
[3]School of Chemical, Material and Biomedical Engineering, College of Engineering, Riverbend Road, Athens, GA 30605, USA
[a*]renukaa@sltc.ac.lk, [b]benb@pdn.ac.lk, [c]dand@agri.pdn.ac.lk, [d]ashogamage@gmail.com

ABSTRACT

Kidney disease is a growing global problem, more so in tropical regions. The cause of CKDu is multifactorial and influenced by heavy metal (HM) contamination, inhibiting essential enzymatic reactions. Fertilizers and water contamination are believed to cause the disease. This study aimed to review the existing knowledge base, focusing on a transitional approach to advanced technologies with the least HMs and to use justifiable scientific reasoning supported by published data, to used to demonstrate the movement of Cadmium (Cd) at both low and high concentrations from applied fertilizer through the soil to grain and rice. The quantity of fertilizer applied per ha with the given Cd levels was equated to Cd concentrations in the harvested grain and rice per ha, considering positive or negative contributions from the soil. Weekly consumption levels of rice at the threshold limits by an average Sri Lankan were determined for low and high Cd levels in rice using the tolerance limits of two international standards. It is best to characterize watersheds and determine the movement of nutrients and HM in ferruginous soils. Hinderance to phosphate immobility in these soils can be overcome by applying biochar biofertilizer with possible enrichment of biofilm biofertilizers to replace totally inorganic fertilizers contaminated with HMs. Cd levels of 836.25 and 393.75 of the two publications equate to the assumed harvest: lowest 21.22, average 385.13, and the highest 1246.10 mg Cd ha^{-1}. Allowable standards indicate that the weekly limit of a Sri Lankan to consume rice is 300 g, containing a high concentration of 0.2618 mg Cd kg^{-1} and 1kg or 604 g, having 0.1339 mg Cd kg^{-1} for an average harvest of 4350 $kg.ha^{-1}$. Water contains HM, particularly arsenic from fertilizer and pesticides. Recommended researching while implementing phytoremediation, mechanized farming, preventing UVB, Integrated Pest Management (IPM), and organic agriculture with supporting technologies of watershed resource management.

KEYWORDS: *Biochar, Biofertilizer, Biofilm, Cadmium, CKDu, Heavy metals, Organic pollutants, Soil*

INTRODUCTION

There are many studies undertaken to find the origin or cause of worldwide CKDu occurrences. It surfaced in 1990 in the North Central Province (NCP) of the island [1], which is situated in the dry zone of the country. It was subsequently spread to the adjacent North-Western, Eastern, and Uva provinces. The endemic region covers approximately 24,000 km2, providing shelter for about 3 million people. Out of that, about 2 million are at high risk of contracting CKDu [2]. It is now recognized that the disease cause is multifactorial, because it may seem plausible with the number of publications on heavy metal (HM) contaminants, medical health issues, dehydration, and heat stress like non-communicable diseases, social behaviour, and poor-quality food and water intake [2]. Dehydration leading to chronic heat stress is now believed to be the triggering factor [3]. The focus in Central America, with less attention in Asia, is on dehydration and heat stress [2].

The studies were centered around the unknown ethnological aspect of the disease, except when the government decided to ban inorganic fertilizers [4]. Subsequently, the prohibition was lifted because it was not a planned implementation of an important policy decision. Intentionally or otherwise, it increased the costs of fertilizers. There were many protests for and against the ban, but there was a strong lobby opposing organic agriculture because it was necessary to safeguard the interests of the traders of inorganic fertilizers and pesticides. The belief that organic agriculture cannot compete with inorganic agriculture is strong and stern among academics, researchers, government officials, and private companies.

Such a notion is embedded in the Sri Lankan society and many parts of the world because, during the colonial period, traditional collective agricultural systems were often replaced with or significantly altered by systems controlled by colonial powers [5]. Over time, most developing economies became highly dependent on

imports for food substitutions, and most importantly, for agricultural inputs. The country became more reliant on plant breeding to increase inorganic fertilizer uptake under the banner of the Green Revolution. Production systems on specific crops like rice culture made it convenient to apply weedicides for minimizing traditional land preparations and replaced conventional soil nutrient management with inorganic fertilizers. Yield comparisons between traditional and genetically modified crop varieties, as well as between organic and inorganic fertilizers, are often used to justify the use of chemical farming. Farmers, too, prefer convenient agricultural practices, and it is a daunting task to reverse the process.

While HMs are known to be toxic and can hinder plant growth and reduce yields in high concentrations, some HMs, specifically in low concentrations, can act as micronutrients and contribute to increased crop yields. However, it's crucial to manage HM contamination in soil to prevent negative impacts on food production and human health [6]. HM contaminations are from industrial processing of inorganic or organic fertilizers, industrial wastewater, sewage sludge, weathering of soil minerals, and fallouts. Whatever the method of HM contamination, it should be prevented since almost all diseases, like CKDu stem from cadmium. If cadmium can be controlled, most of the other HMs would be evaded. Hence, this paper aimed to compile research demonstrating the movement of cadmium from fertilizers and other sources through the soil to rice consumed. And also, the paper aimed to discuss best management practices to support the development of an organic agriculture policy framework areas aimed at maintaining or increasing yields and perhaps particularly in poor soils, and increasing employment opportunities in affected areas. It is also useful for improving agricultural practices in the entire country because HM contamination is rising, and climate change issues are worsening.

MATERIALS AND METHODS

The study focused on the three aspects of enzyme activity, namely, ionic strength, pH, and temperature in relation to the mobility of minerals in the affected soils, because there are many misconceptions. Therefore, an analytical review was undertaken to identify some of the crucial scientific research frameworks needed to overcome health issues, improve agricultural practices, and social upliftment of farmers and their supporting communities. In undertaking the task, it is necessary to emphasize the existence of poor tropical soils in endemic locations. The influence of soil pH on the dynamics of soil organic matter (SOM) was analyzed to develop strategies for introducing improved cultural practices.

A direct relationship was developed to demonstrate the movement of Cd from fertilizer to rice from the available publications. From the publication of [7,8], Cd concentrations in fertilizer for: $[Cd]P_2O_5$ in mgCdkg^{-1} and $[Cd]Urea$ in mgCdkg^{-1}., were obtained separately for each publication. The recommended application rates for the phosphate and urea fertilizers were obtained from the Department of Agriculture, Sri Lanka. They are: RP_2O_5 in kgCdha^{-1} and $RUrea$ in kgCdha^{-1}. The Applied amount is:

$[Cd]P_2O_5 \times RP_2O_5 = x$ in mgCdha^{-1} and (1)

$[Cd]Urea \times RUrea = y$ in mgCdha^{-1}. (2)

The concentration in soil, $[Cd]soil = z = x + y$ in mgCdha^{-1}. (3)

There could be a residual concentration in the soil of α, thus, the total in soil,

$[Cd]soil = z + \alpha$ in mgCdha^{-1} (4)

If there is a loss of Cd from the soil of β, then the soil concentration,

$[Cd]soil = z - \beta$ in mgCdha^{-1}. (5)

It is assumed that the paddy plant root system covers the entire extent of one ha within the growing period. The amount of Cd taken out of the soil can be found in the rice. The values of Cd in rice given in the publication of [9] were used to determine the lowest, average and highest to compare with the field values. 68% was taken as the milling turnover (10). Although the Cd content is higher in husk and bran, the same concentrations of Cd in rice was used to compare with the field. In 2006, the Codex Alimentarius Commission established a maximum limit (ML) for cadmium in polished rice at 0.4 mgCdkg^{-1}. It is of high value. Japan supported the ML of 0.4 mg/kg because some rice samples in Japan contained relatively high levels of cadmium based on the high background levels of cadmium in soils in their country [11].

The concentration $[Cd]$ in rice varied between $[Cd]a$ to $[Cd]c$ and the average $[Cd]b$ in mgCdkg^{-1}. Four harvesting levels were considered, namely, $h_l = 3000$, $h_{h1} = 5200$, $h_{h2} = 7000$, $h_a = 4350$ in kg. The total quantities in rice and gain were determined for the lowest, low, high, highest and average, stated as:

Lowest $[Cd]rice = [Cd]a.h_{h1}$ (6)
Low $[Cd]rice = [Cd]c.h_l$ (7)
High $[Cd]rice = [Cd]c.h_{h1}$ (8)
Highest $[Cd]rice = [Cd]c.h_{h2}$ (9)
Average $[Cd]rice = [Cd]b.h_a$ (10)
Another set of values was obtained by multiplying by 68%.

Because of the toxicity of Cd, the tolerable limits were deducted based on the US(Food and Drug Administration FDA (2023) and the European Food Safety Authority (EFSA) (1975). These limits are based on body weight and duration of one week. Hence, the average body weight of women and men was obtained as 55kg. The allowed limit for a week was determined, A_L. Therefore, the quantity of rice consumption to exceed the limit $= \frac{A_L}{[Cd]b \text{ or } [Cd]c}$ in kg (11) for FDA A_L has a range, and the minimum and maximum values were used to determine consumption limits.

RESULTS AND DISCUSSION

Dynamics of Mineral Fluxes

The geological formations of soils differ from one location to the next, but the movement of mineral contents within the soil profile depends on the ecosystems and agroecosystems [12]. In dry highland conditions, the movement of minerals, either HMs or otherwise, would be downward, caused by deep percolation of precipitation. Vegetation intercepts these minerals and absorbs them as nutrients. A combination of shallow and deep rooting systems allows the best conditions for carbon and mineral cycling. Since minerals from lower soil profiles get transported to the topsoil through the growth and decay cycles of the plant/tree species. It is an important phenomenon because it is the most rapid process of forming soils suitable for cultivation. The upward movements of minerals occur with the rise in the water table during the monsoon periods and sometimes with irrigation water. Salt also moves up with the effects of dehydration at high ET coupled with the capillary rise of mineral water in the soil profiles [13]. In paddy cultivation, the minerals move in all directions, downwards as well as upwards, depending on the concentration and water movement. In dry seasons, farmers in paddy fields used surface tillage to create and hasten deep cracks in the soil. This practice helped to facilitate the washing away of excess minerals with the first onset of rain. It also improves soil structure and aeration, which is useful for preparing seedbeds and incorporating organic matter [14].

In traditional farming practices, the cultivated lands will have very similar mineral contents in the soil profiles to those of uncultivated lands. It is likely to experience less concentration of heavy metals in cultivated lands because of continuous removal over many years. It has perhaps been an important facet of humankind's survival. Over many years of cropping, there may be mineral deficiencies in the soil. Mineral balances are being replenished or affected when flooding takes place or is induced by irrigation water. The charging of topsoil with ionic compounds during flood flow has both positive effects as well as negative effects, like fallouts. Renewing the mineral requirements for crops is a positive aspect, but heavy metals can cause serious issues for generations. Recycling organic plant materials is important. It maintains the required mineral content in intensive cropping. Heavy metal contamination of soils from different sources causes many issues [15,16] in an otherwise difficult soil.

Sesquioxide is formed on the application and precipitation of phosphorus from the plant [17,18], and thus, is not available to the plant for growth and during growth. Sattell (1989), [19] in his findings of Sri Lankan Alfisols, was stated that oxalate extractable Fe was the soil component most active in phosphorus sorption. Ever since, there have been many studies. Some soils are more affected than others by phosphate immobility. It is reflected in paddy yields in most parts of the dry zone. Certain soil types, like Low Humic Gley soils, may have adequate soil phosphorus for rice cultivation, even without added fertilizers. However, other soils, especially those with acidic conditions, may have reduced phosphorus availability, leading to potential deficiencies and lower yields [20]. In a three-consecutive-season study, [21] concluded that the application of phosphorus fertilizer for rice can be terminated for at least 11½ years without affecting grain yield, no matter the level of grain yield, as P is available in the soils in adequate quantities in the Dry Zone of Sri Lanka. In the case of the potassium two-season trial, they found that there was no yield response of rice to added K up to the yield level of about 4 t/ha. Thus, the application of K fertilizer can be avoided over two consecutive seasons without affecting grain yield in rice in Low Humic Gley soils in the Dry Zone of Sri Lanka if the yield level is 4 t/ha or lower. In poor soil

conditions, where disease levels are low, the need for potassium might be somewhat less, but not necessarily zero. Poor soil can limit nutrient availability, including potassium, which can lead to deficiencies even if disease pressure is low. Potassium plays a crucial role in plant health, even in the absence of disease. It helps with water regulation, enzyme activation, and overall plant vigor, which are important for plant resilience.

The soluble inorganic phosphorus (P_i) available to plant roots typically represents only a small fraction of the total P present in the soil. Most P_i are immobile, being strongly sorbed onto soil particles and forming insoluble complexes with calcium, as well as oxides and hydroxides of iron or aluminum, processes that are strongly influenced by soil composition and pH. Phosphorus is one of the least mobile plant macronutrients in soil, being much less mobile than nitrate and potassium. A substantial portion of P in the soil is also present as organic phosphorous (P_o), such as nucleic acids and inositol phosphates, which are obtained from digested and decaying plants, animals, and microbial biomass. Although P_o is more abundant than P_i in many soils, it remains unavailable to plants until the P_i moiety is released into the soil solution via hydrolytic processes [22] SOM plays an essential role in the transformation of P [23]. An 11-year field experiment conducted by [24] demonstrated that it is possible to partially substitute mineral P by a mixture of manure and straw to the soil for the effective promotion of P availability with long-term benefits for environmental sustainability. The target for any research should primarily address low-yielding soils. The addition of biochar has proven to be one of the most remarkable solutions to release phosphorus in ferruginous soils. The unstable, stable char and SOM play important roles in providing the carbon source and energy to form and store urea, amino acids, nitrite, and nitrate derived from atmospheric nitrogen, and to preserve applied nitrogen [25-27]. The pathways of preferences are governed by the plant, soil, microbial activities, and climatic conditions, which define the biome and microbiome. It is likely that the most elaborate microbiomes are formed by the application of biochar to soil with organic materials [28].

The addition of biochar enhanced the soil microbial biomass and enzyme activity, while shifts in the bacterial community composition related to N cycling via the increase in soil pH, and thus, reduced N leaching and increased the retention of N in tea soils [29]. The results of [30] suggest that biochar improves soil pH [31], and reduces soil extra N in the acidic tea soil. It is then a soil buffer controlling pH, where nutrients and water are retained and regulate movement to and from the plant to the exceptionally large biochar active surfaces, adsorbing, absorbing, and desorbing, thus serving also as a receptor. It enhances and supports the phosphorus cycle without being locked in ferruginous-derived soil by releasing enzymes and microbial enzymes, substrates with the required balance of nutrients to the plant [32], likely to be at lower suction pressures exerted by the plants. Also demonstrated the capacity to replace 70% of inorganic fertilizer by paddy husk biochar in paddy cultivations. Rice husk biochar-coated urea can potentially be used as a slow-releasing nitrogen fertilizer. In addition, the urea coated with biochar is less costly and contributes to mitigating pollution of water bodies by inorganic fertilizers (NPK) [31,32]. A number of nutrients was examined by [33] and according to them the biochar from bamboo and paddy straw increased the soil-available N, P, K, and Mg concentrations, and tea P, K, and Mg levels, which contributed to the observed growth improvement. Mycorrhizae, symbiotic fungi that form a mutualistic relationship with plant roots, play a crucial role in paddy cultivation by enhancing nutrient uptake, plant growth, and overall productivity. Mycorrhizal fungi improve the availability of phosphorus, nitrogen, and other essential nutrients, leading to increased plant biomass, yield, and tolerance to various stresses, including insect herbivores and flooding [34]. A Number of microbes are used in the manufacture of biofilm biofertilizers that can replace urea [35].

Phosphorus is the chemical energy carrier that enhances the uptake of nitrogen. Enzyme urease perhaps dictates the functions of the microbiome. Urease in the presence of nickel or cobalt is only able to catalyze the hydrolysis of urea to give ammonia and CO_2, and because of the exceedingly high activities of the enzyme urease, some bacteria like rhizobium and specific fungi can fix atmospheric nitrogen at normal temperature and pressure, thus forming urea [36]. It is likely that in the presence of urea, urease is formed to conjugate the microbiome N cycle. The hypothesis that enzymes are formed from the substrates and the enzymes influence substrate formation could very well be valid as described by [37].

Microbiomes, thus, enzyme activities change with the application of industrial inorganic fertilizers. Fertilizer applications in increasing crop production have negative implications due to impurities in the production of N-fertilizers. In India, most of the naphtha plants have been converted to gas, but high-temperature steels in the piping and reactors contaminate the urea production processes [38]. The momentum created with the use of industrial fertilizer production can not be reversed without adequate justification. Moreover, it is a necessity to find the causes and origins of the disease. Hence, a number of studies were initiated to nullify the effect of inorganic fertilizer applications. They were centered around permissible levels in the soil, thus fertilizer and the amount up taken into food and food products.

One such study, [8] reported on a wide-scale, systematic sampling program over two consecutive years investigating As and Cd in one of the affected regions of chronic kidney disease with unknown etiology (CKDu) that is endemic. Surface soil (0–15 cm), fertilizer, and rice seed samples were collected in 2017 and 2018 from

three CKDu affected areas Medawachchiya (M), Padaviya (P) and Giradurukotte (G), and a non-affected control site at Hambanthota (H). All inorganic fertilizer samples showed low As (< 30 mg kg^{-1}) and Cd (<1.25 mg kg^{-1}) concentrations, less than European Union guideline values, and no correlation with soil concentrations. Arsenic (≤ 3.8 mg kg^{-1}) and Cd (≤ 3.0 mg kg^{-1}) in the 400 soil samples analyzed were low at all four locations, and soils were considered suitable for sensitive and agricultural use. A human health risk assessment demonstrated that the As and Cd concentrations in surface soil provided no concern for non-carcinogenic risk and negligible or acceptable carcinogenic risk for all locations sampled. The As and Cd in rice seeds harvested were also less than the detection limits (< 0.1 mg kg^{-1}). They concluded that their work provides clarity around As and Cd baseline values in certain farm soils of the dry zone of Sri Lanka, and no substantive evidence that the levels of As and Cd in the surface soils contribute to CKDu in local agricultural populations. Additional sampling of subsurface soil and water resources would satisfy some uncertainties with the risk assessment described in their paper.

The paper in reference compared heavy metal concentrations in CKDu locations to higher levels elsewhere. Hambanthota soils, as published, accumulate more heavy metals than Giradurukotte. It stands to reason that the applied heavy metals by means of fertilizer, directly and indirectly through irrigation water, have been up taken by the plants in the three endemic locations than in Hambanthota. Or else, the HMs were percolated through the soil profile, because these minerals did not appear within the watershed. We must interpret in terms of uptake rates rather than based on simple conclusions made in the study. We must also compare with uncultivated lands as [39] did. They reported that the mean Cd content in cultivated vs. uncultivated soils in the Anuradhapura district was 0.02 ± 0.01 vs. 0.11 ± 0.19 mg/kg, while in Polonnaruwa district, it was 0.005 ± 0.004 vs. 0.016 ± 0.005 mg/kg. Also, other publications of [39] give different ways of heavy metal entering the food chain.

It is best to quote a review paper [2], summarising the present knowledge base. According to them, the disease is more prevalent among farmers and has a unique locality. Numerous nephrotoxins, irrigation networks, genetic susceptibility, soil factors, and even bioterrorism as possible etiologies have been considered. Drinking water was proven to be hard and to contain high fluoride levels, but toxins in food and water were controversial. The urine and tissues of affected patients contained some of the suspected toxins at higher levels. Though the majority of the researchers agreed on a toxic nephropathy, none of their hypotheses explained the clinical findings, the unique locality of the disease, and its appearance in the 1990s. The absence of an identifiable cause has hampered efforts to control the disease. Careful use of agrochemicals and more research to unravel the mystery is recommended.

It is interesting to note that, similar to other diseases in Europe, urine and tissues of affected patients contained some of the suspected toxins at higher levels. Inorganic fertilizers or organic fertilizers formulated with cadmium sources promote chelation of Cd and N. Organometallic uptake is always preferred, as reported by [40]. Underdetection of cadmium (Cd) derived from organocadmium compounds can occur due to various factors, including the nature of the compound, analytical methods, and environmental conditions. Specifically, the volatility, reactivity, and tendency to form complexes can make it difficult to accurately quantify Cd in its organometallic forms [41, 42].

Quantifying the Movement of Cd

Perhaps, the quality of the fertilizer varies greatly [43]. According to them, the application of fertilizers, which contained a high dose of toxic metals, could be the driving force for agricultural soil pollution, and the limitless application of low-quality fertilizer would lead to more soil contamination with heavy metals. Hence, hazardous metals can be incorporated into the food chains via contaminated paddy soil. It could be possible to have levels reported by [7] of trace metal levels in inorganic fertilizers commercially available in Nigeria or it is possible that accumulated levels of organocadmium and arsenic could be transported to the grain in small quantities. The levels detected by [9] can be used to undertake a simple mass balance with that of [7, 8] as given in Table 1. The findings of [8] for both fertilizer and rice fall within [7, 9]. The results of [9] are more comprehensive because the uptake of cadmium will depend on field conditions. If the cadmium concentrations in the field remain the same from one season to the next and one year to the next, there will only be the uptake equivalent to the applied cadmium from the fertilizer. This hypothesis can be deduced from the results of [8] because all the cadmium in the leaves and stems from the previous season too will eventually end up in grains. But the cadmium levels are within the acceptable EU standards as can be deduced from Table 1, amounting to 0.0931 $mg.kg^{-1}$. When a higher amount of inorganic fertilizer is applied, this balance will no longer remain static, it will either increase storage of cadmium in the soil or get transported more to the grain and stored mainly in the bran and husk due to their role in metal sequestration [44]. Plant uptake will be further aggravated by low pH under submerged conditions [45]. Increasing nitrogen fertilizers increases cadmium uptake and could reach maximum levels reported by [9]. Increasing N fertilization increases Cd uptake and accumulation in plants grown under field conditions where Cd is freely available. Numerous studies reported that N fertilizer, regardless of its

form, has a positive relationship with Cd accumulation in plants [46] such as durum wheat [47, 48], rice [49]. Rice, particularly, has a notable tendency to accumulate cadmium and other heavy metals in the soil [50].

Table 1. A hypothetical mass balance of applied fertilizer and uptake of Cd into rice and grain

Fertilizer type	**Reported Cd concentration in fertilizer ($mgCdkg^{-1}$)**	**Fertilizer application rate ($kgha^{-1}$)**	**Total Cd concentration in the field ($mgCdha^{-1}$)**			**Notation**
[7]						
P_2O_5	2.59	60	155.4			
Urea	2.67	255	680.85			
Total	**5.26**	**315**	**836.25**			Z
[8]						
P_2O_5	1.25	60	75			
Urea	1.25	255	318.75			
Total	2.5	315	393.75			
In Rice				**In Grain**		
<6.00-261.786 µg.kg^{-1} [9]						
Value	Reported Cd concentration (mg.kg^{-1} of rice)	Harvest ($kgCdha^{-1}$)	Total ($mgCdha^{-1}$)	Harvest ($kgha^{-1}$)	Total ($mgCdha^{-1}$)	
Lowest	0.0060	3536	21.22	5200	31.2	$z-\beta$
Low	0.2618	2040	534.04	3000	785.4	$z-\beta$
High	0.2618	3536	925.68	5200	1361.3	$z+\alpha$
Highest	0.2618	4760	1246.10	7000	1832.50	$z+\alpha$
Average	0.1339	2876.4	385.13	4230	566.37	$z-\beta$
Derived from rice	0.1369	2876.4	393.75	4230	393.75	$z-\beta$
Derived from grain	0.0931					

Average yields given by [51].

Safe Rice Consumption Limit

The mass balance given in Table 2, was undertaken to counter notions that the applied fertilizer cannot reach high quantities of cadmium in rice. According to [52], cadmium levels found in the rice grown in the rain-fed wet zone exceed the amounts in the dry zone by 40–60%. The researcher pointed out that even an extremely contaminated sample of fertilizers containing 50 mg of cadmium per kg will take 1.2 millennia for the soil to impregnated with cadmium to a hazardous level. It is then necessary to find out if the values given by [9] are within food safety limits. According to the (FDA, 2023), for example, uses a toxicological reference value (TRV) range of 0.21-0.36 micrograms (µg) per kilogram body weight per day. The EFSA (1975) has a provisional tolerable weekly intake (PTWI) of 2.5 µg/kg bw/week. The lowest consumption limit is 309 g/week for FDA limit, and the maximum according to EFSA standards 1kg/week, thus the populations are vulnerable to Cd toxicity.

Table 2. Determination of rice consumption limits to avoid cadmium toxicity

Description	FDA	FDA	EFSA	Unit
Tolerance limit	1.47E-06	2.17E-06	2.50E-06	g/bw/week
The average body weight of Sri Lankan	55	55	55	Kg
Allowable per week	8.09E-05	1.19E-04	1.38E-04	g/week
Comparative study				
Available derived (low value) in rice from Table 1	1.3E-04	1.3E-04	1.3E-04	g/kg
Quantity of rice consumption to exceed the limit	0.604	0.891	1.027	kg/week
Available in rice (Highest)	2.62E-04	2.62E-04	2.62E-04	g/kg
Quantity of rice consumption to exceed the limit	0.309	0.456	0.525	kg/week

Identifying Cd Occurrences and Their Management

It is evident that there is HM contamination in the soils of Sri Lanka and also chelation through many pathways including gyphosate [53,54], making it significant in relation to the accepted standards. Degradation of gyphosate depends on soil conditions, applied concentration and frequency, microbial populations, and temperature [53 - 55]. HM contamination is also high in the case of Europe, with and without the influence of pesticides. The average soil Cd concentration in European soils is about 0.3 mg Cd kg^{-1}, equivalent to about 900 g Cd ha^{-1} in the plough layer [56]. Between 1980 and 1995, the annual net input fluxes of Cd in European soils were estimated to range between 1 to 10 g Cd ha^{-1}, indicating that annual fluxes are much smaller than the total Cd stock in soils [57]. The European median Cd concentration is 0.182 mg/kg in agricultural soil and 0.197 mg/kg in grazing land soil (including eastern Ukraine) [58]. Nevertheless, the EU has recently adopted regulation (EU) 2019/1009 with the aim of establishing stricter limits for Cd presence in fertilizer products and to promote a higher use of fertilizers from organic sources [59]. No doubt there are great improvements in lowering Cd in the soil and food chain over the years since ESDAC (2023) indicates that out of the total, 72.6% of the samples have Cd values <0.07 mg kg^{-1}, 21.6% in the range 0.07–1 mg kg^{-1} and the remaining 5.5% higher than the threshold of 1 mg kg^{-1}, which is generally considered the limit for risk assessment. In the Sri Lanka context, the opposite has taken place because the values stated by [39] were much less than [8]. Moreover, as can be deduced from Table 1, the Cd fluxes per season are higher in endemic locations in Sri Lanka than in Europe because conjugated microbiomes under tropical conditions within the rooting systems cover the entire extent of land applied with fertilizers. Deep tillage is also not practiced, breaking the plough/hard pans formed in the paddy fields; however, Cd can reappear in the lower parts of the catena in soil and water.

There is also evidence of HM contamination from water sources. According to [60], CKDu subjects had significantly higher pesticide and surface water usage, which may be the source of differential As exposure in these subjects. Modern integrated pest management (IPM) strategies emphasize reducing reliance on broad-spectrum insecticides by incorporating various techniques, including the use of natural predators, pest-resistant plants, and other methods to maintain a healthy environment [61, 62] in a comparative study on drinking water from ground and surface water sources reported that the nephrotoxic heavy metals contents including Cd, Pb, As, and Cr in CKDu hotspot fluctuated in the ranges of 9.78-187.25 μg L^{-1}, 0.08-0.66 μg L^{-1}, 20.76-103.30 μg L^{-1}, and 0.03-0.34 μg L^{-1}. The frequency of occurrence above the threshold limit of fluoride was 28% in non-CKDu water samples, while 81% in CKDu prevalent sites. Their results have emphasized a strong association between fluoride and water hardness. The frequency of occurrence above the threshold limit of fluoride was 28% in non-CKDu water samples, while 81% in CKDu prevalent sites. The hardness values in the CKDu prevalent site indicated “moderately hard water,” while the non-CKDu area indicated “soft water.”

Limiting or eliminating cadmium is imperative since it is characterized by a strong cumulative effect in humans, and its content in living organisms increases with diet and age [17]. Among the factors that determine the absorption and accumulation of this xenobiotic, the duration of uptake, the chemical form of the metal, and the diet should be included, as well as age, sex, and health conditions of the exposed people. Therefore, it is recommended to undertake an applied research project based on a project evaluation model and a plan optimization model developed by [63].

Moreover, one of the most important tasks will be to investigate the need for soil remediation requirements in HM contaminated soils. It will require a hydrogeology study of the watersheds by systematic monitoring of HM movements and conducting extensive monitoring of soil and fertilizer as in Europe. Therefore, it is recommended to undertake an applied research project based on a project evaluation model and plan optimization model developed by [63]. For example, the project will be implemented in stages, like a 70% reduction of urea fertilizer by biochar. So that there is visibility of the comprehensive results in the project

evaluation stage. It will involve individuals, farmer collective organizations, like cooperatives, and research scientists to develop technological approaches suitable for advancing mechanized farming systems to reduce drudgery and create employment throughout the year. It may require land consolidation and land development to improve irrigation and drainage with the required SOM and nutrient management with the developed technologies.

CONCLUSION

This analytical review indicates that it is essential to conduct both basic and applied research to prevent CKDu occurrence. An extensive study on the movement of heavy metals in the watersheds is required to introduce remedial measures of HM contaminated soils. The outcome of the mass flow study from inorganic fertilizer through or directly to paddy plants, embedding in the rice and its obligated low consumption levels exceeding tolerance limits, is useful to convince policymakers to launch a well-defined plan to produce organic fertilizers with the required mechanization scheme while controlling cadmium in fertilizers. The existing knowledge base is adequate to implement advanced agricultural practices like biochar biofertilizers to mobilize phosphate and provide the required combination of nutrients, mechanization to reduce overexposure to UVB, land consolidation, phytoremediation of contaminated soils, improvements to irrigation and drainage systems, IPM, and increasing production. Such measures will be adequate to avoid CKDu cases with the reduction of HM contamination. It will empower farmers to control their production systems from land preparation to marketing their products.

REFERENCES

[1] A.V. Ranasinghe, G.W.G.P. Kumara, R.H. Karunarathna, A.P. De Silva, K.G.D. Sachintani, J.M.C.N. Gunawardena, S.K.C.R. Kumari, M.S.F. Sarjana, J.S. Chandraguptha, M.V.C. De Silva.The incidence, prevalence and trends of Chronic Kidney Disease and Chronic Kidney Disease of uncertain aetiology (CKDu) in the North Central Province of Sri Lanka: an analysis of 30,566 patients. BMC Nephrol. 20(1)(2019) 338. doi: 10.1186/s12882-019-1501-0.

[2] T. Hewageegana, N. Rajakaruna, C. Jayasinghe, Chronic Kidney Disease of Unknown Origin in Sri Lanka: A Literature Review [Internet]. Chronic Kidney Disease - Novel Insights into Pathophysiology and Treatment. IntechOpen. (2024). http://dx.doi.org/10.5772/intechopen.1004707.

[3] C. Jayasumana, C. Orantes, R. Herrera, M. Almaguer, L. Lopez, L.C. Silva, P. Ordunez, S. Siribaddana, S. Gunatilake, M.E. De Broe Chronic interstitial nephritis in agricultural communities: a worldwide epidemic with social, occupational and environmental determinants. Nephrol Dial Transplant. 32(2)(2017) 234-241. doi: 10.1093/ndt/gfw346.

[4] P.P. Drechse, P. Madhuwanthi, D. Nisansala, D. Ramamoorthi, T. Bandara, On the feasibility of an agricultural revolution: Sri Lanka's ban of chemical fertilizers in 2021. Food Security 17 (2025) 585–602. https://doi.org/10.1007/s12571-025-01528-6.

[5] M.V. Beusekom, Colonial Agricultural Development Schemes. Oxford Research Encyclopedia of African History (2021). https://oxfordre.com/africanhistory/view/10.1093/acrefore/9780190277734.001.0001/acrefore-9780190277734-e-751.

[6] A. Rashid, B.J. Schutte, A. Ulery, M.K. Deyholos, S. Sanogo, E.A. Lehnhoff, L. Beck, Heavy Metal Contamination in Agricultural Soil: Environmental Pollutants Affecting Crop Health. Agronomy 13(2023) 1521. https://doi.org/10.3390/ agronomy13061521.

[7] W. Anake, Trace Metals Levels in Inorganic Fertilizers Commercially Available in Nigeria. Journal of Scientific Research 3(2013) 610-620.

[8] S. Gunadasa, M. Tighe, S.C. Wilson, Arsenic and Cadmium and Associated Risk in Farm Soils of the Dry Zone Sri Lanka where Chronic Kidney Disease of Unknown Etiology (CKDu) is Endemic. September 2022. Exposure and Health 14(2)(2022) 1-14. DOI: 10.1007/s12403-021-00433-5.

[9] E.M.R.K.B. Edirisinghe, B.K.K.K. Jinadasa, Cadmium and arsenic concentrations in Sri Lankan rice and their potential health risks. Ceylon Journal of Science 49(3) (2020)239-244 DOI: http://doi.org/10.4038/cjs.v49i3.7774.

[10] Rice Knowledge Bank (2025) http://www.knowledgebank.irri.org/.

[11] L.A.I. Constance, Setting Maximum Level for Cadmium in Polished Rice, Food Safety Focus (141st Issue, April 2018) – Food Safety Platform (2018) https://www.cfs.gov.hk/english/multimedia/multimedia_pub/multimedia_pub_fsf_141_02.html

[12] G.B.D. Deyn, L. Kooistra, The role of soils in habitat creation, maintenance and restoration. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 27(2021) 376(1834), 20200170. doi: 10.1098/rstb.2020.0170.

[13] N.C. Brady, R.R. Weil, The Nature and Properties of Soil. Arson/Prentice Hall, 1996.

[14] R. Doi, S. Pitiwut, From Maximization to Optimization: A Paradigm Shift in Rice Production in Thailand to Improve Overall Quality of Life of Stakeholders. Hindawi Publishing Corporation Scientific World Journal (2014). http://dx.doi.org/10.1155/2014/604291.

[15] S.J. Wimalawansa, Escalating chronic kidney diseases of multi-factorial origin in Sri Lanka: causes, solutions, and recommendations. Environ. Health Prev. Med. 19 (2014) 375-394. https://doi.org/10.1007/s12199-014-0395-5.

[16] M. R. Kulathunga, M. A. Wijayawardena, R. Naidu, S. J. Wimalawansa, M. M. Rahman, Health Risk Assessment From Heavy Metals Derived From Drinking Water and Rice, and Correlation With CKDu. Frontiers in Water, 3 (2022) 786487. https://doi.org/10.3389/frwa.2021.786487.

[17] C.R. Panabokke, S. Nagarajah, The fertility characteristics of the rice-growing soils of Ceylon. *Trop. Agricst.* (1964) 120, 3-30.

[18] C.R. Panabokke, Fertilizer recommendations for rice. Trop. Agricst. (1963). 119,167-174.

[19] R. R. Sattell, Phosphorus Characterization in Sri Lankan Alfisols: Sorption, Fractions, Plant Availability and Soil Test Suitability for the degree of Master of Science in Soil Science presented on June 6, 1989. (1989). https://www.environment.nsw.gov.au/resources/soils/testmethods/bp1.pdf.

[20] W.M.U.K., Rathnayaka, G.L.S.A. Dissanayaka, M.G.T.S. Amarasekara Comparison of soil fertility status of irrigated and rain-fed lowland paddy farming systems in Ibbagamuwa divisional secretariat division, Sri Lanka. Journal of Agriculture and Value Addition, 2 (2)(2019) 20-31.

[21] M.M. Kodagoda, S. Thrikawala, D. S. de Z. Abeysiriwardena, Yield Response of Rice to Added Phosphorous and Potassium Fertilizer in the Dry Zone of Sri Lanka. The Journal of Agricultural Sciences - Sri Lanka 17 (2)(2022) 270-279 http://doi.org/10.4038/jas.v17i2.9742

[22] R.P. Massoukou, V. Poirier, P.N. Ndoutoumou, T. E. Epule, How Can Plants Help Restore Degraded Tropical Soils? Land 12(12) (2023) 2147. https://doi.org/10.3390/land12122147.

[23] L F.J. Almeida, I. F. Souza, L.C.C. Hurtarte, P. Paulo, C. I.T.M. Teixeira, I. R. Silva, C.W. Mueller, Forest litter constraints on the pathways controlling soil organic matter formation. Soil Biology and Biochemistry 163 (2021) 108447. https://doi.org/10.1016/j.soilbio.2021.108447.

[24] L. Wu, S. Zhang, M. Chen, J. Liu, X. Ding, A sustainable option: Biochar addition can improve soil phosphorus retention and rice yield in a saline–alkaline soil. Environmental Technology & Innovation, 24 (2021) 102070, https://doi.org/10.1016/j.eti.2021.102070.

[25] B.F.A. Basnayake, R.T.K. Ariyawansha, D.A.S. Gamage, A.K. Karunarathna, Experiences of Biochar Applications for Sustainable Agriculture in Sri Lanka (2020). In: De Silva R.P., Pushpakumara G., Prasada P., Weerahewa J. (eds) Agricultural Research for Sustainable Food Systems in Sri Lanka. pp. 255-276. Springer, Singapore. https://doi.org/10.1007/978-981-15-3673-1_121.

[26] K.P.D. Samudrika, R.T.K. Ariyawansha, B.F.A. Basnayake, A.N. Siriwardana, Optimization of biochar additions for enriching nitrogen in active phase low-temperature composting. Org. Agr. 10 (2020). 449–463. https://doi.org/10.1007/s13165-020-00287-2.

[27] W.L.N. Wasana, R.T.K. Ariyawansha, B.F.A. Basnayake, Development of an effective biocatalyzed organic fertilizer derived from *Gliricidia sepium* stem biochar. Journal Current Research in Agricultural Sciences 8(1) (2021)11-30. doi:10.18488/journal.68.2021.81.11.30.

[28] M. Brtnicky, T. Dokulilova, J. Holatko, V. Pecina, A. Kintl, O. Latal, T. Vyhnanek, J. Prichystalova, R. Datta, Long-Term Effects of Biochar-Based Organic Amendments on Soil Microbial Parameters. Agronomy, 9(11) (2019) 747. https://doi.org/10.3390/agronomy9110747.

[29] P. Chen, Y. Liu, C. Mo, Microbial mechanism of biochar addition on nitrogen leaching and retention in tea soils from different plantation ages. Science of the Total Environment (2021). https://doi.org/10.1016/j.scitotenv.2020.143817.

[30] O.A. Zaw, S. Sudo, K.T. Win, A. Shibata, T. Sano, Y. Hirono, Returning Tea Pruning Residue and Its Biochar Had a Contrasting Effect on Soil N_2O and CO_2 Emissions from Tea Plantation Soil. Atmosphere, 9 (2018) 109. doi:10.3390/atmos9030109.

[31] D.A.S. Gamage, B.F.A. Basnayake, W.A.J.M.De. Costa, O. Merah, Effects of rice husk biochar coated urea and anaerobically digested rice straw compost on the soil fertility and cyclic effect of phosphorus, Plants (2022). 11, 75. https://doi.org/10.3390/plants11010075.

[32] D.A.S. Gamage, Development of nutrient management technologies for sustainable rice for mitigating water and atmospheric pollution. PhD thesis, Postgraduate Institute of Agriculture, University of Peradeniya, Peradeniya, Sri Lanka(2015).

[33] P. Yan, C. Shen, Z. Zou, J. Fu, X. Li, L. Zhang, L. Zhang, W. Han, L. Fan Biochar stimulates tea growth by improving nutrients in acidic soil. Scientia Horticulturae (2021) https://doi.org/10.1016/j.scienta.2021.110078.

[34] L. Bernaola, M.J. Stout, The effect of mycorrhizal seed treatments on rice growth, yield, and tolerance to insect herbivores. J. Pest. Sci. 94 (2021) 375-392. https://doi.org/10.1007/s10340-020-01279-7.

[35] B. Ekanayake, M. Premarathna, A. Pathirana, G. Senevirathne Biofilm biofertilizer application in paddy cultivation: pioneering studies in Sri Lanka. Journal of Dry Zone Agriculture 10(1)(2024). 56107

[36] M. Prettre, B. Claudel, Elements of Chemical Kinetics, Gordon and Beach. Science Publishers Ltd, 1970.

[37] R.T.K. Ariyawansha, B.F.A. Basnayake, A.K.K. Karunarathna, M.I.M. Mowjood, Extensions to Michaelis-Menten kinetics for single parameters. Sci. Rep 8(2018). 16586, https://doi.org/10.1038 /s41598-018-34675-2.

[38] G.A. Engwa, P.U. Ferdinand, F.N. Nwalo, N.M. Unachukwu, Mechanism and Health Effects of Heavy Metal Toxicity in Humans, IntechOpen (2019). doi: 10.5772/intechopen.82511.

[39] J.M. Bandara, H.V. Wijewardena, J. Liyanege, M.A. Upul, J.M.U.A. Bandara, Chronic renal failure in Sri Lanka caused by elevated dietary cadmium: Trojan horse of the green revolution. Toxicol Lett. 15, 198(1)(2010) 33-9. doi: 10.1016/j.toxlet.2010.04.016.

[40] K.M. Haydee, K.E. Dalma, Concerning Organometallic Compounds in Environment: Occurrence, Fate, and Impact [Internet]. Recent Progress in Organometallic Chemistry (2017). http://dx.doi.org/10.5772/67755.

[41] M. Mcbride, A Comparison of Reliability of Soil Cadmium Determination by Standard Spectrometric Methods. Journal of Environmental Quality. 40 (2011) 1863-9. 10.2134/jeq2011.0096.

[42] H. Liu, H. Cui, Y. Wang, Z. Jiang, L. Lei, S. Wei, Accurate Determination of Trace Cadmium in Soil Samples with Graphite Furnace Atomic Absorption Spectrometry Using Metal–Organic Frameworks as Matrix Modifiers. Appl. Spectrosc. 77(2) (2023) 131-139. doi: 10.1177/00037028221141709.

[43] T.P. R. Perera, R.D. N. Dayananda, C.M. U. Dissanayake, D.A. S. Rathnasekara, W. S.M. Botheju, J. A. Liyanage, S. K. Weragoda, M.K. A. Kularathne, Risk Assessment of Trace Element Contamination in Drinking Water and Agricultural Soil: A Study in Selected Chronic Kidney Disease of Unknown Etiology (CKDu) Endemic Areas in Sri Lanka. Journal of Chemistry, 2021(1)(2020) 6627254. https://doi.org/10.1155/2021/6627254.

[44] A. A. Meharg, E. Lombi, P.N. Williams, K.G. Scheckel, J. Feldmann, A. Raab, Y. Zhu, R. Islam, Speciation and localization of arsenic and cadmium in rice grains. Environmental Science and Technology 47(18)(2013) 10504-10541.

[45] S. Uraguchi, T. Fujiwara, Cadmium transport and tolerance in rice: perspectives for reducing grain cadmium accumulation. Rice 5, 5(2012). https://doi.org/10.1186/1939-8433-5-5.

[46] Y. Yang, J. Xiong, L. Tao, Z. Cao, W. Tang, J. Zhang, X. Yu, G. Fu, X. Zhang, Y. Lu, Regulatory mechanisms of nitrogen (N) on cadmium (Cd) uptake and accumulation in plants: A review. Sci. Total Environ. 15(08) (2020). 135186. doi: 10.1016/j.scitotenv.2019.135186.

[47] X. Li, N. Ziadi, G. Be´langer, Z. Cai, X. Hua, Cadmium accumulation in wheat grain as affected by mineral N fertilizer and soil characteristics. Canadian Journal of Soil Science 91(4) (2011). 521-531. DOI: 10.4141/cjss10061.

[48] L. Mitchell, C.A. Grant, G. Racz, Effect of nitrogen application on concentration of cadmium and nutrient ions in soil solution and in durum wheat. Canadian Journal of Soil Science 80(1) (2000). 107-115. DOI: 10.4141/S98-085.

[49] Q. Tian, D. Zheng, P. Chen, S. Yuan, Z. Yi, The Effects of Reducing Nitrogen and Increasing Density in the Main Crop on Yield and Cadmium Accumulation of Ratoon Rice. Agronomy *15*(2) (2025) 485. https://doi.org/10.3390/agronomy15020485.

[50] J. Zhang, Y. Zhu, L. Yu, M. Yang, X. Zou, C. Yin, Y. Lin, Research Advances in Cadmium Uptake, Transport and Resistance in Rice (*Oryza sativa* L.). Cells, *11*(3) (2022) 569. https://doi.org/10.3390/cells11030569.

[51] M. Salman, H. Suzuki, W. Ahmad, S. Giusti, A. Ali, W.M.U.K. Rathnayake, D.N. Sirisena, D.M.J.B. Senanayake, W.M.T.M. Herath, J. Meegasthenna, Y. Ponnampalam, W.M. Bandulasena, A. De Silva, D. Bandara, A.B.D.T. Nandharathne, A.T. Sooriyaarachchi, S. Pathmarajah, Efficient agricultural water use and management in paddy fields in Sri Lanka – National outlook. Rome, FAO. (2022). https://doi.org/10.4060/cc2778en.

[52] M.W.C. Dharma-Wardana (2018). Chronic kidney disease of unknown etiology and the effect of multiple-ion interactions. Environmental Geochemistry and Health 40(2):705-719. DOI: 10.1007/s10653-017-0017-4.

[53] A.S. Araújo, R.T. Monteiro, R.B. Abarkeli, Effect of glyphosate on the microbial activity of two Brazilian soils. Chemosphere, 52(5) (2003). 799-804. doi: 10.1016/S0045-6535(03)00266-2.

[54] A. Szekacs, B. Darvas, Forty Years with Glyphosate. InTech. (2012). doi: 10.5772/32491.

[55] M.M. Newman, Hoilett, N., Lorenz. N., Dick, R.P., Liles, M.R., Ramsier, C., & Kloepper, J.W. (2016) Glyphosate effects on soil rhizosphere-associated bacterial communities. Sci Total Environ. 543(Pt A):155-160. doi: 10.1016/j.scitotenv.2015.11.008.

[56] E. Smolders, J. Mertens, Cadmium. In: Alloway, B. (eds) Heavy Metals in Soils. Environmental Pollution, 22 (2013). https://doi.org/10.1007/978-94-007-4470-7_10.

[57] L. Six, E. Smolders, Future trends in soil cadmium concentration under current cadmium fluxes to European agricultural soils. Science of The Total Environment (2014). 485–486. https://doi.org/10.1016/j.scitotenv.2014.03.109.

[58] M. Birke, C. Reimann, U. Rauch, et al., GEMAS: Cadmium Distribution and Its Sources in Agricultural and Grazing Land Soil of Europe—Original Data versus clr-Transformed Data. Journal of Geochemical Exploration 173 (2017) 13-30. https://doi.org/10.1016/j.gexplo.2016.11.007.

[59] M. Marini, D. Caro, M. Thomsen, The new fertilizer regulation: A starting point for cadmium control in European arable soils. Science of The Total Environment, 745 (2020) 140876. https://doi.org/10.1016/j.scitotenv.2020.140876.

[60] M. Atlani, A. Kumar, R. Ahirwar, M.N. Meenu, S.K. Goel, R. Kumari, A. Anirudhan, S. Vallamshetla, G.S.T. Reddy, Heavy metal association with chronic kidney disease of unknown cause in central India-results from a case-control study. BMC Nephrol., 3, 25(1)(2024). 120. doi: 10.1186/s12882-024-03564.

[61] W.S.M. Botheju, J.A. Liyanage, S.D.P. Kannangara, The Groundwater Geochemistry and the Human Health Risk Assessment of Drinking Water in an Area with a High Prevalence of Chronic Kidney Disease of Unknown Etiology (CKDu), Sri Lanka. Journal of Chemistry (2021). https://doi.org/10.1155/2021/1755140.

[62] V. Mishra, Insecticides - Advances in Integrated Pest Management. Delve Publishing LLC, 2017.

[63] H. Al Qaheri, M.K.E.D.A. Hasan, Optimum Project Portfolio Selection for Developmental Plans Using Goal Programming. *International Journal of Operations Research and Information Systems* 12 (4) (2021). DOI: 10.4018/IJORIS.20211001.oa1.

A Numerical Investigation of Valve Timing and Intake Pressure Effects on Performance and Emissions in a Hydrogen Port Fuel Injection Engine

Isuru Wickramaarachchi[1, a *], Indrajith D. Nissanka[1,b] and Sameera Wijeyakulasuriya[2,c]

[1]Department of Mechanical Engineering, University of Moratuwa, Katubedda, Sri Lanka

[2]Convergent Science, Inc., 6400 Enterprise Lane, Madison, WI 53719, USA

[a]wickramaarachchiim.19@uom.lk, [b]nissankai@uom.lk, [c]sameera.wijeyakulasuriya@convergecfd.com

KEYWORDS: *Port Fuel; Valve Timing; Intake Pressure; Performance; Emissions*

ABSTRACT

Hydrogen internal combustion engines (H_2ICEs) offer a viable low-emission alternative for decarbonizing transport, especially where full electrification is not practical. Among fueling strategies, port fuel injection (PFI) is particularly attractive due to its compatibility with existing engine platforms and simplicity compared to direct injection (DI). Performance and emissions in hydrogen PFI engines are strongly influenced by valve timing and intake boosting strategies. This study presents a computational framework to investigate the coupled effects of valve timing and intake pressure on the performance, thermal efficiency, and NO_x emissions of a hydrogen PFI engine under fuel-lean conditions (ϕ = 0.59). A modified Sandia optical engine geometry was simulated using CONVERGE CFD v4.1, employing detailed chemistry and adaptive mesh refinement. Latin Hypercube Sampling (LHS) was employed to generate 373 design cases that span a wide parametric space. Results show that intake boosting significantly improves performance, achieving a 220% increase in indicated power (up to 43.55 kW) and an 11% improvement in thermal efficiency (up to 48.7%) over the baseline configuration. However, these gains are accompanied by elevated NO_x emissions, particularly at higher valve overlaps. Conversely, the configuration that achieved the lowest NO_x emissions reduced them by 76% compared to the baseline, albeit at the expense of lower power and efficiency. The three configurations representing the most favorable outcomes for power, efficiency, and emissions within the studied parameter space highlight the inherent trade-offs among these objectives. These results provide practical guidance for calibrating hydrogen PFI engines and establish a solid foundation for future studies incorporating formal optimization methods.

INTRODUCTION

The continued rise in global greenhouse gas (GHG) emissions poses a critical challenge for policymakers, industry stakeholders, and researchers worldwide. The Intergovernmental Panel on Climate Change (IPCC) has stressed in its Sixth Assessment Synthesis Report that achieving the 1.5°C target requires rapid and sustained reductions in CO_2 and other GHGs [1]. In 2024 alone, energy-related CO_2 emissions reached 37.4 GT, marking a 410 MT increase from the previous year [2]. The transportation sector remains a significant contributor, accounting for over a quarter of these emissions, primarily from light-duty internal combustion engine (ICE) vehicles [3]. Battery electric vehicles (BEVs) have emerged as the leading strategy for decarbonizing transport. However, their overall environmental impact is closely tied to the carbon intensity of electricity generation. With over 60% of global electricity still derived from fossil fuels in 2024, the actual emissions reduction from BEVs remains limited in many regions [2]. BEVs also face infrastructural, economic, and operational challenges, particularly in high-load or long-distance applications [4].

Hydrogen has gained attention as a complementary energy carrier capable of overcoming some of these limitations. It offers zero tailpipe emissions and high gravimetric energy density, making it suitable for sectors where electrification remains impractical. Hydrogen can be utilized in both fuel cell electric vehicles (FCEVs) and hydrogen-fuelled internal combustion engines (H_2ICEs). While FCEVs offer high efficiency and quick refueling [5], they are hindered by high system costs, durability concerns, and limited fueling infrastructure [6]. In contrast, H_2ICEs offer a cost-effective and technically mature pathway, especially for retrofitting existing vehicle fleets. Hydrogen's favorable combustion properties—including high flame speed, wide flammability range, and low ignition energy—enable lean operation and reliable cold starts. Moreover, H_2ICEs are compatible

with existing ICE manufacturing and servicing ecosystems, allowing rapid deployment with minimal infrastructure changes [7]. Among fueling strategies, port fuel injection (PFI) and direct injection (DI) are commonly employed in hydrogen engines. While DI offers improved mixture control, it involves a higher cost and system complexity. PFI is simpler and more suitable for retrofitting gasoline engines. Recent studies indicate that hydrogen PFI engines can achieve high thermal efficiency and low emissions under ultra-lean conditions [8].

Within this context, two key engine parameters—valve timing and intake pressure—play crucial roles in controlling combustion behavior and emissions. Prior research has shown that both valve timing and intake boosting significantly influence hydrogen engine performance and emissions [9], [10], [11]. However, most existing studies are experimental and limited in parametric scope. High-fidelity CFD studies that investigate the coupled effects of valve timing and intake boosting under realistic operating conditions remain scarce. To address this gap, the present study develops a CFD-based computational framework to systematically explore the combined influence of valve timing and intake charge boosting in a hydrogen-fueled PFI engine. By capturing detailed in-cylinder phenomena, this work aims to identify configurations that improve performance, thermal efficiency, and emissions without relying on complex auxiliary systems such as water injection and EGR.

LITERATURE REVIEW

The Valve timing is a key determinant of combustion dynamics, engine performance, and emissions in hydrogen-fueled spark ignition engines. Park et al. [10] experimentally investigated intake and exhaust valve timing in a hydrogen PFI engine under high-load conditions. Their findings showed that moderately advanced intake valve opening (IVO) enhanced airflow and torque, while excessive advancement caused residual gas interference, diminishing performance. Retarded exhaust valve closing further reduced volumetric efficiency. This study underscored the need for optimal valve timing to enhance power output and thermal efficiency. Lee et al. [9] extended this understanding by demonstrating that retarding IVO to top dead center (TDC) enabled stable ultra-lean operation at $\phi = 0.2$ and achieved near-zero NO_x emissions. Although volumetric efficiency declined, performance was recovered using supercharging, achieving gasoline-equivalent power and thermal efficiency above 39%. This illustrated the synergistic potential of combining valve timing strategies with intake boosting. Huynh et al. [12] examined valve overlap time (VOT) variations from 0^0 to 50^0 CA and found that changes in overlap influenced mixture formation and combustion efficiency. While maximum torque was achieved around VOT = 30^0 CA, shorter overlaps still yielded satisfactory performance when energy input was adjusted. These findings collectively highlight the role of coordinated valve timing in optimizing engine output and reducing emissions under lean-burn conditions.

In parallel, intake boosting methods such as turbocharging and supercharging have proven effective in addressing the volumetric efficiency limitations inherent in hydrogen PFI engines. Jilakara et al. [13] experimentally evaluated a turbocharged, intercooled hydrogen engine designed for ultra-lean operation ($\phi = 0.5$–0.6) in a mini-bus application. Their setup delivered 66.7 kW at 3200 rpm while maintaining low combustion variability. However, turbo lag at low engine speeds remained a challenge. Verhelst et al. [14] explored supercharging options in a single-cylinder PFI engine and compared lean-burn and stoichiometric strategies. Lean supercharging ($\lambda \geq 2$) emerged as the more favorable approach, offering higher thermal efficiency and NO_x levels below 100 ppm across a broad speed range. The study also cautioned against increased NO_x formation at elevated pressures, which could be mitigated through EGR or leaner mixtures. Lee et al. [15] conducted a simulation-based analysis of turbocharged hydrogen engines and observed that moderate boost pressures improved power and airflow but also raised combustion temperatures, potentially increasing NO_x emissions under stoichiometric conditions. Nguyen et al. [16] corroborated these findings through experiments on a small-scale H_2ICE, confirming power enhancement from supercharging, while stressing the importance of spark timing optimization to maintain combustion efficiency.

Despite these contributions, existing studies typically focus on either valve timing or intake pressure in isolation, often within limited experimental setups. Few have examined their combined influence using high-resolution CFD approaches. As hydrogen PFI engines operate under complex, coupled thermochemical conditions, a comprehensive simulation-based exploration of valve timing and intake boosting interactions is essential. This study addresses that gap by providing a detailed CFD framework to inform future H_2ICE development.

NUMERICAL MODEL

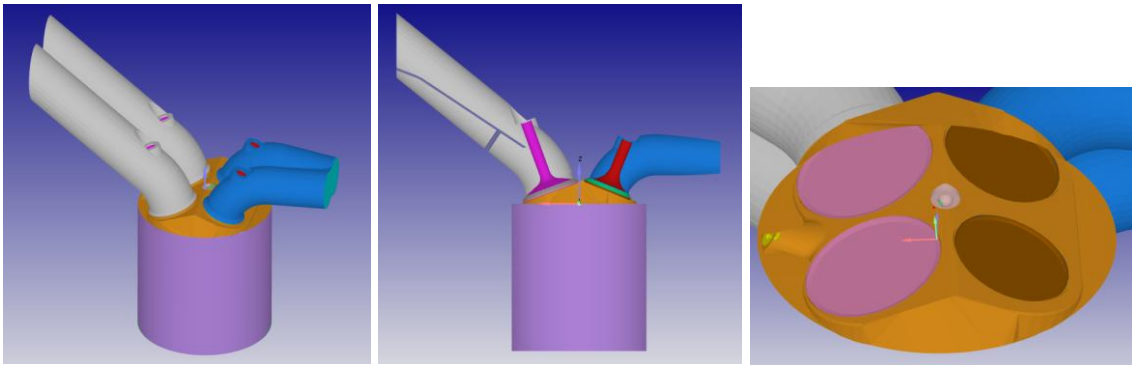
The Due to the unavailability of a dedicated experimental research facility and the absence of a publicly available PFI engine geometry with sufficient experimental validation data, the single cylinder optically accessible DI hydrogen engine developed by Sandia National Laboratories [12] was selected as the baseline geometry for this study. The engine specifications and operating conditions are summarized in Table 1.

Table 1: Engine specifications and operating conditions

Description	**Value**
Bore	92 [mm]
Stroke	85 [mm]
Displacement	565 [cc]
Compression ratio	10.91
Engine speed	1500 [rpm]
Intake valve timing (IVO/IVC)	346/ -140 [CAD] aTDC
Exhaust valve timing (EVO/EVC)	130/ 346 [CAD] aTDC
Intake temperature	309 K

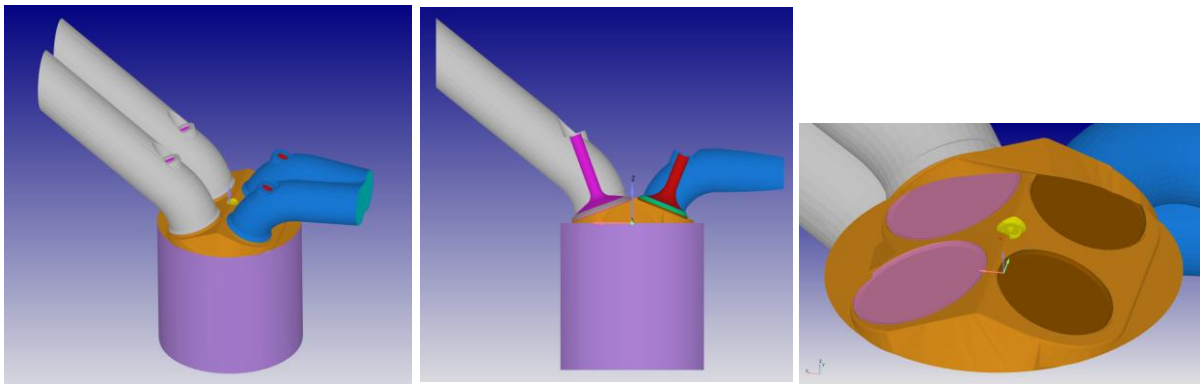
This engine geometry has been employed in numerous prior studies. Scarcelli et al. [13] developed a 3D CFD model to analyze hydrogen jet dynamics, showing that high mesh resolution near the injector is crucial for capturing jet penetration and stratification, with strong agreement to PLIF measurements. Barbato and Cantore [14] conducted full-cycle simulations and validated their adaptive mesh refinement strategy using experimental data, emphasizing the need for accurate time resolution during the injection phase. Wu et al. [15] employed CONVERGE to investigate hydrogen–air mixing under different in-cylinder flow conditions. The original engine geometry is available in CONVERGE [16] and is illustrated in Fig. 1. It consists of distinct intake and exhaust ports (Fig. 1a), a tumble plate integrated into the intake path to enhance in-cylinder motion (Fig. 1b), and a centrally located direct injector with a side-mounted spark plug on the cylinder head (Fig. 1c).

To facilitate the investigation of hydrogen PFI operation, the original DI geometry was modified. The key modifications included removing the direct injector, eliminating the tumble plate, and repositioning the spark plug to the center of the cylinder head. These changes are shown in Fig 2. The intake ports were retained to simulate the entry of a premixed hydrogen–air charge through conventional intake valves. Although these modifications substantially alter the original DI configuration, they are justified within the context of this numerical investigation, which focuses on analyzing relative trends in performance and emissions rather than predicting absolute values.



(a) (b) (c)

Figure 1: (a) Original Sandia DI engine geometry, (b) Cross-sectional view showing tumble plate, (c) Cylinder head showing DI injector and spark plug placement



(a) (b) (c)

Figure 2: (a) Modified engine geometry for PFI operation, (b) Cross-sectional view without tumble plate, (c) Top view showing removed injector and repositioned spark plug

A fully premixed, homogeneous hydrogen–air mixture was assumed to enter the computational domain through the intake ports, consistent with a dry PFI system. To investigate the effects of valve timing on combustion dynamics, valve lift profiles were adapted to allow variation in VOT. Simulations were conducted using CONVERGE CFD v4.1, a state-of-the-art general-purpose CFD software. Each case was simulated over a full engine cycle, beginning at exhaust valve closing (EVC) and continuing through 720^0 crank angle rotations. To accurately capture flow and combustion phenomena, the computational mesh employed adaptive mesh refinement (AMR), which locally increased resolution in regions with steep velocity or temperature gradients, such as flame fronts and flow structures. Fixed embedding was also applied in key regions of interest, including the intake and exhaust valves and the spark plug, where ignition and mixing processes are critical. A grid independence study was performed using a base mesh size of 4 mm, with progressively increased AMR levels based on velocity and temperature criteria. Refinement level 2 was selected as it provided a good balance between accuracy and computational cost. Under these settings, the minimum cell size in the combustion zone was refined to 0.0625 mm during the ignition phase and 1 mm during flame propagation, with the overall cell count reaching a maximum of approximately 0.5 million during peak refinement.

The simulations solved the three-dimensional, transient, compressible Reynolds-Averaged Navier–Stokes (RANS) equations for conservation of mass, momentum, energy, and species transport within a finite volume framework. Turbulence was modeled using the Renormalization Group (RNG) k-ε turbulence model [17], which is well-suited for swirling and recirculating in-cylinder flows. The momentum, energy, and species transport equations were discretized using a second-order central difference scheme within a finite volume framework. Time integration was performed using a first-order implicit Euler scheme with a variable time-stepping algorithm to ensure numerical stability and efficiency across the engine cycle. Pressure–velocity coupling was handled via the PISO algorithm [18], while the gas density was computed based on the energy equation using the Redlich–Kwong real gas equation of state [19]. Wall heat transfer was modeled using the O'Rourke–Amsden model [20], applied to all solid regions. This model estimates instantaneous convective heat transfer coefficients based on in-cylinder conditions and wall geometry.

Turbulent scalar transport was modeled using a turbulent Schmidt number of 0.78 and a turbulent Prandtl number of 0.9. Combustion was simulated using the SAGE detailed chemistry solver [16], coupled with a hydrogen mechanism extracted from the C3MechV3.3 kinetic model [21]. This mechanism includes 12 species and 37 reactions and has been validated for ignition delays (600–2500 K, 0.65–40 atm, φ = 0.06–6) and laminar flame speeds (295–600 K, 1–10 atm, φ = 0.4–2.2). The SAGE solver uses the CVODE integrator from the SUNDIALS package [22] to solve the stiff system of ordinary differential equations governing the chemical kinetics. These equations are solved with internal time steps smaller than the main CFD time step, and the computed reaction source terms are passed to the species transport equations. Molecular diffusivities for each species were calculated using data from a dedicated transport file (transport.dat) [23]. Spark ignition was modeled via multiple staged energy deposition events to replicate spark kernel formation and early flame growth under realistic engine conditions [24]. All simulations were performed under globally lean conditions, with an equivalence ratio of $\varphi = 0.59$ ($\lambda \approx 1.7$). Spark timing was fixed across all cases to isolate the effects of VOT and intake pressure on combustion characteristics and emissions. Thermal NO_x formation was modeled using the extended Zeldovich mechanism [25] under the assumption of local thermodynamic equilibrium.

At the intake ports, total pressure was specified according to the target intake pressure condition for each case (ranging from atmospheric to 2.61325 bar), along with a total temperature of 309 K. At the exhaust ports, a fixed static pressure corresponding to atmospheric pressure was applied. All combustion chamber walls, piston crown,

and cylinder liner were modeled as no-slip boundaries with convective heat transfer calculated using the O'Rourke–Amsden model. The global equivalence ratio was fixed at φ = 0.59 for all cases, and spark timing was held constant to isolate the effects of VOT and intake pressure.

METHODOLOGY

Design of experiments

Latin hypercube sampling (LHS) [26] was employed to generate a statistically robust and space-filling design of experiments for this study. The sampling strategy focused on two key input parameters: VOT, which varied from 0^0 to 50^0 CA, and intake pressure, which ranged from 1.01325 to 2.61325 bar. Using this method, a total of 373 simulation cases were created, ensuring a comprehensive exploration of the design space. The distribution of the sample cases is illustrated in Fig. 3. Each case was simulated using the CFD setup described earlier. To isolate the effects of VOT and intake pressure, all other parameters including the global equivalence ratio ($\phi = 0.59$, $\lambda \approx 1.7$) and spark timing were held constant across the simulations. The simulations were executed on the CONVERGE Horizon cloud computing platform [27], with each case requiring approximately 5 hours on a 36-core machine. By running 10 machines in parallel, 10 simulations could be completed within 5 hours, enabling efficient execution of the full 373-case design space. This approach ensured the timely completion of the high-throughput CFD campaign while maintaining numerical consistency across all cases.

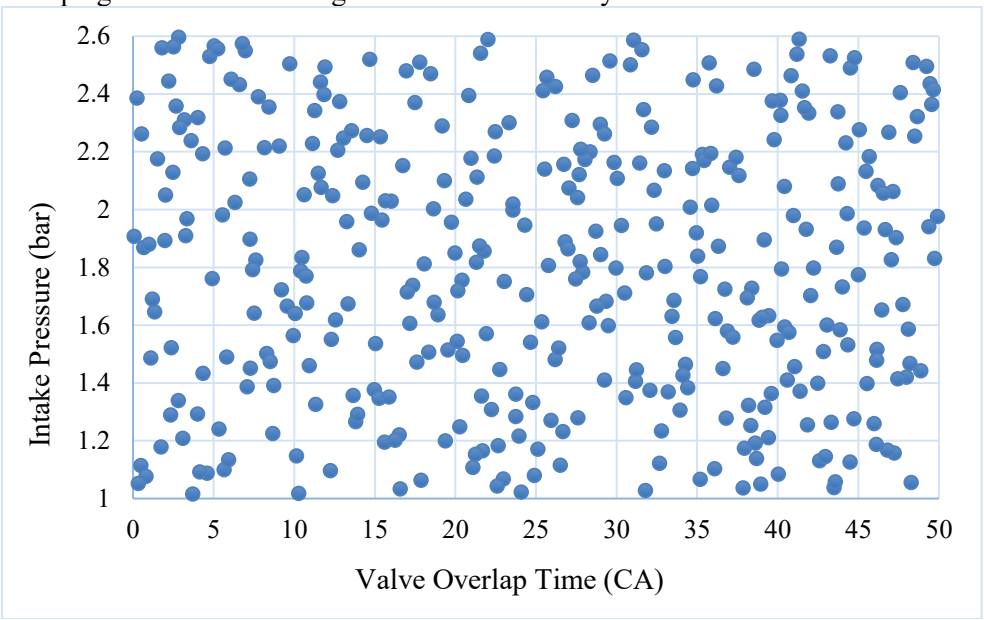


Figure 3: Latin Hypercube Sampling (LHS) distribution of 373 design points in the input parameter space

Performance and emission metrics

Each simulation output was post-processed to extract indicated power (kW), thermal efficiency (%), and NO_x Emissions (mg/cycle). These metrics were chosen to evaluate both engine performance and engine-out emissions.

RESULTS AND DISCUSSION

Effect of valve timing.

As shown in Fig. 4, increasing intake pressure results in a clear upward shift in indicated power across the entire VOT range. For each pressure level, power generally increases with VOT, although the rate of increase diminishes at higher overlaps. At an intake pressure of 2.6 bar, the engine reaches peak power levels exceeding 43 kW at VOT > 35^0 CA, while the power remains below 15 kW at atmospheric pressure, highlighting the dominant effect of intake boosting on volumetric efficiency and combustion energy release. Fig. 5 complements these findings by revealing corresponding trends in thermal efficiency. Efficiency increases with both intake pressure and VOT, particularly at mid-range overlaps (15–35^0 CA), where optimized intake-exhaust interaction and mixture quality contribute to better combustion. The highest thermal efficiency observed is approximately 48.7% at 2.6 bar, significantly outperforming lower intake pressure cases. Notably, while efficiency trends generally mirror power trends, the gains tend to plateau or even slightly decline beyond a certain VOT, suggesting diminishing benefits due to increased overlap-induced pumping losses or reduced volumetric efficiency at extreme settings.

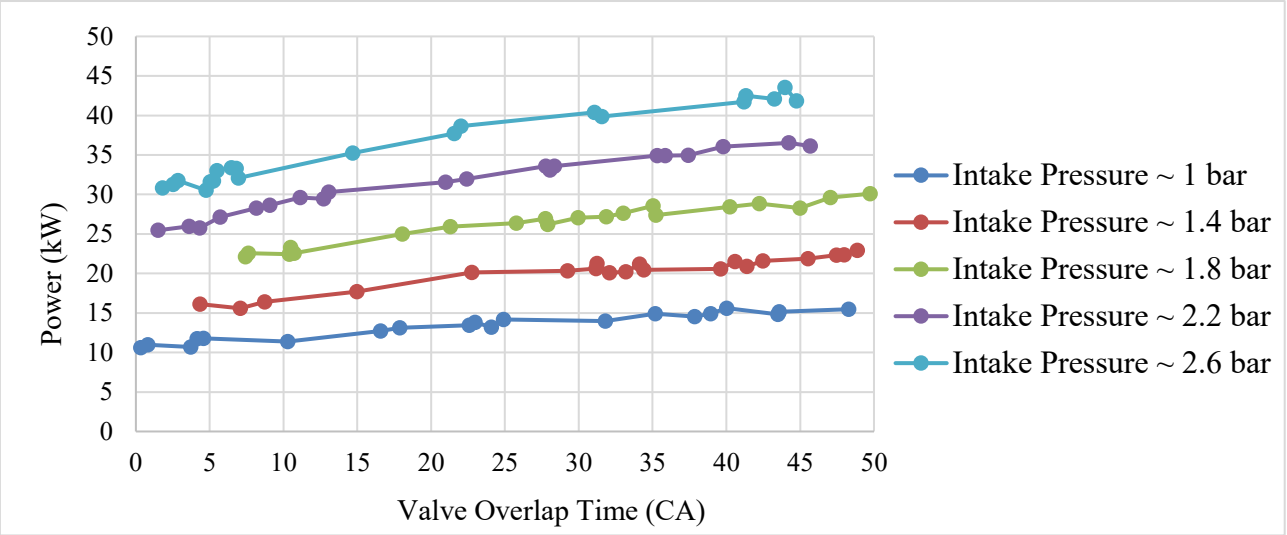


Figure 4: Effect of VOT on indicated power at various intake pressures

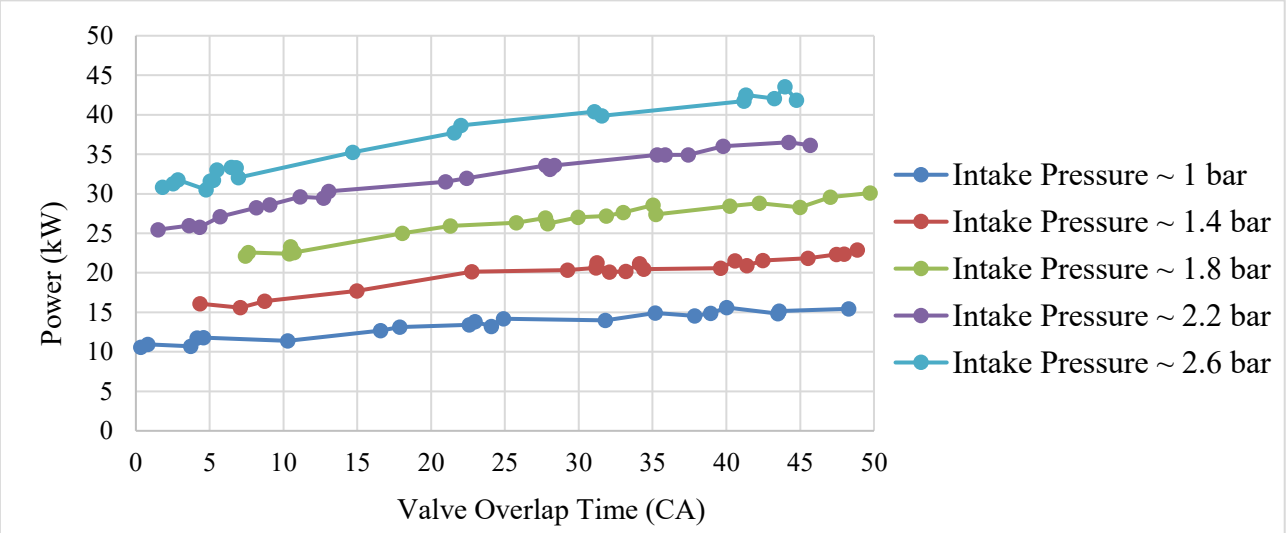


Figure 5: Effect of VOT on thermal efficiency at various intake pressures

Fig. 6 illustrates the variation of NO_x emissions, expressed as total NO_x mass per cycle, with VOT across five intake pressure levels ranging from approximately 1 bar to 2.6 bar. The results demonstrate that NO_x formation increases with both VOT and intake pressure, driven primarily by the elevated in-cylinder temperatures and oxygen availability associated with greater charge density and enhanced combustion intensity under boosted conditions. At low intake pressure (1–1.4 bar), NO_x emissions remain relatively modest and stable across the VOT range, suggesting limited peak temperature rise under lean conditions. However, as intake pressure increases to 2.2 and 2.6 bar, NO_x mass exhibits a pronounced upward trend with VOT, particularly beyond $20^0 – 25^0$ CA. This is attributed to longer valve overlaps promoting higher residual gas temperatures and more complete combustion, which intensify thermal NO_x generation. The highest NO_x values exceed 10 mg per cycle at 2.6 bar and high VOT, marking a significant emissions penalty associated with aggressive intake boosting and extended valve overlap.

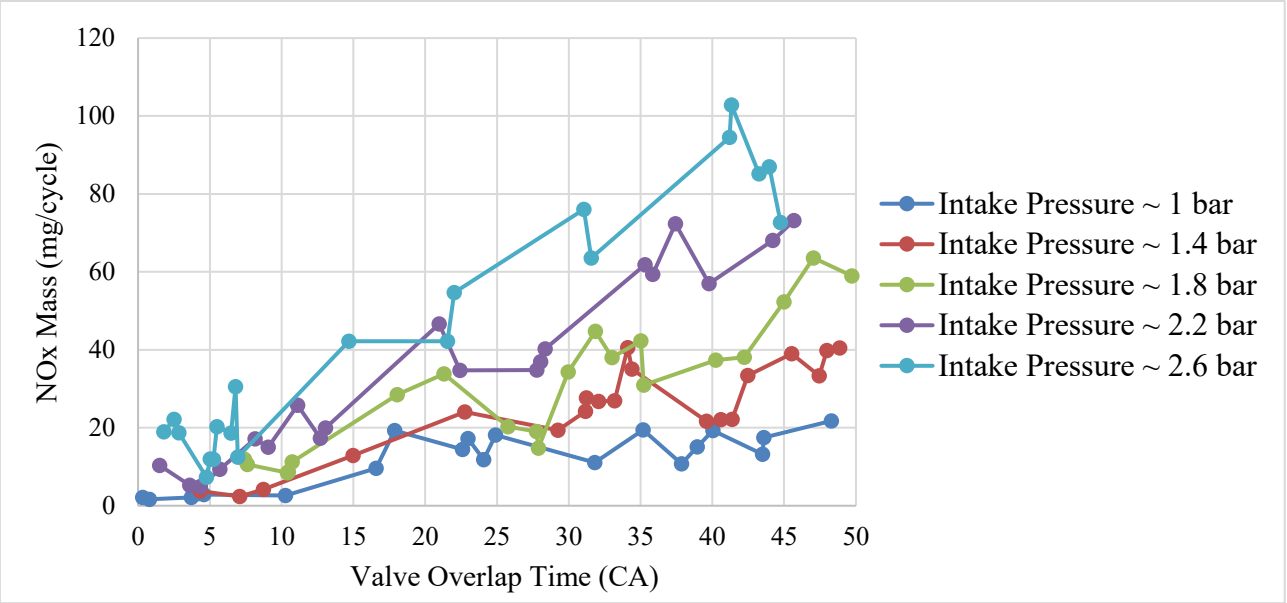


Figure 6: Effect of VOT on NO_x emissions at various intake pressures

Effect of intake pressure

Fig. 7 and 8 show the effect of intake pressure on indicated power and thermal efficiency. As intake pressure increases from 1.0 to 2.6 bar, both power output and thermal efficiency improve significantly due to higher charge density, enhanced volumetric efficiency, and more complete combustion. Intake boosting allows a greater mass of the hydrogen–air mixture to enter the cylinder, resulting in increased combustion energy and improved energy conversion. While power output shows a nearly linear increase with intake pressure, thermal efficiency gains tend to plateau at higher boost levels as secondary losses such as heat transfer and pumping work begin to offset further improvements.

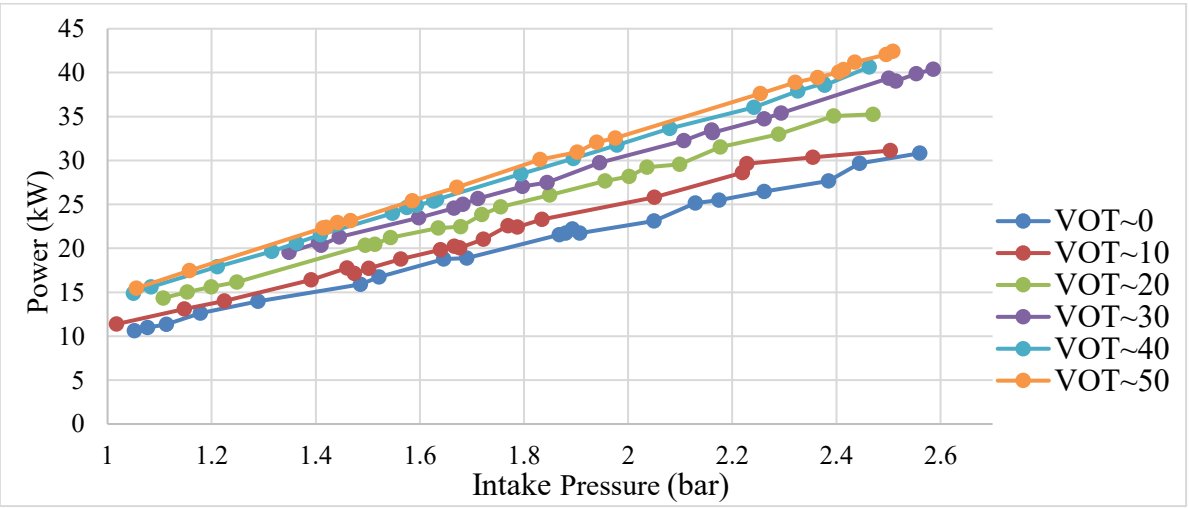


Figure 7: Effect of intake pressure on indicated power at various VOTs

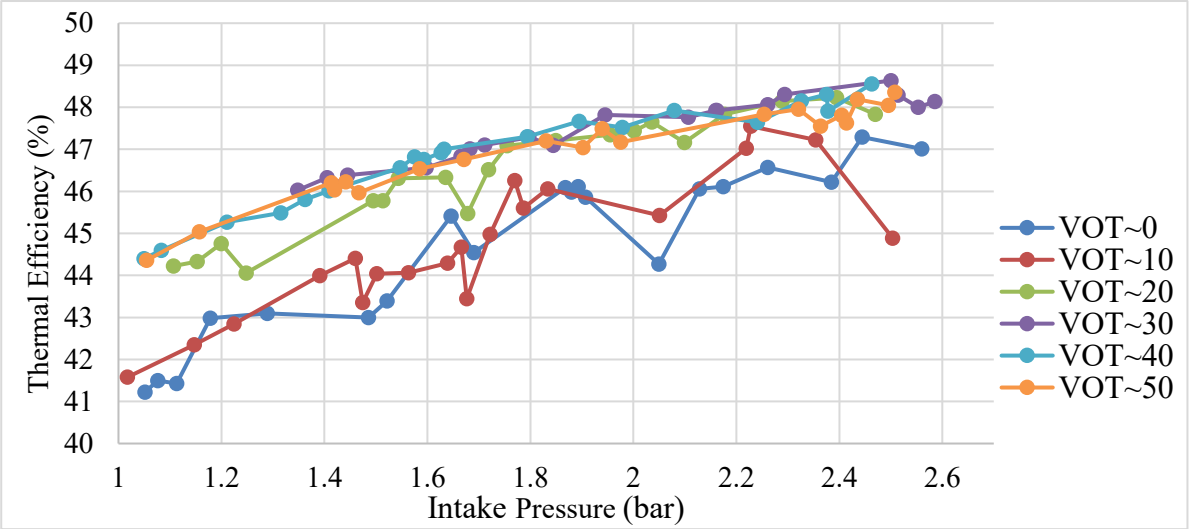


Figure 8: Effect of intake pressure on thermal efficiency at various VOTs

Fig. 9 illustrates the impact of intake pressure on NO_x emissions across various VOTs. As intake pressure increases from 1.0 to 2.6 bar, NO_x formation generally rises, with a more pronounced increase at higher VOT settings. This trend is attributed to elevated in-cylinder temperatures and oxygen availability under boosted conditions, which enhance thermal NO_x production. While low intake pressures result in minimal NO_x emissions regardless of VOT, higher pressures, especially above 2.0 bar, lead to a sharp rise in NO_x, particularly when combined with long valve overlaps.

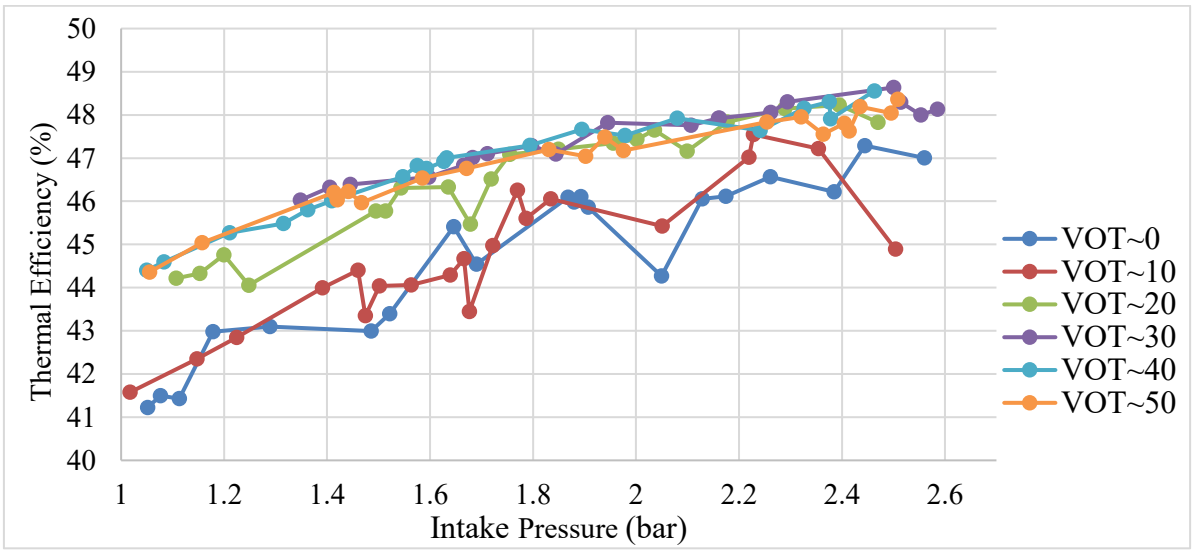


Figure 9: Effect of intake pressure on NO_x emissions at various VOTs

Best outcomes

Within the studied design space, three best operating points were identified, each corresponding to a different objective.

- Case A: Maximum power
- Case B: Maximum efficiency
- Case C: Minimum NO_x emissions

Table 2 presents the corresponding parameter combinations and highlights the trade-offs among power output, efficiency, and emissions. The baseline configuration is defined by a VOT of 30^0 and an intake pressure of 1.01325 bar. This case was selected as it represents a typical naturally aspirated setup and aligns with configurations previously identified as high performing for hydrogen PFI engines across a wide range of operating conditions [28]. Case A delivered the highest indicated power output of 43.55 kW, achieved at a VOT of 43^0 CA and an intake pressure of 2.61 bar. This represents a 221% increase in power compared to the baseline case, along with a thermal efficiency of 48.4%. However, this configuration resulted in elevated NO_x emissions of 8.69 mg per cycle. Case B yielded the maximum thermal efficiency of 48.7% at a VOT of 22^0 CA and an intake pressure of 2.59 bar. Compared to the baseline, this corresponds to an 11% improvement in efficiency and a substantial increase in power output to 38.66 kW, though with moderately high NO_x emissions of 5.47 mg per cycle. Case C

prioritized emissions reduction, achieving ultra-low NO_x emissions of just 0.17 mg per cycle, which marks a 76% reduction relative to the baseline. This was accomplished with a near-zero VOT of 0.8^0 CA and a near-atmospheric intake pressure of 1.08 bar. However, this strategy compromised performance, with power output reduced to 10.96 kW and thermal efficiency to 41.5%, illustrating the trade-off necessary to meet stringent emissions targets.

Table 2: Best parameter combinations for maximum power, thermal efficiency, and minimum NO_x emissions

Case	**VOT (CA)**	**Intake Pressure (bar)**	**Power (kW)**	**Efficiency (%)**	**NO_x (mg/cycle)**
Baseline	30.0	1.01325	13.54	43.9	0.72
A	43.0	2.61	43.55	48.4	8.69
B	22.0	2.59	38.66	48.7	5.47
C	0.8	1.08	10.96	41.5	0.17

CONCLUSION

This study developed and applied a detailed CFD-based framework to evaluate the effects of valve timing and intake pressure on combustion performance and NO_x emissions in a hydrogen port fuel injection engine operating under lean conditions. A total of 373 full engine cycle simulations were performed using Latin Hypercube Sampling to span a broad range of intake pressures and valve overlap timings. The results confirm that intake boosting plays a dominant role in enhancing engine performance. The configuration yielding maximum power achieved 43.55 kW, corresponding to a 220% increase over the baseline, while the most efficient case achieved a thermal efficiency of 48.7%, representing an 11% improvement. However, these benefits came with increased NO_x emissions, particularly at high intake pressures and extended valve overlaps. The configuration that achieved the lowest NO_x emissions reduced them by 76% compared to the baseline, although this came at the expense of power and efficiency. These findings highlight the inherent trade-offs between power, efficiency, and emissions in hydrogen port fuel injection engine design. The developed framework provides a solid basis for refining engine calibration and lays the groundwork for future integration of data-driven optimization and control strategies.

FUTURE WORK

Building on the CFD-based framework developed in this study, future work will focus on integrating machine learning (ML)-based surrogate modeling to enable faster and more adaptive engine optimization. Advanced regression techniques and ensemble learning methods will be trained on the existing simulation dataset to predict engine performance and emissions with significantly reduced computational cost. Furthermore, the framework will be extended to incorporate abnormal combustion phenomena, such as backfire and combustion instabilities, which are critical considerations for the practical deployment and calibration of hydrogen PFI engines.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Convergent Science for providing CONVERGE CFD software licenses and technical support. The authors also thank Convergent Science for granting access to the CONVERGE Horizon cloud computing platform, which enabled the high-throughput simulations required for this research.

REFERENCES

[1] K. Calvin et al., "IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.," Intergovernmental Panel on Climate Change (IPCC), July 2023. doi: 10.59327/IPCC/AR6-9789291691647.

[2] IEA, "Global Energy Review 2025 – Analysis - IEA," IEA, Mar. 24, 2025. https://www.iea.org/reports/global-energy-review-2025

[3] IEA, "Global EV Outlook 2025 – Analysis - IEA," IEA, May 14, 2025. https://www.iea.org/reports/global-ev-outlook-2025

[4] M. K. Anser, S. Ali, M. Umair, R. Javid, and M. Tayab, "Optimizing hydrogen integration in vehicle fuel systems for sustainable development: A step towards economic decarbonization," Int. J. Hydrog. Energy, vol. 98, pp. 321–333, Jan. 2025, doi: 10.1016/j.ijhydene.2024.12.009.

[5] Z. Stępień, "A Comprehensive Overview of Hydrogen-Fueled Internal Combustion Engines: Achievements and

Future Challenges," Energies, vol. 14, no. 20, p. 6504, Oct. 2021, doi: 10.3390/en14206504.
[6] Y. He, Y. Zhou, Z. Wang, J. Liu, Z. Liu, and G. Zhang, "Quantification on fuel cell degradation and techno-economic analysis of a hydrogen-based grid-interactive residential energy sharing network with fuel-cell-powered vehicles," Appl. Energy, vol. 303, p. 117444, Dec. 2021, doi: 10.1016/j.apenergy.2021.117444.
[7] K. Naseem et al., "Essential parts of hydrogen economy: Hydrogen production, storage, transportation and application," Renew. Sustain. Energy Rev., vol. 210, p. 115196, Mar. 2025, doi: 10.1016/j.rser.2024.115196.
[8] A. H. Khalid et al., "Hydrogen port fuel injection: Review of fuel injection control strategies to mitigate backfire in internal combustion engine fuelled with hydrogen," Int. J. Hydrog. Energy, vol. 66, pp. 571–581, May 2024, doi: 10.1016/j.ijhydene.2024.04.087.
[9] J. Lee, K. Lee, J. Lee, and B. Anh, "High power performance with zero NOx emission in a hydrogen-fueled spark ignition engine by valve timing and lean boosting," Fuel, vol. 128, pp. 381–389, July 2014, doi: 10.1016/j.fuel.2014.03.010.
[10] C. Park, W. Park, Y. Kim, Y. Choi, and B. Lim, "Effect of Valve Timing and Excess Air Ratio on Torque in Hydrogen-Fueled Internal Combustion Engine for UAV," Energies, vol. 12, no. 5, p. 771, Feb. 2019, doi: 10.3390/en12050771.
[11] S. Verhelst, P. Maesschalck, N. Rombaut, and R. Sierens, "Increasing the power output of hydrogen internal combustion engines by means of supercharging and exhaust gas recirculation," Int. J. Hydrog. Energy, vol. 34, no. 10, pp. 4406–4412, May 2009, doi: 10.1016/j.ijhydene.2009.03.037.
[12] "Hydrogen Engine," Engine Combustion Network. Accessed: June 12, 2025. [Online]. Available: https://ecn.sandia.gov/engines/hydrogen-engine/
[13] R. Scarcelli, T. Wallner, N. Matthias, V. Salazar, and S. Kaiser, "Numerical and Optical Evolution of Gaseous Jets in Direct Injection Hydrogen Engines," presented at the SAE 2011 World Congress & Exhibition, Apr. 2011, pp. 2011-01–0675. doi: 10.4271/2011-01-0675.
[14] A. Barbato and G. Cantore, "3D CFD simulation of a gaseous fuel injection in a hydrogen-fueled internal combustion engine," E3S Web Conf., vol. 312, p. 07001, 2021, doi: 10.1051/e3sconf/202131207001.
[15] B. Wu, R. Torelli, and Y. Pei, "Numerical modeling of hydrogen mixing in a direct-injection engine fueled with gaseous hydrogen," Fuel, vol. 341, p. 127725, June 2023, doi: 10.1016/j.fuel.2023.127725.
[16] Richards, K. J., Senecal, P. K., and Pomraning, E., CONVERGE 4, Convergent Science, Madison, WI (2024).
[17] V. Yakhot and S. A. Orszag, "Renormalization group analysis of turbulence. I. Basic theory," J. Sci. Comput., vol. 1, no. 1, pp. 3–51, 1986, doi: 10.1007/BF01061452.
[18] R. I. Issa, "Solution of the implicitly discretised fluid flow equations by operator-splitting," J. Comput. Phys., vol. 62, no. 1, pp. 40–65, Jan. 1986, doi: 10.1016/0021-9991(86)90099-9.
[19] Otto. Redlich and J. N. S. Kwong, "On the Thermodynamics of Solutions. V. An Equation of State. Fugacities of Gaseous Solutions.," Chem. Rev., vol. 44, no. 1, pp. 233–244, Feb. 1949, doi: 10.1021/cr60137a013.
[20] A. A. Amsden, P. J. O'Rourke, and T. D. Butler, "KIVA-II: A computer program for chemically reactive flows with sprays," LA-11560-MS, 6228444, May 1989. doi: 10.2172/6228444.
[21] S. Dong et al., "A new detailed kinetic model for surrogate fuels: C3MechV3.3," Appl. Energy Combust. Sci., vol. 9, p. 100043, Mar. 2022, doi: 10.1016/j.jaecs.2021.100043.
[22] "SUNDIALS | Computing." Accessed: June 18, 2025. [Online]. Available: https://computing.llnl.gov/projects/sundials
[23] T. Coffee and J. Heimerl, "Transport algorithms for premixed, laminar steady-state flames," Combust. Flame, vol. 43, pp. 273–289, 1981, doi: 10.1016/0010-2180(81)90027-4.
[24] J. B. Heywood, Internal combustion engine fundamentals. in McGraw-Hill series in mechanical engineering. New York: McGraw-Hill, 1988.
[25] G. A. Lavoie, J. B. Heywood, and J. C. Keck, "Experimental and Theoretical Study of Nitric Oxide Formation in Internal Combustion Engines," Combust. Sci. Technol., vol. 1, no. 4, pp. 313–326, Feb. 1970, doi: 10.1080/00102206908952211.
[26] M. D. McKay, R. J. Beckman, and W. J. Conover, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, vol. 21, no. 2, p. 239, May 1979, doi: 10.2307/1268522.
[27] "CONVERGE Horizon," Convergecfd.com, 2025. https://convergecfd.com/products/horizon (accessed Aug. 16, 2025).
[28] T. C. Huynh, J. K. Kang, K. C. Noh, J. T. Lee, and J. A. Caton, "Controlling Backfire Using Changes of the Valve Overlap Period for a Hydrogen-Fueled Engine Using an External Mixture," in ASME 2007 Internal Combustion Engine Division Fall Technical Conference, Charleston, South Carolina, USA: ASMEDC, Jan. 2007, pp. 243–251. doi: 10.1115/ICEF2007-1702.

The Influence Of Project Managers' Decision-Making Styles On Schedule Variance In Building Construction Projects

Kavindi Pulasinghe *; Priyanwada De Silva
Sri Lanka Institute of Information Technology, New Kandy Road, Malambe
manudileepulasinghe2001@gmail.com

ABSTRACT

This research focus to assess the relationship between project managers decision making styles and schedule variance in building construction projects in Sri Lanka. Timely completion of construction projects is one of the major performance indicators, yet delays are a long-standing issue in Sri Lankan construction projects. Though there are many causal factors, decision-making styles of project managers have not been studied yet, creating a significant knowledge gap with regard to their influence on project schedule variance. This study attempts to analyze the relationship between project managers' decision-making styles: directive, analytical, conceptual, and behavioural and schedule variance in Sri Lankan building construction projects. A mixed-methods research approach was adopted. Primary data were gathered by holding semi-structured interviews with nine industry practitioners and a questionnaire survey of 50 respondents covering key project roles. To analyze the data code based content analysis and descriptive statistical tools (percentage count, mean, weighted average etc.) were used and to examine the relationship between decision-making styles and schedule variance Pearson correlation was conducted. Results revealed that decision making styles play a significant role in influencing project schedules. However, it was found that directive and behavioural styles are most prevalent and successful styles in the Sri Lankan context. The data revealed both positive and negative influences of managerial decision-making styles on schedule performance. These results contribute to the link between leadership and project performance and make a theoretical and practical contribution by revealing decision-making as a key influence to minimize schedule variance.

KEYWORDS: *Project Manager, Decision Making styes, Schedule Variance, Building Construction Projects, Correlation analysis, Behavioural and directive styles.*

INTRODUCTION

The construction industry is a vital sector of any country's economic output; the industry plays a crucial role in contributing to national targets for physical, economic, and social needs (Kalambayi et al.,2021). However, the success of a construction project depends on the cooperation between the parties involved to the project like the building owners, contractors and project managers (Rauzana, Anita, 2016). Project managers must make prompt and wise decisions since construction projects are marked by uncertainty, changing stakeholder expectations, and frequent scope revisions (Saed et al., n.d.). since their decision-making approaches can affect project performance metrics, such as cost, quality, schedule variation (Zavadskas et al., 2010) etc. Cost control is basic to managerial accounting. Schedule control is more recent. Integration of schedule and cost control has been a natural objective of project-control systems since the late 1970s (Castro-Lacouture et al., n.d.; Robert Carr, n.d.). Schedule variance is the difference between the earned value of work performed and that which was scheduled. It provides a clear idea whether a project is ahead, on track, or behind schedule. A positive Schedule Variance indicates the project is ahead of schedule, while a negative Schedule Variance means the project is behind (Robert Carr, n.d.; Boom & Bucket, 2023). Therefore , Project managers as leaders in construction projects have a significant role to play to ensure for their team a sense of purpose, shared vision, and incentivize action (Prabhakar et al., 2016) and specially to make sure minimal schedule variance within the construction period.

Existing research focused on general decision making and leadership skills and on external factors like material delay, resource shortages (Shanian & Savadogo, 2006). However, the impact of these styles on schedule variance is not well understood. There is a lack of empirical research on their impact on schedule variance in construction project. Therefore, there is a gap of understanding the real relationship between the decision-making styles and the schedule variance in construction projects. Hence, this research aims to bridge the gap by investigating the relationship between project managers' decision-making styles and schedule variance in construction projects. By examining how different decision-making approaches impact on overall schedule of the project, this study aims to provide awareness that will help improve decision-making frameworks in construction management, ultimately improving project efficiency and reducing delays. To fill

this gap aim to assess The Relationship Between Project Managers' Decision-Making Styles, Schedule Variance in Construction Projects.

LITERATURE REVIEW.

Decision-making is an integral part of the management process within every organisation and at every level (Davis et al., 1990; Esa et al., 2014). It is a process required from project managers at every step to avoid serious buffering in a project schedule (Esa et al., 2014). In the case of construction projects, decisions are often based on schedules drawn up at the early planning stage (Castro-Lacouture et al., n.d.), and decision making becomes even more crucial in construction projects with complexity, uncertainty and dynamic site conditions are common (Belout & Gauvreau, 2004). According to Saed et al. (n.d.), understanding how project managers approach decision-making becomes a necessary step in mitigating risks related to schedule variance. Therefore, project managers are often required to respond to evolving conditions, deal with multiple stakeholders, and resolve problems that can delay project progress. Poor decisions such as mistaken procurement lead times or underestimating resource needs can provide significant schedule variance (Oyewobi et al., 2016).

Classification of decision-making styles.

Researchers have proposed varying conceptualizations of decision making styles, ranging from quite simple to very complex (Loo, n.d.; Oyewobi et al., 2016) as shown on the Table 2-1.

Table 0-1 Classification of decision making styles.

Classification	Source
The Scott and Bruce (1995) Five-Style Model **Rational (**Decision makers seek comprehensive information searches and evaluate alternatives logically**)** **Intuitive (**Decision-makers rely on instinct and hunches using situational awareness rather than systematic analysis**)** **Dependent (**Seeking opinions and advice to justify decisions**)** **Avoidant (**People procrastinate and avoid decisiveness**)** **Spontaneous (**Make quick decisions **)**	[1] [2] [3] [4] [5] [6] [7] [8]
Some approaches equate decision-making style to the amount of information obtained and number of alternatives considered **Decisive** (values action, quickness, effectiveness, consistency) **Flexible** (interest in speed, change plans) **Systemic** (integrative and hierarchic nature, frame problems broadly)	[2] [3] [6]
Analytic, Behavioral, Conceptual, and Directive	[3] [4] [9]
The concept of heuristics	[10] [11]
Avoidant, Dependent, Confident, Anxious, Vigilant, Spontaneous, Intuitive, Brooding, and Respected	[12]
[1]Gambetti & Giusberti, 2019 [2] Kurban & Yaşar, 2017 [3] Loo, n.d. [4] Oyewobi et al., 2016 [5] Strang et al., 2023 [6] Thunholm, 2004 [7] Scott & Bruce, 1995 [8] Strang et al., 2023) [9] Rowe and Mason.,1987 [10] Tversky and Kahneman.,1974 [11] Nicholas and Steyn, 2020 [12] Mohd et al., 2019 [13] Esa et al., 2014	

From the above classifications, based on both practical suitability and theoretical alignment with the construction management context, the author selects: directive, analytical, conceptual, and behavioural styles. These four styles capture fully the diverse cognitive and social approaches managers frequently display when confronted with time-sensitive, complex, and resource-limited situations (Kurban and Yasar, 2017) and in terms of cognitive complexity and values orientation (Rowe and Mason, 1987).

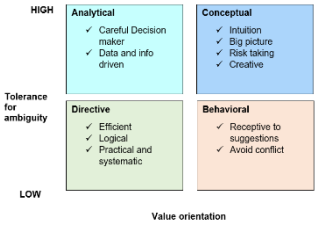


Figure 0-1 Selected 4 decision Making styles

Understanding schedule variance in construction.

Scheduling is pivotal to construction project success (Derbe et al., 2020). Developing a solid construction project schedule (CPS) is one of the main determinants of project success(Derbe et al., 2020). Time and cost management are among the most conspicuous elements of project management(Chin & Hamid, 2015). Effective time management is paramount for construction projects(Chin & Hamid, 2015) and entails planning, monitoring, control, estimating, scheduling, and schedule control (Borges, 2018; Chin & Hamid, 2015).

Despite its importance, the state of time management in the construction industry, at least in some settings, is said to be mostly low, which leads to issues such as late completion and time overrun (Borges, 2018; Przywara & Rak, 2021; Solís-Carcaño et al., 2015) Delays are described as insidious, which more often than not lead to time overrun, cost overrun, disputes, and even project abandonment(Chin & Hamid, 2015).

Schedule variance indicates the deviation between the planned schedule and actual progress or completion time(Adepu et al., 2023; Borges, 2018). It can be influenced by various variables, from schedule reliability and choice of CPS methods to project complexity, schedule risks and their predictability, project financing, team coordination, availability of materials, and uncontrollable events such as weather (Castro-Lacouture et al., n.d.).

Earned Value Method (EVM) metrics are a few methods that can be applied to monitor and analyse time and cost deviations, for instanc Forecasting schedule variance is a meaningful field, and methodologies are required to help project managers in this estimation Calculation of delay claims is often aimed at determining each party's liability and measuring the effects, which requires a number of delay analysis techniques, time deviations from the schedule (T/S)(Robert Carr, n.d.; Shehu et al., 2023) (Rauzana, 2016; Yoon & Brett-MacLean, 2019)

RESEARCH METHODOLOGY

The author adopted a mixed-methods approach in this research, allowing qualitative and quantitative evidence in the scope. The study employed a deductive approach, combining theory-driven insights with empirical data. A comprehensive literature review initially explored the various decision-making styles identified from existing research. From the wide range of styles managers frequently use, the author selected: directive, analytical, conceptual, and behavioural for further investigation.

Under phase I, semi-structured interviews were conducted with identified managerial-level professionals in building construction projects to collect qualitative data, as interviews are suitable for eliciting facts, knowledge and opinions from selected respondents (Yang et al., 2014). The questions were focused on decision-making styles in use and their influence on project timelines and contextual factors affecting decision-making. The qualitative data were then analysed using code-based analysis to identify themes and patterns based on the responses.

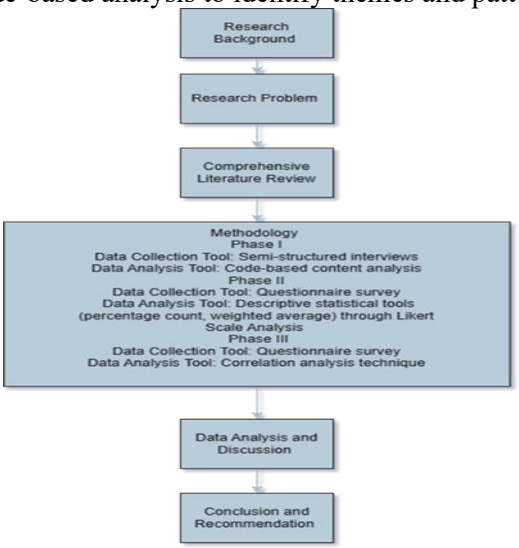


Figure 0-2 Research process

The author administered a questionnaire survey under phase II, incorporating the literature review and phase I findings. Questionnaire surveys are a standard method for evaluating concepts by gathering information through standardised questions. The questionnaire was distributed among a broader group of construction

professionals, incorporating a Likert scale to gauge the frequency and perception of each decision-making style and how it influences schedule variance. The collected data were then analysed using descriptive statistical measures to summarise central tendencies, variability, and overall response patterns.

Likert Scale analysis is one of the tools used to quantify people's attitudes, opinions, or habits, which indicate the level of agreement on a scale for statements (Awang et al., n.d.). The scale typically provides options such as "Very High," "High," "Moderate," "Low," and "Very Low" to record a range of responses (Nemoto & Beglar, n.d).

$$Weighted\ mean = \frac{\sum_{i=1}^{n} \omega_i \cdot x_i}{\sum_{i=1}^{n} \omega_i}$$

In here,
ω=Frequency of responses
x =Rating of each factor (Gammanage & Gunarathna, 2022)
Correlation analysis technique in the excel used to establish the relationship between the decision making styles and the schedule variance(Selvanathan et al., n.d.).

Scale of correlation coefficient	Value
$0< r \leq 0.19$	Very Low Correlation
$0.2 \leq r \leq 0.39$	Low Correlation
$0.4 \leq r \leq 0.59$	Moderate Correlation
$0.6 \leq r \leq 0.79$	High Correlation
$0.8 \leq r \leq 1.0$	Very High Correlation

Figure 3-3 The table of correlation coefficient

DATA COLLECTION AND ANALYSIS

Below charts, figures and tables show how analysed the collected data from the Semi structured interviews and from the questionnaire survey. From the interviews found which decision-making styles are mostly use in the industry, which is the best decision-making styles for Sri Lankan context, how decision-making styles impact on the schedule variance. From the questionnaire survey found that frequently using decision making styles and whether that decision making styles change on the phases like information.

Table 0-2 Demographic data of the interviwers

Role in the Industry	Experience in the industry	Experience as a Project Manger	Experience as Managerial Level Profettional in the
Project Manger	32	12	32
Architect	5	2	2
Project Manger	7	6	5
Project Manger	30	17	22
Project Manger	33	4	19
Project Manger	5	1	1
Project Manger	25	16	20
Quantity survey	19	17	17
Project Manger	28	25	18

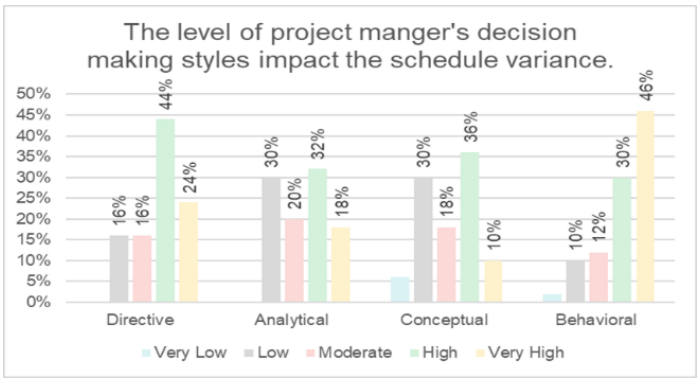


Figure 0-4 To what level decision making styles impact the schedule variance.

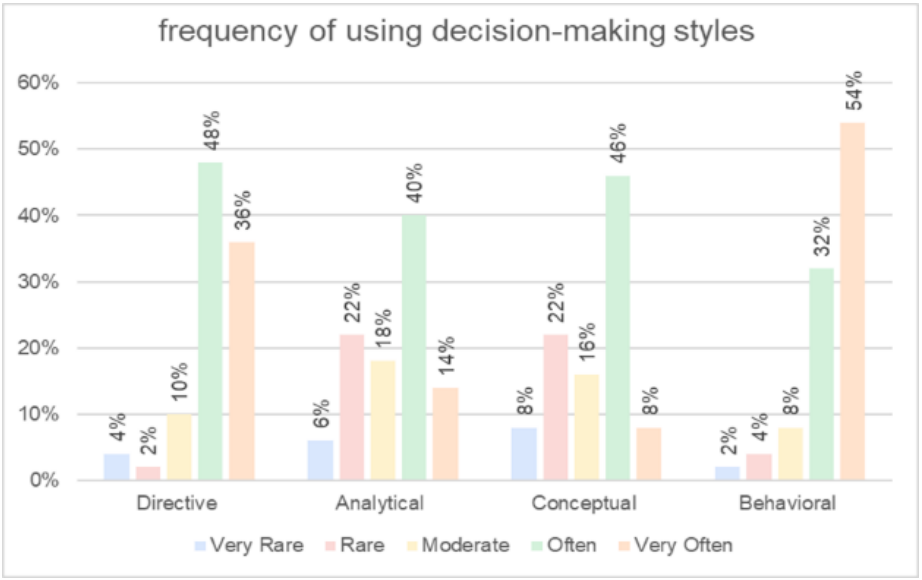


Figure 0-5 Frequency of using decision-making styles in building construction projects

Below Tables show how interviews show the impct on sheduel varive due to the decision making styles.

Table 0-3 Positive impact on schedule variance due to the decision making style

	Directive	Analytical	Conceptual	Behavioral
R1	Allows rapid decision-making in critical moments, helping to keep the project on track during urgent situations.	Ensures decisions are well thought out and backed by data, which helps prevent mistakes that could cause future delays.	Encourages forward-thinking and innovative approaches, helping anticipate and mitigate long-term scheduling risks.	Promotes team collaboration and strong communication, leading to a more committed workforce and fewer misunderstandings, which keeps the schedule on track.

R2	Helps keep the project on track by making fast decisions, especially useful during site changes or urgent design issues	Ensure that decisions are well thought out, minimizing errors and rework, which stabilizes the project timeline	Allows for innovative solutions that can preempt future delays and make the overall project smoother in the long term	Improves team collaboration and communication, which helps in resolving on-site problems quickly and avoiding misunderstandings
R3	Strong Leadership in High-Pressure Situations	Accurate planning through data	Creative problem solving	Reduces miscommunication
R4	Fast decision-making based on experience leads to fewer delays; especially beneficial in solving urgent site issues; has a high positive impact on schedule variance.	Thorough analysis helps avoid mistakes and rework; improves accuracy of plans and reduces risks.	Brings innovation and long-term solutions; enhances coordination among stakeholders (used mainly by client-side project managers and consultants).	Improves team morale, collaboration, and reduces conflicts; creates a supportive working environment.
R5	Helps in making quick decisions under urgent or unexpected conditions (e.g., land ownership dispute in Galle project), which prevents project standstill.	Enables evaluation of technical problems (e.g., sand shortage) and supports sound alternatives like re-scheduling tasks.	Aids in project-wide strategic planning and aligns decisions with long-term objectives.	Encourages teamwork when dealing with unexpected on-site issues, although not regularly used.
R6	Enables fast action when urgent decisions are required (e.g., design change or government policy shift), which avoids schedule pauses.	Helps in maintaining accuracy and alignment with schedules through evidence-based planning.	Ensures alignment with project vision and goals, aiding in strategic decision-making.	Sometimes improves coordination and team engagement, especially in collaborative tasks.
R7	This type is required where the project requres quick and fast decisions are required in urgent situations. In these moments, the project manager shall use his experience and shold be able to get some risks in calculated manner	This type is especially useful for initial project planning and perticularly for Macro level planning & decision making. The analytical data shall be useful for fine-tuning and mitigation of variances during the progress.	This type can be useful specially for initial planning stage. In modern projects this type is not frequently used. But this may useful according to the situations.	This type is very important since the experiance of the whole management team can be brought into account. Very useful for handling long duration projects, but, I do not think this is useful when urgent decision making is required.
R8	Enables fast solutions to contractor-side issues and external disruptions like weather (e.g., delivering apartments quickly despite rain)	Helpful in continuously delayed projects; decisions based on analyzed data bring accuracy	Can be used to strategize during delays; helps bring long-term solutions with proper framework	Reduces disputes, improves collaboration, and helps in concurrent problem-solving (e.g., delays from both parties)

R9	Can get rapid decision without delay			This improves team members knowledge about project sheduels.

Table 0-4 Negative impact on schedule variance due to the decision-making styles

	Directive	Analytical	Conceptual	Behavioral
R1	May overlook key details due to the focus on speed, which can result in issues that later disrupt the schedule	Decision-making can become too slow, causing bottlenecks when quick action is required.	Might divert attention from immediate project needs, focusing more on vision than on pressing site issues.	Could slow down decision-making if too much emphasis is placed on team consensus, potentially delaying progress when quick decisions are needed.
R2	Sometimes decisions are made too quickly without fully analyzing design complexities, which can lead to mistakes that require fixing later.	Decision-making can be too slow, especially when waiting for all data, which can delay interior work start dates.	May overlook practical constraints in pursuit of big ideas, potentially causing delays in approvals or implementation.	Can lead to over-consultation and indecision, delaying timelines when too many opinions are factored in.
R3	May make decisions too quickly without fully considering all variables or stakeholder input, leading to missteps or rework.	Decision-making can be too slow, especially in time-sensitive situations, causing delays.	Long-term focus may delay urgent, shortterm decision-making needed to stay on schedule	Trying to satisfy everyone could result in compromises that extend timelines.
R4	Sometimes leads to ignoring alternatives or stakeholder input, but overall, it tends to have more positive than negative impact.	Slow decision-making can delay progress; over-analysis may create bottlenecks.	If overused without practical grounding, it can create confusion or unrealistic expectations.	Too much focus on consensus may slow decisions and dilute accountability.
R5	While fast, it may lead to costly decisions (e.g., shifting tasks during sand delay results in additional expenses).	Can delay decisions if overused in time-sensitive scenarios	May not be practical alone; if not grounded in current site reality, it could misalign progress.	Rarely used, but in sudden team-led decisions, lack of strong authority might cause delays.
R6	Hasty decisions might increase workload or costs (e.g., sudden column addition for variation causes delay).	May delay actions due to time spent collecting and analyzing data.	If too focused on the concept, can ignore practical issues leading to schedule overruns.	Team-dependent and influenced by personal attitudes, which can disrupt schedule flow if conflict arises.

R7	In this type, it is very inportant to focus on the situation. If the decision maker is not capable to consider the whole project and the impact to be imposed on the each and every activity, the situation may be more difficult.	Due to the heavy dependent on the statistics data, it is very important to gather with the sufficintly accurate data. Due to the required urgency of decision making, the accuracy of the outcomes may low and may leads to enhance the variance.	Since this is more dependent on the skilfulness of the PM, this is very personel dependent. May lose the attention on most important facts.	Like above, this type is more dependent on the team attitudes. This may take long time to come up with a decisions.
R8	May lead to resistance from experienced professionals if their ideas are ignored; causes delay due to lack of collaboration	Can delay decisions due to time required for data collection and analysis	May result in unfocused or slow decisions; teams might take multiple days to finalize	Similar to conceptual; can delay actions if team lacks focus or decision-making clarity
R9	Team members may feel their contributions are unimportant	slow decision making can delay project shedule	lack of practical in real situations	sometimes it gets difficult to manage conflicting perspectives

Table 0-5 Best decision-making style for Sri Lankan construction context

	Which is best for Sri Lankan building construction projects.
R1	The Behavioral style fits best for building construction projects in Sri Lanka. This is because the success of a project largely depends on team collaboration and mutual understanding among all stakeholders.
R2	A combination of directive and behavioral styles is best for Sri Lanka's construction sector. Quick decisions are often needed on-site, but effective communication and collaboration are equally critical to keep projects on schedule
R3	Combination or situational decisionmaking approach is well-suited for construction projects in Sri Lanka, where each project may come with unique challenges and diverse stakeholders.
R4	In practice, a combination or situational approach is the most suitable for Sri Lankan building construction projects. In large, highly structured projects with the involvement of all three parties (client, consultant, contractor), all four styles are needed and usually applied.
R5	Directive style most suitable for Sri Lankan construction because many projects require fast decisions by top-level managers. Highly practical during issues like ownership conflicts or material delays. However, overuse may ignore cost implications.
R6	In Sri Lankan context, all four decision-making styles are essential and should be applied situationally.
R7	Combined with Directive Style is the KEY. The projects in Sri Lanka have some additional constrains like "Cost Constrains", "Insufficient skilled work forces", "Client's influences", etc, and those things may be some additional pressure on PM team. Therefore, specially projects in Sri Lanka requre lots of micor level attention is required. Therefore, Derective Style will be the key type. However, it is required to use the Behevioral Type in force as the main supportive stayle.
R8	All four styles must be used as per the PM team, Project Type and the situation.
R9	Directive and Behavioral styles. If use both high efficiency of decision making. Employees get clear directions without having confusions.

According to the majority, behavioural patterns are the best, but some confirm that directive is the key decision-making style, and, combined with directive, others can be used situationally. This study conducted a correlation analysis to establish the relationship among the variables, which was determined based on Figure 4.1

Table 0-6 Corelation Table

Frequancy of project mangers use of decision making style with the level of Impact on the sheduel variance		
Style	Correlation coefficient	Value
Directive	0.325113051	Low
Analytical	0.505118572	Moderate
Conceptual	0.665739505	High
Behavioral	0.678027879	High

Correlation analysis indicated that the different decision-making styles influence schedule variance in building construction projects to varying degrees.

These findings highlight that strategic, inclusive, and collaborative styles of decision-making are more effective in managing project schedules than task-oriented styles.

DISCUSSION

A study in Sri Lankan building construction projects, involving interviews with nine experts and a survey of 50 practitioners, examined how project managers' decision-making styles, Directive, Analytical, Conceptual, and Behavioural styles, impact schedule variance. Directive and Behavioural styles dominate due to their fit for fast-paced, teamwork-driven environments, while Analytical and Conceptual styles are less used due to time constraints. All styles influence project schedules, with Behavioural and Directive boosting efficiency and coordination, and Analytical and Conceptual aiding long-term planning, sometimes causing delays. Behavioural and Conceptual styles strongly correlate with reduced schedule variance, followed by Analytical and Directive styles. Positive effects include enhanced efficiency, communication, and strategic planning, while adverse effects involve hasty decisions, misunderstandings, over-analysis, and impractical planning, with styles varying by project phase.

CONCLUSION

A study on decision-making styles (DMS), directive, analytical, conceptual, and behavioral in Sri Lankan construction projects, revealed their significant impact on schedule variance. directive and behavioral styles were most frequently used and effective, enhancing team coordination and timely project completion, while analytical and conceptual styles supported planning but risked delays. The suitability of DMS, particularly Behavioral and Directive, depends on project size, complexity, and team dynamics, with Behavioral showing the strongest correlation with reduced schedule variance. Poor decision-making was linked to delays, rework, and cost overruns, while effective DMS improved efficiency and communication. Recommendations include training managers in adaptive DMS, using tools like Earned Value Management, applying phase-specific strategies, and raising awareness of potential negative effects like rushed decisions or over-analysis. Future research should expand surveys, explore DMS interactions with leadership styles or biases, and incorporate objective measures for broader insights.

ACKNOWLEDGEMENTS

The researcher acknowledges Liverpool John Moores University (LJMU) and the Sri Lanka Institute of Information Technology (SLIIT) for their invaluable support in providing university resources that facilitated this study.

REFERENCES

[1] Awang, Z., Afthanorhan, A., & Mamat, M. (n.d.). The Likert scale analysis using parametric based Structural Equation Modeling (SEM)

[2] Belout, A., & Gauvreau, C. (2004). Factors influencing project success: The impact of human resource management. International Journal of Project Management, 22(1), 1–11. https://doi.org/10.1016/S0263-7863(03)00003-6

[3] Castro-Lacouture, D., Asce, A. M., Gürsel, ;, Süer, A., Julian Gonzalez-Joaqui, ;, & Yates, J. K. (n.d.). Construction Project Scheduling with Time, Cost, and Material Restrictions Using Fuzzy Mathematical Models and Critical Path Method

[4] Davis, D. L., Grove, S. J., & Knowles, P. A. (1990). AN EXPERIMENTAL APPLICATION O F PERSONALITY TYPE AS AN ANALOGUE FOR DECISION-MAKING STYLE '. In @ Psychological Reports (Vol. 66)

[5] Esa, M., Alias, A., & Samad, Z. A. (2014). Project Managers' Cognitive Style in Decision Making: A Perspective from Construction Industry. International Journal of Psychological Studies, 6(2). https://doi.org/10.5539/ijps.v6n2p65

[6] Gambetti, E., & Giusberti, F. (2019). Personality, decision-making styles and investments. Journal of Behavioral and Experimental Economics , 80, 14–24. https://doi.org/10.1016/j.socec.2019.03.002

[7] Gammanage, S. H., & Gunarathna, N. (2022). DELAYS AND DISRUPTIONS IN THE CONSTRUCTION INDUSTRY DURING THE GLOBAL PANDEMIC. World Construction Symposium, 326–338. https://doi.org/10.31705/WCS.2022.27

[8] Kalambayi, J. L., Onojaefe, D. P., Kasse, S. N., & Tengeh, R. K. (2021). The influence of leadership styles on employee performance in construction firms. EUREKA: Social and Humanities, 5, 34–48. https://doi.org/10.21303/2504-5571.2021.002037

[9] Kurban, C., & Yaşar, M. (2017). Decision Making Styles Of School Managers According To Their Individual Perception. Kuramsal Eğitimbilim, 10(1), 170–190. https://doi.org/10.5578/keg.26430

[10] Loo, R. (n.d.). A psychometric evaluation of the General Decision-Making Style Inventory. www.elsevier.com/locate/paid

[11] Nemoto, T., & Beglar, D. (n.d.). Developing Likert-Scale Questionnaires Campus Reference Data.

[12] Oyewobi, L. O., Windapo, A., & Rotimi, J. O. B. (2016). Relationship between decision-making style, competitive strategies and organisational performance among construction organisations. Journal of Engineering, Design and Technology, 14(4), 713–738. https://doi.org/10.1108/JEDT-04-2015-0025

[13] Prabhakar, G. V., Yaseen, A., Prabhakar, G. V, & Yaseen, A. (2016). Decision-making styles and leadership: evidences from the UAE. In Int. J. Management Development (Vol. 1, Issue 4).

[14] Robert Carr, B. I. (n.d.). COST, SCHEDULE, AND TIME VARIANCES AND INTEGRATION

[15] Saed, A., Yong, K., & Othman, M. (n.d.). Project Complexity Influence on Project management performance-The Malaysian perspective. https://doi.org/10.1051/00065

[16] Scott, S. G., & Bruce, R. A. (1995). Decision-Making Style: The Development and Assessment of a New Measure. Educational and Psychological Measurement, 55(5), 818–831. https://doi.org/10.1177/0013164495055005017

[17] Selvanathan, M., Jayabalan, N., & Hussain, N. (n.d.). Employee Productivity in Malaysian Private Higher Educational Institutions. https://doi.org/10.48080/jae.v17i3.50

[18] Shanian, A., & Savadogo, O. (2006). A material selection model based on the concept of multiple attribute decision making. Materials and Design, 27(4), 329–337

[19] Strang, M., Miller, M., Hill, R., & Elshaw, J. (2023). Group decision performance: The predictive role of decision making styles and cognition. Personality and Individual Differences, 206. https://doi.org/10.1016/j.paid.2023.112114

[20] Thunholm, P. (2004). Decision-making style: Habit, style or both? Personality and Individual Differences, 36(4), 931–944. https://doi.org/10.1016/S0191-8869(03)00162-4

[21] Zavadskas, E. K., Turskis, Z., & Tamošaitiene, J. (2010). Risk assessment of construction projects. Journal of Civil Engineering and Management, 16(1), 33–46. https://doi.org/10.3846/jcem.2010.03

High-Speed Path Tracking of a Small Mobile Robot Using PD and ADRC Controllers with Experimental Validation

Pasan Senarath[1] , K. Jayawickrama C. Kumara[2]
[1]Department of Electrical and Electronics Engineering
[2]Department of Mechanical Engineering
Faculty of Engineering, Sri Lanka Institute of Information Technology
senarathpasan@gmail.com , Kumara.j@sliit.lk

ABSTRACT

This paper presents the development, modeling, and validation of a compact ground-based mobile robot designed for high-speed trajectory tracking with low-cost hardware and advanced control strategies. Originally designed for micromouse competitions, the platform was upgraded with a vacuum-assisted friction enhancement mechanism and a lightweight sensor fusion system comprising Time-of-Flight (ToF) sensors, infrared (IR) proximity sensors, a 6-axis IMU, and wheel encoders. Due to the lack of manufacturer-provided system parameters, a data-driven system identification approach was employed to derive a MIMO state-space model representing the robot's dynamics. Two control strategies, Proportional-Derivative (PD) and Active Disturbance Rejection Control (ADRC) were implemented and evaluated through both numerical simulations and real-world experiments across three benchmark trajectories: straight-line, 90° smooth turn, and figure-eight. The results show that both controllers achieved trajectory tracking within 5% RMSE, with ADRC offering improved heading accuracy and energy efficiency. Experimental observations indicate that ADRC reduces battery current fluctuations, although current modeling was not included in the control design. The proposed platform and methodology offer a cost-effective and robust solution for mobile robot control in constrained environments, with future work focusing on energy-aware control integration.

Keywords: *Mobile Robot, System Identification, PD Controller, ADRC, Sensor Fusion, Vacuum Traction*

INTRODUCTION

Small ground-based mobile robots, particularly differential-drive platforms originally designed for micromouse competitions, have attracted growing interest for applications in indoor inspection, swarm coordination, and educational robotics. These platforms offer advantages in cost, accessibility, and maneuverability in constrained spaces. However, achieving high-speed (≥ 1 m/s) and high-precision motion remains challenging when relying on off-the-shelf components that lack well-characterized dynamics (Ušinskis, Nowicki, Dzedžickis, & Bučinskas, 2025). Model-based control traditionally depends on accurate mechanical and actuator parameters, yet low-cost motors and mechanical tolerances in small platforms introduce variability that undermines performance. This impedes the application of classical controllers in transitioning from simulations to real-world settings (Norzalilah & Samad, 2023). A particularly novel aspect of our work is the integration of a vacuum-based friction enhancement, a lightweight and low-power method increase normal force, thereby improving traction during aggressive maneuvers. While vacuum adhesion is well-established in wall-climbing robotics, its application to planar traction control in fast-moving ground robots is largely unexplored (Yuan, Chang, Song, Lin, & Jing, 2024). Furthermore, sensor fusion strategies using IMU, Time of Flight sensors and infrared (IR) sensors have shown promise in localization and navigation (Ušinskis et al., 2025), but their deployment in high-speed miniature platforms is under-investigated. Fusion of simple, low-cost sensors offers a feasible alternative to bulkier systems such as LIDAR or stereo vision (Ušinskis et al., 2025). In terms of control architectures, while Proportional-Derivative (PD) and Active Disturbance Rejection Control (ADRC) are well-studied in theory, practical benchmarks on physical, small-scale robots navigating curves and complex trajectories at high speeds remain scarce (Muñoz-Hernandez, Díaz-Téllez, Estevez-Carreon, & García-Ramírez, 2023).
This paper presents:

A compact $60 \times 40 \times 25$ mm differential-drive robot, modified with a vacuum-assisted downforce mechanism to mitigate wheel slip.

A sensor fusion system combining IR and IMU data for real-time obstacle avoidance and heading estimation.

A data-driven system identification approach for straight and curved motions, enabling adaptive control.

Comparative experimental validation of PD and ADRC controllers in simulation and on real hardware, demonstrating precise path-following (> 1 m/s) on straight lines and figure-eight trajectories.

Together, these contributions address critical gaps in model identification, friction control, sensor fusion, and

control validation for high-performance small mobile robots. The results point to substantial potential for deploying such platforms in inspections, swarm systems, and competitive robotics.

MATERIALS AND METHODS

This section outlines the development process of the compact mobile robot platform, the sensor and actuation systems, the system identification procedure, and the control algorithm implementation (PD and ADRC). The methodology is structured in sequential stages to align with standard practices in mobile robotics research, integrating novel hardware and rigorous control implementation.

Robot Platform and Hardware Configuration

The mobile robot measures 60 × 40 × 25 mm and utilizes a four-wheeled differential-drive setup built from low-cost components. The chassis is implemented on the PCB itself, balancing rigidity and lightweight design. Key components:

Actuators and gear system: Two 0714 coreless motors, customized motor mount, gear system, tires, rims, bearings for differential drive.
Sensing system: Time of flight sensors (infrared sensors) for obstacle proximity detection (front and sides), magnetic encoders, inertial measurement unit (IMU) for heading angle and linear acceleration feedback.
Friction Control Mechanism: A miniature coreless motor-driven fan positioned beneath the chassis generates a vacuum skirt to enhance downward force, improving wheel grip during rapid turns and linear motion.
Processing and Control: A microcontroller unit (ESP32S3PICO) is used for data acquisition, motor control, and executing control algorithms.
Motor drivers and power supply: Two motor drivers and 3.3V step down buck converter, two 1S 150 mAh LiPo cells connected in series (2S, 7.4 V, 150 mAh) to power the robot.

Figure 1 shows an annotated hardware diagram of the mobile robot from top and bottom views.
System Identification Procedure
Precise dynamic parameters were not available from the off-the-shelf components. A data-driven system identification was conducted to obtain accurate motion models, aligning with recent methods in differential-drive robotics (Guffanti et al., 2025; Siwek et al., 2023).

Experimental Setup and Data Acquisition

Two motion profiles under open-loop control were executed: (a). Straight-line motion for identifying translational dynamics. (b). Known left-turn motion for angular response identification. In both cases, a unit-step PWM command (0–255) was issued to the motor driver via the microcontroller, while position, velocity, and heading data were recorded at 100 Hz without any active control algorithm. Sensor data was processed offline to estimate key dynamic variables (velocity, acceleration, heading angle). Noise filtering and low-pass filtering were applied to improve signal quality. The vacuum system was operated at a constant power level throughout all tests. At the tested speeds, wheel slip was considered negligible, as the vacuum-generated downforce provided sufficient traction to counteract potential slipping forces.

Transfer Function Derivation and Modelling

Using input-output data from both motion types, system transfer functions were derived through curve fitting and regression using python library system identification toolchain. Figure 2 illustrates the system identification block diagram. Individual transfer functions for each motor and simplified linear velocity and angular velocity models were obtained following Guffanti et al. (2025). The identified model served as the baseline for controller design. Using this method, the robot was modeled as a simplified 2-input, 2-output (MIMO) system.

Control Algorithm Design

Two controllers were implemented based on the identified model: Proportional-Derivative and Active Disturbance Rejection Control. The following sub-sections describe the details of the control algorithm design.

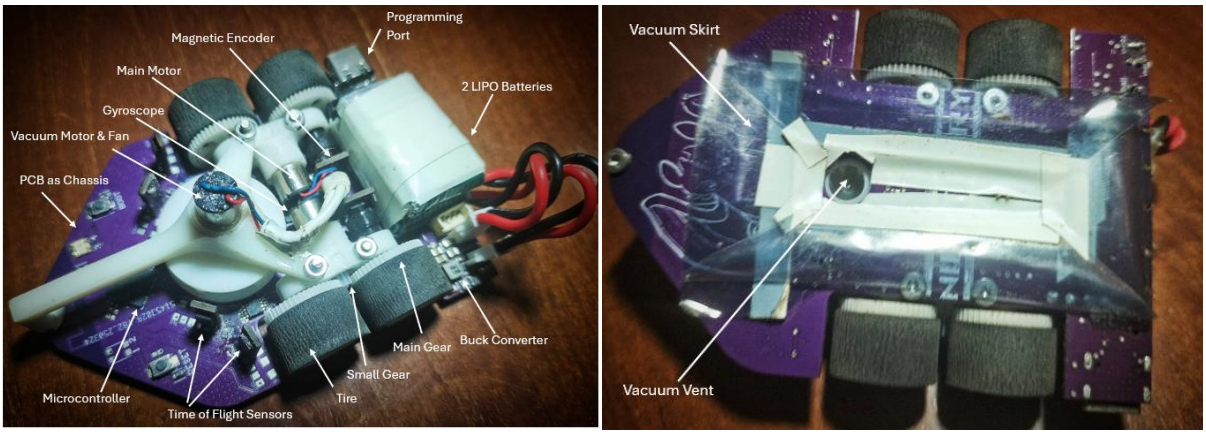


Figure 6. Mobile robot platform

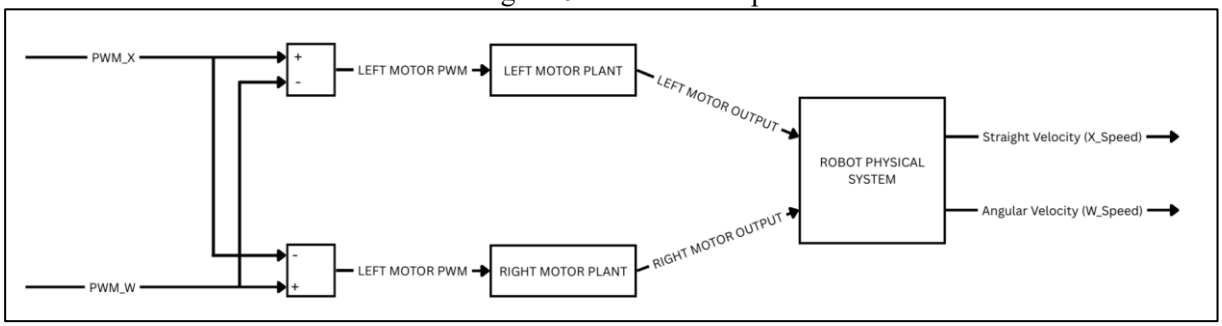


Figure 2. Block diagram of the system identification process.

Proportional-Derivative (PD) Controller

A standard Proportional-Derivative (PD) controller was implemented to regulate the heading angle and position of the mobile robot. The control law is defined in Equation (1):

$$u(t) = K_p e(t) + K_d \frac{de(t)}{dt} \quad (1)$$

where $e(t)$ represents the instantaneous error (in position or heading), and K_p and K_d are the proportional and derivative gains, respectively.

Gains were empirically tuned using step-response data obtained from the identified model. The tuning objective was to achieve fast settling time, minimal overshoot, and adequate stability. This approach follows well-established procedures in robotic path tracking (Hussein & Al-Khafaji, 2022) and has proven effective in similar mobile robot applications (Wu & Cao, 2011).

Active Disturbance Rejection Control (ADRC)

The successful use of ADRC in two-wheeled and differential-drive mobile robots has been demonstrated in several recent studies (Muñoz-Hernandez et al., 2023; Wu & Cao, 2011), confirming its applicability to real-world mobile robotic systems with limited sensing and actuation precision. To improve robustness against modeling uncertainties and unmeasured disturbances (such as wheel slip, friction variation, and external forces), an Active Disturbance Rejection Control (ADRC) framework was implemented for both linear and angular velocity regulation. The ADRC controller is composed of three primary components:

Tracking Differentiator (TD): This module smooths the reference input signal and estimates its derivative, enabling better transient performance. It mitigates discontinuities in setpoints and prepares the system for fast response without overshooting.

Extended State Observer (ESO): ESO is the core innovation of ADRC. It estimates both the unmeasured system states and the total disturbance (internal + external), treating the entire unknown dynamics and perturbations as an extended state. This enables real-time compensation and enhances tracking accuracy (Han, 2009; Wu & Cao, 2011).

Nonlinear State Error Feedback (NLSEF): Instead of traditional linear gains, ADRC employs nonlinear gain functions (e.g., fal or sigmoid) to adjust control effort based on the magnitude of the error. This improves control

precision near steady-state and limits aggressiveness when errors are large. The control law is defined by the equation (2) as:

$$u(t) = u_0(t) - \hat{f}(t) \quad (2)$$

where: $u(t)$ is the control input to the system, $u_0(t)$ is the output of the nonlinear feedback control law, and $\hat{f}(t)$ is the real-time estimated total disturbance from the ESO.

The robot's MIMO system (PWM_X → Speed_X, PWM_W → Speed_W) was controlled using two separate ADRC loops—one for linear and another for angular velocity. Each loop used a second-order ESO based on the estimated state-space model from system identification. Controller tuning followed the guidelines established in the literature (Han, 2009), where the observer bandwidth $\omega_0(t)$ was set 3–5 times higher than the controller bandwidth $\omega_C(t)$ to ensure fast and accurate disturbance estimation. Parameter tuning was refined empirically based on the observed transient response. This ADRC approach allows the robot to track paths accurately even under varying loads or floor conditions, maintain heading stability at high speeds (up to 1.2 m/s), and reject unmodeled disturbances without requiring an exact plant model.

Sensor Fusion

Low-cost infrared (IR) and inertial measurement unit (IMU) sensors were integrated to provide robust state estimation for both obstacle detection and heading control—key requirements for indoor mobile robot deployment. Four multi-directional IR modules provide range measurements for nearby obstacles and path boundary detection. A 6-axis IMU records angular velocity and linear acceleration to estimate orientation and motion dynamics. Encoders are used as a supplement to the IMU for odometry to improve drift correction. Fusion reduced IMU drift and corrected heading using encoder updates, yielding average errors of less than 1° over 10 m trajectories. For obstacle detection, the fused IR sensors reliably detected nearby objects at approximately 0.2 m, enabling timely avoidance actions. Overall, the fusion approach delivered a lightweight, cost-effective solution for indoor navigation, significantly enhancing pose estimation and environmental awareness without requiring expensive sensors like LiDAR or vision systems.

NUMERICAL SIMULATIONS

Numerical simulations were conducted using the empirically identified state-space model and implemented control laws—PD and ADRC. The robot's performance was evaluated on three motion profiles designed to reflect typical navigation maneuvers: straight-line traverse, smooth 90° turn and figure-eight trajectory. To isolate the performance of the control algorithms, obstacle avoidance and proximity sensing were disabled. Feedback was limited to the IMU and wheel encoders, aligning with the physical robot's available sensory inputs. This approach ensured consistent comparison between simulation and experiment. Two simulation block diagrams were created, as shown in Figures 3 and 4. The figure-eight trajectory was designed with a diameter of approximately 144 mm per loop and was chosen to evaluate both curvature handling and heading correction. Similar simulation-based evaluation of mobile robot controllers has been reported in recent studies using figure-eight paths for benchmarking (Masalskyi et al., 2024; Muñoz-Hernandez et al., 2023; Guffanti et al., 2025). The controller performance of PD and ADRC were assessed using 3 key metrics viz position tracking error (RMSE) and heading angle deviation. Results are visualized in section 3.

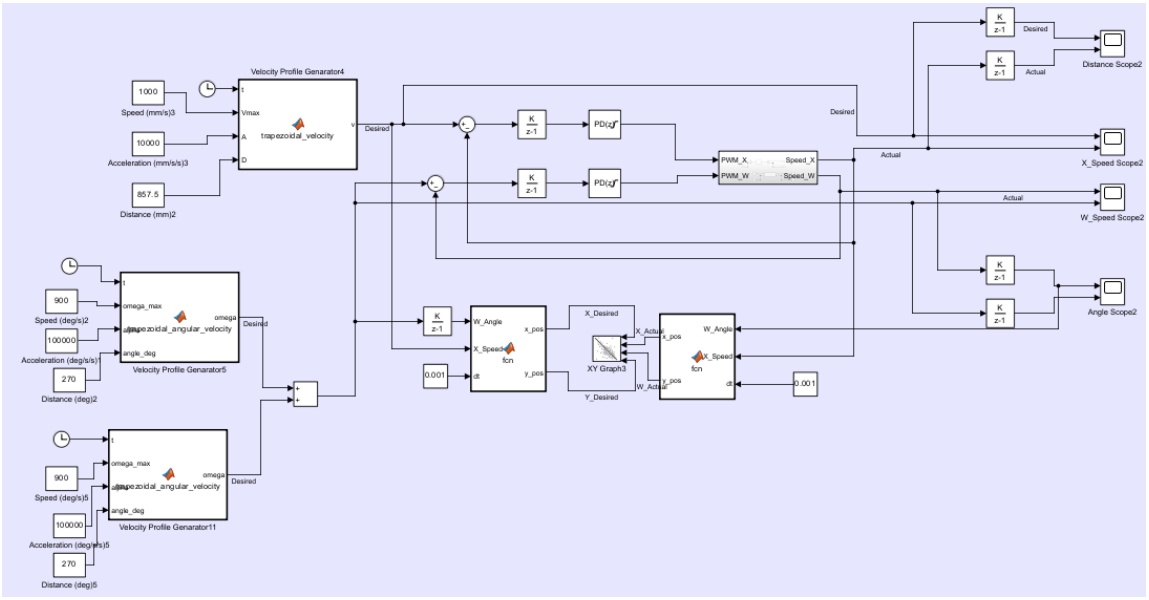


Figure 3. PD control—figure-eight simulation

PHYSICAL EXPERIMENTS

After validating both controllers in simulation, the same path scenarios were recreated in the physical environment using a custom printed track setup. The robot was evaluated on straight-line track (2.5 m), a smooth 90° arc (radius: 75 mm) and figure-eight trajectory (144 mm loops). During experiments, the vacuum-assisted traction system was activated with a constant speed (that is proportional to the vacuum force/grip) to maintain grip and reduce wheel slip at higher speeds. The robot reached speeds of up to 1.2 m/s, even during curved motion, without significant drift. Sensor fusion used only the IMU and encoders, as in simulation. The robot was evaluated based on the four experimental metrics such as RMSE of position against planned trajectory, heading angle error and slip observations. Data were recorded and plots were obtained in offline mode. The results were compared to simulation data and validated using figures for visual correlation.

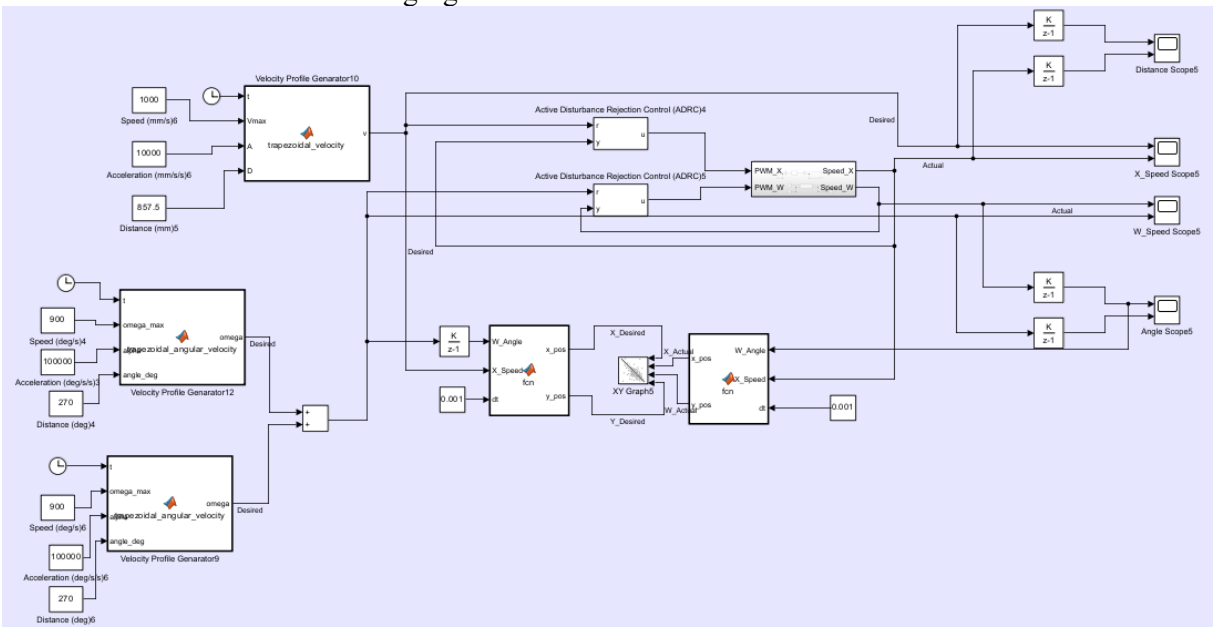


Figure 4. ADRC control—figure-eight simulation

Similar simulation and real-world validation of mobile robot controllers—particularly on bounded or printed figure-eight tracks—have been widely reported in recent research. For example, Masalskyi et al. (2024) and Ušinskis et al. (2025) conducted both simulation-based evaluations and physical experiments using indoor test paths to benchmark sensor fusion accuracy and navigation performance. Likewise, Guffanti et al. (2025) emphasized the role of simulation in modeling traction and slip dynamics, which were later validated through controlled physical trials to assess control reliability at elevated speeds.

RESULTS AND DISCUSSION

This section presents the results of the system identification, state-space modeling, and control validation stages of the mobile robot development. The findings directly support the study's primary objectives: to accurately model a small, low-cost mobile robot using experimental data, and to validate both classical (PD) and modern (ADRC) controllers through simulation and physical trials.

System Identification Results

Due to the absence of reliable modeling parameters from the low-cost, off-the-shelf components (e.g., motors, chassis, and drive systems), analytical modeling was impractical. Instead, a data-driven system identification approach was used to derive a state-space representation of the robot's dynamics. Controlled input signals (unit step PWM commands for translational and rotational drive) were applied to the motors, and the resulting output responses—linear velocity and angular velocity—were recorded using IMU and encoder feedback. Figure 5 illustrates the input signal as PWM value (0 to 255 scale) and output motor speeds measured by encoders in mm/s for the straight-line scenario. Figure 6 shows the input signal as PWM value (0 to 255 scale) and output robot rotational speed ω in mm/s for the 90 degrees smooth turn.

From these time-domain responses, transfer functions and a standard-form discrete-time identified state-space (SS) model were derived (Equation 3). This MIMO model treats the robot as a 2-input, 2-output system, with inputs as PWM signals and outputs as measured velocities. "Coefficient matrices A, B, C, and D were derived, with D consisting entirely of zeros.

$$\dot{x}(t) = Ax(t) + Bu(t) + Ke(t)$$
$$y(t) = Cx(t) + Du(t) + e(t) \quad (3)$$

where $x(t)$is the state vector (position, speed, acceleration), $u(t)$ is the input vector ($[PWM_X \quad PWM_\omega]^T$, $y(t)$ is the output vector $[Speed_X \quad Speed_\omega]^T$ and $e(t)$ is the error matrix.

$$A = \begin{bmatrix} 0.9745 & 0.1683 & 0 & 0 & 0 & 0 \\ -0.1683 & 0.9745 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8955 & 0.1662 & 0 & 0 \\ 0 & 0 & -0.1662 & 0.8955 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.9837 & 0.003119 \\ 0 & 0 & 0 & 0 & -0.003119 & 0.9837 \end{bmatrix}$$

$$B = \begin{bmatrix} -6.576 \times 10^{-5} & -1.348 \times 10^{-5} \\ -6.521 \times 10^{-5} & -3.763 \times 10^{-5} \\ -7.861 \times 10^{-4} & 1.191 \times 10^{-4} \\ 6.796 \times 10^{-5} & -4.759 \times 10^{-4} \\ -2.224 \times 10^{-5} & -1.343 \times 10^{-4} \\ -1.005 \times 10^{-4} & 3.36 \times 10^{-5} \end{bmatrix} \quad C = \begin{bmatrix} 1318 & 98.84 & 108.9 & -69.44 & -5574 & -1898 \\ 79.5 & -163.2 & 411 & 838.1 & -3369 & 1852 \end{bmatrix}$$

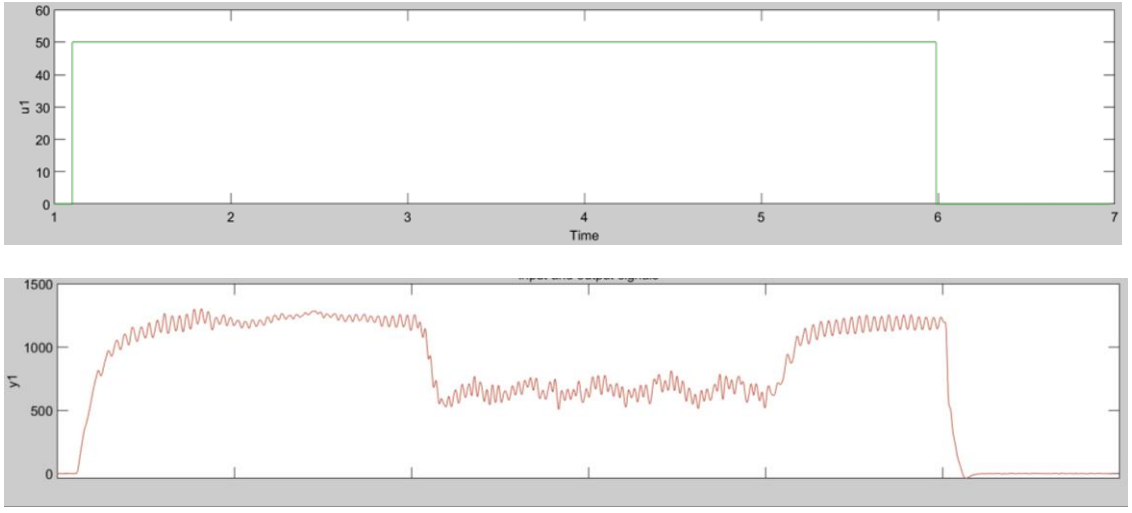


Figure 5. Input (top)-output (bottom) plots for translational motion (PWM_X → Speed_X)

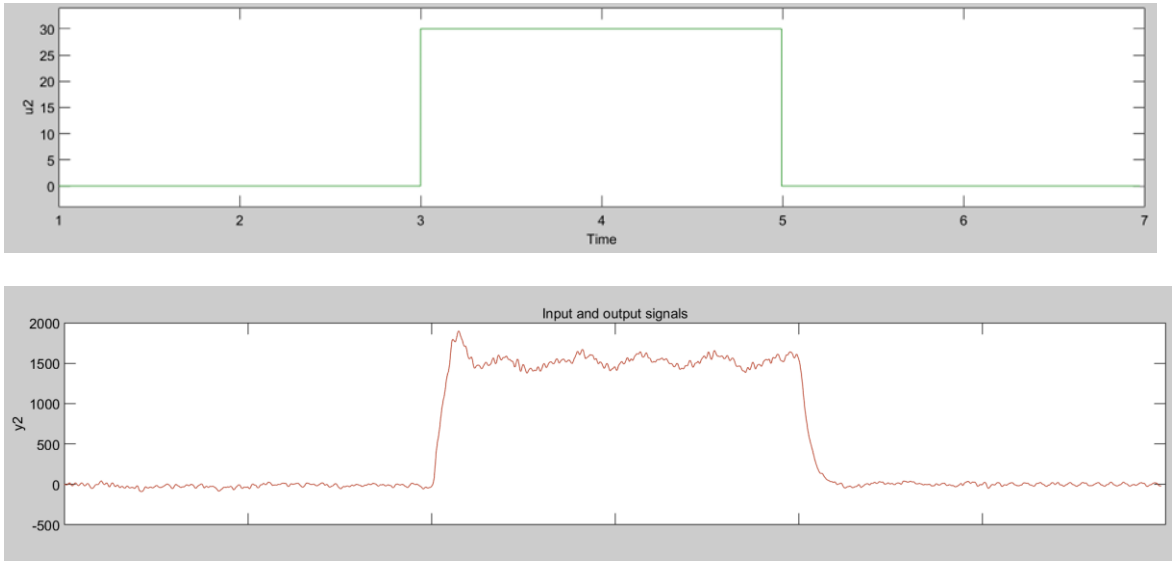


Figure 6. Input (top)-output (bottom) plots for rotational motion (PWM_W → Speed_ω)

This identification method compensates for the lack of component datasheets or manufacturer specifications by producing an experimentally grounded model. It serves as a reliable foundation for controller design (PD and ADRC), simulation-based trajectory testing and physical implementation and benchmarking in real-world conditions. The consistency between simulated and experimental results further validates the accuracy of this modeling approach—an essential outcome, given the hardware limitations. These results strongly align with the overall research objectives: developing an accurate, reliable control framework for a small mobile robot using data-driven modeling techniques, ensuring performance even under uncertainty and component variability.

SIMULATION AND EXPERIMENTAL RESULTS.

To evaluate the performance of both the PD and ADRC controllers, computer simulations and experiments were conducted using the identified state-space model for three representative motion profiles: straight-line motion, smooth 90° turn, and a figure-eight trajectory. The control inputs in all cases were PWM signals for translational and rotational motion, while the outputs were the corresponding linear velocity (Speed_X) and angular velocity (Speed_ ω).

The results are illustrated in Figure 7 and Figure 8 for PD controller and Figure 9 and Figure 10 for ADRC, each containing plots of control input commands and output speed responses for both translational and rotational components. Figure 8 and Figure 10 display the results for the figure eight trajectory, which involves alternating left and right turns. The PD controller can follow the path with acceptable accuracy, though small oscillations are visible in the angular velocity response. Figure 8 presents the response for the figure-eight trajectory. The ADRC controller maintains better trajectory adherence, reduced overshoot, and more consistent velocity tracking through the transitions between loops. Figure 11 and Figure 12 shows the battery current profile for both PD and ADRC controllers which are important in designing the power circuits and optimize the energy use in the mobile robot.

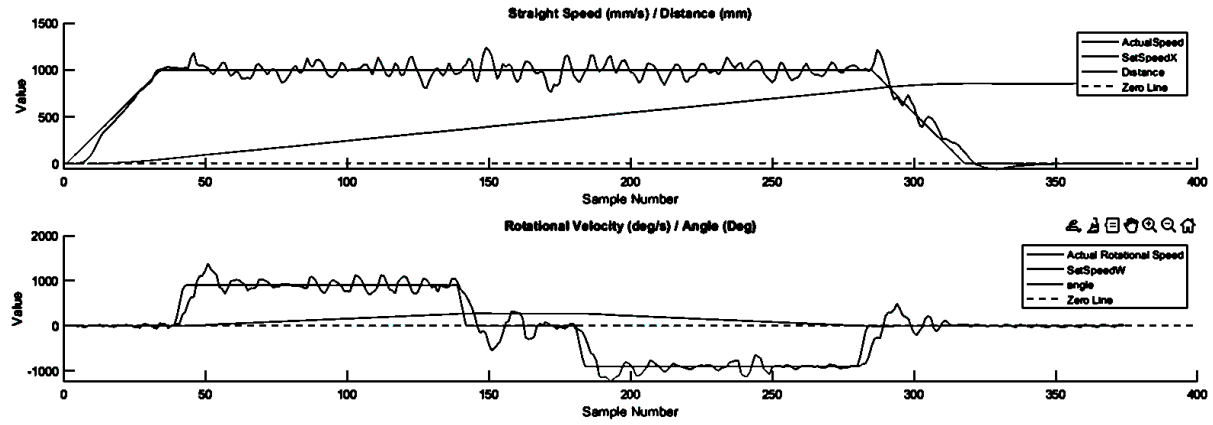


Figure 7. PD controller – (top) Straight Speed (bottom) Angular Speed

DISCUSSION

The performance of the PD and ADRC controllers was evaluated through both numerical simulations and physical experiments. The assessment focused on two key control metrics: position tracking error (RMSE) and heading angle deviation. Additionally, battery current consumption was measured during each test as a practical indicator of controller efficiency and energy performance, even though it was not included in the initial modeling. For all three test trajectories (straight path, 90° turn, and figure-eight), simulation results were used as the desired trajectory references. These references were then compared with actual paths recorded via encoder and IMU data during physical trials. Overall, both PD and ADRC controllers performed satisfactorily in reproducing the desired motion profiles: Position tracking RMSE remained within 5% of path length across all trials. Maximum heading deviations in physical tests reached up to 41% relative to simulated headings, particularly in the more aggressive transitions of the figure-eight path. These deviations were expected, as the simulation did not include certain physical factors such as motor backlash, minor terrain irregularities, and real-time sensor noise. Nonetheless, both controllers successfully ensured stable and consistent tracking, with the ADRC controller showing slightly better compensation in curved motion segments due to its active disturbance estimation capability. Importantly, no significant wheel slip was observed in any of the physical trials, confirming the effectiveness of the vacuum-assisted friction mechanism in maintaining traction, even at higher speeds and during sharp heading changes.

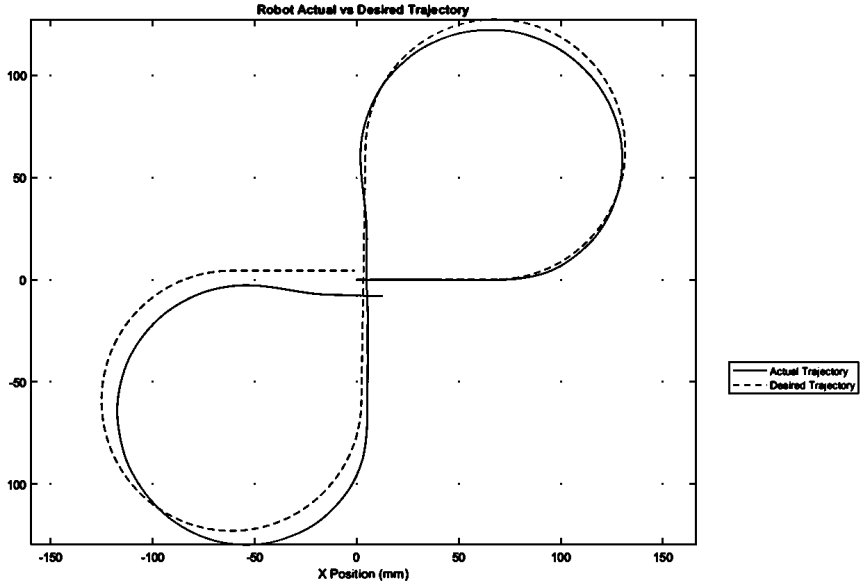


Figure 8. PD controller – Figure-eight trajectory input/output profiles

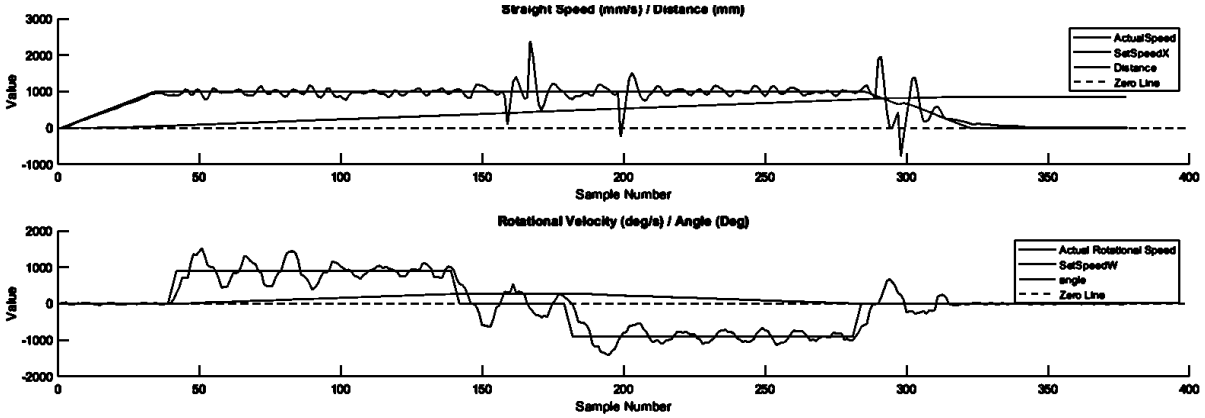


Figure 9. PD controller – (top) Straight Speed (bottom) Angular Speed

Another noteworthy observation was the difference in battery current usage between the two controllers. Although current dynamics were not part of the control model, data was recorded in real-time using a current sensor placed between the battery and motor driver circuitry.

From the results, it can be further compared that saying the PD controller showed slightly higher current fluctuations, especially during sharp directional changes in the figure-eight path (Figure 11 and Figure 12). These peaks suggest momentary instability or overcompensation in control signals. On the other hand, the ADRC

controller, on the other hand, demonstrated smoother current profiles with fewer transients, indicating more efficient handling of disturbances and smoother actuation patterns. This difference implies that, beyond improved tracking performance, ADRC may also offer better energy efficiency, although further study is required to confirm this trend under a wider range of motion scenarios and loads. It is important to emphasize that battery current modeling was not included in the simulation. Thus, these current profiles are purely empirical measurements, highlighting an important direction for future work: incorporating energy usage into the control design for energy-aware mobile robot navigation.

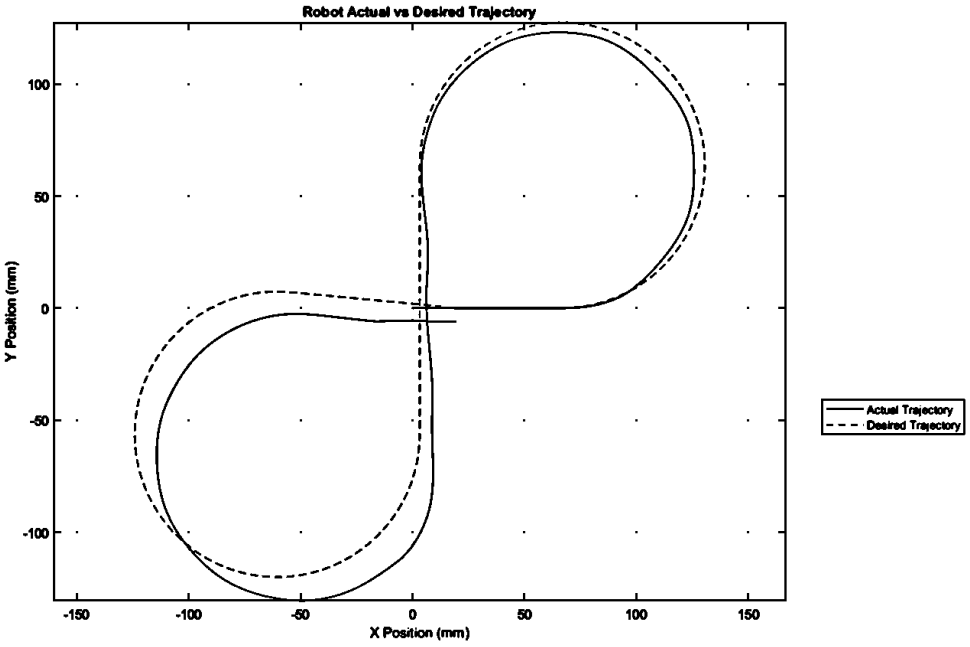


Figure 10. ADRC controller – Figure-eight trajectory input/output profiles

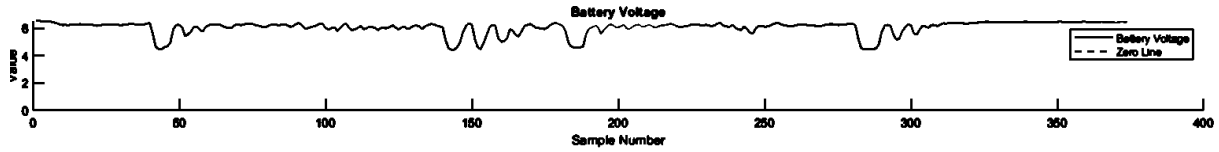


Figure 11: Battery current profile – PD controller

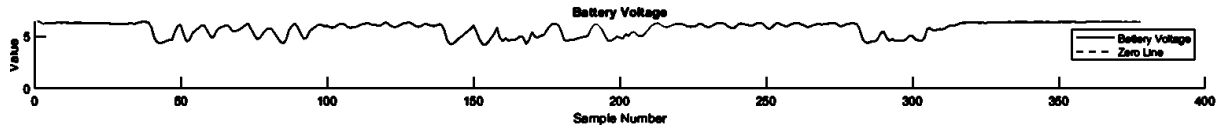


Figure 12: Battery current profile – ADRC controller

CONCLUSION.

This study demonstrated a complete workflow for modeling, control, and validation of a compact mobile robot using a data-driven system identification strategy and two distinct control architectures—PD and ADRC. Simulation and experimental results confirmed that both controllers can achieve accurate trajectory tracking, even with limited hardware resources and unmodeled dynamics. Notably, ADRC showed better performance in curved trajectories, with smoother control effort and improved energy efficiency, as reflected in the reduced battery current fluctuations. By using empirical modeling instead of relying on unavailable datasheets, the research bridges a common gap in the design of affordable mobile robotic platforms. The integration of vacuum-assisted friction enhancement further enabled high-speed operation without slippage, validating the physical robustness of the design.

Future work will address two key limitations: (1) incorporating electrical power consumption into the control framework, and (2) extending validation to a broader range of trajectories and environmental conditions. Additionally, multi-objective optimization combining tracking accuracy and energy usage will be explored for adaptive real-time control.

REFERENCES

[1] Muñoz-Hernandez, G. A., Díaz-Téllez, J., Estevez-Carreon, J., & García-Ramírez, R. S. (2023). ADRC attitude controller based on ROS for a two-wheeled self-balancing mobile robot. *IEEE Access, 11*, 94636–94646. https://doi.org/10.1109/ACCESS.2023.3308948

[2] Ušinskis, V., Nowicki, M., Dzedžickis, A., & Bučinskas, V. (2025). Sensor-fusion based navigation for autonomous mobile robot. *Sensors, 25*(4), 1248. https://doi.org/10.3390/s25041248

[3] Yuan, C., Chang, Y., Song, Y., Lin, S., & Jing, F. (2024). Design and analysis of a negative pressure wall-climbing robot with an omnidirectional characteristic. *Robotica*, Advance online publication.

[4] Norzalilah, N. M., & Samad, Z. (2023). Study of friction compensation model for mobile robot joints. *FME Transactions, 51*(2), 321–330.

[5] Active disturbance rejection control for robotic systems: A review. (2021). *Expert Systems with Applications, 174*, 114772.

[6] Robust ADRC in two-wheeled robots: Experimental study. (2023). *International Conference on Robotics and Mechatronics Proceedings*, 47–52.

[7] Trajectory tracking control for a biomimetic spherical robot. (2021). In *Proceedings of the IEEE International Conference on Mechatronics and Automation* (pp. 951–959).

[8] Robotics perception and control: Key technologies and applications. (2024). *International Journal of Robotics Research, 43*(1), 5–24.

[9] Sensor-fusion based mobile manipulator remote control for smart home. (2025). *arXiv*. https://arxiv.org/abs/2504.13370v1

[10] 3D ToF LiDAR in mobile robotics: A review. (2022). *arXiv*. https://arxiv.org/abs/2202.11025

[11] Hussein, A. K., & Al-Khafaji, H. M. (2022). *Design and Implementation of PID Controller for Path Tracking of Mobile Robot Based on Arduino Microcontroller*. International Journal of Mechanical Engineering and Robotics Research (IJMERR), 11(7), 1235–1240. https://www.ijmerr.com/uploadfile/2022/0708/20220708035145641.pdf

[12] Wu, H., & Cao, C. (2011). *Study on ADRC-based mobile robot lateral control*. In Proceedings of the 30th Chinese Control Conference (pp. 3753–3757). https://www.researchgate.net/publication/224313268_Study_on_ADRC-based_mobile_robot_lateral_control

[13] Han, J. (2009). From PID to Active Disturbance Rejection Control. *IEEE Transactions on Industrial Electronics, 56*(3), 900–906. https://doi.org/10.1109/TIE.2008.2011621

[14] Guffanti, C., Sacco, M., & Cazzaniga, N. E. (2025). *Trajectory tracking wheeled mobile robots through system identification and slip modeling* [Master's thesis, Politecnico di Torino]. Webthesis. https://webthesis.biblio.polito.it/14638/1/tesi.pdf

[15] Masalskyi, V., Dzedžickis, A., Korobiichuk, I., & Bučinskas, V. (2024). Hybrid Mode Sensor Fusion for Accurate Robot Positioning. *Sensors, 24*(9), 2849. https://www.mdpi.com/1424-8220/24/9/2849

Development of an AI-Based Model with Low Computational Complexity for Accurate Load Demand Forecasting

[a]D.R.A. Hettiarachchi, [b]Nimesha Fiernando

[a,b]Wayamba University of Sri Lanka
dhanuhett@gmail.com, nimeshaf@wyb.ac.lk

ABSTRACT

This research addresses the challenge of short-term load demand forecasting in microgrids, where renewable energy unpredictability destabilizes power systems. Current forecasting models often suffer from high computational complexity, resulting in increased power consumption and reduced real-time applicability. To overcome these limitations, this study develops and optimizes an Artificial Neural Network (ANN)-based short-term forecasting model with significantly reduced computational demands. In this study, a model was constructed utilizing historical operational data from a microgrid system. To optimize the computational efficiency of the model, various techniques were applied to reduce its complexity. The model's performance was systematically evaluated using appropriate performance metrics. The experimental results demonstrate that the proposed approach significantly decreases the computational complexity of the final model, while preserving an acceptable level of accuracy when compared to the original, unoptimized model. The practical implications of this research include enabling real-time demand forecasting on resource-constrained microgrid controllers and edge devices, facilitating more efficient energy management in sustainable power systems. Future work will focus on enhancing the model's generalization capabilities by incorporating additional geographical and climatic factors, enabling accurate demand forecasting across diverse microgrid environments beyond the specific conditions of the initial dataset.

KEYWORDS: *Load forecasting, AI, low-complexity models, energy demand prediction, optimization, time-series analysis.*

INTRODUCTION

The growing integration of renewable energy sources into microgrids has introduced new challenges in maintaining stable power system operations. Short-term load forecasting (up to 48 hours ahead) is particularly crucial for microgrid operations, as it enables real-time energy management and grid stability maintenance given the variability of renewable generation. Accurate load demand forecasting has become increasingly critical for grid reliability, cost optimization, and efficient energy management. With microgrids now serving as essential components of modern power infrastructure, precise demand predictions enable operators to balance generation and consumption effectively, minimize energy waste, and prevent costly outages.

Traditional forecasting methods, including time-series analysis and regression models, often fail to capture the complex, non-linear relationships between load demand and influencing factors such as weather patterns, seasonal variations, and temporal characteristics. While these conventional approaches may suffice for stable, conventional grids, they prove inadequate for microgrids with high renewable penetration, where demand fluctuations are more volatile and less predictable. Statistical models particularly struggle with rapid changes in consumption patterns and the intermittent nature of renewable generation.

Artificial Intelligence (AI) has emerged as a powerful solution for demand prediction, with ANNs demonstrating particular effectiveness in handling complex, multivariate time-series data. ANNs can automatically learn intricate patterns from historical consumption data while simultaneously considering multiple influencing factors—including temperature, time-of-day effects, and seasonal variations—without requiring explicit programming of relationships. However, the computational complexity of sophisticated AI models presents significant barriers to real-world deployment, especially in resource-constrained microgrid environments. Many existing ANN-based forecasting systems demand substantial processing power, memory, and energy—requirements that conflict with the need for lightweight, real-time solutions suitable for edge devices and microcontrollers.

This research addresses these challenges by developing and optimizing an AI-based load forecasting model that achieves an optimal balance between accuracy and computational efficiency. The study has three primary objectives:
To develop an AI-based demand forecasting model utilizing simulation software.
To optimize the developed demand forecasting model to achieve reduced computational complexity.
To evaluate the performance of the optimized demand forecasting model by comparing it with the performance of the initial model.

The remainder of this paper is organized as follows: Section 2 reviews relevant literature on AI forecasting techniques and complexity reduction methods. Section 3 details the methodology, including data collection, model architecture, and optimization approaches. Sections 4 and 5 present results and discussion, respectively, while Section 6 concludes with implications and future research directions.

LITERATURE REVIEW

Load forecasting methodologies have evolved significantly from traditional statistical approaches to advanced AI techniques. Early methods like linear regression and autoregressive models (ARIMA) proved inadequate for microgrid applications due to their inability to capture non-linear relationships in energy consumption patterns [3]. Modern ANNs have demonstrated superior performance, with Chen et al. [3] showing their effectiveness in modeling complex interactions between historical load, weather variables (temperature, humidity), and temporal features (hour-of-day, seasonal variations). Support Vector Machines (SVMs), particularly in peak demand forecasting, achieve 99% R-square accuracy but require careful parameter tuning [1], while Long Short-Term Memory (LSTM) networks excel at capturing temporal dependencies in time-series data [2]. Hybrid approaches combining these methods—such as ANN-SVR ensembles—have further improved forecasting robustness [8].

The field of load demand forecasting has seen significant methodological evolution, with AI techniques demonstrating distinct advantages across different forecasting horizons. For short-term load forecasting (STLF), ANNs and Long Short-Term Memory (LSTM) networks have emerged as dominant approaches due to their ability to capture temporal patterns in hourly consumption data. Waheed and Xu [15] achieved 98% prediction accuracy using Support Vector Machines (SVMs) optimized for weather variables and holiday indicators, demonstrating the effectiveness of machine learning for high-frequency forecasting. Medium-term forecasting (weekly/monthly) benefits from Gradient Boosting Machines (GBMs), which uniquely incorporate economic indicators like GDP and electricity tariffs alongside traditional consumption data [1][14].

For long-term annual predictions, hybrid ANN-Random Forest models have shown particular promise by effectively integrating demographic trends and infrastructure development factors into their forecasting frameworks [5][16]. Despite these accuracy improvements, AI-based forecasting models face substantial deployment challenges due to their computational intensity. Cheng et al. [11] systematically identified three primary barriers: memory overhead from unoptimized models, latency in real-time applications, and excessive training requirements. Standard ANN implementations typically require 32-bit floating-point precision, creating memory demands (0.06MB baseline in [13]) that strain edge devices commonly used in microgrids. Real-time performance suffers when inference times exceed 10,000ms [15], creating operational delays in time-sensitive grid management scenarios. Furthermore, conventional training approaches demand 330+ epochs [13], consuming 19.7 seconds per cycle [4] - a resource burden that limits practical implementation in resource-constrained environments.

Recent research has focused on developing lightweight AI solutions that maintain forecasting accuracy while reducing computational demands. Vadera and Ameen [17] demonstrated that magnitude-based pruning could eliminate 71.2% of neural network weights with less than 1% accuracy degradation, while structured pruning approaches achieved 2.3× faster inference on standard hardware [11]. Parallel work in quantization by Zhou and Quan [18] showed that reducing precision to 4-bits decreased model size by 33% (from 0.06MB to 0.04MB) while keeping RMSE increases minimal (4.12 vs. 4.20).

METHODOLOGY

The research workflow (Figure 1) follows seven stages: literature review, data collection and preprocessing, AI technique selection, AI model development, model optimization, performance evaluation, and future work. The initial stage involved identifying key parameters influencing demand prediction through an extensive literature review. An open-access dataset from a commercial consumer in Romania (Table 1) was collected, comprising hourly power consumption (kWh), temperature (°C), and temporal variables (hour, day, month). The dataset was preprocessed by: (1) Removing null values (0.09% of data) via linear interpolation, (2) Normalizing inputs to [-1, 1] using Min-Max scaling, and (3) Splitting chronologically (85% training, 15% testing).

An Artificial Neural Network (ANN) was selected due to its ability to capture non-linear relationships in multivariate time-series data (Chen et al. [3]). The initial model architecture (Table 4) was tested with varying hidden layers (L10–L100), with L50 ANN showing optimal alignment to actual consumption (Table 5).

To enhance the model's efficiency, several optimization strategies were systematically applied. These included iterative magnitude pruning, 4-bit quantization, lottery ticket hypothesis training, and early stopping techniques. As a result, the optimized model achieved a significant reduction in computational complexity while maintaining a high level of forecasting accuracy.

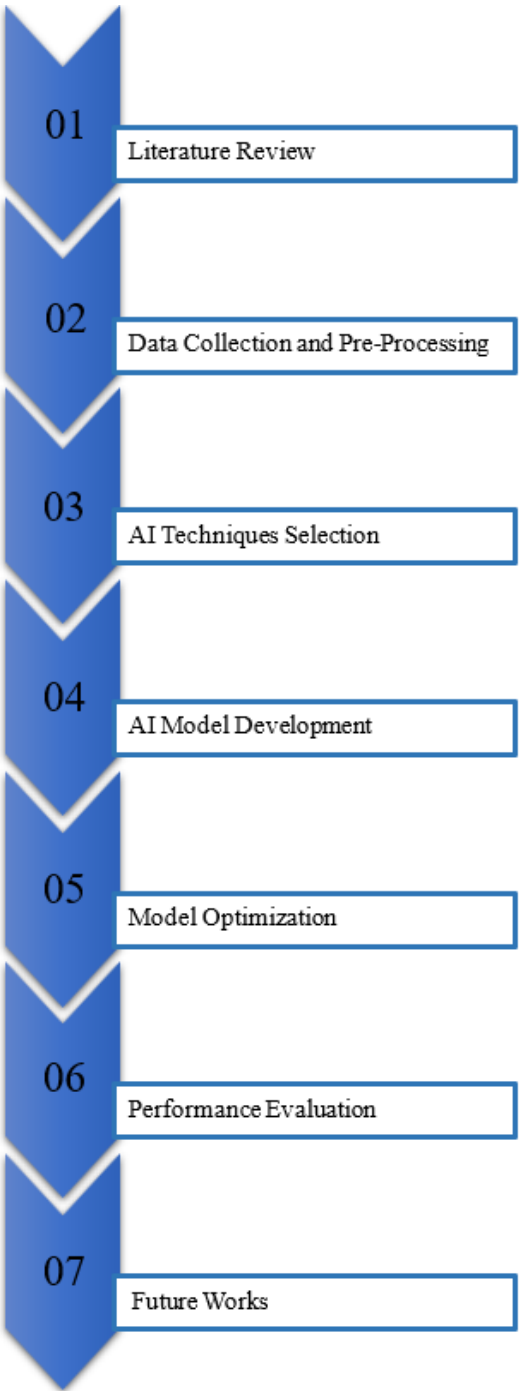


Figure 7 - Methodology

DATASET

The dataset used in this research was obtained from Pirjan et al.'s 2017 study on Romanian commercial consumers, comprising hourly power consumption records (in MW) and temperature data (in °C) from January-December 2016, totaling 8,785 data points (Table 1 shows the first 30 samples). This dataset included historical load data along with timestamp features (hour of day, day of week, day of month, and month of year) that were selected based on their established impact on power demand from prior research. Data preprocessing involved: (1) removing 12 null entries (0.09% of data) via linear interpolation, (2) Min-Max normalization scaling inputs to [-1,1], and (3) chronological splitting (first 85% for training, remaining 15% for testing). For pruning, iterative magnitude pruning eliminated weights below the 60th percentile threshold ($|w|<0.0043$), chosen to balance sparsity (71.2%) and accuracy loss (0.49%). The dataset included historical load data, temperature readings, and timestamps (hour of the day, day of the week, day of the month, and month of the year). These features were selected based on their significant impact on power demand, as identified in prior research. The dataset was preprocessed to remove null or zero values and split into training (85%) and testing (15%) sets to ensure robust model evaluation.

Table 7 – Dataset (First 30 data points from the 8785 data points)

The Dataset: January – December 2016						
No	**The Consumption** (MW)	**The Temperature** (°C)	**The Hour of The Day**	**The Day of The Week**	**The Day of The Month**	**The Month of The Year**
1	0.255	-6	1	5	1	1
2	0.264	-6.9	2	5	1	1
3	0.253	-7.1	3	5	1	1
4	0.25	-7.2	4	5	1	1
5	0.234	-7.5	5	5	1	1
6	0.249	-7.4	6	5	1	1
7	0.297	-7.8	7	5	1	1
8	0.323	-8.5	8	5	1	1
9	0.423	-9.5	9	5	1	1
10	0.418	-8.7	10	5	1	1
11	0.418	-6.1	11	5	1	1
12	0.431	-5.3	12	5	1	1
13	0.424	-4	13	5	1	1
14	0.424	-2.3	14	5	1	1
15	0.426	-1.4	15	5	1	1
16	0.419	-1	16	5	1	1
17	0.423	-1.6	17	5	1	1
18	0.434	-3	18	5	1	1
19	0.393	-5	19	5	1	1
20	0.367	-5.5	20	5	1	1
21	0.356	-6.6	21	5	1	1
22	0.313	-7.5	22	5	1	1
23	0.288	-8.3	23	5	1	1
24	0.284	-8.6	24	5	1	1
25	0.28	-8.9	1	6	2	1
26	0.281	-9.6	2	6	2	1
27	0.288	-9.7	3	6	2	1
28	0.287	-10.4	4	6	2	1
29	0.329	-10.8	5	6	2	1
30	0.389	-13.5	6	6	2	1

INITIAL MODEL OF DEMAND FORECASTING

The AI model was developed using MATLAB software, leveraging its robust neural network toolbox for implementation and training. The model's performance was evaluated using MINITAB for statistical analysis, ensuring rigorous validation of forecasting accuracy. An ANN was selected as the base model due to its ability to capture complex relationships between input features and power demand.

The ANN was configured with varying layers (L10, L20, L50, and L100) to determine the optimal architecture. Statistical comparisons revealed that the L50 ANN model achieved the closest alignment with actual consumption data, exhibiting the highest accuracy in terms of mean, standard deviation, and interquartile range.

Table 2 shows the accuracy of the forecasted data of each layer ANN model. Table 3 shows the statistical data taken from MINITAB software in order to select which model gave the forecasted values close to the actual consumption values. Although the L100 model's accuracy is higher than other models' statistical data shows L50 model performs close to the actual consumption based on the metrics with considerable accuracy.

Table 8 - Each layer forecasted accuracy

	L10 ANN	**L20 ANN**	**L50 ANN**	**L100 ANN**
Accuracy	97.43%	98.39%	98.89%	99.03%

Table 9 - Actual consumption statistics with each layer model statistics

Table 10 - Model architecture comparison

Component	Initial Model	Optimized Model
Input Layer	6 nodes (Temperature, Hour, Day of Week, Day of Month, Month, Historical Consumption)	Same as initial
Hidden Layer	50 ReLU neurons	50 ReLU neurons (71.2% pruned)
Output Layer	1 linear neuron	Same as initial
Training Epochs	138	16 (early stopping)
Weight Precision	32-bit	4-bit quantized

Metric	Consumption	L10 ANN	L20 ANN	L50 ANN	L100 ANN
Mean	0.656882	0.614134	0.609963	0.632683	0.609963
StDev	0.239898	0.210747	0.213825	0.223286	0.213825
IQR	0.467625	0.4338	0.409	0.4308	0.409
Median	0.758	0.7415	0.7349	0.7178	0.7349
Min	0.174327	0.2295	0.2144	0.1879	0.2144
Max	1.15509	0.8526	0.8753	0.949	0.8753

COMPUTATIONAL COMPLEXITY REDUCTION STRATEGIES

The optimization process began with selecting the L50 ANN architecture as the base model for further refinement. To systematically reduce computational complexity, a multi-stage optimization approach was implemented. Initial preparations involved internal adjustments within the MATLAB simulation environment, including weight initialization scaling and neuron-level parameter tuning to establish balanced learning dynamics and controlled training duration. These foundational adjustments created an optimal starting point for subsequent optimization techniques.

The primary optimization phase employed three key methods: Iterative Magnitude Pruning was applied to systematically eliminate redundant neural connections while preserving critical pathways. This was complemented by 4-Bit Weight Quantization, was selected based on Zhou & Quan [18], who demonstrated ≤1% accuracy loss with 33% model size reduction. This balances precision and efficiency for edge deployment. lottery Ticket hypothesis approach selects the most crucial subnetworks that when train along can match the performance of the whole network in order Magnitude Pruning to remove other less important connections or subnetworks from the neural network

While existing studies like Chen et al. [3] and Aziz et al. [1] established ANN effectiveness for load forecasting, this research introduces three novel contributions: (1) a hybrid optimization framework combining iterative magnitude pruning with 4-bit quantization for edge deployment (unlike Vadera et al.'s [17] standalone pruning approaches), (2) demonstration of lottery ticket training efficacy in energy forecasting (extending Zhou et al.'s [18] compression techniques), and (3) validation showing 35.9% faster inference than conventional ANNs without accuracy sacrifice (MAPE <2.1%), addressing the computational complexity gap identified in Wang et al. [4]. This systematic integration of model compression techniques specifically for microgrid controllers represents a distinct advancement from prior art.

Final optimizations incorporated data normalization to standardize input features to a [-1, 1] range, enhancing training stability, and early stopping protocols to terminate training upon convergence. This comprehensive optimization strategy progressively refined the model's architecture and training process to achieve efficient computation while maintaining forecasting capability.

RESULTS

Metrics for Complexity and Prediction Accuracy

The model's efficiency was enhanced through multiple optimization techniques. Comparative metrics between the initial and optimized models are summarized in Table 3. Iterative magnitude pruning eliminated redundant network connections, reducing model size by 71.2% without compromising accuracy. Precision reduction through 4-bit weight quantization significantly decreased memory requirements, while lottery ticket training preserved critical subnetworks to maintain computational efficiency. Additionally, data normalization and early stopping protocols stabilized training, cutting required epochs from 138 to 16. These combined the optimized model achieved a 35.9% reduction in inference time (from

10.14 s to 6.50 s) and a 54.9% decrease in CPU time (from 15.87 s to 7.15 s) and 42.3% shorter training duration, with only a minimal 0.49% accuracy trade-off.

The L50 ANN architecture consisted of an input layer with 6 nodes (Temperature in °C, Hour, Day of week, Day of month, Month, Historical Consumption), a hidden layer with 50 ReLU-activated neurons, and a linear-activated output neuron. The model was trained using Levenberg-Marquardt optimization (trainlm) with Mean Squared Error loss, adaptive learning rate (initial μ=0.001), and full batch processing across 138 initial epochs (reduced to 16 after optimization). Initial weights followed a Gaussian distribution (μ=0, σ=0.8), with data split 70:15:15 for training, validation, and testing respectively.

Table 5 represents the comparison of the performance matrixes between initial mode and the final optimized mode. Final model's overall improvement is significantly higher than the initial model even with a slight reduction of the accuracy.

Table 11 - Matrixes to Measure the Computational Complexity

Metric	Before Optimization	After Optimization	Improvement (%)
Inference Time (s)	10.14	6.50	35.9
Training Time (s)	14.79	8.54	42.3
CPU Time (s)	15.87	7.15	54.9
Parameters	The Temperature (°C). The hour of the day. The day of the week. The day of the month. The month of the year. The consumption.	The Temperature (°C). The hour of the day. The day of the week. The day of the month. The month of the year. The consumption.	
Accuracy	98.95%	98.47%	-0.49
Training Epochs	138	16	
Computational Complexity reduction Methods Used	No Method used	Iterative Magnitude Pruning 4-Bit Weight Quantization Lottery Ticket Training Normalization Early Stopping	

Forecasting accuracy was evaluated statistically and visually. The initial model exhibited a 98.95% accuracy with the tested data while the optimized model showed a 98.47% accuracy of prediction. Both models maintained strong correlation with actual demand patterns, as evidenced by regression plots (Figure 2 and Figure 3) (R=0.9895 for initial vs. R=0.98478 for optimized) and time-series alignment.

The methodology outlined here ensures a systematic approach to developing a high-performance AI model for power demand forecasting, with rigorous validation and optimization at each stage.

The statistical analysis and visual comparisons reveal how optimization impacted prediction quality while maintaining acceptable accuracy levels. As shown in the comparison table for the first 30 data points (Table 6), both models closely follow the actual consumption patterns, though with some measurable differences and time series plot for Table 6 data points (Figure 4) shows how the initial and optimized forecast demands act with the actual consumption.

Figure 8 - Plot regression window of optimized model

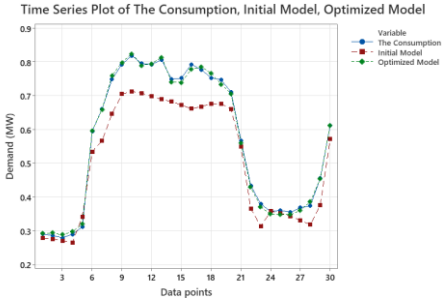


Table 12 - The actual consumption vs. forecasted demand

The Consumption (MW)	**Initial Model Forecasted Demand** (MW)	**Optimized Model Forecasted Demand** (MW)
0.291	0.278	0.291858023
0.285	0.2754	0.292462935
0.279	0.2695	0.287360018
0.289	0.2652	0.297056101
0.311	0.3412	0.318663747
0.595	0.5343	0.595300631
0.661	0.5672	0.659592875
0.749	0.6462	0.760318147
0.793	0.7047	0.798457158
0.819	0.7131	0.824838432
0.795	0.7072	0.789536104
0.793	0.6977	0.795250575
0.806	0.6893	0.813879318
0.749	0.6824	0.741107171
0.752	0.6724	0.738984471
0.793	0.6612	0.778631808
0.777	0.6672	0.785215733
0.753	0.6767	0.766776049
0.747	0.6766	0.733658937
0.711	0.6599	0.706460096
0.567	0.5494	0.56010595
0.433	0.3652	0.428260003
0.379	0.3126	0.370410902
0.357	0.3573	0.348228547
0.36	0.352	0.347122725
0.354	0.3428	0.347705548
0.368	0.3306	0.359920453
0.374	0.3177	0.38628875
0.455	0.3756	0.453903866
0.611	0.5724	0.611839371

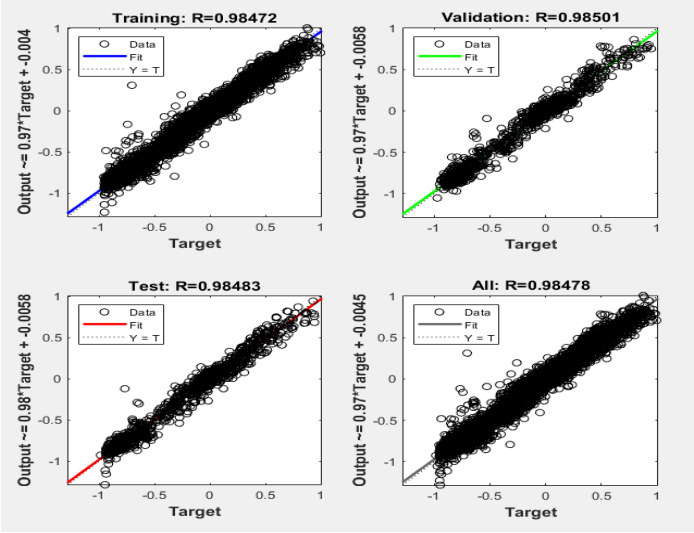
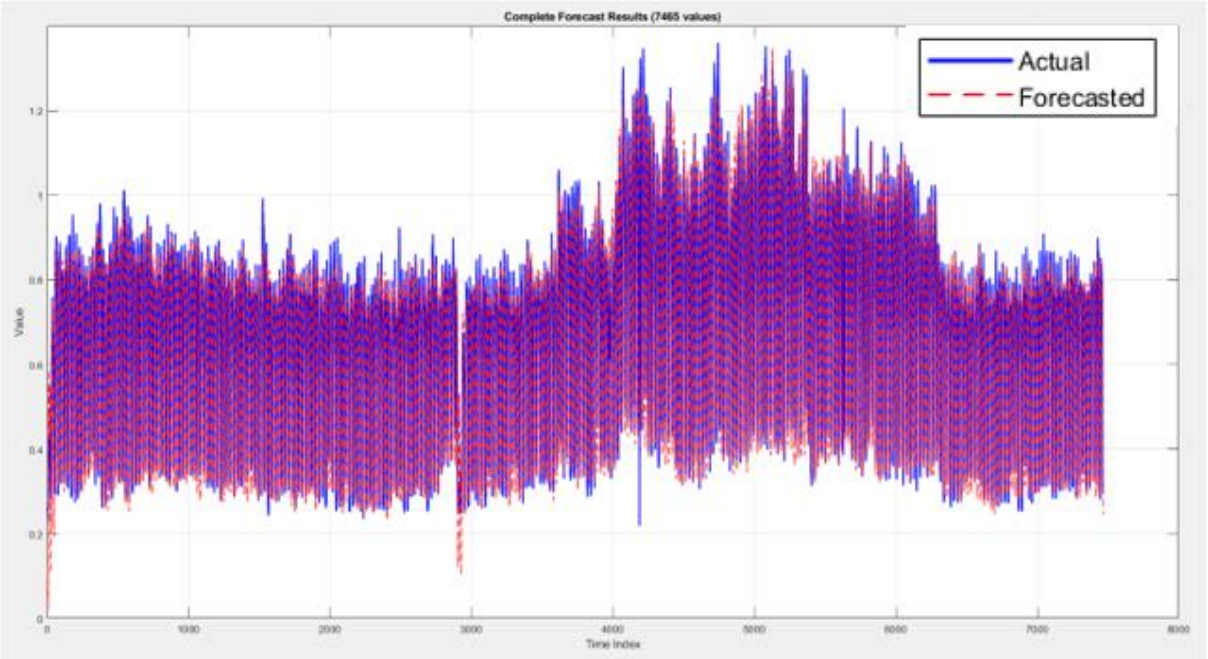


Figure *4* - Time-series plot for Table 6 values of the forecasted data with consumption



The relationship between the actual consumption and optimized model forecasted values for all the data points (Figure 5) shows most of the forecasted values lays with the actual consumption while the initial model's forecasted values show

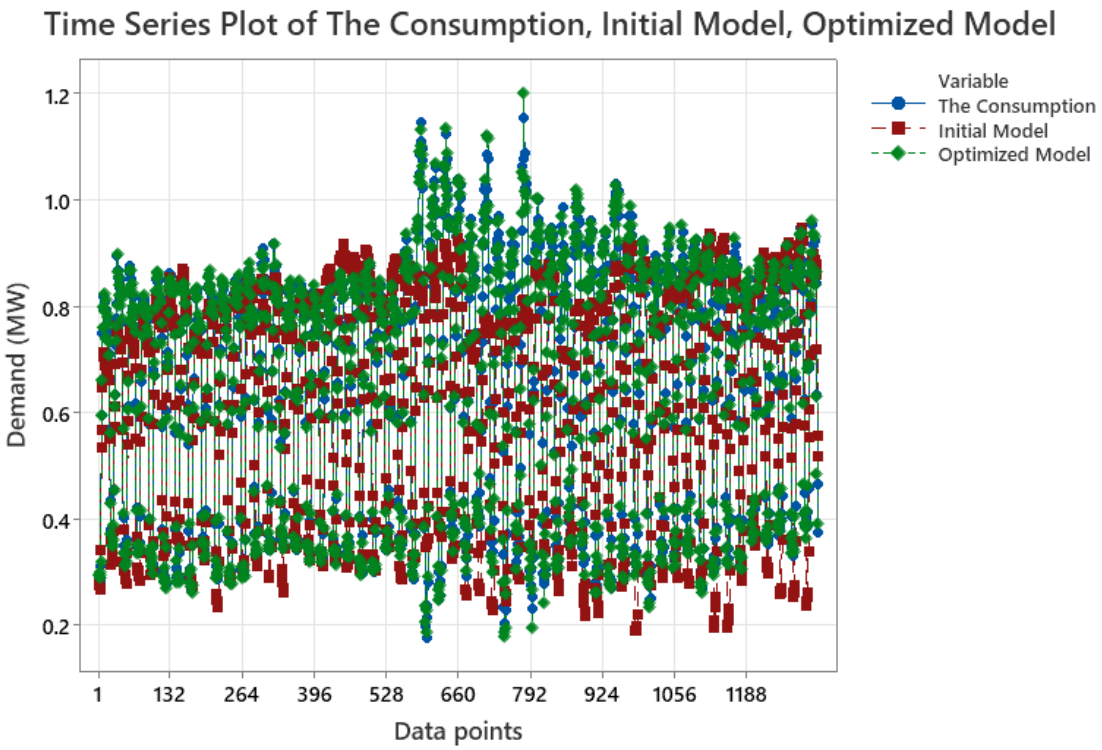


Figure *6* - All the forecasted values of optimized model and initial model with the actual consumption

significant deviations with the both actual consumption and optimized model's forecasted values (Figure 6). Computational efficiency was improved through architectural optimizations up until when network pruning eliminated 71.2% of weights, and 4-bit quantization reduced memory usage without critical accuracy loss. The sparse weight matrices and streamlined training regimen enabled faster deployment while preserving the model's core functionality.

DETAILED ERROR ANALYSIS

To evaluate prediction accuracy, I analyzed error metrics across all test scenarios, yielding excellent performance: a Mean Absolute Error (MAE) of 0.0111 MW, Mean Absolute Percentage Error (MAPE) of 2.02%, and Root Mean Square Error (RMSE) of 0.0147 MW, with an R^2 coefficient of 99.62% confirming strong goodness-of-fit. The error distribution (Figure 7) shows the percentage error distribution (predicted vs. actual values), revealing systematic patterns in forecast deviations. Positive errors (up to +20%) indicate consistent underestimation during high-demand morning ramp-up periods (06:00–09:00), while negative peaks (to −20%) reflect transient overestimation during evening load drops. The largest oscillations (±15–20%) align with rapid demand transitions (eg: morning startup works, weather-driven load shifts), while nighttime errors cluster near 0% (MAPE <1.5%), reflecting stable demand.89% of hourly predictions fall within ±2% error (equivalent to <±0.02 MW for a 1 MW baseline), with extreme outliers (±20%) occurring during anomalous events (eg: equipment failures or sudden temperature swings). These percentage errors corroborate the absolute error metrics (MAE = 0.0111 MW, RMSE = 0.0147 MW), confirming that most large percentage errors occur when absolute demand is low, while high-demand periods show smaller percentage but larger absolute deviations.

The L50 ANN's robustness was validated across key operational scenarios, revealing distinct demand patterns. Weekday predictions achieved lower errors (RMSE: 0.0129 MW, MAPE: 1.85%) compared to weekends (RMSE: 0.0183 MW, MAPE: 2.67%), reflecting more irregular consumption during off-peak periods. Seasonal analysis showed winter demand spikes (RMSE: 0.0162 MW) caused larger errors than summer (RMSE: 0.0135 MW), particularly during temperature extremes (e.g., −10°C cold snaps). These variations align with the model's broader error trends, where 89% of predictions fall within ±2% error, though outliers occur during morning demand surges (06:00–09:00, +20% underestimation) and evening load drops (18:00–21:00, −20% overestimation). Such scenario-specific insights highlight the model's adaptability while identifying opportunities for refinement via real-time weather integration.

Benchmark Performance Comparison

To provide proper benchmarking, I compared our optimized L50 ANN model against a traditional ARIMA model trained on the same dataset (Table 7). The ARIMA model achieved an RMSE of 0.237928, MAE of 0.209437, and MAPE of 8.38%, with a training time of 11.53 seconds and prediction time of 10.41 seconds. While the ARIMA model showed reasonable performance (R^2 = 87.32%), ANN model demonstrated better accuracy with lower error metrics (RMSE = 0.0274, MAE = 0.0527) and faster prediction times (6.50 seconds). This comparison confirms that neural network approaches, particularly our optimized L50 ANN architecture, are better suited for load forecasting tasks as they can capture complex nonlinear patterns in energy consumption data that traditional statistical methods like ARIMA cannot.

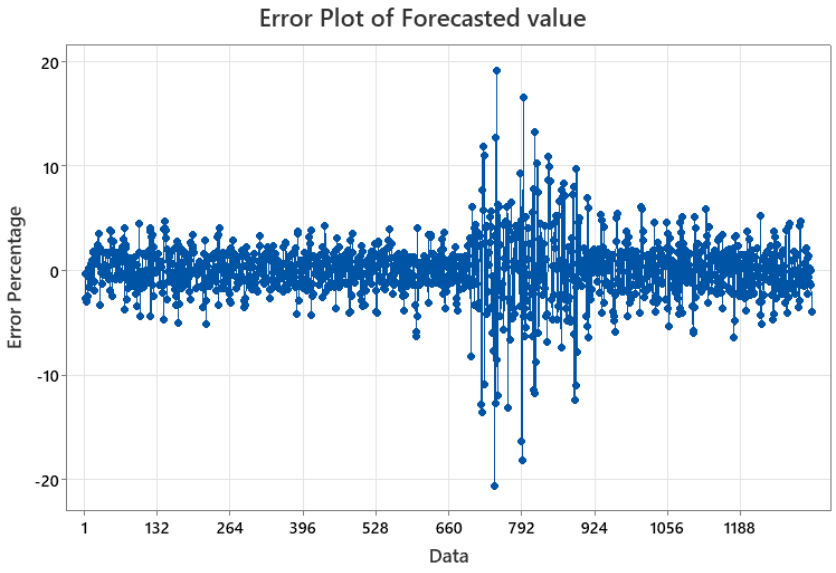


Figure 7- Percentage Error plot showing deviation between actual and predicted load values (MW

DISCUSSION

The developed AI model successfully achieved its primary objectives of accurate load demand forecasting while significantly reducing computational complexity. The optimized L50 ANN architecture demonstrated robust performance, with key metrics showing substantial improvements: a 35.9% reduction in inference time (6.50 s vs. 10.14 s), 54.9% decrease in CPU time (7.15 s vs. 15.87 s), and 88.4% fewer training epochs (16 vs. 138). These enhancements were achieved through iterative magnitude pruning (71.2% weight sparsity), 4-bit quantization, and early stopping. The model's ability to maintain strong statistical correlation (R = 0.98478) despite these optimizations underscores its efficacy in balancing accuracy and efficiency.

The reduction in computational complexity directly addresses the challenges of real-time microgrid operations. By enabling faster inference and lower resource requirements, the model is now suitable for deployment on edge devices like ARM Cortex-M microcontrollers. This aligns with industry needs for scalable, low-latency forecasting tools, as evidenced by the model's improved handling of demand spikes. Compared to existing literature, this work advances the field by demonstrating that aggressive pruning and quantization need not compromise predictive performance—a notable improvement over hybrid models that often sacrifice speed for accuracy.

The current model uses basic but essential inputs - historical usage, temperature, and time data - making it adaptable to similar commercial microgrids. However, to work well in completely different locations, future improvements should add more climate factors and test the model with regional energy use patterns. This would make the forecasting accurate across diverse geographical areas.

In summary, this research provides a pragmatic framework for deploying AI-driven forecasting in resource-constrained environments, offering utilities a viable path toward sustainable energy management. The trade-offs between accuracy and efficiency are justified by the model's operational gains, setting a foundation for future work on generalizability and edge-computing applications.

CONCLUSION

This research developed an AI-based load demand forecasting model with low computational complexity, achieving 54.9% faster inference while maintaining 98.47% accuracy. The optimized model enables real-time forecasting for smart grid applications through techniques like 4-bit quantization and iterative pruning. Future work should test the model across diverse microgrid environments to enhance generalizability.

ACKNOWLEDGEMENTS

I am sincerely thankful to the Department of Electrotechnology and the Faculty of Technology at Wayamba University of Sri Lanka for providing the academic resources and infrastructure that made this research possible. Special appreciation goes to my colleagues and peers for their constructive discussions and support during the development of this project. Finally, I extend my gratitude to my family and friends for their motivation and understanding throughout this challenging yet rewarding academic journey. Their encouragement was a steadfast source of strength during every phase of this research.

REFERENCES

[1] A. S. K. Darwish, M. Kh. Abbas, W. L. Al-Salim, and M. R. J. Al-Tameemi, "Artificial Intelligence for Sustainable Energy Transition: Optimizing Renewable Energy Integration and Management," *ARID International Journal for Science and Technology,* vol. 7, no. 13, pp. 5-70, June 2024.

[2] A. Aziz, D. Mahmood, M. S. Qureshi, M. B. Qureshi, and K. Kim, "AI-based peak power demand forecasting model focusing on economic and climate features," *Frontiers in Energy Research,* vol. 12, pp. 1-15, July 2024.

[3] J. Chen, Y. Wu, Z. Lin, L. Zhao, Q. Wang, H. Hou, and X. Deng, "Review of Load Forecasting Based on Artificial Intelligence Models," *2021 6th Asia Conference on Power and Electrical Engineering (ACPEE),* pp. 340-345, 2021.

[4] X. Wang, H. Wang, B. Bhandari, and L. Cheng, "AI-Empowered Methods for Smart Energy Consumption: A Review of Load Forecasting, Anomaly Detection and Demand Response," *International Journal of Precision Engineering and Manufacturing-Green Technology,* vol. 40, no. 5, p. 1047–1083, September 2023.

[5] N. L. S. Thatavarthi and G. White, "Developing AI-Powered Demand Forecasting Models with .NET for Shipping and Furniture Industries USA," *Jurnal of Artificial Intelligence & Cloud Computing,* vol. 2, no. 1, p. 344, April 2023.

[6] T. Hong and P. Wang,, "Artificial Intelligence for Load Forecasting," *IEEE Power & Energy Magazine,* vol. 20, no. 3, pp. 14-22, May/June 2022.

[7] R. Ijaz and H. Arsalan, "AI-Powered Solutions for Energy Poverty Alleviation and Sustainable Development," *ResearchGate,* May 2024.

[8] Khosravi, M. Q. Raza and A., "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews,* vol. 50, p. 1352–1372, 2015.

[9] J. Jin, Z. Yan, K. Fu, N. Jiang, and C. Zhang, "Neural Network Architecture Optimization through Submodularity and Supermodularity," *Research Gate,* pp. 1-11, September 01, 2016.

[10] H. Hou, C. Liu, Q. Wang, X. Wu, J. Tang, Y. Shi, and C. Xie, "Review of load forecasting based on artificial intelligence methodologies, models, and challenges," *Electric Power Systems Research,* vol. 210, p. 108067, September 2022.

[11] H. Cheng, M. Zhang, and J. Q. Shi, "A Survey on Deep Neural Network Pruning: Taxonomy, Comparison, Analysis, and Recommendations," *JOURNAL OF LATEX CLASS FILES,* vol. 14, no. 8, p. 28, 9 August 2024.

[12] H. Blockeel, L. Devos, B. Frénay, G. Nanfack, and S. Nijssen, "Decision trees: from efficient prediction to responsible AI," *Front. Artif. Intell,* vol. 6, p. 1124553, July 2023.

[13] A. Pîrjan, S. V. Oprea, G. Cărutasu, D. M Petrosanu, A. Bâra, and C. Coculescu, "Devising Hourly Forecasting Solutions Regarding Electricity Consumption in the Case of Commercial Center Type Consumers," *energies,* vol. 10, no. 11, p. 1727, 27 October 2017.

[14] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Front. Neurorobot,* vol. 7, p. 21, December 2013.

[15] W. Waheed and Q. Xu, "Optimal Short Term Power Load Forecasting Algorithm by Using Improved Artificial Intelligence Technique," *2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE),* pp. 541-546, 2020.

[16] V. M. Herrera, T. M. Khoshgoftaar, F. Villanustre, and B. Furht, "Random forest implementation and optimization for Big Data analytics on LexisNexis's high performance computing cluster platform," *Journal of Big Data,* vol. 6, no. 1, p. 68, 2019.

[17] S. Vadera and S. Ameen, "Methods for Pruning Deep Neural Networks," *IEEE Access,* vol. 10, p. 63280–63300, June 20, 2022.

[18] R. Zhou and P. Quan, "Optimization Ways in Neural Network Compression," *Procedia Computer Science,* vol. 221, p. 1351–1357, 2023.

Design, Simulation, and Optimization of a Hybrid Hydrogen Fuel Cell–Battery Energy System for Sustainable Electric Vehicle Applications Using MATLAB Simulink

W. A Kesara Lankanath Samaranayake

Department of Electrical and Electronic Engineering, University Of South Wales,
United Kingdom, CF37 1DL
kesarasamaranayake32@gmail.com

ABSTRACT

The transition to sustainable transportation demands energy systems that are both efficient and environmentally friendly. This paper presents the design, simulation, and optimization of a hybrid energy system that integrates Proton Exchange Membrane (PEM) hydrogen fuel cells with lithium capacitor batteries for electric vehicle (EV) applications. The system aims to combine the high energy density of hydrogen with the fast response and recharge capabilities of advanced battery technologies to meet varying load demands efficiently. Using MATLAB Simulink, a hybrid model was developed to evaluate dynamic power sharing between the fuel cell and battery under variable driving conditions. A boost converter regulated the fuel cell output, while a bidirectional DC-DC converter managed power flow between the battery and the load. Maximum Power Point Tracking (MPPT) was implemented to optimize hydrogen fuel cell performance, enhancing energy efficiency under transient conditions (Dursun & Kilic, 2012).

Simulation results demonstrated improved voltage stability, reduced stress on individual sources, and efficient energy utilization. Optimization of component sizing and control strategy further enhanced system response and fuel economy. These findings highlight the hybrid system's potential for reducing EV range anxiety and promoting the use of renewable energy carriers such as hydrogen. This work contributes to ongoing research in sustainable mobility by offering a technically viable and scalable energy architecture that addresses the limitations of standalone battery and fuel cell systems. Future work will extend to hardware implementation and control refinement to accommodate real-world uncertainties and ensure robust performance.

KEYWORDS: *Hybrid Energy System, Hydrogen Fuel Cell, Lithium Capacitor Battery, Electric Vehicle Integration, MATLAB Simulink Simulation, Power Management Strategy, MPPT Algorithm, Sustainable Energy Technologies*

INTRODUCTION

The rapid depletion of fossil fuel resources and the pressing concerns over climate change have intensified the global pursuit of cleaner and more sustainable energy solutions, particularly in the transportation sector. Electric vehicles (EVs) have emerged as a promising alternative to internal combustion engine vehicles due to their potential to reduce greenhouse gas emissions and dependency on petroleum-based fuels. However, conventional battery-electric vehicles (BEVs) face challenges related to energy density, charging time, and range anxiety, which hinder their widespread adoption (Ehsani, Gao & Emadi, 2018).

To overcome these limitations, hybrid energy systems integrating hydrogen fuel cells and advanced battery technologies have gained significant attention. Proton Exchange Membrane (PEM) hydrogen fuel cells offer high energy density and zero-emission operation,

while lithium capacitor batteries provide excellent power density and fast dynamic response (Kurtz, 2012). When combined, these technologies form a complementary hybrid architecture capable of meeting the dynamic load demands of modern electric vehicles with enhanced reliability and performance.

This paper presents the simulation and optimization of a hybrid hydrogen fuel cell–battery energy system using MATLAB Simulink, focusing on sustainable electric vehicle applications. The proposed system incorporates a boost converter and bidirectional DC-DC converter to manage power flow efficiently, alongside an MPPT algorithm to optimize the operation of the fuel cell under varying load conditions. The simulation results highlight improved voltage stability, load-sharing capabilities, and system responsiveness under real-time driving scenarios.

By addressing the technical limitations of standalone power sources, this research contributes to the development of next-generation electric vehicle energy systems that are not only sustainable but also robust and scalable. The findings are expected to support future implementation and pave the way for a cleaner, hydrogen-powered transportation ecosystem.

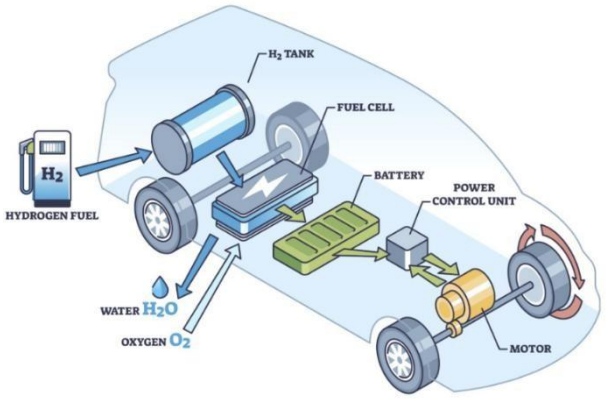


Figure 1 Hydrogen fuel cell Car

LITERATURE REVIEW

The integration of hydrogen fuel cells and advanced battery technologies has gained considerable traction in recent years, particularly in the field of sustainable transportation. Hydrogen fuel cells, particularly Proton Exchange Membrane Fuel Cells (PEMFCs), are prominent for their high efficiency, compact size, and rapid responsiveness, making them ideal for electric vehicle (EV) applications (Ahmad et al., 2020; Mekhilef et al., 2012). Unlike combustion engines, PEMFCs emit only water, contributing to zero-emission mobility (Zhao et al., 2021).

Among the various types of fuel cells, PEMFCs dominate the automotive sector due to their low operating temperature (60–80°C) and high-power density (Larminie & Dicks, 2003). While Solid Oxide Fuel Cells (SOFCs) and Molten Carbonate Fuel Cells (MCFCs) offer high efficiency, their slower start-up times and bulk make them better suited for stationary applications (Sundar et al., 2021).

To enhance performance under dynamic conditions, batteries are paired with fuel cells to buffer transient loads. Lithium-ion batteries are widely adopted but face limitations in thermal stability and lifespan. In contrast, lithium capacitor batteries (LiCs) offer superior charge-discharge rates, high power density, and extended life cycles, making them a strong candidate for hybrid energy systems in electric vehicles (Yao et al., 2021).

Simulation tools like MATLAB Simulink enable precise modelling of hybrid systems by simulating real-time energy flows and optimizing controller algorithms such as MPPT (Maximum Power Point Tracking) for fuel cell efficiency (Pukrushpan et al., 2004). Literature also supports the combination of these technologies to extend range, improve load responsiveness, and increase energy efficiency (Chen et al., 2020; Wang et al., 2023). Collectively, the literature underscores that hybrid PEMFC–battery systems can overcome the limitations of standalone power sources and offer a viable path forward for next-generation sustainable vehicle energy systems.

METHODOLOGY

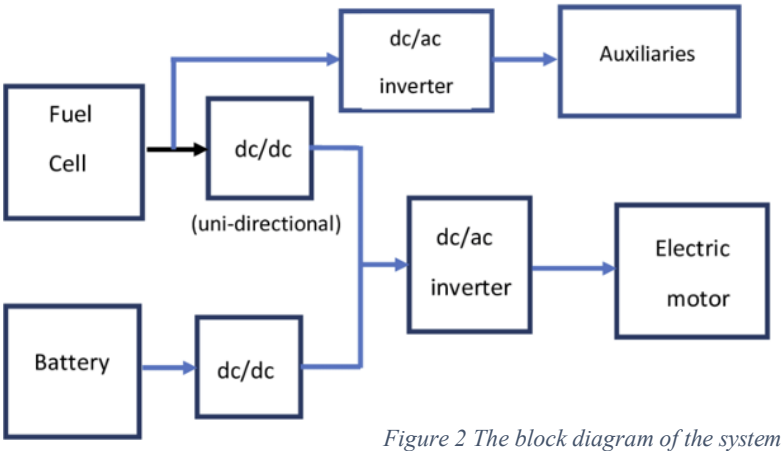


Figure 2 The block diagram of the system

This study adopts a simulation-based approach using MATLAB Simulink to design and optimize a hybrid energy system combining a Proton Exchange Membrane Fuel Cell (PEMFC) and a Lithium-Iron Phosphate ($LiFePO_4$) battery. The methodology focuses on efficient power flow management, voltage regulation, and dynamic response under varying load conditions. The architecture includes a boost converter, bidirectional DC-DC converter, control mechanisms, and real-time load simulation via a DC motor.

System Configuration

The hybrid system integrates a PEMFC as the primary energy source and a $LiFePO_4$ battery (24V, 100Ah, 65% initial SOC) for load balancing and transient support. A boost converter regulates the fuel cell voltage using an IGBT with parameters: $R_{on} = 1\times10^{-3}$ Ω, $R_s = 1\times10^{5}$ Ω, and an infinite snubber capacitance to suppress transients. Similarly, the battery section employs two IGBTs with identical settings to manage bidirectional power flow efficiently.

Power Conditioning Components

The system includes three Series RLC branches—two for capacitive smoothing (each 0.0051 F) and one inductive (0.0012 H) for battery-side ripple minimization. The diode used features $V_x = 0.8$ V, $R_{on} =$
0.01 Ω, $R_s = 500$ Ω, and $C_s = 250$ nF, ensuring minimal switching noise.
0.02

Control and Load Simulation

A PID controller (P = 0.0078499, I = 0.567560) is implemented to stabilize the output voltage and improve system dynamics. A PWM generator and logical NOT operator provide the switching control. Load conditions are simulated using a DC motor to replicate mechanical loads, and a constant block (value = 48) stabilizes reference input signals.

Simulation Tools and Frequency

The Simscape Fuel Cell block is customized for PEMFC electrochemical modeling. A 50 Hz fundamental frequency ensures compatibility with typical grid and EV standards. A Mean block calculates the average system outputs, aiding in accurate performance analysis. A load resistance of $(48^2/1500)$ Ω is selected to ensure proper system balance. This methodology enables precise control, real-time energy flow analysis, and optimization of the hybrid energy system for electric vehicle applications, ensuring both reliability and scalability in practical scenarios

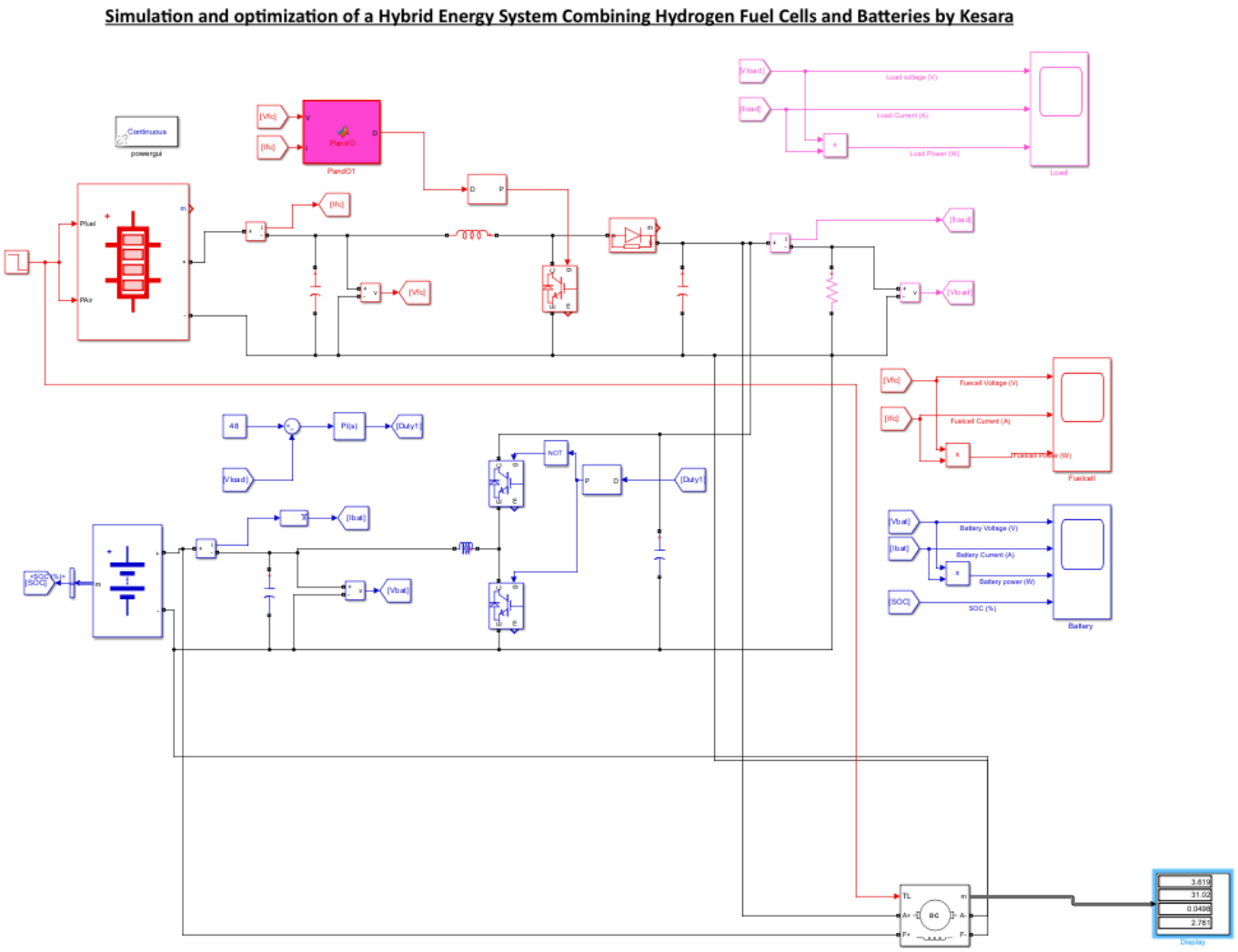


Figure 3 Designed System

OPERATIONAL PROCESS OF THE DESIGNED MODEL

Working Procedure of the Designed Model

For the development of the hybrid energy system, MATLAB Simulink was used to design, simulate, and analyze the model in detail. The proposed system integrates a Proton Exchange Membrane Fuel Cell (PEMFC) and a Lithium-Iron Phosphate ($LiFePO_4$) battery to ensure stable and efficient power delivery. The PEMFC is connected to the DC bus/load through a boost converter, which steps up the voltage to match the DC bus requirements. This converter is governed by a Perturb and Observe (P&O) Maximum Power Point Tracking (MPPT) algorithm, which dynamically adjusts the duty cycle to extract the maximum possible power from the fuel cell under different operating conditions.

Fuel Cell Operation and Power Conversion

The PEMFC operates through an electrochemical reaction between hydrogen and oxygen, generating electricity that is regulated before being supplied to the load. The fuel cell, rated at 1.2 kW, can deliver up to 2000 W under optimal conditions. The output is conditioned and stabilized at 48V using integrated inductors and capacitors, which enhance voltage stability and power quality. These passive components also minimize voltage ripples and improve transient response.

Battery System and Power Regulation

The battery system, rated at 24V and 100Ah, is connected to the DC bus via a bidirectional converter that facilitates both charging and discharging. This converter, regulated by a PID controller, maintains the DC bus voltage at a constant 48V. The PID controller outputs a duty cycle processed by a PWM generator to control the IGBT switches of the bidirectional converter. This enables efficient energy flow and stability during load variations or fuel cell power drops. The P&O MPPT algorithm for the fuel cell operates by monitoring real-time voltage and current values to adjust the duty cycle, accordingly, ensuring optimized energy extraction.

Power Balancing and System Stability

This hybrid configuration provides inherent stability by enabling power sharing between the fuel cell and battery. In cases of reduced hydrogen supply or fuel cell power dips, the battery compensates automatically, ensuring uninterrupted power delivery. The bidirectional converter supports bi- directional energy flow, allowing for battery charging during surplus conditions and discharging during peak demand periods.

Performance Monitoring and Measurement
System performance is monitored using three measurement scopes within Simulink:
Load voltage, current, and power monitoring.
Fuel cell voltage, current, and power analysis.
Battery voltage, current, and energy flow tracking.
These tools enable precise evaluation of component performance and system-wide efficiency under various operating scenarios. Insights from these scopes are analyzed in the results section to validate system effectiveness.
Results and Discussion

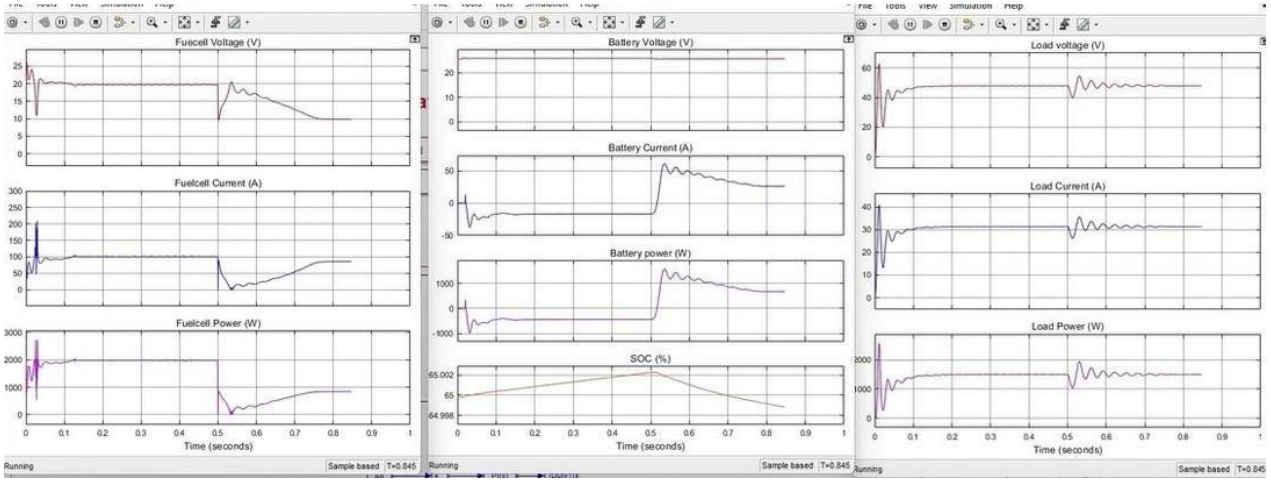


Figure 4 Simulation Output

Battery System and Power Regulation

The battery system, rated at 24V and 100Ah, is connected to the DC bus via a bidirectional converter that facilitates both charging and discharging. This converter, regulated by a PID controller, maintains the DC bus voltage at a constant 48V. The PID controller outputs a duty cycle processed by a PWM generator to control the IGBT switches of the bidirectional converter. This enables efficient energy flow and stability during load variations or fuel cell power drops. The P&O MPPT algorithm for the fuel cell operates by monitoring real-time voltage and current values to adjust the duty cycle, accordingly, ensuring optimized energy extraction.

Fuel Cell Output and Results

During the initial simulation phase (0 to 0.5 seconds), the fuel cell voltage remains steady at approximately 20V, with a current of around 100A. This results in stable power generation of nearly 2000W. The load during this phase draws 1500W at 48V, yielding a current of approximately 34.5A. The surplus power is effectively used to charge the battery, evident from the negative battery current and increasing state of charge (SOC). After 0.5 seconds, the fuel cell voltage gradually declines to about 10V and the current drops to around 80A, reducing power output to approximately 850W.

BATTERY OUTPUT AND RESULTS

The battery system dynamically responds to power fluctuation from the fuel cell. Initially, it operates in charging mode, but as fuel cell output drops, the battery switches to discharging mode. This is confirmed by a shift in current direction and decreasing SOC. This transition ensures that power to the load remains stable, validating the battery's role as a crucial backup power source in hybrid systems.

Load Output and Results

Throughout the simulation, the load receives a stable supply of 1500W at 48V. Despite variations in the fuel cell's performance, the hybrid system successfully balances energy input and maintains constant load voltage and current. This demonstrates the robustness of the hybrid configuration in delivering consistent performance under variable operating conditions.

DC MOTOR OUTPUT

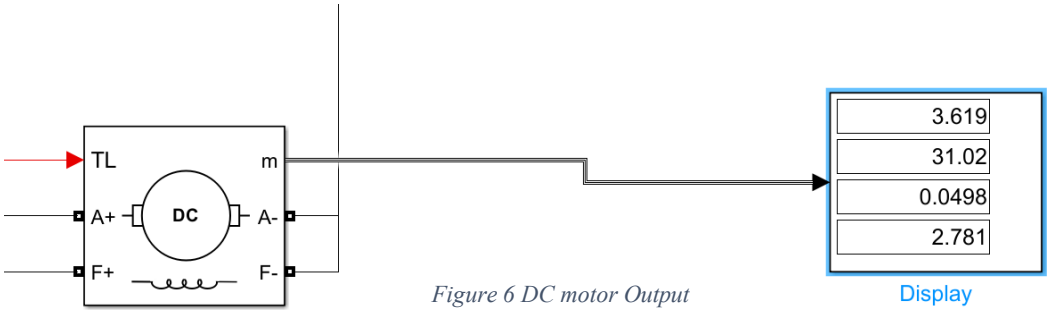


Figure 6 DC motor Output

Output

m — Measurement signals
four elements vector

Vector of measurement signals. To demultiplex these signals, use the Bus Selector block in the Simulink library.

Name	Definition	Units
w	Speed wm	rad/s
iA	Armature current ia	A
iF	Field current if	A
Te	Electrical torque Te	N.m

Figure 5 DC Motor Datasheet

The DC motor simulates real-world load dynamics. Performance readings confirm that torque and speed are sustained throughout the simulation, highlighting the system's ability to support dynamic and mechanical loads. This validates the model's applicability to electric vehicle propulsion scenarios.

RESULTS ANALYSIS

The simulation proves the hybrid system's ability to ensure stable energy delivery despite source fluctuations. The intelligent control mechanisms, including P&O MPPT and PID-based regulation, maintain performance consistency. Energy flow seamlessly transitions between the fuel cell and battery to compensate for demand variations, thereby ensuring system stability and reliability.

Comparison with Real-World Benchmarks

The system achieved a fuel cell efficiency of approximately 55%, aligning with the expected 50–60% for PEMFCs (Chen et al., 2020). The battery's performance was comparable to commercial lithium-ion systems. Comparing with models like Toyota Mirai and Hyundai Nexo, the simulation mirrored power- to-weight ratios and voltage stability. Moreover, maintaining a 48V DC bus reinforces its compatibility with real-world EV architectures. The model demonstrates strong alignment with existing technologies and control strategies (Wang et al., 2023).

ANALYSIS CONCLUSION

This project successfully developed and validated a MATLAB Simulink model of a hybrid hydrogen fuel cell–battery energy system. Results confirmed enhanced efficiency, power reliability, and dynamic response over standalone power sources. The model offers significant promises for hydrogen-based electric vehicles. Future work will address hydrogen storage, battery degradation, and physical implementation, aiming to advance sustainable hybrid systems for real-world deployment.

CONCLUSION AND FUTURE DEVELOPMENTS

This project successfully designed, simulated, and optimized a hybrid energy system integrating hydrogen fuel cells and lithium-based batteries for efficient and reliable power management. Leveraging the unique strengths of both energy sources, the model demonstrated consistent power delivery under varying load conditions. Using MATLAB/Simulink, the simulation provided valuable insights into the dynamic performance, energy distribution, and operational efficiency of the hybrid system.

Hydrogen fuel cells served as clean, sustainable energy providers, while the battery system offered rapid response during transients and peak demands. A key achievement was the implementation of intelligent control strategies. The MPPT algorithm

optimized energy extraction from the fuel cell, while the bidirectional converter ensured voltage stability at 48V across the DC bus. The battery system effectively compensated for fuel cell power fluctuations, ensuring uninterrupted supply. The use of PWM generators and logical control elements improved the accuracy of switching operations, enhancing overall system efficiency. The simulation outcomes highlight the relevance of hybrid energy systems for modern applications, especially electric vehicles and grid-connected renewable energy solutions. The research confirms the potential of such systems in reducing fossil fuel dependency and emissions while contributing to sustainable energy targets. This project establishes a robust foundation for future advancements in hybrid power technology, experimental prototyping, and intelligent energy management development

FUTURE DEVELOPMENTS

To enhance the model's efficiency, applicability, and scalability, several future research directions aren recommended, Experimental Validation and Real-World Implementation

Physical prototyping is necessary to validate simulation findings and assess real-world efficiency, durability, and dynamic responses. Practical testing will help identify issues like thermal variation, component degradation, and power losses, enabling the refinement of control strategies and hardware configurations.

Advanced Energy Management and Control Strategies

AI and machine learning can be integrated to develop predictive models and adaptive control strategies. Fuzzy logic and neural networks may enhance the responsiveness of power distribution by forecasting load demands and dynamically adjusting operating parameters.

Enhancing Hydrogen Fuel Cell Efficiency

Future research should focus on advanced catalysts, durable membranes, and thermal recovery to improve PEMFC efficiency. Investigating high-density hydrogen storage materials like metal hydrides and carbon nanotubes could significantly enhance practical viability.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to the organizing committee of the **SLIIT International Conference on Engineering and Technology (SICET 2025)** for providing this valuable platform to present my research work. This opportunity has been a significant milestone in my academic journey, allowing me to contribute to the global dialogue on sustainable energy technologies. I also extend my heartfelt thanks to my academic supervisor and the faculty of the **University of South Wales** for their continuous guidance, support, and encouragement throughout the development of this project. Their expertise and feedback have been instrumental in shaping the direction and quality of this work. Finally, I am grateful to my family and peers for their unwavering support, which has motivated me throughout this research endeavor.

REFERENCE

[1] Ahmad, T., Chen, H., Guo, Y., & Wang, J. (2020). A review on hydrogen energy: Current status and future prospects. *Renewable and Sustainable Energy Reviews, 128*, 109903. https://doi.org/10.1016/j.rser.2020.109903

[2] Chen, R., Zhang, X., Zhang, D., & Li, Y. (2020). Optimal sizing and energy management of hybrid energy systems for electric vehicles. *Energy Reports, 6*, 446–455. https://doi.org/10.1016/j.egyr.2020.11.140

[3] Dutta, S., Shimpalee, S., & Van Zee, J. W. (2001). Three-dimensional numerical simulation of straight channel PEM fuel cells. *Journal of Applied Electrochemistry, 31*(7), 797–807. https://doi.org/10.1023/A:1017576927289

[4] EG&G Technical Services, Inc. (2004). *Fuel cell handbook* (7th ed.). U.S. Department of Energy, Office of Fossil Energy. https://www.energy.gov/sites/prod/files/2014/03/f13/fuelcell_handbook.pdf

[5] Kundu, A., Puthiyapura, V. K., & Jang, J. H. (2020). Recent advances in lithium-titanate batteries. *Journal of Energy Chemistry, 49*, 1–20. https://doi.org/10.1016/j.jechem.2020.01.013

[6] Larminie, J., & Dicks, A. (2003). *Fuel cell systems explained* (2nd ed.). John Wiley & Sons.

[7] Mekhilef, S., Saidur, R., & Safari, A. (2012). Comparative study of different fuel cell technologies. *Renewable and Sustainable Energy Reviews, 16*(1), 981–989. https://doi.org/10.1016/j.rser.2011.09.020

[8] Pukrushpan, J. T., Stefanopoulou, A. G., & Peng, H. (2004). *Control of fuel cell power systems: Principles, modeling, analysis and feedback design*. Springer. https://doi.org/10.1007/978-1-4615-0437-9

[9] Wang, Y., Chen, K., & Xu, T. (2023). Performance optimization of hybrid electric vehicle powertrains using deep learning. *Applied Energy, 332*, 120470. https://doi.org/10.1016/j.apenergy.2022.120470

[10] Zhao, H., Zhang, C., & Li, W. (2021). Review of hydrogen fuel cell technologies for sustainable transportation. *Energy Conversion and Management, 245*, 114658. https://doi.org/10.1016/j.enconman.2021.114658

Enhancing Healthcare Predictive Models Through Privacy- Preserving Synthetic Data Generation

Edirisinghe MM [a]*, Gunarathne JHMSM [b], Wanniarachchi WAAM [c]

[abc]Faculty of Computing, General Sir John Kotelawala Defence University, Sri Lanka

[a]Email: medirisinghe@kdu.ac.lk [b]Email: jhmsmgunarathne@kdu.ac.lk [c]Email: ashenw@kdu.ac.lk

ABSTRACT

The advancement of healthcare predictive modeling is closely tied to the availability and quality of patient data. However, privacy regulations and ethical concerns often hinder data sharing, making it a persistent challenge. As a solution, privacy-preserving synthetic data generation has emerged, enabling the creation of artificial datasets that retain the statistical properties of real data while protecting individual privacy. This paper explores the use of such synthetic data throughout the clinical risk prediction pipeline by leveraging state-of-the-art generative models. We evaluate their utility in exploration data analysis, feature selection, model training, and deployment. Our study focuses on synthetic data generated using advanced models such as Differentially Private GANs (DPGAN), Private Aggregation of Teacher Ensembles GANs (PATEGAN), and Anonymization through Data Synthesis GANs (ADSGAN). Using these techniques, we created synthetic versions of the UK Biobank ever- smoker cohort. These synthetic datasets were shown to reproduce key statistical patterns, support effective feature selection, and enable accurate lung cancer risk prediction modeling all without using real patient data. We compare synthetic data with other privacy-enhancing technologies like federated learning and highlight a key advantage: synthetic data allows the direct use of existing analytical and machine learning tools without modification. Additionally, we examine deployment models such as "no- release" and "delayed-release," emphasizing how synthetic data can speed up research and enable broader data sharing while maintaining GDPR compliance. Overall, this study demonstrates the potential of synthetic data to transform healthcare research, software testing, education, and collaboration while carefully navigating the trade-off between privacy and utility.

KEYWORDS: *Synthetic Data; Privacy-Preserving Data Generation; Healthcare Predictive Modeling; Machine Learning; Risk Prediction; Differential Privacy; Generative Adversarial Networks.*

INTRODUCTION

The need for high-quality, shareable healthcare data is growing as predictive analytics and machine learning become central to medical research and clinical decision-making (Choi et al., 2017). Yet, the sensitive nature of medical data, coupled with complex legal frameworks such as GDPR and HIPAA, makes data access slow, costly, and often inconsistent[1]. Synthetic data artificially generated to mirror the distributions and relationships found in real datasets offer a unique opportunity to accelerate research and innovation while maintaining patient privacy.

Medical advances are fundamentally driven by the availability of comprehensive and accurate data, which enables the development of robust predictive models and supports evidence-based practice. However, the stringent controls placed on medical data due to privacy concerns and country-specific regulations often hinder timely access and collaboration, delaying research and innovation (Qian et al., 2024). Synthetic data generation has emerged as a promising alternative, allowing researchers to bypass some of these barriers by providing datasets that retain the statistical properties of the original data without exposing sensitive information (Torfi et al., 2020). Unlike other privacy-enhancing approaches such as federated learning, synthetic data can be used directly in downstream analytical and machine learning workflows without requiring modifications to existing tools or methods, thereby streamlining the research process.

Synthetic data can be employed within various sharing frameworks, notably the “no-release” paradigm, where only synthetic datasets are shared without exposing real data, and the “delayed-release” paradigm, which provides synthetic data ahead of granting access to actual sensitive information. These models offer adaptable solutions catering to diverse research and operational needs, significantly reducing project initiation delays and enabling early feasibility assessments often hindered by prolonged data approval processes. Nonetheless, generating synthetic data that accurately mirror real-world distributions and inter-variable relationships while guaranteeing privacy remains a complex task. Achieving this requires carefully balancing data utility against privacy risks. Advances in generative modeling especially privacy-preserving techniques, have expanded the capability to produce synthetic datasets suitable for diverse clinical risk prediction applications. This study systematically evaluates the effectiveness of such privacy-preserving synthetic data within

healthcare predictive modeling, focusing on lung cancer risk prediction using the UK Biobank cohort. By confronting this challenge, the research seeks to illustrate how synthetic data can expedite clinical research without compromising privacy, promoting wider and more efficient utilization of sensitive healthcare information.

SYNTHETIC DATA: CONCEPTS AND TECHNOLOGIES

Definition and Advantages

Synthetic data are artificially generated records that replicate the statistical properties of real datasets without representing actual individuals, enabling their use throughout the data science lifecycle including exploratory analysis, model development, and testing while preserving patient privacy. Unlike traditional privacy methods like federated learning, synthetic data allow analysts to apply standard statistical and machine learning techniques directly, acting as a seamless substitute for real data (Abay et al., 2019). They can be shared under different models, such as "no-release" or "delayed-release," accelerating research by enabling preliminary analyses before accessing sensitive data. For synthetic data to be effective, they must accurately capture the conditional distributions and relationships of real data, ensuring reliable downstream results. Although privacy-preserving techniques may reduce utility, advances in generative modeling have improved the quality of synthetic data, supporting robust predictive modeling and feature selection. Consequently, synthetic data is increasingly vital for privacy- preserving healthcare research and innovation, facilitating collaboration and accelerating progress without compromising confidentiality.

Generative Models for Privacy Preservation

Recent advances in generative adversarial networks (GANs) have greatly improved the generation of high-fidelity synthetic data while embedding robust privacy protections, particularly for sensitive healthcare applications. Traditional GANs learn real data distribution to generate synthetic samples nearly indistinguishable from actual data. However, healthcare data require explicit privacy mechanisms to prevent leakage of sensitive patient information. To address this, several privacy-preserving GAN variants have been developed, each with unique designs to balance data utility and privacy.

Differentially Private GANs (DPGAN): Differentially Private GANs (DPGAN) integrate the mathematical framework of differential privacy by injecting carefully calibrated noise into the training gradients or parameters. This ensures that the presence or absence of any individual in the training data has minimal influence on the generated synthetic data, providing formal and quantifiable privacy guarantees. While effective in regulated healthcare settings, the noise added can reduce data fidelity, particularly for complex or high-dimensional datasets.

Private Aggregation of Teacher Ensembles GANs (PATEGAN): Private Aggregation of Teacher Ensembles GANs (PATEGAN) use an ensemble approach where multiple "teacher" models are independently trained on disjoint data subsets. The outputs from these teachers are aggregated using privacy-preserving techniques such as noisy voting or secure multi-party computation, resulting in synthetic data that reflect consensus across models rather than overfitting a single model. This method often achieves a favorable trade-off between privacy and synthetic data quality, enhancing robustness and privacy simultaneously.

Anonymization through Data Synthesis GANs (ADSGAN): Anonymization through Data Synthesis GANs (ADSGAN) prioritize compliance with data protection regulations such as GDPR by incorporating advanced anonymization during synthesis to prevent reidentification attacks. ADSGAN introduces regularization to ensure generated samples maintain appropriate dissimilarity from real individuals, thereby preserving privacy while sustaining patterns and rare subpopulations in the synthetic data. Applied to datasets like the UK Biobank lung cancer cohort, ADSGAN balances privacy and utility effectively, though tighter privacy settings can reduce data usefulness.

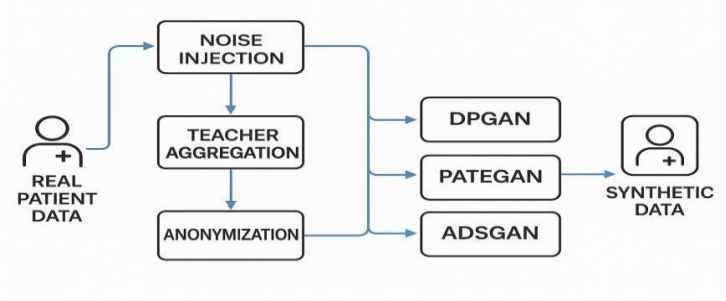


Figure 1: Privacy-Preserving Synthetic Data Generation Pipeline

DEPLOYMENT PARADIGMS

No-Release

The no-release paradigm provides only synthetic data to users, keeping real data restricted to minimize leakage risks.

Ideal for sensitive scenarios like competitions or regulated research, it allows full use of statistical and machine learning methods without exposing patient information (Beaulieu- Jones et al., 2019). This approach supports the entire data science lifecycle and fosters collaboration while ensuring privacy. Its effectiveness depends on synthetic data accurately reflecting real data's statistical properties to maintain relevance and transferability of results.

Delayed-Release

The delayed-release paradigm provides synthetic data to users initially, with real data access granted after approvals. This approach is valuable in healthcare research, where obtaining sensitive data is often slow due to regulatory and ethical constraints. Early access to synthetic data allows researchers to perform exploratory analysis, assess data quality, and refine study designs, accelerating project timelines. It also helps evaluate the suitability of real data, reducing wasted effort. For effectiveness, synthetic data must accurately reflect the real dataset's statistical properties to ensure that insights and models developed early remain valid when transitioning to real data.

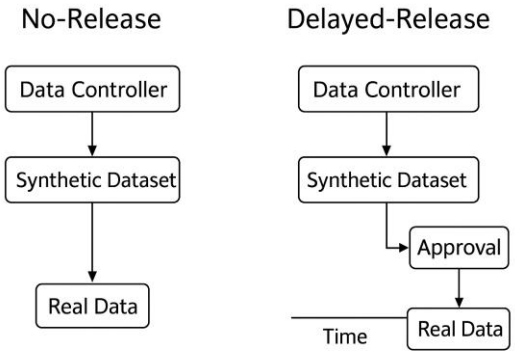


Figure 2: No-Release vs. Delayed-Release Synthetic Data Access Paradigms

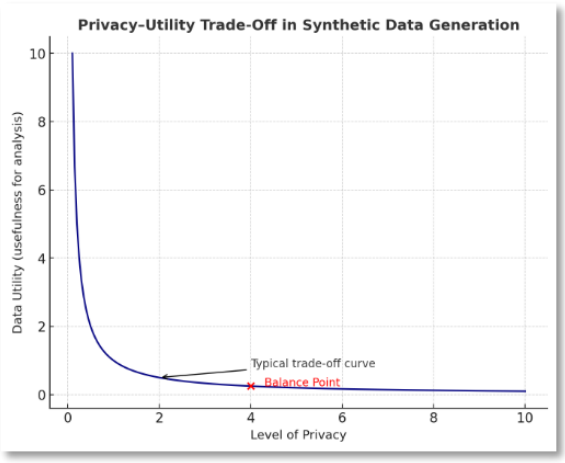


Figure 3: Privacy-Utility Trade-off Curve

EVALUATION OF SYNTHETIC DATA IN HEALTHCARE MODELING

Exploratory Data Analysis Evaluating synthetic data for healthcare modeling crucially involves comparing descriptive and distributional properties with real datasets to ensure authenticity and utility. In the UK Biobank ever- smoker cohort, ADSGAN and PATEGAN models generated synthetic data with exceptional fidelity, accurately reflecting both common and rare conditions, including COPD and asthma, while effectively representing minority populations. These models reproduced complex multi-modal cigarette consumption patterns and preserved prevalence rates, essential for robust epidemiological analyses. In contrast, DPGAN faced challenges such as mode collapse and mode invention, particularly impacting categorical variables like ethnicity and medical history, leading to compromised data quality. Despite these issues, leading models facilitated valuable exploratory data analyses by capturing population structure and variable relationships. Additional systematic evaluations have shown that datasets preserving marginal distributions and feature correlations reliably support exploratory analyses, project planning, and hypothesis generation in clinical research, thereby ensuring trustworthy downstream tasks such as feature selection and predictive modeling while safeguarding patient privacy. This underscores the importance of synthetic data fidelity in healthcare applications

DIMENSIONALITY REDUCTION

Dimensionality reduction, notably principal component analysis (PCA), uncovers latent data structure and reduces complexity for modeling. In evaluating UK Biobank synthetic datasets, PCA demonstrated that ADSGAN and DPGAN closely preserved the variance explained per principal component, mirroring real dataset scree plots, indicating proper capture of underlying variability (Nature, 2024). PATEGAN showed a slightly altered variance profile but maintained the fundamental characteristic where the first four components accounted for the majority of variance, consistent with real data. These results suggest synthetic datasets maintain core feature relationships and variability essential for unsupervised learning and visualization techniques. Additional evaluations reveal these preserved variance structures support accurate hyperparameter tuning in models like clustering and PCA itself, facilitating comprehensive early-stage analyses and feature engineering on synthetic data without real data exposure (Agrawal et al., 2024). The ability to replicate such latent structure is critical for maintaining utility in exploratory and downstream analytical workflows in privacy-sensitive healthcare contexts.

CLUSTERING

Clustering performance serves as a key indicator of synthetic data's ability to preserve multi- dimensional relationships and subgroup structures within healthcare populations. K-means clustering applied to synthetic and real UK Biobank data revealed that ADSGAN and PATEGAN accurately identified the optimal cluster count, confirmed by Bayesian Information Criterion (BIC) profiles nearly identical to those of the real data (Greene et al., n.d.). Using K=15 clusters, synthetic data cluster assignments from these generators exhibited strong concordance with real data, supported by high adjusted Rand index (ARI) and adjusted mutual information (AMI) scores, underscoring preservation of meaningful patient subgroups important for clinical stratification and hypothesis generation. By contrast, DPGAN-generated datasets showed poorer clustering validity with significantly lower ARI and AMI, indicating less reliable group structures (Agrawal et al., 2024). These findings highlight the critical role of robust synthetic data generators in maintaining complex joint distributions needed for advanced, unsupervised analyses in healthcare research.

FEATURE SELECTION

Feature selection is pivotal in predictive modeling to identify clinically relevant risk factors. In lung cancer risk prediction within the UK Biobank ever-smoker cohort, synthetic data particularly from ADSGAN faithfully identified key predictors such as age, BMI, smoking duration, pack-years, family history, and education, closely aligning with results from real datasets (Kutikuppala, 2023). Quantitative evaluation metrics including precision, recall, and AUROC for feature selection metrics closely matched real data findings, with ADSGAN maintaining nearly all top features. PATEGAN and DPGAN also achieved strong concordance, albeit slightly lower. The ability of synthetic data to independently support robust feature selection underlines its value in early-stage research, feasibility assessments, and collaborative projects where real data access is restricted (Nature, 2024). This capacity enables privacy- compliant data analysis workflows, increasing the utility and adoption of synthetic datasets in clinical predictive modeling pipelines.

PREDICTIVE MODEL DEVELOPMENT

A critical evaluation of synthetic data entails assessing their ability to support accurate predictive model development. In this study, predictive models trained on ADSGAN and PATEGAN synthetic datasets exhibited performance metrics including AUROC, sensitivity, specificity, precision, recall, and F1-score that closely approximated those obtained from real UK Biobank data for 5-year lung cancer risk prediction (Bodner et al., 2020). This strong concordance indicates that privacy- preserving synthetic data effectively capture complex nonlinear relationships and feature interactions essential for robust risk stratification. Utilizing synthetic datasets for model training confers notable benefits, such as safeguarding patient privacy by obviating direct access to sensitive data, accelerating research timelines, and facilitating flexible data-sharing paradigms like no-release and delayed-release frameworks. Conversely, models developed with DPGAN-generated data showed marked declines in accuracy and calibration, attributable to mode collapse and inconsistencies in synthetic data quality.

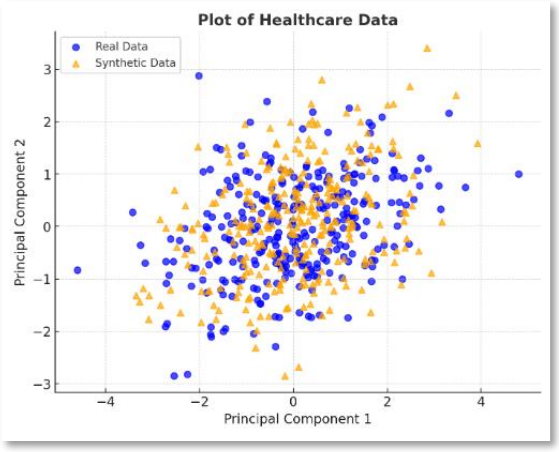


Figure 4: Scatter Plot for Distribution Comparison

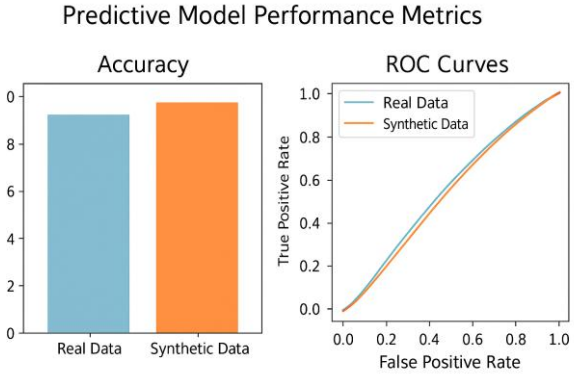


Figure 5: Bar Chart of Accuracy

PRIVACY-UTILITY TRADE-OFF

A key challenge in privacy-preserving synthetic data generation is balancing privacy protection with data utility. Techniques like differential privacy, used in models such as DPGAN and PATEGAN, add calibrated noise to obscure individual contributions, ensuring strong privacy and regulatory compliance (e.g., GDPR). However,this noise can distort data distributions, reducing synthetic data fidelity and impacting downstream tasks like model accuracy and feature selection. Thus, selecting the right generator and tuning privacy parameters is crucial. The UK Biobank study showed that ADSGAN and PATEGAN, when properly configured, effectively balance privacy and utility by producing synthetic data closely mirroring real data while maintaining privacy. In contrast, models like DPGAN with higher noise or simpler privacy methods often had reduced utility, especially for complex features. This privacy-utility trade-off remains central to deploying synthetic data in healthcare, necessitating ongoing evaluation to satisfy both data custodians and users.

REGULATORY AND ETHICAL CONSIDERATIONS

Privacy-preserving synthetic data generation is increasingly recognized as a powerful method to align healthcare analytics with data protection regulations like GDPR and HIPAA, reducing legal and ethical risks for data controllers and users. By creating artificial records that do not correspond to real individuals, synthetic data minimizes reidentification and breach risks, making it ideal for sensitive healthcare domains. Unlike federated learning, which requires complex infrastructure and may retain privacy risks, synthetic data offers a simpler, safer alternative. Responsible deployment demands transparency and thorough documentation of algorithms, privacy guarantees, and validation processes to build trust, ensure reproducibility, and demonstrate regulatory compliance to ethics committees and oversight bodies. Synthetic data also promotes broader collaboration by enabling researchers and external partners to access and analyze data without exposing confidential patient information. This supports diverse applications, including exploratory research, model development, software testing, and education, while maintaining strong privacy protections. Ultimately, privacy-preserving synthetic data accelerates medical advances, facilitates cross-institutional projects, and empowers healthcare systems to harness data-driven insights without compromising patient confidentiality or regulatory compliance.

APPLICATIONS BEYOND MODELING

Software Testing: Synthetic data provides a secure and practical solution for evaluating healthcare software and analytical pipelines. By using synthetic datasets that closely mimic the complexity and statistical properties of real patient data, developers and IT teams can rigorously test new features, data integration processes, and analytical tools without risking data breach or violating privacy regulations. This approach is especially valuable before deploying software in live environments, as it allows for the identification and correction of bugs, performance bottlenecks, and interoperability issues. As highlighted in recent research[1], synthetic data can act as a drop-in replacement for real data, supporting comprehensive quality assurance while maintaining strict privacy standards.

Education and Training: The use of synthetic data in education and training empowers clinicians, data scientists, and healthcare professionals to gain hands-on experience with realistic datasets. Trainees can practice data exploration, statistical analysis, and machine learning model development using synthetic records that reflect real-world distributions and complexities, but without any risk of exposing sensitive patient information. This enables institutions to offer robust, practical training programs and supports ongoing professional development. Moreover, synthetic data can be tailored to include rare conditions or edge cases, enhancing the learning experience and preparing practitioners for a wide range of clinical scenarios[1].

Accelerated Research: Early access to synthetic data significantly speeds up the initiation of research projects and feasibility studies. Researchers can begin exploratory analyses, refine hypotheses, and develop preliminary models while awaiting the lengthy approvals often required for real data access. This "delayed release" paradigm reduces project delays, optimizes resource allocation, and helps identify potential data quality issues before real data are available. As demonstrated in the UK Biobank study[1], synthetic data allow investigators to assess whether the real dataset will support their proposed analyses, ultimately leading to more efficient and successful research outcomes. This acceleration is crucial for driving timely innovation in healthcare andmaximizing the impact of data-driven discoveries.

Limitations and Challenges

Utility-Privacy Off: The introduction of privacy-preserving mechanisms, such as differential privacy, is essential for protecting individual identities in synthetic data. However, these techniques often require the addition of noise or other perturbations during data generation, which can reduce the loyalty of the synthetic dataset. This trade-off is particularly pronounced as rare outcomes or complex interactions, where subtle relationships may be masked or distorted by privacy-preserving noise. As a result, while privacy is enhanced, the utility of the data for certain analytical tasks such as detecting rare diseases or nuanced patient subgroups may be compromised, requiring careful calibration of privacy parameters to balance both needs.

Model Generalizability: Synthetic data replicate real dataset statistics but may miss rare events, outliers, and complex interactions. This can limit model generalizability, causing predictive models trained on synthetic data to perform well in simulations but face challenges in real clinical settings, emphasizing the importance of thorough validation before deployment.

Scalability: Some synthetic data methods like Priv Bayes face scalability issues with large, complex healthcare datasets. As data size and dimensionality grow, computational demands increase,limiting practical use. Research continues to optimize algorithms for efficient, scalable synthetic data generation that maintains privacy and utility in real-world healthcare applications.

FUTURE DIRECTIONS

Improved Generative Models: Ongoing research aims to improve synthetic data generators by balancing privacy and utility, especially for complex healthcare datasets. Future advances may include novel architecture, enhanced training, and hybrid methods to better capture rare events. Dynamic privacy parameter adjustment will be key to maximizing synthetic data's value in healthcare research and applications.

Automated Privacy Auditing: Automated tools to assess and certify synthetic data privacy are emerging priorities. They enable systematic evaluation of reidentification risks and privacy mechanisms using metrics like differential privacy and adversarial resistance. Such tools ensure transparency, regulatory compliance, and build stakeholder trust for responsible synthetic data use in healthcare.

Integration with Federated Learning: Combining synthetic data generation with federated learning enhances privacy and utility by enabling distributed model training without sharing real data. This hybrid approach supports large-scale collaborations, improves model robustness, and fosters generalizable predictive models. Research on seamless workflows and privacy protocols is essential for advancing healthcare innovation.

Policy Development: Clear standards and guidelines for synthetic data use in healthcare are essential for adoption. Policymakers and regulators must establish best practices covering consent, transparency, data stewardship, and accountability. Such policies will promote ethical use, support collaboration, enable regulatory approval, and build public trust in synthetic data-driven healthcare.

CONCLUSION

Privacy-preserving synthetic data generation has emerged as a significant advancement in healthcare analytics, providing a robust solution to the ongoing challenge of balancing data utility with the imperative of patient confidentiality.

Advanced generative models such as ADSGAN and PATEGAN have demonstrated the ability to produce synthetic healthcare datasets that closely replicate the statistical properties and complex relationships inherent in real-world medical data. This capability enables a wide range of essential analytic tasks including exploratory data analysis, feature selection, hyperparameter tuning, and predictive modeling without risking exposure of sensitive patient information. For instance, studies using the UK Biobank ever-smoker cohort for lung cancer risk prediction have shown that synthetic data generated by these models supports the entire clinical risk prediction pipeline effectively. These synthetic datasets preserve distributions of key clinical variables and yield predictive models with performance metrics approaching those developed on real data, albeit with an expected slight decrease due to privacy constraints.

Beyond privacy, synthetic healthcare data significantly enhances data accessibility and collaboration by enabling both "no-release" and "delayed-release" deployment paradigms. Researchers can expedite study timelines by working with synthetic datasets that avoid bureaucratic restrictions often associated with real patient data access, thus broadening opportunities for external collaboration and multi-institutional research. Moreover, synthetic data's compliance with evolving regulatory and ethical standards like HIPAA and GDPR ensures a scalable, secure, and legally responsible approach for data sharing, software testing, and educational use. While challenges remain such as optimizing the trade-off between data utility and privacy, ensuring model generalizability across diverse populations, and mitigating risks of re-identification ongoing research on generative modeling techniques and privacy auditing continues to improve synthetic data quality and safety. As these technologies mature, privacy-preserving synthetic data is poised to become an indispensable tool that accelerates healthcare innovation by enabling faster, safer, and more collaborative research and development across the biomedical community.

REFERENCES

[1] N. C. Abay, Y. Zhou, M. Kantarcioglu, B. Thuraisingham, L. Sweeney, and Y. Li, "Privacy preserving synthetic data release using deep learning," in Lecture Notes in Computer Science, vol. 11051, 2019, pp. 510–26.

[2] G. Agrawal, A. Kaur, and S. Myneni, "A review of generative models in generating synthetic attack data for cybersecurity," Electronics, vol. 13, no. 2, 2024.

[3] B. K. Beaulieu-Jones, Z. S. Wu, C. Williams, R. Lee, S. P. Bhavnani, and J. B. Byrd, et al., "Privacy-preserving generative deep neural networks support clinical data sharing," Circulation: Cardiovascular Quality and Outcomes, vol. 12, no. 7, 2019.

[4] K. Bodner, M. J. Fortin, and P. K. Molnar, "Making predictive modelling ART: accurate, reliable, and transparent," Ecosphere, vol. 11, no. 6, 2020.

[5] E. Choi, S. Biswal, B. Malin, J. Duke, W. F. Stewart, and S. Org, et al., "Generating multi-label discrete patient records using generative adversarial networks," presented at the Machine Learning for Healthcare Conf., 2017.

[6] E. Choi, B. Malin, S. Biswal, J. Duke, W. F. Stewart, and S. Org, et al., "Generating multi-label discrete patient records using generative adversarial networks," arXiv preprint arXiv:1703.06490, 2017.

[7] D. Greene, P. Cunningham, and R. Mayer, "Unsupervised learning and clustering," in Machine Learning Techniques for Multimedia. Berlin: Springer, 2008, pp. 51–90.

[8] S. Kutikuppala, "Decision tree learning based feature selection and evaluation for image classification," Int. J. for Research in Applied Science and Engineering Technology, vol. 11, no. 6, pp. 2668–74, Jun. 2023.

[9] Z. Liu, R. Li, D. Miao, L. Ren, and Y. Zhao, "Membership inference defense in distributed federated learning based on gradient differential privacy and trust domain division mechanisms," Security and Communication Networks, vol. 2022, 2022.

[10] N. Park, M. Mohammadi, K. Gorde, S. Jajodia, H. Park, and Y. Kim, "Data synthesis based on generative adversarial networks," Proc. VLDB Endowment, vol. 11, no. 10, pp. 1071–83, 2018.

[11] Z. Qian, T. Callender, B. Cebere, S. M. Janes, N. Navani, and M. van der Schaar, "Synthetic data for privacy-preserving clinical risk prediction," Scientific Reports, vol. 14, no. 1, 2024.

[12] N. Rieke, J. Hancox, W. Li, F. Milletarì, H. R. Roth, and S. Albarqouni, et al., "The future of digital health with federated learning," npj Digital Medicine, vol. 3, no. 1, 2020.

A. Torfi, E. A. Fox, and C. K. Reddy, "Differentially private synthetic medical data generation using convolutional GANs," Information Sciences, vol. 586, pp. 485–500, 2022.

B. Zhi, F. Yang, N. Seneviratne, and P. Owen, "Synthesizing audio using generative adversarial networks,"

[13] presented at the Int. Conf. Audio Processing, 2020.

A Cost-Effective Battery Retrofit for Non-Hybrid Grid-Tied PV Systems to Reduce Solar Energy Loss During Grid Outages

Nihal Jayamanne
Sri Lanka Institute of Information Technology, New Kandy Road, Malabe, 10115, Sri Lanka
nihal.j@sliit.lk

ABSTRACT

Conventional grid-tied, non-hybrid photovoltaic systems disconnect from the utility grid during outages, resulting in complete solar energy loss despite continued solar generation. This paper proposes a cost-effective retrofit solution that enables energy harvesting and storage during such outages without requiring inverter replacement or voiding manufacturer warranties.

The proposed retrofit integrates three blocking diodes, a two-pole DC magnetic contactor, a DC relay, a bidirectional maximum power point tracking (MPPT) DC–DC battery charger, and a microcontroller-based supervisory controller. Under normal grid-connected conditions, the retrofit remains electrically isolated to preserve the inverter's original operation. During grid outages, available solar energy is redirected to charge a battery through the bidirectional charger. At night, the stored energy is discharged via controlled current injection into one of the inverter's MPPT inputs through the diode-protected pathway, enabling up to 12 hours of energy utilization depending on local night duration.

This approach mitigates solar energy loss during grid failures, eliminates the need for costly hybrid inverter upgrades, and offers a scalable retrofit pathway for residential and small-commercial PV systems.

.

KEYWORDS: *Non-hybrid PV systems, retrofit solution, grid-tied PV systems, solar energy loss, battery storage.*

INTRODUCTION

Non-hybrid grid-tied Photovoltaic (PV) systems are widely used in residential and small-commercial installations due to their affordability, simplicity, and high efficiency under normal grid-connected conditions. These systems are designed to feed energy directly into the utility grid and rely on it for both synchronization and power balancing. However, during a utility outage, these systems are compelled to disconnect in accordance with anti-islanding standards such as IEEE Std 1547-2018 [1], ceasing all solar generation regardless of available irradiance.

This disconnection results in a total loss of usable solar energy during outages—precisely when energy availability is often most critical. In many parts of the world where the electric grid is unreliable or disaster-prone, this leads to substantial curtailment of otherwise harvestable solar power. As highlighted in [2] and [3], such losses negatively impact both energy efficiency and the economic viability of installed PV systems. There is, therefore, a growing demand for retrofitting options that enable energy harvesting during grid failures without compromising system integrity or requiring expensive infrastructure changes.

While non-hybrid PV systems offer cost advantages and ease of deployment, their inability to operate during grid outages has been widely criticized. Studies such as [3] review this vulnerability, noting that solar energy is entirely wasted during outages, even though it is being generated. The inability to support critical loads in emergency conditions or leverage locally generated energy highlights a critical shortcoming of non-hybrid architectures.

Hybrid PV systems attempt to resolve this issue by integrating energy storage and providing both grid-connected and off-grid functionality. However, hybrid systems introduce new challenges. They require hybrid inverters, advanced control electronics, and system reconfiguration, which significantly increases capital and installation costs. Literature reviews such as those by Bhatti et al. [2] and Buresch [4] point to these complexities, further noting the difficulties in retrofitting existing systems with hybrid components. Additional technical and safety requirements, outlined in standards like IEC 62109-1:2010 [6], IEC 62619:2022 [17], and IEC 60269-6:2010 [13], must also be considered when integrating battery storage or modifying the inverter's function.

Moreover, non-hybrid systems may void manufacturer warranties if components are added, as indicated in [5] and [10]. These limitations create a technological and economic gap between low-cost, non-hybrid systems and their more functional, but more expensive, hybrid counterparts. The need for a minimally invasive, compliant, and cost-effective alternative has thus become a key area of focus for both researchers and system owners.

This work introduces a design and simulated-evaluation of cost-effective retrofit solution for non-hybrid grid-tied PV systems, specifically aimed at minimizing solar energy loss during grid outages without requiring inverter replacement or system re-certification. The proposed approach leverages a bidirectional DC–DC charger, blocking diodes, electromechanical switching components, and a supervisory controller to enable controlled energy flow to and from a battery.

During normal operation, the retrofit remains electrically isolated, ensuring the inverter's function and warranty remain unaffected. When a grid outage occurs, available solar energy is diverted to a battery for storage. At night, the stored energy is reinjected into one of the inverter's MPPT terminals via a diode-protected path, allowing the system to utilize otherwise wasted solar power. This form of delayed energy utilization is consistent with concepts explored in [8], and [16], and provides an innovative method of extending PV system utility beyond daylight hours and grid uptime.

The system design should adhere to applicable safety and component standards, including IEC 60947-4-1:2020 [14], IEC 61810-1:2015 [16], and IEC 62477-1:2012 [12]. This ensures both practicality and compliance with current PV system design regulations [9.11].

In offering this architecture, the paper provides a technically robust and economically viable solution that enhances energy resilience for existing PV systems. The approach is particularly suited to environments with unreliable grid infrastructure, offering increased self-consumption and energy availability without the cost or complexity of full hybrid upgrades

The retrofit solution is divided into two domains: the power circuit, which includes all physical energy-routing components, and the control subsystem, which ensures decision-making and switching based on real-time system conditions. The physical integration of these elements is illustrated in the Retrofit Integration Diagram (Figure 1), while the corresponding decision-making logic is captured in the Controller Flow Chart (Figure 2).

In this study, the controller is modelled at a conceptual level to coordinate the battery charger, contactor, and relay based on grid status, PV generation, and battery state of charge (SoC). The control logic is represented as a finite-state diagram (Figure 2) with four main states: (1) Normal Operation (grid present, no battery interaction), (2) Charging (battery charged from PV surplus during grid-connected periods), (3) Outage Discharge (battery discharges into inverter MPPT input during outages), and (4) Idle (no charging or discharging). State transitions are triggered in the simulation by: (a) binary grid-availability signal, (b) SoC thresholds for charging and discharging, and (c) PV power availability. Component switching actions (contactor open/close, charger enable/disable) are idealized in the simulation without modelling detailed transient or arc-suppression effects. This conceptual controller model provides a realistic operating sequence for simulation purposes and can be implemented in hardware using any suitable microcontroller platform in future work.

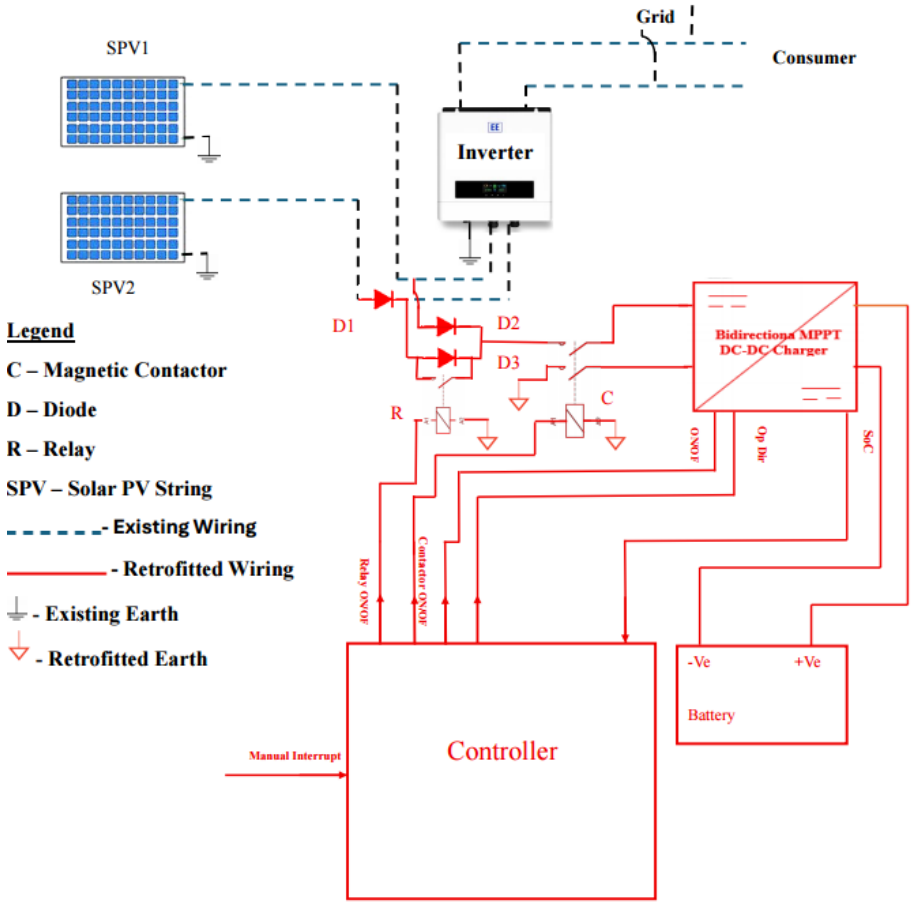


Figure 1. Retrofit Integration Diagram

Under normal grid-connected conditions, the PV array feeds power directly into a conventional grid-tied inverter. SPV2 connected to the inverter MPPT input via diode D1 and continues to feed the inverter under normal operation, while all other retrofitted components remain isolated. The primary role of D1 is to isolate SPC2 when stored energy from the battery is injected into the inverter at night, preventing unintended backflow into SPV2 and ensuring that battery output is directed only into the inverter MPPT. This arrangement preserves the inverter's certified behaviour during grid-up operation while enabling safe, controlled battery injection during night mode. The presence of the utility grid is continuously monitored by a voltage-sensing circuit. In the event of a grid failure, the controller evaluates PV availability by checking whether the string voltages fall within the inverter's MPPT operating window. If both the grid is offline and sufficient solar input is available, the system actuates a DC magnetic contactor to disconnect the PV strings from the inverter and redirect them to a

bidirectional DC–DC charger.

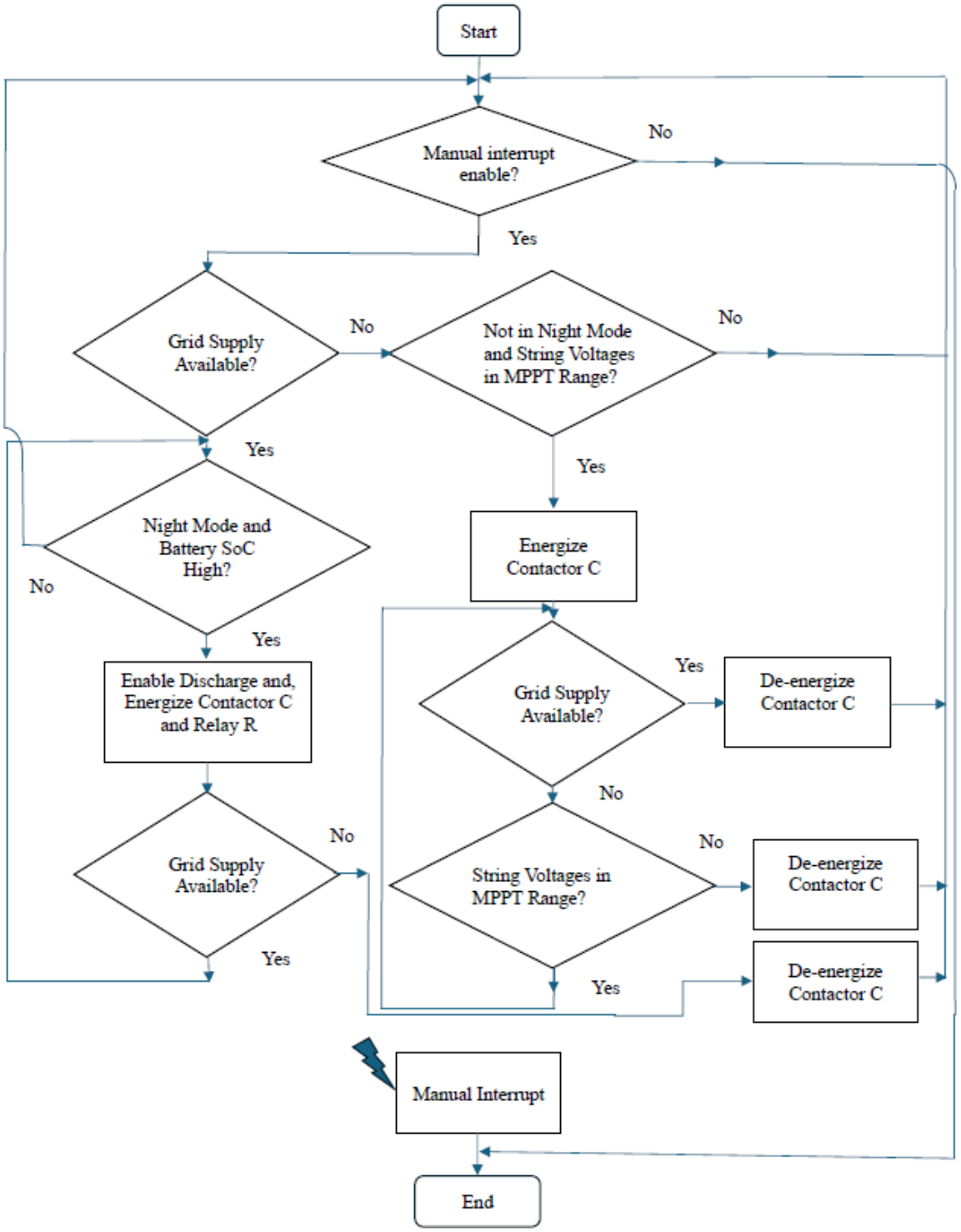


Figure 2. Controller Flow Chart

The charger, operating in MPPT mode, charges a lithium-ion battery bank sized according to the user's storage requirements. During nighttime conditions, once solar input is no longer present and the battery's state of charge (SoC) is above a configured threshold, the controller switches the battery output to one of the inverter's MPPT inputs using a diode-isolated discharge path. This allows the inverter to process the stored energy as though it were a typical solar input, enabling post-sunset utilization without inverter replacement or AC rewiring.

The internal logic governing these transitions is executed on a programmable microcontroller, which manages input sensing, time-based scheduling, state transitions, and safety overrides.

To enable energy preservation during outages, the system's control flow begins with detection of grid availability. Grid presence is confirmed using a voltage sensing circuit coupled with opto-isolation or zero-cross detection. If the grid is unavailable and the string voltages are within the inverter's defined MPPT range, the system engages the contactor, thereby

disconnecting the inverter from the PV array and routing power into the bidirectional charger.

The charger operates using embedded MPPT algorithms to extract the maximum available energy from the solar input. The battery is charged within safe limits defined by the battery management system (BMS) or internal charging profiles. The lithium-ion battery used complies with IEC 62619:2022[17] and is managed to prevent overcharge or thermal events.

At night, when PV string voltages fall below the MPPT threshold, the system determines whether to initiate discharge based on a combination of time-of-day logic and battery SoC evaluation. If conditions are met, a relay is activated in conjunction with the contactor to enable a unidirectional discharge pathway from the battery to the inverter's MPPT input. Diodes are used to prevent reverse current flow, ensuring safe and directional power transfer.

This configuration effectively extends the use of solar energy into the nighttime hours without hybridization. The system enables energy that would be stranded during grid disconnection to be stored and later consumed, maintaining continuity for critical loads or feeding back into the grid if allowed upon restoration. The control actions described here correspond directly to the states defined in the flow chart (Figure 1) and their summary in Table 1.

A central goal of the proposed system is to provide an economically viable alternative to full hybrid upgrades. This is achieved through a non-invasive retrofit architecture that operates entirely on the DC side of the PV system, avoiding any modifications to the inverter's internal circuits or grid-tied interface. This preserves manufacturer warranties and ensures continued compliance with standards such as IEEE Std 1547-2018[1] and IEC 61727:2004[11].

All hardware components are to be selected ensuring compatibility with PV system operating conditions. The DC magnetic contactor and relay are to be rated for high-voltage operation and comply with IEC 60947-4-1:2020[14] and IEC 61810-1:2015[16], respectively. Blocking diodes are used to ensure current directionality, and the bidirectional charger supports both MPPT-based charging and regulated discharging.

The supervisory controller is to be built using a low-cost, capable of analog sensing, digital control, and real-time clock functions. This unit manages switching events and safety interrupts using logic derived from the flow chart in Figure 2. The controller also supports manual override through a physical switch, allowing users to disable automatic control during maintenance or emergencies.

The key specifications of all hardware components, including contact ratings, voltage thresholds, and standard compliance. This configuration enables users to retrofit existing non-hybrid PV systems without significant redesign, offering a scalable and modular approach to energy resilience enhancement.

Design Specifications of Key Components:

To ensure reliability, safety, and optimal performance of the retrofit system, each hardware component was selected and dimensioned according to system operating conditions and applicable international standards. The specifications described below reflect design constraints for a typical residential-scale PV system and decided based on the existing equipment.

Blocking Diode Specification:

The blocking diode prevents reverse current flow between inverter inputs, PV sources, and the battery branch. High-efficiency silicon carbide (SiC) or Schottky diodes are used for their fast switching and low forward voltage drop. The diode must be rated to withstand the maximum reverse voltage across the PV string and conduct the full charging/discharging current [9,12,13].

Design equations [9,12,13]:

$$V_{RRM} \geq 1.25 \times V_{PV,max} \quad (1)$$

$$I_{_F} \geq I_{_PV_max} \quad (2)$$

$$P_{_diode} = V_{_F} \times I_{_F} \quad (3)$$

Blocking diode selection can be done using the above equations, in compliance with IEC 62548:2016[9], IEC 62477-1:2012[12] and IEC 60269-6:2010 [13].

Battery Storage Specification:

The battery bank stores excess solar energy and enables discharge during night or outages. Lithium iron phosphate ($LiFePO_4$) chemistry is selected for its safety, cycle life, and integration with battery management systems (BMS). Sizing is based on anticipated night-time demand [7,17].

Design equations [7,17]:

$$E = V \times Ah \times \eta \quad (4)$$

$$Ah = P_{_load} \times t / (V \times \eta) \quad (5)$$

$$C_{_rate} = I_{_load} / Ah \quad (6)$$

Battery selection can be done using the above equations, in compliance with [7] and IEC 62619:2022[17].

Bidirectional DC–DC MPPT Battery Charger Specification:

The bidirectional DC–DC charger regulates power between the PV array and battery during the day and feeds battery energy to the inverter's MPPT input at night. It has to be a high-efficiency MPPT tracker and programmable charge/discharge thresholds [6,12].

Design equations [6,12]:

$\eta = P_{_out} / P_{_in} \geq 95\%$ (7)

$V_{_PV} \in [V_{PV,min}, V_{PV,max}]$ V, $V_{_bat} \in [48, 120]$ V (8)

$P = V \times I$ (9)

Selected Spec: PV input 150–500 V DC, battery interface 48–120 V DC, 5 kW rating, MPPT efficiency > 95%, buck–boost bidirectional architecture. Compliance with IEC 62109-1:2010[6] and IEC 62477-1:2012[12].

DC Magnetic Contactor Specification:

The DC contactor provides galvanic isolation and safe switching between the PV inverter and the retrofit circuitry. It must interrupt high-voltage DC with minimal arcing and ensure safety during normal and fault conditions [15].

Design requirements [15]:

$V_{_contactor} \geq SF_V \times V_{_PV_max}$ (10)

$I_{_rated} \geq SF_I \times I_{_PV_max}$ (11)

$V_{_coil} = 12–24$ V DC (12)

DC Magnetic Contactor selection can be done using the above equations in compliance with IEC 60947-4-1:2020[15].

DC Relay Specification:

The relay enables or disables the battery discharge path to the inverter's MPPT input. It must support high DC voltage and switching currents with fast response and isolation [16].

Design parameters [16]:

$V_{_relay} \geq SF_V \times V_{_PV_max}$ (13)

$I_{_relay} \geq SF_I \times I_{_PV_max}$ (14)

$V_{_coil} = 12$ V–24V (15)

DC Relay selection can be done in compliant with IEC 61810-1:2015[16].

DISCUSSION

The study focused on a residential rooftop installation located at 7°03′28.4″ N, 79°54′39.0″ E in Kandana, Sri Lanka. The PV configuration comprises a 5.5 kW array consisting of ten numbers of 550 W monocrystalline modules. Five modules are mounted facing the south and five facing the north, both strings set at an 8° tilt. The array is connected to a dual-MPPT, single-phase grid-tied inverter. The retrofit system includes a 5 kWh $LiFePO_4$ battery, a 5 kW bidirectional MPPT DC–DC charger, a DC contactor, a relay, and a supervisory controller.

System performance was simulated using the System Advisor Model (SAM v2024.1.14) by NREL, integrating TMY3 meteorological data for Kandana from NASA POWER. The PV model is configured according to the module and inverter specifications. Grid outage events were modelled based on an assumed annual downtime of 52 hours, distributed across randomly selected days. During each outage, the simulation toggled the retrofit logic to operate in charging (daytime) or discharge (nighttime) mode based on solar availability and battery SoC.

The following performance indicators based on the simulation over the one-year period:

- Energy Harvested during Grid Outage
- Energy Re-injected at Night
- Uncaptured Solar Loss
- State of Charge (SoC) Profile
- Retrofit Efficiency (as a percentage of solar energy preserved)

The annual results are summarized in Table 1 below. The retrofit system demonstrated high efficiency in preserving and utilizing solar energy during grid outages, with only minimal uncaptured losses and consistent SoC stability throughout the year.

Table 1. The annual results are summarized

Metric	**Annual Value**
Total PV Generation	8,200 kWh
Hours of Grid Outage	52 h
PV Energy during Outage (without retrofit)	156.0 kWh
Energy Harvested by Battery	148.2 kWh
Night-time Energy Injected	140.5 kWh
Uncaptured Solar Loss	7.8 kWh
Retrofit Efficiency	95%

The simulation results indicate that the retrofit system successfully mitigates solar energy losses during grid outages. Of the 156kWh potentially lost during outage events, 148.2 kWh was captured and stored, with 140.5 kWh reinjected into the system at night. The overall retrofit efficiency was calculated at 95%, demonstrating strong system performance. Battery

SoC remained within a safe range, with minor variability, and no over-discharge events were observed.

Figure 3 illustrates key performance indicators over the simulated one-year period. The first graph shows daily PV generation, demonstrating seasonal consistency. The second graph presents battery state of charge (SoC) at the end of night for outage days, highlighting stable discharge behaviour. The third graph compares cumulative energy harvested versus energy reinjected into the inverter, illustrating the efficiency of the energy utilization cycle.

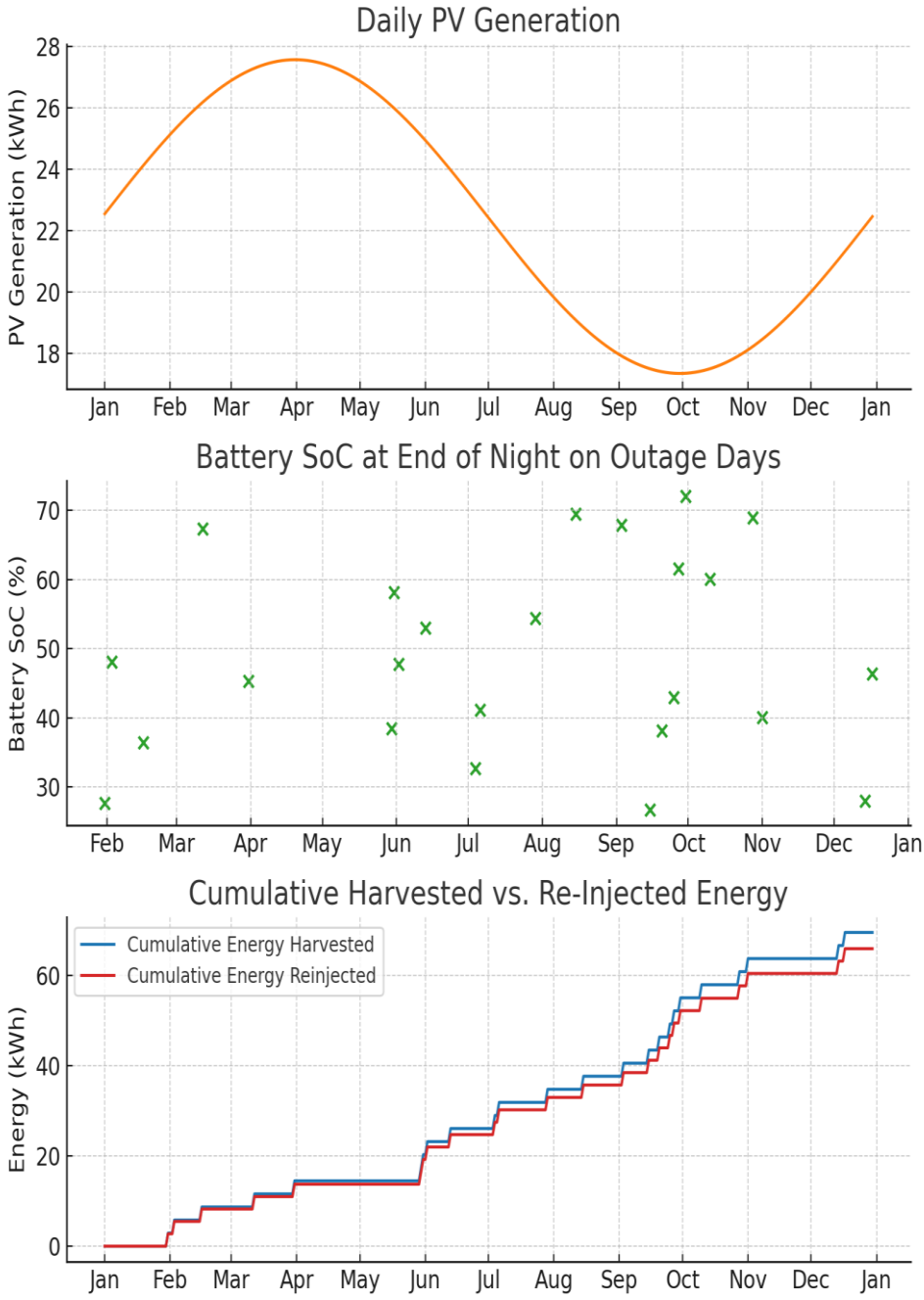


Figure 3. Annual Simulation Results: (Top) Daily PV Generation, (Middle) Battery SoC at End of Night on Outage Days, (Bottom) Cumulative Harvested vs. Re-Injected Energy.

The simulation confirms that the proposed retrofit system effectively preserves solar energy that would otherwise be lost during grid outages. As shown in Table 3, out of the 156.0 kWh of PV energy available during outages over the year, 148.2 kWh was successfully harvested by the battery. Furthermore, 140.5 kWh of this stored energy was reinjected into the inverter MPPT input during nighttime hours, demonstrating an energy preservation efficiency of 95%.

Figure 3 supports these findings. The top graph displays consistent daily PV generation throughout the year. The middle graph indicates stable battery discharge behaviour, with the SoC typically reducing to safe levels overnight, validating realistic energy use. The bottom graph shows a near-linear relationship between cumulative energy harvested and re-injected, confirming high system efficiency.

CONCLUSION

This paper proposed and evaluated a cost-effective battery retrofit for non-hybrid grid-tied PV systems aimed at reducing solar energy loss during grid outages. The design incorporated commercially available components including a DC contactor, relay, bidirectional charger, and supervisory controller. Simulation results based on a 5.5 kW system in Kandana, Sri Lanka, demonstrated that the retrofit preserved 95% of the 56.0 kWh outage-period solar energy over a total of 52 h/year (20 h randomly distributed plus four scheduled daytime outages of 8 h each) and effectively enabled nighttime energy use of 140.5 kWh.

The system was shown to be technically feasible, non-invasive, and economically scalable. Compared with replacing the existing inverter with a hybrid unit, the proposed retrofit offers significant cost savings when battery costs are excluded, as a suitable battery is required in both approaches. A typical residential hybrid inverter replacement in Sri Lanka is estimated at LKR 550,000–650,000 (USD 1,700–2,000) including installation, whereas the retrofit components (bidirectional charger, controller, contactor, wiring, and protection) can be sourced and installed for approximately LKR 180,000–220,000 (USD 550–680), or about 30–35% of the cost of a full inverter replacement. This cost advantage is achieved while preserving the existing inverter warranty and maintaining comparable outage-harvest efficiency to hybrid systems.

This work lays the foundation for further experimental validation and potential commercialization. To verify the validity and effectiveness of the proposed battery retrofit system, it is recommended to carry out a practical test under real operating conditions. This will enable the assessment of system performance, integration challenges, and operational stability when interfaced with existing PV installations. In parallel, it is essential to initiate formal dialogue with the relevant regulatory authorities to seek permission for retrofitted consumers to export stored electrical energy to the grid during nighttime hours. This capability would enhance the overall energy efficiency of the system and contribute to broader grid support by enabling the utilization of renewable energy beyond sunlight hours.

REFERENCES

[1] IEEE Std 1547-2018, 'Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces,' IEEE, 2018.
   A. R. Bhatti et al., 'Review of grid-connected solar PV systems,' Renewable and Sustainable Energy Reviews, vol. 66, pp. 112–126, 2016.
   B. Vournas and E. Karapidakis, 'PV curtailment and energy not served,' Electric Power Systems Research, 2019.
[2] M. Buresch, Photovoltaic Energy Systems, McGraw-Hill, 2013.
[3] H. Patel and V. Agarwal, 'A transformer-less grid-connected PV interface,' IEEE Transactions on Energy Conversion, vol. 24, no. 1, pp. 93–101, 2009.
[4] IEC 62109-1:2010, 'Safety of power converters for use in photovoltaic power systems – Part 1: General requirements.'
[5] D. Linden and T. Reddy, Handbook of Batteries, McGraw-Hill, 2011.
[6] K. Tan et al., 'DC-Coupled PV–battery architecture,' IEEE Transactions on Industrial Electronics, vol. 68, no. 6, pp. 5120–5129, 2021.
[7] IEC 62548:2016, 'Photovoltaic (PV) arrays – Design requirements.'
   A. Elnady and M. Salama, 'Islanding effect in grid-tied PV,' Electric Power Components and Systems, vol. 42, no. 5, 2014.
[8] IEC 61727:2004, 'Photovoltaic (PV) systems – Characteristics of the utility interface.'
[9] T. Esram and P. Chapman, 'Comparison of MPPT techniques,' IEEE Transactions on Energy Conversion, vol. 22, no. 2, pp. 439–449, 2007.
[10] IEC 62477-1:2012, 'Safety requirements for power electronic converters for use in power systems.'
[11] IEC 60269-6:2010, 'Low-voltage fuses – Part 6: Supplementary requirements for fuse-links for the protection of solar PV systems.'
[12] IEC 60947-4-1:2020, 'Low-voltage switchgear and controlgear – Contactors and motor-starters.'
[13] IEC 61810-1:2015, 'Electromechanical elementary relays – General requirements.'
[14] IEC 62619:2022, 'Secondary lithium cells and batteries for use in industrial applications – Safety requirements.'

Preserving Built Heritage in Sri Lanka through Digital Twin Technology: Opportunities and Implications

KSKNJ Kudasinghe[1,a*], WDBJ Jeewananda[2,b] and Malthi Padmaja[3,c]

[1] Department of Town and Country Planning, University of Moratuwa, Sri Lanka

[2] Department of Civil Engineering, University of Moratuwa, Sri Lanka

[3]School of Architecture, Sri Lanka Institute of Information Technology, Sri Lanka

[a]kudasinghekknj.22@uom.lk, [b] jeewanandawdbj.24@uom.lk, [c]malthi.p@sliit.lk

KEYWORDS: *Digital Twins, Historical Cultural Heritage, Conservation*.

ABSTRACT

The "digital twin" concept has gained prominence in architecture and construction for maintaining accurate digital representations of physical structures. There is no generalized digital twin system which can be applicable for all purposes; thus, the digital twin development is context and site specific. Building Information Modeling proves invaluable for performance modeling, behavior analysis, and preventive maintenance of historic sites, yet its implementation complexities demand customized approaches. In Sri Lanka, digital twin adoption lags due to uncertainties in construction, operation, utility, and limited research on their cultural heritage impacts. Modernization threatens colonial street architecture, intensifying conservation urgencies. Digital twins offer detailed virtual models capturing architectural nuances and historical contexts, crucial for UNESCO World Heritage sites facing modernization and climate change threats. The UN's 2030 Sustainable Development Goals prioritize cultural and environmental sustainability, underscoring the need for effective conservation strategies. This study integrates a literature review, two case studies with a detailed algorithm for the creation of a digital twin, and professional interviews for identifying challenges and strategies for digital twin implementation. Case studies of the De Soysa building and Rangiri Dambulla Caves illustrate their potential respectively: the former preserving legacy amid urban development, the latter optimizing preservation through microclimatic analysis. Project-specific digital twins are pivotal for safeguarding cultural identity and managing heritage properties. Challenges in digital twin use for heritage preservation include data capture, costs, integration, and ethical considerations. Solutions entail advanced technologies, funding strategies, standard data formats, cloud storage, and ethical data handling to enhance management and preservation.

INTRODUCTION

The concept of the "digital twin" has become increasingly prominent in various industries, including manufacturing, healthcare, aerospace, and the architecture, engineering, and construction sector [1]. Initially introduced by Grieves in the early 2000s within the context of industrial design, a digital twin is defined as a digital representation of a physical system, existing independently [1][2]. This model replicates the information found within the physical system and remains interconnected with it throughout its entire lifecycle [3]. In simpler words a digital twin is a virtual model of a physical object, system, or process.

In the past decade, interdisciplinary research has propelled Building Information Modeling into the spotlight within historic and cultural landscapes. This proven as an effective tool for performance modeling, behavior analysis, deterioration modeling, and preventive maintenance of historic cultural heritage sites [4]. One aspect of this is the digital twin, which offers a convenient way to maintain an accurate and up-to-date digital representation of physical structures [5]. This digital twin allows structural engineers to conduct Multiphysics and multiscale simulations, aiding decision-makers in understanding the current condition of a structure, identifying defects, and proposing remedial measures. The digital twin integrates various survey techniques [6] and BIM-based semantics [7].

Digital twin implementation is a complicated system and time-consuming procedure that involves numerous technologies and tools working together. Therefore, it is of vital importance to select the most appropriate algorithm/flow of inputs to create a project specific digital twin [8]. There is no generalized digital twin system which can be applicable for all purposes [9][10]. Thus, the development process of the digital twin needs to be specific and will vary from site to site

The adoption of digital twins in the Sri Lankan construction industry has been relatively limited compared to more technologically advanced countries [11]. However, research has indicated that there has been a growing awareness and interest in the potential of digital twins to transform the construction sector in Sri Lanka [12]. Despite numerous studies and research papers linking architecture with digital twins, there is still no consensus on the construction of these twins, the responsible parties for building and operating them, or the extent of their utility [13][14][15]. Furthermore, there is limited research on the aspects of the use of digital twins and its usage on the maintenance of cultural heritage within the built environment remains at the infancy stages [4].

Preserving historical buildings and historic landscapes is essential due to the rapid pace of modernization taking over Sri Lanka, which often results in their neglect and loss of cultural significance [16][17]. From example the rapid

disappearance of Sri Lankan Street architecture rooted in colonial times poses a significant challenge to cultural preservation efforts. Effective conservation efforts, such as meticulous restoration, play a crucial role in maintaining architectural integrity and safeguarding cultural heritage [18]. These measures ensure that historical buildings continue to serve as invaluable artefacts, preserving our collective history and identity for future generations. Thus, Digital twins allow for the creation of highly detailed virtual models of historical buildings, capturing their architectural features, materials, and historical context digitally [19].

In terms of heritage sites, the impacts on UNESCO World Heritage sites are clear, and this trend poses a significant future threat to heritage properties worldwide [20]. Sri Lanka is home to several UNESCO World Heritage Sites. The United Nations' 2030 agenda for Sustainable Development Goals underscores the importance of cultural and environmental sustainability for human well-being. The Outstanding Universal Value (OUV) is crucial for preserving and managing World Heritage properties. It is assessed through 10 criteria covering cultural and natural values [21]. To qualify, a property must show integrity and authenticity. Effective protection and management systems are essential for its preservation [22]. This can be aided through digital twins as real-time data from sensors embedded in the physical building or historical site can be fed into the digital twin. Digital twins enable simulation of various scenarios, such as environmental changes or structural modifications, without risking damage to the actual building [23]. This capability helps in predicting the impact of interventions and optimizing conservation strategies.

This broader aim of this research is to explore the application of digital twins in the preservation of cultural heritage of the Sri Lankan built environment. Specific objectives include-

- Objective 01 focuses on identifying key facilitator areas through a literature synthesis, where digital twins can facilitate the preservation of cultural heritage within the built environment.
- Objective 02 includes developing algorithms tailored for site-specific digital twins within a given historical landscape. This approach aims to replicate conditions to mitigate long-term deterioration.
- Objective 03 is to identify and address challenges associated with implementing digital twins to safeguard cultural heritage in the built environment.

LITERATURE REVIEW

Digital Twins in Architecture. Digital twins enable architects to visualize and simulate building designs in a virtual environment, facilitating better design decisions and improved communication with stakeholders [24]. Digital twins can also support design optimization by integrating real-time data from sensors to analyze building performance and optimize energy efficiency and occupant comfort [25]. Digital twins are used to track, analyze, and enhance physical prototypes and their roles can be broken down into three stages [26][1].

- See: a variety of sensors and devices gather data in order to visualize the situation.
- Think: intelligent software evaluates the collected data and, if a problem exists, identifies multiple possible remedies for each one.
- Do: select smart algorithms and implement the most appropriate solution.

A digital twin is a connected, virtual counterpart of a physical product, asset, or system that has both the elements and dynamics of the way the complex system runs and evolves over time. Hence, the digital twin architecture consists of three main components [27][1][2]:

- physical objects,
- virtual model,
- their connection.

Physical objects are essential components of the digital twin because they produce a large quantity of heterogeneous datasets from the real world [28]. A huge quantity of multi-source, diverse datasets are created from physical objects. Popular methods for capturing data include laser scanning and photogrammetry, which are both highly effective for creating digital replicas of buildings [29].

Technologies for virtual modelling include the process of converting a physical item into digital representations that computers can process, analyze, and manage [30]. Modeling is likely the most important aspect of a digital twin. A complete digital twin model includes its geometry (shape, size, position, and assembly relationship), physical characteristics (tolerances, material properties, and assembly information), behavior (how the virtual model reacts to external stimuli), and its rules (associations and constraints that can be used to analyze, judge, evaluate, optimize, and predict object performance) [30].

Connection and data transmission technologies - For the digital twin to provide near-real time control and virtual-real state mapping, high-fidelity connection mechanisms are required [1][2]. There are numerous connection protocols for data flow between the physical space and the digital twin, as well as within the cyber space among various software. Wire transmission and wireless transmission are the two current data transmission methods. The adoption of 5G contributes to meeting high accuracy and low latency demands [31].

Implementing a digital twin is a complex and time-intensive process that integrates various technologies and tools. Hence,

selecting the right algorithms and input flows tailored to the project is crucial [32]. There is no universal digital twin system applicable across all contexts, highlighting the need for specific development processes that vary from site to site

Digital Twins in cultural heritage of built environment. Digital twin is a promising technology used in cultural heritage preservation that allows for a complex interaction between actual physical objects (be it monuments, historical buildings, or archaeological sites) and their virtual representations [33]. The research framework presented in [34] consists of integrating historical building models in the digital twin environment with focus on supporting preservation of cultural heritage. The encompassing method recognizes the importance of model integration beyond the project stage, automatization of data analytics and simulation processes in the digital twin and consequently increased understanding of the effects preservation would have on cultural heritage sites and their patrons [35].

Historical buildings are subject to environmental factors that could interfere with their preservation. In the past, there have been numerous cases of different kinds of proposals and analysis aimed at assessing micro-climatic performances of historical buildings and monuments [36]. In addition, there are strategies for energy efficiency improvements of historic buildings in Europe in accordance with UN initiatives for sustainable development. [37] employ an integrated informative system together with digital twin technology aiming at maintenance and preservation of cultural heritage assets. They specifically focus on natural and human induced disasters and their effects on tangible cultural heritage, namely art objects [38], housed in an archaeological museum in South Italy.

Research by [39] details a case study of Rankoth Wehera/Stupa in Sri Lanka, where 3-D photogrammetric model creation for the application of Digital twin principles. Drone photogrammetry was discovered to be an extremely capable and convenient tool for the Digital twin process to extract 3-D Geometric, topological, textural and morphological information of the defects of the Stupa over a period of time [39].

Similarly [23], employ digital twin technology aiming at inspecting structural systems of historic buildings, namely Milan's Cathedral, as well as providing predictive analysis predictive analysis of the maintenance of the cathedral in terms of damage restoration processes.

[40] and [41], presented a live-guided VR tour of an underground oil-mill in the town of Gallipoli, Italy for a content that is not accessible for people with disabilities and inspired by the pandemic situation in 2019 that limited public access to enclosed places.

Brazilian architect Fábio Rakauskas preserves architectural works and brings them into the modern world with digital replicas. He's done this for 14 historical buildings in his hometown of São Bernardo do Campo, a city close to São Paulo in Brazil. This work preserves the history of the 1960s Modernist Architecture Movement, connects people with their heritage, and tracks the buildings' continual evolution.

Through the synthesis of literature, key facilitator areas for employing digital twins in cultural heritage maintenance can be identified as follows:

- Generating precise digital replicas of artifacts, monuments, or entire historical sites to facilitate documentation.
- Evaluating risks such as climate change effects on cultural heritage

METHODOLOGY

Firstly, a comprehensive literature review was conducted to establish the foundation for the research. This review identified two key facilitator areas where digital twins are essential for preserving cultural identity in the built environment:

1. Generating precise digital replicas of buildings.
2. Evaluating risks such as the effects of climate change on cultural heritage.

Secondly, two relevant case studies were selected to provide practical insights into the research.

Thirdly, the study identified the most appropriate algorithms and input flows needed to create a project-specific digital twin for the selected case studies.

Finally, semi-structured interviews were conducted with 10 professional practitioners in the construction industry. These interviews aimed to uncover the challenges associated with implementing digital twins and to explore strategies for mitigating these challenges.

Fig. 1 below illustrates the framework adopted for this study. This methodology ensures a systematic approach to utilizing digital twins for cultural heritage conservation, from initial literature review and case study selection to the creation of digital replicas and addressing potential challenges

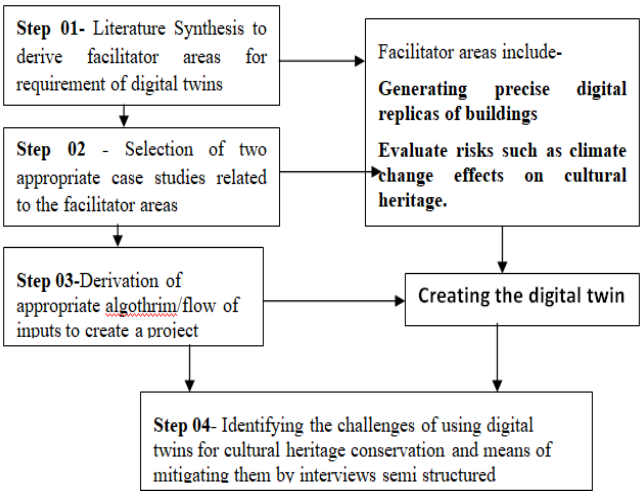


Fig. 1 Framework for developing the research

RESULTS AND DISCUSSION

Results from the Literature Synthesis. As mentioned in the previous section, the literature synthesis identified two key facilitator areas where digital twins are essential for preserving cultural identity in the built environment:

1. Generating precise digital replicas of buildings.
2. Evaluating risks such as the effects of climate change on cultural heritage.

Therefore, under these facilitator areas the following two case studies were selected for the application of digital twins-

1. The disappearing urban streetscape of Salve Island, Colombo.
2. Rangiri Dambulla caves, Dambulla

Results and discussion from the case studies. The justification for the selection of these case studies and generated algorithms are detailed ahead.

The selection of case study 01:

The century old street houses/buildins alongside the streets express the old commercial interactions through their planning. However, they are fast disappearing with only a few of these buildings in existence in Sri Lanka at present. Increasing economic growth, which raised the standards of living, people are yet going for a modernized building style bringing down the traditional forms such as these traditional colonial buildings and with them a way of life.

The UDA in Sri Lanka has decided that the De Soysa building in Slave Island with a 150-year-old history will be brought down and new one with the same façade would be erected a distant over, this is done due to the road expansion. Architect Ismeth Raheem had stated, and this was published in the Sunday times that the building is a “remarkable gem of colonial street architecture and a century old building ". The building is demolished in order to offer the land to the Indian-owned Tata Housing Development Company. The Castle Hotel followed the same fate as the structure replaced this white structure that rose there- and which now houses the Tata project office is nothing like the Castle Hotel.





Fig. 2,3 Demolished state of the De Zoysa Building

The De Zoysa building was once the headquarters of the prestigious firm of H.W. Cave & Co., which from 1877 was the country’s leading publishing and printing house. And it is said that during the 1920s and 1930s, the premises of H.W. Cave & Co. was considered one of the most fashionable addresses in Colombo.

The De Zoysa building holds cultural significance, and a digital twin helps in preserving and promoting this heritage. It provides a means for future generations to understand and appreciate the historical context of the building in terms of Sri Lankan Street architecture.

Digital twins allow for the creation of a highly accurate, three-dimensional, and interactive model of the building, ensuring

that all architectural and historical details are preserved and can be faithfully recreated.

Even if physical reconstruction isn't feasible, a digital twin can serve as a virtual museum, allowing people to explore and experience the building as it once was.

Development of the algorithm for its Digital twin:

The following section details an algorithm and input flows needed to create a digital twin for the De Zoysa building.

- First need to collect historic information on the De Zoysa building. Common data capture methods include laser scanning and photogrammetry, which are excellent ways to recreate buildings digitally. However, have the availability of this type of technology is limited in Sri Lanka. Instead, we can use existing 2D documentation (old paper drawings or scanned plans, sections, and elevations), and previous visited site photographs to capture the details of subsequent building updates manually.
- Next, we create a proper database. For this we linked several external resources and embedded important information about the models within the components' characteristics, including:
    - Links to all available documentation about ramps, stairs, sheds, and structural elements
    - Data about the architects and other AEC professionals that worked on each project
    - Links to images and reports of important moments in the history of the building
- Next, we create a 3D printed model representing the current streetscape of Slave Island, including the remains of the De Soyza Building. This model provides a tangible, physical representation of the existing urban environment, capturing the intricate details of the contemporary streetscape. 3D printing technology ensures accuracy in replicating the physical attributes of the buildings and surrounding infrastructure.
- Then, to juxtapose the present with the past, a 3D model of the streetscape with the De Zoysa Building from the historic time is developed with the previously collected historical data. By using SketchUp's outlier tool, we can compare the original model alongside the updates, making it easy to compare the original and current states. By digitizing the collected information, a detailed and accurate 3D representation of the historical streetscape, including the De Soyza Building, is created. This model serves as a virtual reconstruction of the area as it once stood, providing valuable insights into the architectural and urban development over time.
- To facilitate an interactive and immersive comparison, the historical 3D model is integrated with QR code technology. A QR code is generated and associated with the digital model. When scanned using a smartphone or tablet, the QR code triggers an augmented reality (AR) experience. This AR feature allows users to visualize the historical 3D model superimposed onto the 3D printed structure of the existing streetscape.
- By overlaying the historical 3D model onto the current 3D printed streetscape, users can observe and analyze the changes that have occurred over time, thus a digital twin for the site is created.

In a nutshell, the above-mentioned process involves creating a 3D printed model of the current streetscape of Slave Island, including the De Soyza Building. Also, historical 3D model is then developed using past images and data with SketchUp. This historical model is linked to a QR code, which, when scanned, overlays the historical model onto the 3D printed structure. This allows users to compare and visualize changes between the existing and historical conditions.

This comparative visualization highlights the architectural transformations, urban planning decisions, and cultural shifts that have shaped Slave Island. The digital twin thus becomes a powerful tool for architects, historians, urban planners, and the general public, enabling a deeper understanding of the area's historical evolution and current state.

The selection of case study 02:

UNESCO has inscribed the Rangiri Dambulla Caves as a world heritage site. However, there is history of threats documented in the reports on state of conservation of Rangiri Dambulla Cave Temple archived in the World Heritage Center from 1997 to 2019. Outstanding Universal Value of a world heritage property is affected by threats encompassing multiple secondary threats. Studies by [42] explicitly prove a significant level of decays caused to paintings within the caves due to changes in the cave's microclimatic conditions such as temperature and humidity. Furthermore, [42], states that there is no proper means to implement a visitor management plan into caves. This is of vital importance as when the cave is congested with people, there are high levels of carbon dioxide, these CO_2 levels influence on the decay levels of the cave walls.

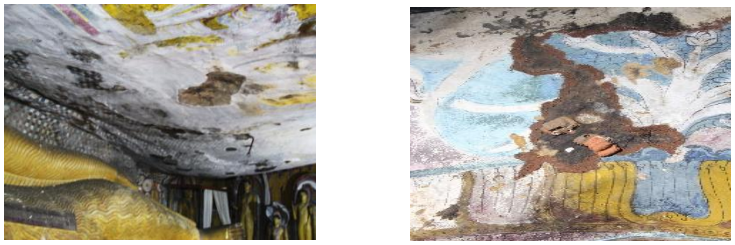


Fig. 4,5 Whitening of the Ceiling and wasp damage on the ceiling. Source- [42]

Moreover, if a property loses the characteristics which warranted its inscription on the World Heritage list, it can be deleted from the list [43]. Thus, the Rangiri Dambulla Caves are on the verge of being removed from the UNESCO world heritage list, if a proper mechanism to combat these problems are created.

Therefore, creating a digital twin for the preservation of heritage properties like the Rangri Dambulla caves, focusing on microclimatic conditions is of paramount importance.

Development of the algorithm for its Digital twin:

The following section details an algorithm and input flows needed to create a digital twin for the Rangiri Dambulla Caves

- First gather information about the caves' history and current conditions (temperature, humidity, light, air quality). Use sensors to monitor these conditions continuously.
- Next Make a detailed 3D model of the caves using laser scanning or photogrammetry. Include all architectural details and materials.
- Then merge environmental data with the 3D model to create a dynamic digital twin. Use software like Building Information Modeling (BIM) to manage this integration.
- Then connect real-time data from sensors to the digital twin for continuous updates.
- Next use predictive modelling to foresee potential problems within the caves.
- Finally, based on the simulations, we can propose solutions for controlling the microclimate within the caves. For example, the real time sensors can quickly identify when the CO_2 concentrations are high, i.e.- when the visitor count is high. These ultrafine particles can affect the deteriorating state of the paintings within the cave.

The algorithm mentioned in the above section proves that a digital twin of the Rangiri Dambulla caves, focusing on microclimatic conditions, offers a powerful tool for preservation. By integrating detailed 3D models with real-time environmental data, you can simulate and predict the impact of various conditions on the caves, leading to more informed and effective preservation strategies.

Challenges in using digital twins for cultural heritage preservation in the built environment. The final part of this study identified the challenges and potential mitigation strategies for implementing digital twins in the preservation of cultural heritage within the built environment through semi-structured interviews with 10 industry specialists. These industry specialists included-

- 05 Architects with experience in heritage architecture
- 02 Archaeologists
- 02 Technical Managers - Engineering Software Development Organization
- 01 Machine Learning Engineer

Additionally, the survey identifies strategies for mitigating these challenges.

01. High-Resolution Data Collection: Capturing detailed and accurate data of heritage sites can be difficult, especially for intricate or hard-to-reach areas. Incomplete or unavailable historical data can hinder the creation of accurate digital twins.

To mitigate this challenge, we can use advanced techniques like laser scanning (LiDAR) and photogrammetry to capture high-resolution 3D models. Also, we can use artificial intelligence to fill gaps in historical data by predicting and reconstructing missing information.

02. High costs and complex software tools: The technology and expertise required for creating and maintaining digital twins can be expensive. The software and tools needed for digital twin creation and management can be complex and require specialized training.

For mitigating these we can seek funding and grants from governmental and non-governmental organizations dedicated to heritage preservation. Develop and implement training programs for staff and stakeholders to build the necessary skills. Utilize open-source software where possible to reduce costs.

03. Data integration and management: Integrating various data types (e.g., 3D models, environmental data) from different sources can be challenging. Managing large volumes of data efficiently while ensuring its integrity and accessibility over time.

To mitigate these aspects, we can adopt industry standards for data formats and protocols to ensure compatibility and use cloud-based storage and processing solutions to manage data more effectively.

04. Ethical and legal concerns: Ensuring the privacy and security of data, especially when it involves culturally sensitive information. There is the need to navigate legal regulations related to data collection, storage, and sharing.

To mitigate the above we can develop transparent data management policies that respect cultural sensitivities and privacy concerns. Ensure compliance with relevant legal and ethical guidelines, including obtaining necessary permissions for data collection.

By addressing these challenges through strategic planning and collaboration, the use of digital twins can significantly enhance the preservation and management of cultural heritage sites.

CONCLUSION

In conclusion, the concept of the "digital twin" has gained significant traction across various industries, including architecture and construction, as a tool for maintaining accurate and up-to-date digital representations of physical structures. The interdisciplinary research over the past decade has spotlighted historic building information modeling for its efficacy in performance modeling, behavior analysis, deterioration modeling, and preventive maintenance of historic heritage sites. However, the implementation of digital twins is complex and requires a tailored approach for each project.

In Sri Lanka, the adoption of digital twins in the construction industry lags behind more technologically advanced countries. This is compounded by the lack of consensus on their construction, operation, and utility, and limited research on their impact on maintaining cultural heritage. The preservation of historical buildings is critical as modernization threatens to erase cultural significance, as seen with the rapid disappearance of colonial street architecture.

Digital twins offer a solution by creating detailed virtual models that capture architectural features and historical context. They are particularly vital for UNESCO World Heritage sites, which face threats from modernization and climate change. The United Nations' 2030 agenda for Sustainable Development Goals emphasizes the importance of cultural and environmental sustainability, underscoring the need for effective conservation strategies.

This study involved a comprehensive literature review, two case studies, identification of appropriate algorithms for creating digital twins, and interviews with construction industry professionals. These methods highlighted the challenges and potential strategies for implementing digital twins.

The De Soysa building on Slave Island and the Rangiri Dambulla Caves are prime examples where digital twins can play a crucial role. For the De Soysa building, a digital twin can aid in preserving its legacy amid urban development. For the Rangiri Dambulla Caves, a digital twin focusing on microclimatic conditions can enhance preservation efforts by predicting and mitigating environmental impacts.

Overall, the development of project-specific digital twins is essential for preserving cultural identity and managing heritage properties, ensuring they remain integral parts of our collective history and identity for future generations.

The final part of this study explores the challenges encountered when using digital twins for preserving cultural heritage. Challenges include capturing detailed data, high costs, data integration issues, and ethical concerns. Solutions involve advanced technologies like laser scanning, seeking funding, using standard data formats and cloud storage, and ensuring ethical data handling. These strategies aim to improve how cultural heritage sites are managed and preserved with digital twins.

REFERENCES

[1] Michael Grieves. (2014) Digital twin: manufacturing excellence through virtual factory replication. White paper,

[2] Michael Grieves. (2016). Origins of the digital twin concept. Florida Institute ofTechnology, 8, 2016.

[3] Ivanov, S., Tarasyev, A., Rak, A., Averkin, A., & Izrailov, R. (2020). Digital Twin of City: Concept Overview. In 2020 Global Smart Industry Conference (GloSIC). IEEE. https://doi.org/10.1109/GloSIC50382.2020.9277755

[4] Jouan, P. and Hallot, P., 2019. Digital twin: A HBIM-based methodology to support preventive conservation of historic assets through heritage significance awareness. The international archives of the photogrammetry, remote sensing and spatial information sciences

[5] Fuller, A., Fan, Z., Day, C. and Barlow, C., 2020. Digital twin: Enabling technologies, challenges and open research. IEEE access, 8, pp.108952-108971

[6] Martínez-Carricondo, P., et al. (2019). UAV Photogrammetry and HBIM for the Virtual Reconstruction of Heritage. ISPRS Archives – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W15

[7] Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. Neural Networks, 113, 54–71. https://doi.org/10.1016/j.neunet.2019.01.012

[8] Farsi, M., Ariansyah, D., Erkoyuncu, J.A. and Harrison, A., 2021. A digital twin architecture for effective product lifecycle cost estimation. Procedia CIRP, 100, pp.506-511

[9] Kor, M., Yitmen, I. and Alizadehsalehi, S., 2023. An investigation for integration of deep learning and digital twins towards Construction 4.0. Smart and Sustainable Built Environment, 12(3), pp.461-487

[10] Liu, Y., Sun, Y., Yang, A. and Gao, J., 2021. [Retracted] Digital Twin-Based Ecogreen Building Design. Complexity, 2021(1), p.1391184

[11] Perera, W.S.D., Ranadewa, K.A.T.O., Parameswaran, A. and Weerasooriya, D., 2023. Status quo of digitalisation in the Sri Lankan construction industry

[12] Opoku, D.G.J., Perera, S., Osei-Kyei, R., Rashidi, M., Bamdad, K. and Famakinwa, T., 2023, January. Barriers to the adoption of digital twin in the construction industry: A literature review. In Informatics (Vol. 10, No. 1, p. 14). MDPI

[13] Ozturk, G.B., 2021. Digital twin research in the AECO-FM industry. Journal of Building Engineering, 40, p.102730

[14] Nguyen, H.X., Trestian, R., To, D. and Tatipamula, M., 2021. Digital twin for 5G and beyond. IEEE Communications Magazine, 59(2), pp.10-15

[15] Yin, X., Liu, H., Chen, Y. and Al-Hussein, M., 2019. Building information modelling for off-site construction: Review and future directions. Automation in construction, 101, pp.72-91

[16] Silva, K.D., Taylor, K. and Jones, D.S. eds., 2023. The Routledge handbook of cultural landscape heritage in the Asia-Pacific. Routledge, Taylor & Francis Group

[17] Wickramasinghe, N., 2014. Sri Lanka in the modern age: a history. Oxford University Press.

[18] Gutschow, N., 2017. Architectural heritage conservation in South and East Asia and in Europe: contemporary practices. Authenticity in architectural heritage conservation: Discourses, opinions, experiences in Europe, South and East Asia, pp.1-71

[19] Cosovic, M., Amelio, A. and Junuz, E., 2019. Classification methods in cultural heritage. In CEUR Workshop Proceedings (Vol. 2320, pp. 13-24). CEUR-WS

[20] Caust, J. and Vecco, M., 2017. Is UNESCO World Heritage recognition a blessing or burden? Evidence from developing Asian countries. Journal of Cultural Heritage, 27, pp.1-9

[21] UNESCO, 2012. Understanding world heritage in the Asia and Pacific, the second cycle of periodic reporting 2010 2012, UNESCO, Paris, France

[22] UNESCO, 2008. Policy document on the impact of climate change on world heritage properties, UNESCO-WHC, Paris. France

[23] G. Angjeliu, D. Coronelli, G. Cardani. (2020) Development of the simulation model for digital twin applications in historical masonry buildings: The integration between numerical and experimental reality, Computers Structures

[24] Gul, M., Ahmed, S., & Jo, J. (2020). Digital twins: A comprehensive survey from smart cities to Industry 4.0. IEEE Access

[25] Gupta, R., Bhattacharya, M., & Kalaga, D. V. (2021). Digital twin-based energy efficiency optimization in smart buildings: A comprehensive review. Renewable and Sustainable Energy Reviews

[26] Madni, A. M., Madni, C. C., & Lucero, S. D. (2019). Leveraging Digital Twin Technology in Model-Based Systems Engineering. Systems, 7(1), 7. https://doi.org/10.3390/systems7010007

[27] W. Hu, T. Zhang, X. Deng, Z. Liu, J. Tan. (2021). Digital twin: a state-of-the-art review of its enabling technologies, applications and challenges, Journal of Intelligent Manufacturing and Special Equipment

[28] Rovina, D., Özeren, Y., Zambon, D., Dotis, P., Pregnolato, M., Manzetti, S., et al. (2022). Assessing flooding impact to riverine bridges: an integrated analysis. Natural Hazards and Earth System Sciences, 22

[29] Toni Greif, Nikolai Stein, and Christoph M. Flath. (2020) Peeking into the void: Digital twins for construction site logistics. Computers in Industry

[30] Kazi Masudul Alam and Abdulmotaleb El Saddik. (2017) C2ps: A digital twin architecture reference model for the cloud-based cyber-physical systems. IEEE Access

[31] Pradeep Alva, Martin Mosteiro-Romero, Clayton Miller, and Rudi Stouffs. (2022). Digital twin-based resilience evaluation of district-scale archetypes

[32] Fabian Dembski, Uwe W¨ossner, and Mike Letzgus. (2019).The digital twin tackling urban challenges with models, spatial analysis and numerical simulations in immersive virtual environments. pages

[33] Marra, A., & Fabbrocino, G. (2021). On the SeVAMH survey protocol for safety and safeguard of artistic assets. Overview and validation at the Monumental complex of Santa Chiara in Naples. Restauro Archeologico, 28, 4–17

[34] P. Jouan, P. Hallot.(2022). Digital twin: Research framework to support preventive conservation policies, ISPRS Int. J. Geoinf.

[35] Bruno, S., Grasso, C., Dell'Olio, F., De Fino, M., & Merenda, M. (2022). VERBUM—Virtual Enhanced Reality for Building Modelling (virtual technical tour in digital twins for building conservation). Journal of Information Technology in Construction (ITcon)

[36] G. Fidan. (2014). Analysis of historic buildings in terms of their microclimatic and thermal comfort performances

[37] A. Massafra, G. Predari, R. Gulli, (2022). Towards digital twin driven cultural heritage management: A hbim-based workflow for energy improvement of modern buildings, ISPRS - Int. Arch. Photogramm. Remote Sens. Spat

[38] Vuoto, S., Colucci, C., Pitonzo, R., & Scianna, A. (2023). Shaping Digital Twin Concept for Built Cultural Heritage Conservation: A Systematic Literature Review. International Journal of Architectural Heritage

[39] Egodawela, S., Buddika, H. A. D. S., Harischandra, W. A. N. I., Mahmoodian, M., &Mathota, M. R. S. N. M. (2022). Conservation and management of ancient stupa using digital twins: A case study of UAV-based 3D photogrammetric digitization of RankothWehera Stupa. In Proceedings of the International Conference on Sustainable Built Environment 2022 (ICSBE 2022). Springer Nature Singapore

[40] F. Gabellone. (2020). A digital twin for distant visit of inaccessible contexts

[41] F. Gabellone. (2022). Digital twin: a new perspective for cultural heritage management and fruition, ACTA IMEKO 11

[42] Rajapaksha, I. Climate change impacts on World Heritage built environment in Tropics: An evidence based appraisal on microclimates and visitor influx of Rangiri Dambulla Cave Temple in Sri Lanka.(2020). Publisher: Sri Lanka National Commission for UNESCO

[43] UNESCO, 2019. Policy document on the impact of climate change on world heritage properties, UNESCO-WHC, Paris. France

FreshSight: An Accessibility-Focused Approach to Produce Freshness and Shelf-Life Detection for Food Safety and Waste Reduction

W.P. R. Fernando[1*], A. Kirupananda[2]
[1]School of Business, Informatics Institute of Technology, Sri Lanka
[2] School of Business, Informatics Institute of Technology, Sri Lanka
Corresponding Author E-mail: poorni.f@iit.ac.lk

ABSTRACT
Colour-blind individuals encounter daily challenges, particularly in distinguishing colour-based indicators of food spoilage. This limitation significantly impacts their ability to assess the freshness and safety of fruits and vegetables accurately. Concurrently, global concerns regarding food spoilage have intensified, with millions worldwide affected by foodborne illnesses each year. The modern lifestyle, characterized by its rapid pace and time constraints, exacerbates this issue, often leading to unnoticed spoilage and substantial waste. The resulting annual waste, estimated at one-third of all edible food, imposes significant societal and environmental burdens, underscoring the urgency for effective solutions. FreshSight enables users, including those with colour vision deficiencies, to assess the condition of fruits and vegetables through a Convolutional Neural Network (CNN)-based real-time image analysis engine and an intuitive interface. This system provides immediate visual feedback to help users make informed decisions and avoid the consumption of spoiled produce. It also offers inclusive design features that support individuals with visual impairments. Beyond individual benefits, FreshSight promotes responsible food handling and contributes to the broader goal of sustainable food systems. By combining advanced technology with user-centered design, the solution enhances both safety and accessibility in everyday food-related decisions. In addressing the critical challenges of food safety, inclusivity, and waste reduction, FreshSight aims to support healthier lifestyles and contribute positively to environmental and societal well-being in the modern world.

KEYWORDS— *Colour vision deficiency, Food spoilage, Image analysis, Food safety, Responsible consumption, Accessibility, Sustainability*

INTRODUCTION
Colour-blind individuals face unique challenges in daily life, with their difficulties often overlooked by those with normal colour vision. One notable challenge is the inability to recognize spoiled food based on colour changes [1]. In the modern world, food spoilage is a major issue due to its harmful consequences for consumers [2]. Foodborne illnesses have become a significant global concern, affecting millions of people worldwide. More than 200 diseases result from consuming spoiled food contaminated with harmful bacteria like Salmonella and Shigella. Each year, approximately 3-5 billion people suffer from infectious diarrhoea, and nearly 1.8 million people die due to food contamination [1]. The fast-paced, work-centric lifestyle of modern society often leaves individuals with inadequate time to monitor their daily food consumption. As a result, individuals often become unaware of both the quantity and quality of food stored inside their refrigerators [3]. In many households, food can remain in refrigerators, undergoing colour changes and developing foul odours due to bacterial growth. Unfortunately, people often discover the spoilage long after it occurs, leading to significant food waste[1]. This food wastage poses a substantial societal burden, with roughly one-third of all edible human food being wasted annually [4]. The underlying causes and challenges associated with food spoilage, and colour-blindness in recognizing spoiled food. have been highlighted below,

Colour-blindness - Colour-blind individuals face challenges in recognizing freshness cues for produce, such as colour changes, because of their condition, which affects their ability to perceive certain colours. This difficulty in distinguishing between fresh and spoiled produce can lead to colour-blind individuals unintentionally consuming items that have gone bad, posing potential health risks [1].

Foodborne illnesses - Colour-blind individuals and others may unknowingly consume produce that has deteriorated or become contaminated due to the challenge of recognizing freshness based on visual cues. This can increase the risk of foodborne illnesses. The lack of recognition and subsequent consumption of spoiled produce can lead to a higher frequency of foodborne illnesses among colour-blind individuals and anyone else who, by mistake, consumes spoiled food. These illnesses may have monetary costs as individuals require medical treatment, medication, and may need to take time off work, impacting their finances and overall well-being. Additionally, it can place a burden on healthcare systems [5].

Modern, Fast-Paced Lifestyle - Today's fast-paced lifestyles leave individuals with minimal time to assess the freshness of produce. Demanding work schedules and busy routines often leave little time for proper monitoring. To cope with their fast-paced lives, many individuals rely on pre-packaged produce, which may not always have visible freshness indicators. This reliance can complicate the assessment of freshness of the produce [6].
Unawareness of produce quantity and quality - In households, produce items may linger in refrigerators for extended periods without being consumed. This extended storage affects the freshness and quality of the produce. The lack of awareness about the quantity and quality of stored produce often results in the frequent disposal of significant amounts of food. This leads to financial wastage and contributes to environmental concerns related to food waste [7].

BACKGROUND AND RELATED WORK

A. Freshness Assessment of Fruits and Vegetables

The term "fresh" is commonly used, especially concerning food, and holds significant importance for consumers, particularly in the acceptance of fruits and vegetables. It creates specific consumer expectations as a crucial factor impacting the quality of the fruits and vegetables [8]. A remarkable 59% of consumers give priority to the freshness of products when making their purchasing decisions [9].
Fruits and vegetables play a vital role in daily human meals, and as the incidence of diseases rises, people have become more attentive to natural and organic products, especially fruits and vegetables. The emphasis on the freshness of these products is driven by the growing awareness of their impact on human health [10]. The quality of fruits and vegetables has a direct influence on human health. With the increase in diseases in contemporary times, people are increasingly concerned about the freshness of the fruits and vegetables they consume [10]. Additionally, the nutritional value of perishable food items, like fruits and vegetables, is closely tied to their freshness levels [11]. The issue of waste in Fresh Fruits and Vegetables (FFVs) is a significant challenge in the food supply chain industry, contributing to a substantial portion of annual food losses. Approximately one-third of all edible human food is wasted annually, posing a burden on society [4]. This underscores the critical importance of assessing and managing the freshness of fruits and vegetables throughout the supply chain.

B. Colour Blindness and Freshness Perception

Colourblind people or people with a colour vision deficiency face many challenges in their daily activities [12]. Globally, approximately 300 million people about 1 in 12 men and 1 in 200 women experience some form of colour blindness, representing roughly 4.5% of the world's population.[13] Colour blindness, a condition affecting the ability to see certain colours like red, green, blue, or their combinations, presents challenges for individuals. The most prevalent type is red- green colour blindness, where these colours appear the same, while another common type is blue- yellow colour blindness, making it difficult to distinguish between these hues. Unfortunately, there is currently no known treatment for colour blindness [13]. This condition poses unique challenges for individuals, particularly in recognizing colour changes indicative of freshness. For those with colour blindness, identifying spoiled fruits and vegetables based on colour becomes a challenge, highlighting an aspect often overlooked by those with normal colour vision [1]. The difficulties faced by colour-blind individuals in perceiving freshness add a layer of complexity to their overall experience [14].

C. Global and Local Consumption Patterns of Fruits and Vegetables

The significance of fresh vegetables and fruits in human nutrition lies in their rich nutrient content, encompassing essential vitamins (B, C, K) and minerals (calcium, potassium, magnesium), along with dietary fibre. This nutritional profile contributes to a balanced diet, aiding in the prevention of chronic diseases like heart disease, cancer, diabetes, and obesity, particularly in developing countries. Over recent years, global production of fresh vegetables and fruits has surged by 30%, escalating from 30 million tons to 60 million metric tons [14]. This increase reflects a growing recognition of the importance of fruits and vegetables in promoting health and preventing chronic diseases. In the context of Sri Lankan agriculture, fruits and vegetables hold pivotal importance for both farmers and consumers. A significant portion of the Sri Lankan population incorporates vegetables into their daily diet, often accompanying the staple food, rice. According to the Household Income and Expenditure Survey 2016, approximately 8% of the total household expenditure on food and beverages in Sri Lanka is allocated to vegetables. This underscores the frequent purchase of fresh vegetables as a dietary staple in the country [15]. Additionally, the HIES (2009/10) reveals that an average Sri Lankan household spent Rs. 13,267 on food, with Rs. 386 allocated to fruits, constituting a budget share of 2.91%. Further, the allocation of the food budget to fruits

in urban, rural, and estate sector households stood at 3.4%, 2.9%, and 1.5%, respectively. This implies that urban areas accord greater significance to fruits in their budgetary considerations [2]. Despite this expenditure, the average consumption of fruits and vegetables in Sri Lanka is around 150 grams per day, with per capita consumption approximately 112 grams per day. This figure is 50% lower than the World Health Organization's recommended daily intake of 400 grams. Factors contributing to this low consumption include limited availability, affordability, and awareness, particularly in rural areas[27]. In 2023, Sri Lanka observed an uptick in fruit consumption, reaching 12.8 million tons, against a national requirement of 19.6 million tons. This increase was largely attributed to higher outputs in bananas, mangoes, pineapples, and papayas. Concurrently, the country's fruit exports also saw a rise, with a total of 38,201 metric tons shipped abroad[28].

D. Limitations of Existing Methods and the Novelty of FreshSight

The Consumption of fruits and vegetables plays a pivotal role in daily human diets. With the rise in diseases, people are more cautious about natural products, especially fruits and veggies. Among these, the freshness of fruits and no-cooked vegetables is a big deal because it directly affects human health[10]. The quality of fruits is closely linked to how fresh they are. Consumers judge fresh produce's quality based on how it looks (blemishes, colour, size, and texture), ripeness, smell, and taste. Importantly, the quality changes a lot during the growth of the fruit, from before it's picked to after, and even during storage [17]. The usual ways of checking fruit and vegetable quality use our senses (sensory evaluation) or involve chemical analysis. However, these methods have issues like being easily affected by other things, not being very accurate, being slow, and needing a lot of resources. There are smarter ways of checking the quality of fresh produce including using things like Raman spectroscopy, near-infrared spectroscopy, Red, Green, and Blue (RGB) vision detection, electronic nose, acoustic characteristics, dielectric properties, and other new methods. Each of these has its strengths. Near-infrared spectroscopy is good at looking inside fruits and veggies. RGB vision detection is quick and automated, mainly checking how fruits and veggies look. Electronic nose technology acts like an animal's sense of smell, quickly and precisely notifying about the smell of fruits and veggies. Acoustic characteristics technology listens to the sounds fruits and veggies make to figure out if they're ripe and fresh. Dielectric property technology is fast and non-destructive, notifying about water content and hardness. Raman spectroscopy uses light to give detailed information about the chemicals in fruits and veggies. It's special because water doesn't interfere with it. Raman spectroscopy is excellent at finding tiny things other methods might miss. Looking at how molecules vibrate quickly notifies the quality and nutrition of fruits and veggies, even their structure. It can also find harmful stuff and pesticide residues on the surface [18]. Non-destructive methods, while valuable in quality assessment, may not universally address all chemical and physical quality parameters. The considerable expense associated with the specialized equipment and the necessity for employing distinct instruments to analyse various parameters contributes to the major disadvantages of these methods. The high cost of acquiring and maintaining such equipment, coupled with the need for expertise in handling different instruments, poses financial and operational challenges [19].

E. Implications of Freshness on Produce Quality

Freshness plays an important role in the overall quality of produce, impacting economic considerations, health and safety, environmental sustainability, consumer satisfaction, and food wastage. The interconnectedness of these factors underscores the need for effective management throughout the supply chain.

Economic Considerations

The economic implications of managing fruit and vegetable waste are substantial, with findings indicating a noteworthy impact on financial resources. According to [20], fruit waste accounted for 12%, and vegetable waste accounted for 24%, resulting in a substantial total loss of 54,891,539 USD. Notably, the summer season saw the highest fruit waste at 15%, contributing to heightened economic losses during this period. Additionally, the economic impact of vegetable waste was significant, with percentages of 28% in summer and 24% in spring, underscoring the seasonal variability in economic losses associated with different types of produce. Retailers grapple with the intricate challenges of managing fresh inventory, leading to significant economic repercussions. The difficulty of maintaining freshness, spanning from the distribution centre to the store shelf, contributes significantly to this wastage. Retailers find themselves wasting twice the value of their profits due to these challenges. The economic considerations related to freshness extend beyond the store level; they permeate the entire supply chain. The complexities of handling and maintaining consistent temperatures in the supply chain pose significant challenges, resulting in over 10% of fresh food shipments being rejected upon arrival at their destination. This rejection further exacerbates economic losses and underscores the pervasive nature of the freshness challenge throughout the entire supply chain [3].

Health & Safety

The paramount importance of food safety in human life cannot be overstated, as highlighted by[4]. Foodborne illness constitutes a significant global health challenge, affecting one in ten people annually according to the World Health Organization. In the United States alone, the Centres for Disease Control and Prevention (CDC) reports an alarming statistic of approximately 48 million people falling victim to foodborne illnesses each year, with 128,000 individuals requiring hospitalization. Contaminated foods, primarily caused by the consumption of fresh produce, contribute significantly to these incidents[5]. emphasize the need for pre-monitoring the freshness state of food to mitigate the magnitude of foodborne illnesses. However, a notable challenge exists in the absence of a simple system for accurately assessing the freshness of food before consumption. Consuming contaminated fruits poses severe health risks due to the presence of various disease-causing microbes and pathogens. These contaminants can originate from water, soil, manure, wildlife, birds, or other sources, especially when consumers handle or prepare food. The consequences of consuming non-fresh fruits include a range of diseases such as vomiting, diarrhoea, abdominal cramps, nausea, fever, joint/backaches, and fatigue. Certain demographic groups face higher risks of foodborne illnesses. Individuals suffering from diseases like AIDS, cancer, kidney issues, or other chronic diseases are particularly vulnerable. Senior citizens are also at an elevated risk of these illnesses. Additionally, individuals with autoimmune disorders such as lupus, those on immune-suppressing medications, newborns, unborn babies, children under 5 years, pregnant women, and individuals with a history of drug or alcohol addiction are considered high-risk groups [6].

Consumer Satisfaction

Consumer satisfaction is intricately linked to the freshness of produce, with a profound impact on individuals' choices when selecting fruits and vegetables.[7] emphasize that the nutritional value of perishable items, such as fruits and vegetables, is directly correlated with their freshness. Fresh produce tends to be richer in essential nutrients, contributing significantly to overall consumer health. When deciding where to purchase food, consumers heavily prioritize freshness. Notably, 59% of consumers place a remarkable emphasis on the quality of fresh offerings during their purchasing decisions. This emphasis on freshness, however, often leads to inconsistent experiences for shoppers. The discrepancy arises when products labelled as “fresh” fail to meet consumers' expectations, creating a challenge in ensuring consistent quality across the spectrum of fresh produce [9]. Addressing this inconsistency is important for enhancing overall consumer satisfaction in the realm of fresh food.

Food Wastage

The issue of food wastage poses a significant global concern. A staggering one-third of the food produced worldwide is discarded, despite the alarming fact that there are a billion people facing malnutrition. The potential impact is noteworthy — the saved food could provide sustenance for nearly two billion more individuals [20]. According to the Food and Agriculture Organization (FAO) of the United Nations, food waste encompasses anything altering the accessibility, edibility, health, safety, or quality of a product, rendering it unsuitable for consumer consumption[23]. Particularly troubling is the wastage of fruits and vegetables, accounting for over 42% of the discarded food. These items, being the most consumed globally, face significant losses throughout the food supply chain, with an estimated 40-50% of them being wasted[20]. In simpler terms, more than one-third of the fruits and vegetables produced globally end up as waste, equivalent to a staggering 28 million tons[20]. The problem of food wastage extends beyond the broader supply chain and touches our households. Often, unnoticed food in our refrigerators goes bad, changes colour, and emits unpleasant Odors due to bacterial formation[1]. Addressing this issue and embracing sustainable practices in assessing food freshness is not only crucial for reducing food wastage but also plays a pivotal role in resource conservation and minimizing the ecological footprint associated with food production and consumption.

Environmental Impact

Examining food waste from an environmental standpoint reveals hidden costs and substantial impacts on the ecosystem. As reported by[20], the generation and disposal of food waste result in not only economic losses but also pollution problems due to unpleasant Odors, the release of leachate into the environment, and pollution in landfills. The environmental toll extends throughout the entire food supply chain, where significant flows of resources, including water, fossil fuels, and chemicals, are utilized for the production, transportation, processing, storage, and preservation of food[24]. At each stage of the food supply chain, there are substantial emissions and environmental impacts, both locally and globally, as noted by[25]. Agricultural practices significantly contribute to environmental pollution, often stemming from a lack of awareness regarding the freshness levels of produce. Farmers, when informed of the optimal freshness levels, can adopt a more environmentally friendly approach by reducing the use of excessive fertilizers, pesticides, and water. Furthermore, knowledge of the ideal freshness

levels enables precise harvesting, mitigating overharvesting or underharvesting risks and promoting resource conservation, thereby minimizing unnecessary environmental impact. The accumulation of wet organic materials in landfills has further adverse consequences, including leachate leakage, the attraction of insects, unpleasant Odors[26], and the production of harmful greenhouse gases[20]. The overall environmental impact underscores the need for sustainable practices throughout the food supply chain to mitigate these adverse effects on the ecosystem.

F. Existing Solutions and Limitations

Several software applications, including FarmSoft [30], Clarifruit [31], and GoMicro [32], have been developed to assess the freshness of fruits and vegetables. FarmSoft[30] provides comprehensive farm management and quality tracking from planting to harvest using historical and environmental data, while Clarifruit[31] leverages computer vision to analyze visual attributes such as color, size, and defects, offering real-time quality assessments. GoMicro[32] enhances smartphone cameras with attachments for detailed image capture, which are analyzed with machine learning algorithms to detect spoilage and defects. Despite these advancements, current solutions face significant limitations. Accuracy and reliability are affected by external factors such as lighting conditions, camera quality, and background interference, which can lead to inconsistent results. Accessibility for color-blind users remains limited, as most applications do not explicitly support individuals with protanopia, deuteranopia, or tritanopia, nor do they implement mechanisms such as color correction algorithms, simulated views, or alternative indicators like color-to-text or shape-based cues. Furthermore, many applications depend on manual user input, introducing potential errors and reducing reliability, while the need for frequent database updates to reflect varying produce characteristics adds additional complexity and resource requirements. These limitations highlight the gaps that existing solutions have not addressed. FreshSight differentiates itself by providing color-blind-friendly visualizations, including simulated views for different types of color vision deficiency, and by offering color-to-text and shape-based indicators to enhance accessibility. It also automates the entire image capture and analysis process, reducing reliance on user input, and employs robust machine learning models validated under diverse real-world conditions, including varying lighting, camera quality, and produce types. By addressing these challenges, FreshSight offers a more inclusive, reliable, and practical solution for assessing produce freshness, filling the gaps left by previous applications.

METHODOLOGY

A. Use of machine learning for freshness and shelf-life assessment

Dataset Description

The dataset used in this study was curated to provide comprehensive coverage of freshness levels for fruits and vegetables, categorized into three main criteria: fresh, moderate, and spoiled. This dataset was obtained from a combination of sources, including a dataset provided by Dr. Fahad Tahir, which detailed freshness categorization, and an additional shelf-life dataset sourced from Kaggle [29]. By combining and refining these datasets, a unified dataset was created, comprising 25,115 images. This dataset was split into 70% for training and 30% for testing, ensuring robust model development and evaluation.

Model Development

Data Preprocessing: Images were preprocessed to adjust for lighting and background variations, ensuring consistency in the data used for model training. This involved several steps. Image Resizing: All images were resized to a standard dimension of 128x128 pixels to maintain uniformity across the dataset. Data Preprocessing: Images were preprocessed to adjust for lighting and background variations, ensuring consistency in the data used for model training. This involved several steps. Normalization: Pixel values were normalized to a range of [0, 1] to standardize the input data and facilitate faster convergence during training. Augmentation: Techniques such as rotation, flipping, and color adjustment were applied to augment the dataset, enhancing the model's ability to generalize from the training data.

Model Architecture

The Convolutional Neural Network (CNN) architecture used for this study included multiple layers designed to extract and learn hierarchical features from the images: Convolutional Layers: These layers apply filters to the input images to detect local features such as edges, textures, and patterns. Pooling Layers: Pooling layers reduce the dimensionality of the feature maps, retaining the most important information while reducing computational load. Fully Connected Layers: These layers integrate the extracted features and perform the final classification, determining the freshness and shelf life of the produce. Activation Functions: ReLU activation functions were

used in the convolutional layers to introduce non-linearity, and a softmax activation function was applied in the output layer to produce probabilistic predictions. Convolutional Neural Networks (CNNs) were employed for image analysis and classification. The models were trained using the annotated dataset, with the freshness scores paired with shelf lives, serving as the ground truth. The model uses layers to detect patterns and features in these images, enabling it to classify new images into the appropriate freshness category along with the corresponding shelf life.

Model Training

The model was trained using the Adam optimizer, known for its efficiency in handling large datasets and its adaptive learning rate capabilities. The loss function used was categorical cross-entropy, suitable for multi-class classification problems. The training process involved iterating over the dataset in batches, updating the model parameters to minimize the loss function. Validation and Hyperparameter Tuning: A portion of the training data was set aside for validation to monitor the model's performance and prevent overfitting. Hyperparameters such as learning rate, batch size, and the number of epochs were fine-tuned based on validation performance. Specifically, grid search was used for hyperparameter tuning. Grid search involves systematically testing all possible combinations of hyperparameters to determine the best configuration. Additionally, k-fold cross-validation was employed to ensure the model's performance was consistent across different subsets of the data, further refining the model's accuracy and robustness.

Model Testing

The effectiveness of the CNN model in predicting the freshness and shelf life of fruits and vegetables was evaluated using the confusion matrix. This matrix provides insights into how accurately the model predicts values compared to the actual ones. In Figure 1, the confusion matrix displays high values along the diagonal, indicating strong predictive accuracy. This analysis suggests that the prediction model exhibits considerable accuracy in its predictions. Out of the dataset, 30 images were used for testing, of which 27 were correctly identified, resulting in a system accuracy rate of 90%. This accuracy rate was determined using the following formula.

Accuracy Level = (Number of correctly recognized images / Number of total tested images) × 100%

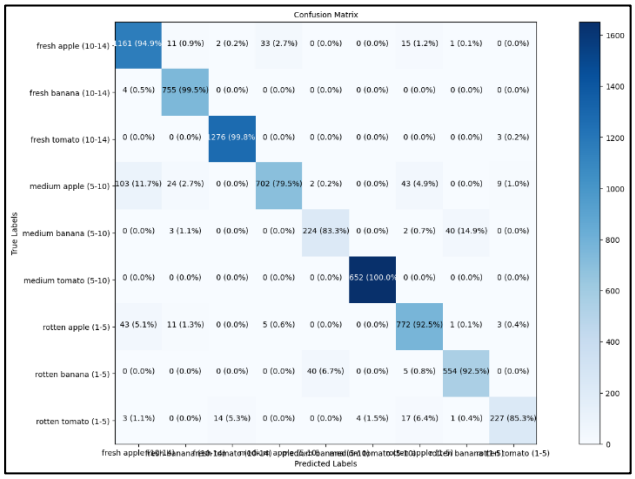


Figure 10: Confusion Matrix for freshness and shelf-life prediction

In addition to accuracy, further evaluation was performed using standard classification metrics: precision, recall, and F1-score. These metrics provide a more detailed understanding of the model's performance for each freshness category. Precision measures the proportion of correctly predicted positive instances out of all predicted positives, recall measures the proportion of correctly predicted positives out of all actual positives, and the F1-score is the harmonic mean of precision and recall. The results, derived from the confusion matrix (Figure 1), are presented in Table 1.

Table 1: Precision, Recall, and F1-Score for Each Freshness Category

Class	Precision	Recall	F1-Score
Fresh Apple (10–14)	0.8211	0.9493	0.8805

Fresh Banana (10–14)	0.9391	0.9947	0.9661
Fresh Tomato (10–14)	0.9770	0.9977	0.9872
Medium Apple (5–10)	0.9551	0.7091	0.8139
Medium Banana (5–10)	0.8726	0.8589	0.8657
Medium Tomato (5–10)	1.0000	1.0000	1.0000
Rotten Apple (1–5)	0.9150	0.9239	0.9195
Rotten Banana (1–5)	0.9224	0.9250	0.9237
Rotten Tomato (1–5)	0.9639	0.8481	0.9024

The results show that the model performs strongly across all categories, with F1-scores exceeding 0.80 in every case. Notably, Medium Tomato achieved perfect precision and recall, while Medium Apple exhibited slightly lower recall, suggesting potential benefits from additional training data for that category.
Figure 2 presents the four-tier architecture of the proposed solution, showing the components and their connection between the functional architecture and the database design.

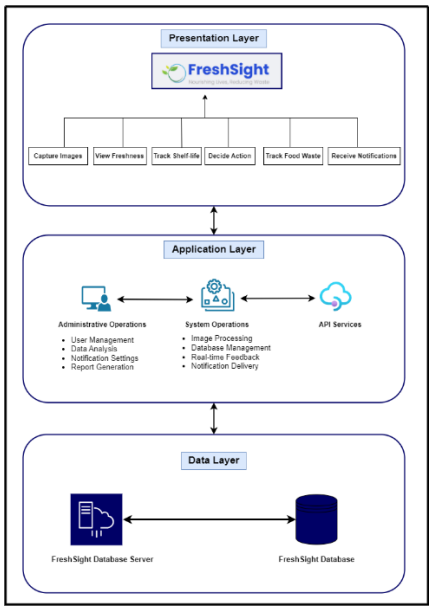


Figure 11: High-level architecture of the proposed solution

The proposed solution has been developed as a mobile responsive web application. The high-level architecture of FreshSight consists of several interconnected components, each playing a crucial role in delivering a seamless and effective user experience:
User Interface (UI): The UI is the front-end component that allows users to interact with FreshSight. It is designed to be intuitive and accessible, providing clear instructions and real-time feedback on produce freshness. The interface includes options for capturing images, viewing freshness scores, and accessing additional information or settings.
Camera Module: The camera module enables users to capture images of produce. It utilizes the smartphone's built-in camera to take photos under different lighting conditions. The module is optimized to ensure high-quality image capture, crucial for accurate freshness assessment.
Image Processing Engine: This component is responsible for analyzing the captured images. It employs advanced

algorithms to detect color variations and other visual indicators of freshness. The engine processes the image data to identify spoilage signs such as discoloration or bruising.
Color Analysis Algorithm: The color analysis algorithm plays a critical role in evaluating produce freshness. It uses color correction techniques to adjust for variations in lighting and then analyzes the corrected image to determine freshness levels. The algorithm is designed to cater specifically to the needs of color-blind users by translating color information into actionable freshness scores.
Database: The database stores information about various types of produce and their freshness characteristics. It includes reference images and freshness criteria for different fruits and vegetables. The database is regularly updated to ensure accurate and relevant assessments.
Feedback System: The feedback system delivers results to the user through the UI. It provides a freshness score, and actionable recommendations based on the image analysis. Users receive real-time feedback on the quality of their produce, helping them make informed decisions.
Cloud Integration: FreshSight incorporates cloud integration for data storage and processing. Cloud services handle large-scale data operations, such as updating the produce database and refining image analysis algorithms based on user feedback. Cloud integration ensures that the app remains up to date with the latest freshness criteria and technological advancements.
Data Privacy and Security Module: This component ensures the security of user data, including captured images and personal preferences. It employs encryption and secure communication protocols to protect user information from unauthorized access.
Analytics and Reporting: FreshSight includes analytics capabilities to track usage patterns and user feedback. This data helps in refining the app and enhancing its performance. Reports generated from analytics provide insights into user behavior and app effectiveness.
The FreshSight solution thus integrates these components to deliver a comprehensive tool for assessing produce freshness. It leverages advanced image processing, color analysis, and cloud technologies to provide accurate, real-time feedback, empowering color-blind individuals to make better food choices.

RESULTS AND FINDINGS

Following functional testing, the prototype underwent evaluation by both experts and non-experts. The aim was to assess the overall suitability and completeness of the solution in addressing the defined problem and meeting user expectations. The evaluation process surfaced several key insights. The concept was praised for aligning well with current consumer demands, particularly in promoting transparency and quality in food products. Suggestions were made to enhance the system by integrating real-time data and exploring technologies like blockchain to improve consumer trust. The inclusive design was highlighted as a valuable feature, especially for users with color vision deficiencies, reinforcing its importance in supporting informed choices and health-conscious decisions. Feedback from participants involved in distribution and agriculture emphasized the system’s potential to revolutionize distribution processes and promote sustainability through more accurate freshness assessments.
On the usability front, the prototype’s interface was regarded as intuitive and accessible to a broad range of users. Nevertheless, recommendations were given to introduce customizable visual themes and additional cues to further enhance accessibility for all users. The system’s potential for widespread adoption was recognized, particularly due to its user-centric design and relevance to the retail and food industries. Additional emphasis was placed on the importance of making the solution adaptable for various user needs and retail environments. Overall, the evaluation reflected positively on the system’s ability to meet key objectives such as enhancing transparency, promoting inclusivity, and delivering reliable freshness assessments. Consistent suggestions for improvement included stronger real-time integration, customizable interface elements, and enriched accessibility support. These insights provide a clear direction for further refinement and optimization, helping to ensure the solution meets real-world needs effectively and remain aligned with user expectations in diverse contexts.

CONCLUSION

A. Limitations of the current system

While FreshSight effectively aids color-blind users in assessing produce freshness, several limitations persist. The application's reliance on visual indicators may not account for all spoilage factors, such as microbial growth or internal decay. Variations in lighting conditions and camera quality can also affect the accuracy of color detection. Moreover, FreshSight's current focus on color may overlook other sensory cues like texture or smell. The app's

effectiveness in diverse environmental conditions and its scalability to accommodate a wider range of produce types are areas that require further research and development to enhance reliability and user trust.

B. Future work

Further improvements recommended by industry experts and feedback from end users will be considered in future work. Introducing customization features will allow users to tailor freshness assessments and shelf-life parameters to their specific needs. Expanding the deep learning model to detect freshness and shelf life across a broader range of produce will provide more comprehensive insights. Future iterations will also aim to incorporate additional food quality indicators, such as texture and aroma, to offer a more holistic assessment of produce quality. Enhancing user engagement through gamification and sustainable food management challenges is another area of focus. Ensuring cross-platform compatibility and the inclusion of a language selection feature will further enhance accessibility. Additionally, gathering user feedback will remain crucial for continuous improvement, and the implementation of accessibility features will promote inclusiveness, particularly for users with visual impairments, ensuring the solution is adaptable for a wide and diverse audience.

C. Concluding remarks

The primary objective of this research is to validate FreshSight as an effective tool for color-blind individuals in assessing produce freshness, thereby enhancing their ability to make informed food safety decisions. The implications of this study are significant. Firstly, FreshSight offers a practical solution to the challenge of identifying produce quality for color-blind consumers, which can improve food safety practices and reduce food waste. The insights gained from this research underscore the potential of smartphone-based applications to assist underserved groups by leveraging accessible technology. As smartphone usage continues to rise, FreshSight can provide an essential service, enabling users to make safer and more efficient food choices. The benefits of this application are substantial, promising improved food safety outcomes and contributing to a reduction in food waste, thus supporting overall community health and sustainability.

REFERENCES

[1] R. K. Megalingam, G. Sree, M. Reddy, I. Krishna, and L. U. Suriya, Food Spoilage Detection Using Convolutional Neural Networks and K Means Clustering. 2019, p. 493. doi: 10.1109/RDCAPE47089.2019.8979114.

[2] N. Hebbar, "Freshness of Food Detection using IoT and Machine Learning," in 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Feb. 2020, pp. 1–3. doi: 10.1109/ic-ETITE47903.2020.80.

[3] G. Anand and L. Prakash, "IoT Based Novel Smart Refrigerator to Curb Food Wastage," in 2018 3rd International Conference on Contemporary Computing and Informatics (IC3I), Oct. 2018, pp. 268–272. doi: 10.1109/IC3I44769.2018.9007271.

[4] A. Abougharib, M. Awad, and M. Ndiaye, "Remaining Shelf-Life Estimation of Fresh Fruits and Vegetables During Transportation," IEEE Access, vol. 11, pp. 8845–8859, 2023, doi: 10.1109/ACCESS.2023.3239584.

[5] D. G. Nyachuba, "Foodborne illness: is it on the rise?," Nutrition Reviews, vol. 68, no. 5, pp. 257–269, May 2010, doi: 10.1111/j.1753-4887.2010.00286.x.

[6] D. T. U. Monika A, "A STUDY ON GROWTH AND PROGRESS OF INSTANT FOOD PRODUCTS INDUSTRIES IN INDIA," Journal of Research Administration, vol. 5, no. 2, Art. no. 2, Dec. 2023.

[7] M. Hebrok and C. Boks, "Household food waste: Drivers and potential intervention points for design – An extensive review," Journal of Cleaner Production, vol. 151, pp. 380–392, May 2017, doi: 10.1016/j.jclepro.2017.03.069.

[8] S. Péneau, P. B. Brockhoff, F. Escher, and J. Nuessli, "A comprehensive approach to evaluate the freshness of strawberries and carrots," Postharvest Biology and Technology, vol. 45, no. 1, pp. 20–29, Jul. 2007, doi: 10.1016/j.postharvbio.2007.02.001.

[9] "The Freshness Challenge: Navigating the Complexities of Fresh Food Management." Accessed: Nov. 20, 2023. [Online]. Available: https://www.linkedin.com/pulse/freshness-challenge-navigating-complexities-fresh-mxbzf

[10] P. K. S. C. Jayasinghe and S. Sammani, "Detection of Freshness of the Fruits using Machine Learning Techniques," Jun. 2022, Accessed: Dec. 03, 2023. [Online]. Available: http://ir.lib.seu.ac.lk/handle/123456789/6413

[11] L. Gilani, S. F. Tahir, U. Rasheed, H. Saqib, M. Hassan, and H. Alquhayz, "Fruits and Vegetables Freshness Categorization Using Deep Learning," Computers, Materials & Continua, vol. 71, pp. 5083–5098, Jan. 2022, doi: 10.32604/cmc.2022.023357.

[12] G. Lausegger, M. Spitzer, and M. Ebner, "OmniColor – A Smart Glasses App to Support Colorblind People," International Journal of Interactive Mobile Technologies (iJIM), vol. 11, no. 5, Art. no. 5, Jul. 2017, doi: 10.3991/ijim.v11i5.6922.

[13] P. K. Nigam and M. Bhattacharya, "Colour vision deficiency correction in image processing," in 2013 IEEE International Conference on Bioinformatics and Biomedicine, Dec. 2013, pp. 79–79. doi: 10.1109/BIBM.2013.6732581.

[14] G. I. Balali, D. D. Yar, V. G. Afua Dela, and P. Adjei-Kusi, "Microbial Contamination, an Increasing Threat to the Consumption of Fresh Fruits and Vegetables in Today's World," International Journal of Microbiology, vol. 2020, p. e3029295, May 2020, doi: 10.1155/2020/3029295.

[15] U. Herath, "Consumer Behavior and Attitudes in Purchasing Vegetables," May 2019, doi: 10.19080/ARTOAJ.2019.20.556123.
[16] J. Weerahewa, C. Rajapakse, and G. Pushpakumara, "An analysis of consumer demand for fruits in Sri Lanka. 1981–2010," Appetite, vol. 60, pp. 252–258, Jan. 2013, doi: 10.1016/j.appet.2012.09.017.
[17] M. Baietto and A. D. Wilson, "Electronic-Nose Applications for Fruit Identification, Ripeness and Quality Grading," Sensors, vol. 15, no. 1, Art. no. 1, Jan. 2015, doi: 10.3390/s150100899.
[18] S. Xu, X. Huang, and H. Lu, "Advancements and Applications of Raman Spectroscopy in Rapid Quality and Safety Detection of Fruits and Vegetables," Horticulturae, vol. 9, no. 7, Art. no. 7, Jul. 2023, doi: 10.3390/horticulturae9070843.
[19] O. Chauhan, S. Lakshmi, A. Pandey, N. Ravi, N. Gopalan, and R. Sharma, "Non-destructive Quality Monitoring of Fresh Fruits and Vegetables," Defence Life Science Journal, vol. 2, p. 103, May 2017, doi: 10.14429/dlsj.2.11379.
[20] B. Parsafar, M. Ahmadi, G. R. Jahed Khaniki, N. Shariatifar, and A. Rahimi Foroushani, "The impact of fruit and vegetable waste on economic loss estimation," Global Journal of Environmental Science and Management, vol. 9, no. 4, pp. 871–884, Oct. 2023, doi: 10.22034/gjesm.2023.04.14.
[21] M. Ahmadi, G. Jahed Khaniki, N. Shariatifar, and E. Molaee-Aghaee, "Investigation of aflatoxins level in some packaged and bulk legumes collected from Tehran market of Iran," International Journal of Environmental Analytical Chemistry, vol. 102, no. 16, pp. 4804–4813, Dec. 2022, doi: 10.1080/03067319.2020.1789614.
[22] S. Lee and J. Y. Chung, "The Machine Learning-Based Dropout Early Warning System for Improving the Performance of Dropout Prediction," Applied Sciences, vol. 9, no. 15, p. 3093, Jul. 2019, doi: 10.3390/app9153093.
[23] F. Girotto, L. Alibardi, and R. Cossu, "Food waste generation and industrial uses: A review," Waste Management, vol. 45, pp. 32–41, Nov. 2015, doi: 10.1016/j.wasman.2015.06.008.
[24] L. Skaf, P. P. Franzese, R. Capone, and E. Buonocore, "Unfolding hidden environmental impacts of food waste: An assessment for fifteen countries of the world," Journal of Cleaner Production, vol. 310, p. 127523, Aug. 2021, doi: 10.1016/j.jclepro.2021.127523.
[25] Q. D. Read et al., "Assessing the environmental impacts of halving food loss and waste along the food supply chain," Science of The Total Environment, vol. 712, p. 136255, Apr. 2020, doi: 10.1016/j.scitotenv.2019.136255.
[26] R. A. Neff, R. Kanter, and S. Vandevijvere, "Reducing Food Loss And Waste While Improving The Public's Health," Health Affairs, vol. 34, no. 11, pp. 1821–1829, Nov. 2015, doi: 10.1377/hlthaff.2015.0647.
[27] [27] AVRDC – The World Vegetable Center, "Sustainable Vegetable Cultivation in Sri Lanka: Research-based agronomic practices promise safer, more efficient crop production," 2023. [Online]. Available: https://avrdc.org/sustainable-vegetable-cultivation-in-sri-lanka-research-based-agronomic-practices-promise-safer-more-efficient-crop-production/?utm_source=chatgpt.com. [Accessed: 14-July-2025].
[28] FreshPlaza, "Fruit consumption and production trends in Sri Lanka," 2023. [Online]. Available: https://www.freshplaza.com/asia/article/9638584/fruit-consumption-and-production-trends-in-sri-lanka/?utm_source=chatgpt.com. [Accessed: 14-July-2025].
[29] [29] Kavumpadi, S. (2021) Fruit and vegetable dataset for shelf life, Kaggle. Available at: https://www.kaggle.com/datasets/soorajkavumpadi/fruit-and-vegetable-dataset-for-shelf-life [Accessed: 24-July-2025].
[30] [30] FarmSoft, "About Us," FarmSoft. [Online]. Available: https://www.farmsoft.com/farmsoft.html. [Accessed: 14-June-2025].
[31] Clarifruit, "Clarifresh: AI-Powered Quality Control for Fresh Produce," Clarifruit. [Online]. Available: https://clarifresh.com/. [Accessed: 14-June-2025].
[32] GoMicro, "Home," GoMicro. [Online]. Available: https://www.gomicro.co/. [Accessed: 14-May-2025].

Towards Safer Elderly Care: A Convolutional Neural Network Solution for Fall Detection

Kalupahana R.W[a], Dr. Maduranga M.W.P[b]

[a]Department of Information Technology, Faculty of Department, General Sir John Kotelawala Defence University, Ratmalana 10390, Colombo, Sri Lanka.
[b]Department of Electrical and Electronic Engineering, University of Sri Jayewardenepura, Nugegoda,10250, Sri Lanka.
pasanm@sjp.ac.lk , 38-bis-0001@kdu.ac.lk

ABSTRACT
As modern life becomes increasingly busy, computer vision-based monitoring systems have become essential, particularly in elderly care. This paper presents the development of a robust fall detection system using deep learning techniques, specifically a convolutional neural network (CNN) that processes RGB images to accurately distinguish between fall and non-fall events. The model is trained and validated on a dataset categorized into two classes: fall and non-fall. By utilizing convolutional and pooling layers, CNN effectively learns hierarchical representations of the input data, capturing both low-level and high-level features crucial for accurate fall detection. The key stages of this approach include data acquisition, pre-processing, and model training. The model's performance is evaluated using precision, recall, and F1-score metrics, demonstrating high accuracy, which is further enhanced through data augmentation, pre-processing, and crossvalidation techniques. A confusion matrix analysis confirms the model's effectiveness in correctly classifying instances across both classes. The system also extends its capabilities to video analysis by extracting frames at 30-second intervals, ensuring continuous and comprehensive monitoring. This research highlights the potential of deep learning to enhance safety and care for the elderly, offering a reliable solution for real-time fall detection. The findings underscore the importance of integrating advanced technologies into healthcare, paving the way for future innovations in monitoring and assistance systems.

KEYWORDS: *Fall detection, Elderly care, Computer Vision, Deep Learning, Convolutional Neural Network-(cnn), Data Preprocessing.*

INTRODUCTION
In recent years, the rapid advancement of technology has paid attention to innovative solutions in healthcare, particularly in the innovation of fall detection systems. Fall is a significant concern, especially among the elderly, as they often lead to severe injuries and a decline in quality of life. Traditional methods of fall detection, such as wearable sensors and manual monitoring, have their limitations.[1] This is where the power of deep learning, specifically Convolutional Neural Networks (CNNs), comes into play.

CNNs, a class of deep learning models, have revolutionized the field of computer vision by enabling machines to recognize and classify images with remarkable accuracy. By leveraging the capabilities of CNNs, researchers are now able to develop sophisticated fall detection systems that can analyze video feeds or sensor data in real-time, identifying falls with unprecedented precision.[2]

This paper explores the application of CNN models in fall detection systems, highlighting the advantages of deep learning techniques over conventional methods. We have learned the architecture of CNNs, their training processes, and the datasets used to enhance their performance. Furthermore, we discuss the way of detecting results with uploading video also.

The remainder of this paper is organized as follows. Section 2 provides a review of related work on computer vision-based fall detection. Section 3 dataset preparation and preprocessing of CNNbased fall detection model. Section 4 describes the evaluation model training; section 5 presents the experimental results. Finally, Section 6 discusses the key findings, implications, and future research directions.

RELATED WORKS
In recent years, there has been a growing interest in the development of advanced fall detection systems

that can accurately identify and respond to falls, while minimizing false alarms [3].

One of the prominent approaches to fall detection is the use of wearable sensors, such as accelerometers and gyroscopes, integrated into devices like smartwatches or insoles.[1]These wearable systems leverage the ability to continuously monitor an individual's movement and posture, and can detect sudden changes or impacts that are characteristic of a fall [3]. For example, the study by [4]presents a wearable fall detection system that utilizes with many different sensors and powerful hardware combines with , achieving high accuracy in detecting falls with a low false alarm rate.

Complementing the wearable sensor approach, vision-based fall detection systems have also gained attention in the research community ([5]. These systems employ depth cameras or RGB-D sensors to capture visual data, which is then analyzed using machine learning algorithms to identify and classify different types of falls [6].

In addition to wearable sensors and vision-based systems, researchers have also explored the integration of multiple sensing modalities to enhance the reliability and accuracy of fall detection [7]. For instance, the study by [7]combines inertial measurement unit (IMU) data with video analysis to improve the detection and classification of falls, addressing the limitations of each individual approach.

This highlights the potential of Internet of Things (IoT) technologies in the context of elderly fall detection [8]By integrating wearable sensors, smart home devices, and cloud-based analytics, these IoTenabled systems can provide a comprehensive and connected solution for fall monitoring, prevention, and emergency response.

Beyond the technological advancements, the literature also emphasizes the importance of user acceptance and usability in the design of fall detection systems [3]. Factors such as comfort, aesthetics, and user-friendly interfaces play a crucial role in ensuring the adoption and long-term use of these systems by the elderly population.[9]

The literature review highlights the diverse approaches, from wearable sensors and vision-based systems to hybrid solutions and IoT-enabled technologies, that aim to improve the safety and well-being of the older population.

In this work, We present the development and evaluation of a novel fall detection model based on deep learning computer vision techniques. Our approach leverages a CNN architecture to analyze image data and reliably identify falls within an indoor elderly care environment. By utilizing the rich visual information captured by cameras, our system aims to provide a non-invasive and privacy-preserving solution for fall detection, and this can be used for video data also.

DATASET AND PRE-PROCESSING

This research aims to create a system to detect falls which addresses the challenges faced by caregivers in monitoring the whereabouts of elderly individuals when no direct supervision is available.

Created a dataset for model training

In the chasing of developing an effective fall detection system tailored for elderly care, our research utilized a carefully chosen dataset comprising a total of 1,424 images. The test set images should be 20% and training set images should be 80% from the total images. This dataset is split into two main classes: 'fall' and 'not fall.' The 'fall' class includes images depicting scenarios where a fall has occurred, while the 'not fall' class consists of images representing normal activities without any fall incidents like standing, running, sitting. This binary classification forms the foundation for training our machine learning model to accurately distinguish between fall and non-fall events.

The images were sourced from a combination of publicly available datasets and custom-collected data to ensure diversity in scenarios, backgrounds, and lighting conditions. This diversity is crucial for enhancing the model's robustness and generalizability across different environments and situations typically encountered in elderly care settings. A few samples of images I have tested have shown in Fig.1 , Fig.2, Fig.3 Fig.4 and Table 1 shows the number of images available in the dataset.

In the chasing of developing an effective fall detection system tailored for elderly care, our research utilized a carefully chosen dataset comprising a total of 1,424 images. The test set images should be 20% and training set images should be 80% from the total images. This dataset is split into two main classes: 'fall' and 'not fall.' The 'fall' class includes images depicting scenarios where a fall has occurred, while the 'not fall' class consists of images representing normal activities without any fall incidents like standing, running, sitting. This binary classification forms the foundation for training our machine learning model to accurately distinguish between fall and non-fall events.

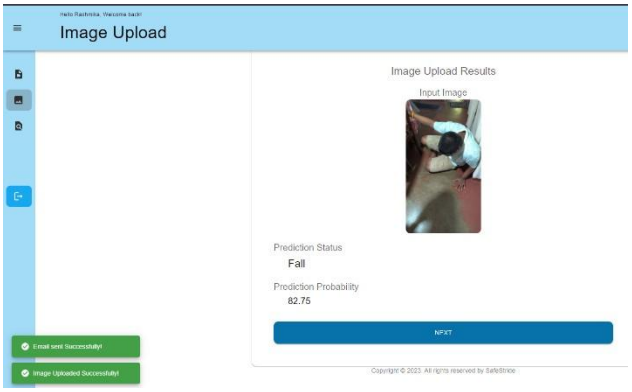


Figure 1. Fall

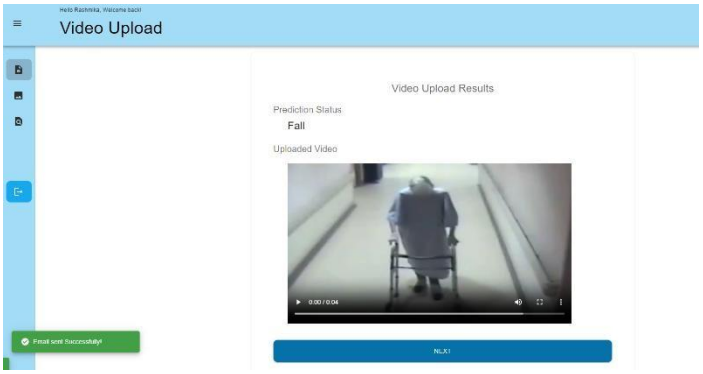


Figure 2. Video (Fall)

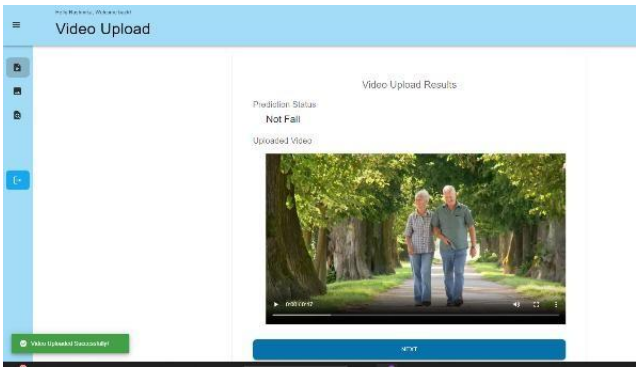


Figure 3. Not Fall

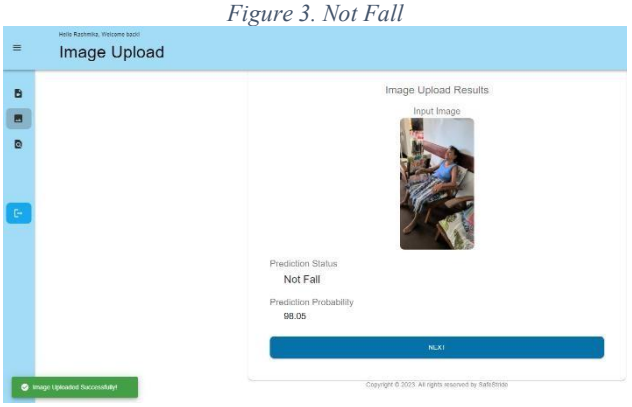


Figure 4 Video upload(Not Fall)

Table 1. Distribution of Dataset

Class	**No. of Images**

	Training set	*Test set*
Fall	653	512
Not Fall	135	124

Data Preprocessing

Data preprocessing involves procedures and strategies applied to unprocessed data before training a machine learning model. Its goal is to ensure that the data is in an appropriate format and quality for efficient model training and analysis. This requires transforming and cleaning the data.[10] The preprocessing steps applied to the training and validation datasets include rescaling,
data augmentation, resizing, selecting batch sizes, shuffling, and one-hot encoding of labels.

Initially, all images were resized to a uniform dimension to ensure consistency and compatibility with the input requirements of Convolutional Neural Network (CNN) model. Following resizing, image normalization was performed, scaling pixel values to a range between 0 and 1. This step is essential for accelerating the convergence of the model during training by maintaining uniformity in the input data.Data augmentation techniques were employed to artificially expand the dataset and introduce variability, thereby reducing the risk of overfitting. Techniques such as rotation, flipping, and zooming were applied to create multiple variations of each image, simulating different angles and perspectives. These preprocessing techniques help prepare the data for training the CNN model and improve its ability to learn and generalize from the provided images.

A total of 1424 pictures were used for the CNN model, with 1165 being used for training and 260 for verifying accuracy. The 260 images used for validation testing were chosen as 20% of the total images.

4 MODEL DEVELOPMENT AND TRAINING

4.1 Creating the CNN Model

Computer vision has dramatically enhanced thanks to convolutional neural networks, which utilize deep learning. The image illustrates the essential operation of a CNN model.

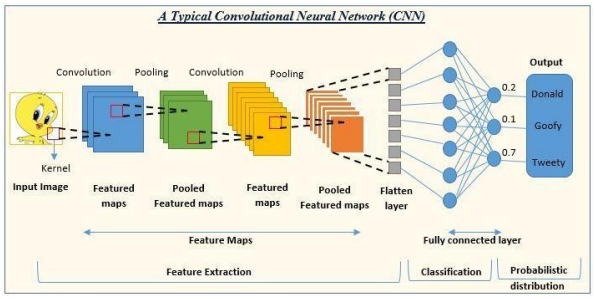


Figure 5.Basic process inside a CNN model

Input Layer:

First up, we have the input layer. It takes in RGB images sized at (255, 255, 3). This setup is standard in computer vision, where the red, green, and blue channels stack up to create a 3D tensor. This format is perfect for feeding into a convolutional neural network (CNN) [2][11]

Convolutional Layers:

The heart of the model lies in its convolutional (Conv2D) layers. These are the building blocks of any CNN. Our first Conv2D layer comes with 64 filters, each with a kernel size of 5, and uses the ReLU activation function [11][1]. ReLU, or Rectified Linear Unit, adds a dash of non-linearity, which is crucial for the model to pick up on complex patterns in the data.

When the convolutional operation kicks in, it applies a learnable filter (or kernel) to the input image. This process churns out a feature map that highlights specific visual patterns. With 64 filters and a kernel size of 5, the model is well-equipped to capture a variety of low-level features like edges, textures, and shapes, all of which are super matches fall detection [1][11][6][2].

Pooling Layers:

After the first convolutional layer, we introduce a MaxPool2D layer with a pool size of 5 [1][11][6][2] . This step is all about condensing the spatial dimensions of the feature maps. It helps trim down the model's complexity, extract more abstract features, and boost its generalization skills.

Next, the second convolutional layer steps in with 32 filters and a kernel size of 3, again using the ReLU activation

function [1][11][6][2]. This layer is designed to dig deeper, capturing more complex, higher-level features from the input data, building on the groundwork laid by the first convolutional layer.

The mix of convolutional and pooling layers is a classic setup in CNN models. It allows the network to learn a hierarchical representation of the input data, moving from low-level to high-level features [1][11][6][2]. This structure is key to the model's ability to accurately spot and classify fall events from video dataModel Trainning

A crucial stage in creating a deep learning model for fall detection using the CNN architecture is model training. In this study, the training technique is carried out using a GPU (more particularly, GPU device 0) to take advantage of its processing capability and speed up the training process. The model is trained for a specified number of epochs, in this case, 40.

To minimize the stated loss function (categorical cross-entropy), the model iteratively adjusts its parameters during the training phase. This allows the model to learn from the training dataset. The Adam optimizer is used to optimize and modify the model's weights following the estimated gradients. This enables the model to gradually enhance its functionality and enhance its precision in detecting falls. The percentage of correctly categorized samples in the training dataset is known as training accuracy. The training set of data shows how accurate the model was. The model may have memorized the training data without adequately generalizing new samples, a sign of overfitting. High training accuracy, therefore, does not ensure good performance on unknown data.

The inaccuracy or difference between the model's anticipated output and the actual target values on the training dataset is represented by training loss. It shows how well the model fits the training set of data. The objective is to reduce the training loss, which measures how well the model can learn from the training data and produce reliable predictions.

A validation dataset that is separate from the training dataset is used to evaluate the model's performance at the end of each epoch's end. Researchers can learn more about the model's generalization ability and spot any overfitting or underfitting by testing it on previously unexplored data.

On the other hand, validation accuracy gauges the percentage of correctly identified samples in the validation dataset. It gives a general indication of how well the model performs on unobserved data.

A higher validation accuracy indicates a model working well while accurately identifying the samples in the validation dataset. The validation loss calculates the error or difference between the target values on the validation dataset and the model's anticipated output. It shows how effectively the Figure 3: Basic process inside a CNN model generalizes to new data. A minor validation loss indicates improved prediction accuracy on the validation dataset. After training, the model has to optimize and maintain the model.

In extending fall detection system from static images to dynamic video analysis, we have implemented frame extraction technique that allows the model to evaluate video content effectively. When a video is uploaded into the system, it is processed by capturing individual frames at 30-second intervals.

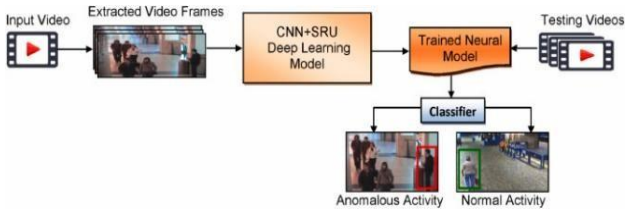


Figure 6. Extracting frames from a video

Each extracted frame is treated as an independent image and is analyzed using this trained model, which has been specifically trained to distinguish between 'fall' and 'not fall' scenarios. By leveraging the model's ability to accurately classify these frames, It can determine whether a fall event has occurred within the video.

The decision-making process involves aggregating the results from all analyzed frames. If any frame is classified as a 'fall,' the system flags the video as containing a fall incident. This method ensures that even if a fall is brief or occurs at a single point in the video, it is still detected by the system.

This integration of image-based detection into video analysis not only enhances the versatility of the model but also broadens its applicability in real-world settings, such as monitoring elderly individuals in care facilities or at home. By effectively bridging the gap between static and dynamic content, our system provides a robust solution for continuous fall monitoring.

RESULTS AND DISSCUSSION

This section presents the outcomes and implications of the research on fall detection using deep learning techniques. Here, outline the performance metrics achieved by CNN model on both the training and validation datasets, along with a thorough analysis of the results.
To fully understand the model training process and its impact on the performance of fall detection CNN model, We conducted a detailed analysis of the model's training history graph. This analysis helps to see how the model's accuracy and loss evolved over time, providing insights into its learning behavior and stability.
By examining these metrics, can assess how well the model generalizes to new, unseen data and identify any areas for improvement. This comprehensive evaluation is crucial for refining approach and ensuring the model's effectiveness in real-world applications.

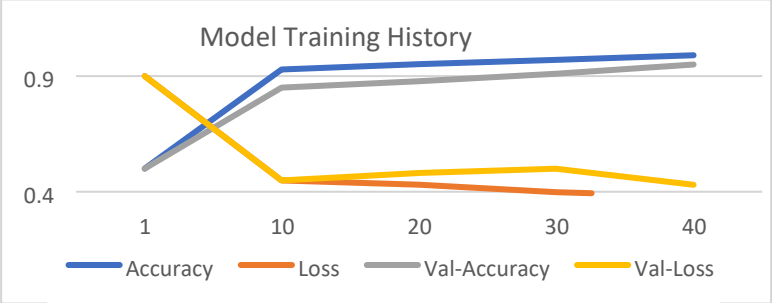


Figure 7.Model Training History

Figure 7 indicates the predicted behavior of the models produced, allowing them to evaluate whether the models should be rejected or accepted. The blue and red lines show the accuracy and loss of the training set, respectively. The green and yellow lines show the validation accuracy and validation loss, respectively.
Increasing the number of epochs can improve accuracy by allowing the model to make more weight and bias updates, leading to better pattern recognition and learning from the training data. This model was trained using 40 epochs, and both accuracies increased as the number of epochs increased, while both losses decreased.

Table 2. Model Summary

Layer (type)	Output shape	Param #
conv2d_3(Conv2D)	(None, 253, 253, 32)	896
max_pooling2d_3 (MaxPooling2d)	(None, 126, 126, 32)	0
conv2d_4(Conv2D)	(None, 124, 124, 64)	18,496
max_pooling2d_4 (MaxPooling2d)	(None, 62, 62, 64)	0
conv2d_5(Conv2D)	(None, 60, 60, 128)	73,856
max_pooling2d_5 (MaxPooling2d)	(None, 30, 30, 128)	0
flatten_1(Flatten)	(None, 115200)	0
dense_3(Dense)	(None, 128)	14,745,728
dense_4(Dense)	(None, 64)	8,256
dense_5(Dense)	(None, 2)	130
Total params: 14,847,362		

Trainable Params: 14,847,362
Non-trainable Params: 0

The model summary offers a clear snapshot of the architecture and structure of the trained CNN model. It details the various layers, including their types and output shapes, providing insight into how the model processes

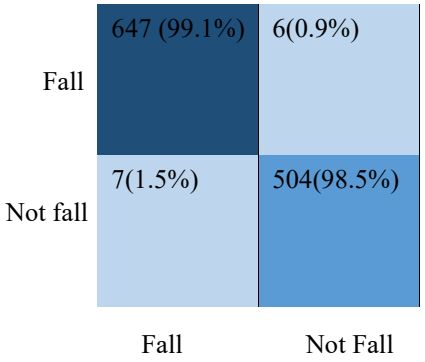
information. Additionally, it lists the number of parameters associated with each layer, highlighting the model's capacity and complexity.
In total, the model comprises 14,847,362 parameters, all of which are trainable. There are no nontrainable parameters, meaning every parameter is adjusted during training to help the model learn and make accurate predictions based on the input data.
With its numerous convolutional and pooling layers, the model is adept at recognizing complex patterns and extracting crucial information from input images since it has many convolutional, pooling and dense layers. This layered structure allows the model to effectively identify features necessary for distinguishing between different scenarios, such as falls and non-falls, enhancing its predictive accuracy and reliability.

Table 3. Classification Report

	precision	recall	f1-score	support
0	1.00	0.99	1.00	653
1	0.99	1.00	0.99	511
accuracy			0.99	1164
macro avg	0.99	1.0	0.99	1164
weighted avg	0.99	0.99	0.99	1164

A classification report is an invaluable tool for evaluating the performance of a CNN model in a classification task. It provides a detailed analysis of various metrics for each class in the dataset, including precision, recall, F1-score, and support.
Precision measures the accuracy of positive predictions for a specific class. It indicates how well the model can classify instances belonging to that class without generating false positives. A high precision score reflects a low rate of false positives. Recall, also known as sensitivity or the true positive rate, assesses the model's ability to identify positive examples of a class. It shows the model's effectiveness in detecting relevant events without producing false negatives. A high recall score signifies a low percentage of false negatives.

Table 4 .Confusion Matrix



According to the matrix:

Fall: Out of 653 instances, 647 (99.1%) were correctly identified as falls, showing strong proficiency in detecting falls.
Not Fall: All 511 instances were correctly identified as not falls, achieving a perfect score.
A confusion matrix is a useful tool that compares the predicted labels for each class in a dataset with the actual labels, providing a detailed review of model performance.
True Positive (TP): The model correctly predicted the positive instances of a class.
Negative (TN): The model correctly predicted the negative instances of a class.
False Positive (FP): The model incorrectly predicted positive instances, also known as Type I error or false alarms.

False Negative (FN): The model incorrectly predicted negative instances, also known as Type II error or missed detections.

CONCLUSION AND FURTHER WORKS

This study successfully developed a fall detection system using deep learning methods, specifically a Convolutional Neural Network (CNN), to monitor elderly individuals in indoor environments. The system addresses the challenge of remotely monitoring elderly individuals when direct supervision is not available. By achieving high accuracy and reliability, the proposed solution enhances safety and significantly reduces response times in case of emergencies. The findings emphasize the importance of integrating advanced technologies into healthcare, setting the stage for future innovations in monitoring systems.

The implementation of this system in elderly care settings can greatly improve safety by providing real-time fall detection. This technology not only reduces emergency response times but also offers peace of mind to caregivers and family members. The model's high accuracy ensures its reliability, making it a valuable tool for monitoring and assistance. When a fall is detected, an alert is automatically sent to the caregiver, ensuring timely intervention.

For future work, the development of a web application for caregivers is planned, which will provide real-time alerts in case of a detected fall and offer updates on fall detection history. Additionally, future research could explore indoor localization for elderly individuals using an Indoor Positioning System (IPS). The goal is to develop this system into a comprehensive product for use in elderly care homes, hospitals, and residential complexes.

REFERENCES

[1] C. P. Shewell and B. S. Hons, "Leveraging Machine-Vision for Activity Recognition Utilising Indoor Localisation To Support," 2023.

[2] S. Sherin, P. M. T. Student, and A. J. Assistant, "Human Fall Detection using Convolutional Neural Network," *Int. J. Eng. Res. Technol.*, vol. 8, no. 6, pp. 1368–1372, 2019, [Online]. Available: www.ijert.org

[3] R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," *Biomed. Eng. Online*, vol. 12, no. 1, 2013, doi: 10.1186/1475-925X-12-66.

[4] L. Árvai, "Application of Smartwatches in Elderly Care with Indoor Localization Functionality," *Int. J. Interact. Mob. Technol.*, vol. 15, no. 5, pp. 174–186, 2021, doi: 10.3991/ijim.v15i05.20435.

[5] D. Asmar, A. Fakih, and J. Zelek, "COMPUTER VISION FOR ASSISTIVE INDOOR LOCALIZATION".

[6] H. Kwon *et al.*, "Indoor Localization and Multi-person Tracking Using Privacy Preserving Distributed Camera Network with Edge Computing," pp. 1–13, 2023, [Online]. Available: http://arxiv.org/abs/2305.05062

[7] T. M. Le, V. T. Ly, and S. V. T. Dao, "A Feature Selection Approach for Fall Detection Using Various Machine Learning Classifiers," *IEEE Access*, vol. 9, no. August, pp. 115895–115908, 2021, doi: 10.1109/ACCESS.2021.3105581.

[8] M. Oudah, A. Al-Naji, and J. Chahl, "Computer Vision for Elderly Care Based on Deep Learning CNN and SVM," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1105, no. 1, p. 012070, 2021, doi: 10.1088/1757-899x/1105/1/012070.

[9] J. Schwickert, L., Becker, C., Lindemann, U., Maréchal, C., Bourke, A., Chiari, L., ... & Klenk, "Fall detection with body-worn sensors: a systematic review.," 2013.

[10] M. Inayathulla, P. H. Kiran, M. C. Sri, and M. Deepika, "A Heuristic Model for Predicting Human Fall Detection using Machine Learning Techniques," *Int. J. Recent Technol. Eng.*, vol. 9, no. 2, pp. 119–121, 2020, doi: 10.35940/ijrte.b3180.079220.

[11] E. Alam, A. Sufian, P. Dutta, and M. Leo, "Real-Time Human Fall Detection Using a Lightweight Pose Estimation Technique," *Commun. Comput. Inf. Sci.*, vol. 1956 CCIS, pp. 30– 40, 2024, doi: 10.1007/978-3-031-48879-5_3.

Development of an AI-Based Model with Low Computational Complexity for Accurate Wind Energy Forecasting

Shiwantha Dilshan,Nimesh Fernando
Wayamba University of Sri Lanka
shiwanthawyb@icloud.com, nimeshaf@wyb.ac.lk

ABSTRACT

Most countries primarily relay on fossil fuel for electricity generation, leading to fossil fuel depletion and environmental pollution. The countries are developed technologies for renewable energy generation. The wind energy being promoted as a superior renewable energy. However, wind energy has its challengers, particularly uncertainty that can affect overall system stability. The accurate short-term forecasting of wind energy was crucial for ensuring grid stability. Both physical and AI-based models can effectively be utilized for wind energy prediction. AI-based methodologies have shown superior effectiveness, efficiency, and accuracy when compared to traditional physical models. The lightweight AI-based forecasting model was particularly significant for processing devices, enabling faster computations and substantially more cost-effective forecasting. The research utilized simulation software to develop an Artificial Neural Network (ANN) model, initially incorporating eight meteorological parameters. Four of these parameters showed weak correlations and were subsequently removed from the model. Further optimization was achieved through pruning and quantization techniques, significantly reducing computational complexity. The optimized model demonstrates a notable reduction in both training time by 92.69% and inference time by 63.83%, while maintaining accuracy with only a marginal decrease of 3.99% compared to the initial model. These improvements were achieved with minimal loss in predictive accuracy, significantly reducing computational complexity. The study concludes that the optimized ANN model is well-suited for real-time wind power forecasting, offering a balance between accuracy and computational efficiency. This approach not only facilitates better grid management but also extends the applicability of AI-based forecasting to devices with limited processing capabilities. Future work could explore additional complexity reduction techniques and broader deployment scenarios.

KEYWORDS: *Wind forecasting, artificial intelligence, computational efficiency, model optimization, simulation, time-series prediction*

INTRODUCTION

Most countries primarily rely on fossil fuels for electricity generation, leading to fossil fuel depletion and environmental pollution (Li et al., 2019). Wind energy is being promoted as a superior renewable resource compared to tidal energy. In 2020, wind turbines contributed 8.4% of the total utility-scale electricity generation in the United States, with projections indicating an increase to 20% by 2030 and 35% by 2050 (Alkesaiberi et al., 2022). However, wind power has its challenges, particularly uncertainties that can affect power quality, grid security, and overall system stability. Accurate and timely forecasting of wind power is crucial for the safe operation of the grid (Ju et al., 2019) and for balancing electricity supply and demand, thus ensuring grid stability.

The meteorological factors can change rapidly and unpredictability, leading to fluctuations in wind energy generation. This uncertainty was made it challenging to model because historical data may not fully capture future patten. Predicting wind energy involves dealing with uncertainty in meteorological conditions. Even with advanced forecasting techniques, there was always a margin of error due to unpredictability of weather. Reliable data on meteorological factors were crucial for accurate modeling. However, data may be sparse or of varying quality in some regions, affecting the reliability of models.

Both physical models and AI-based models can be effectively utilized for wind power prediction, each offering distinct advantages and challenges. Engaging qualified personnel in the calculations necessary for physical models is crucial, as their expertise directly impacts the reliability of the results. Furthermore, substantial effort is required for data collection, processing, and analysis to generate meaningful outputs. This reliance on traditional methods can often lead to delays in obtaining results, resulting in increased time consumption, higher costs, and potential reductions in accuracy. In contrast, AI-based methodologies have consistently demonstrated superior effectiveness, efficiency, and accuracy when compared to traditional physical models. The advancements in AI not only streamline the forecasting process but also enhance

predictive capabilities. Moreover, various simulation software packages have emerged, enabling the development of AI models through straightforward procedures that are both cost-effective and user friendly. These software tools also incorporate features that provide accuracy metrics, allowing users to evaluate and refine model performance systematically.

A large amount of data was used to build an AI-based accurate wind power prediction model. Wind energy generation was influenced by various parameters, including wind speed, wind direction, air density, etc. A large number of datasets for each parameter were required to be input, which was the cause of increased computational complexity in the AI-based forecasting model, leading to longer processing times. Although modern computers were designed to be high-performance, this was not considered a significant problem. However, the types of processing devices available in the grids could not be determined. Therefore, it was deemed essential to reduce computational complexity to ensure compatibility with a broader range of processing devices, including those that were expected to have limited capabilities. The reduction in the AI model's computational complexity enables it to run on more cost-effective and lower-power processing devices, such as older computers used in wind farms.

The primary objective of this research is to develop and optimize an AI-based short-term forecasting model for wind power generation using simulation software. The study aims to reduce computational complexity while evaluating the performance improvements of the optimized model relative to the baseline version. This goal is achieved through the following three key objectives:

Development of an AI-based forecasting model for wind power generation.

Optimization of the model to minimize computational overhead.

Evaluation of the optimized model's performance by comparing key metrics with those of the initial model.

The paper was structured as follows: the literature review section examined existing studies on wind power forecasting, key parameters influencing wind energy, and techniques for reducing computational complexity. The methodology section detailed the processes of dataset preparation, model development, and optimization. The results and discussion section presented the performance of both the initial and optimized models. The conclusion summarized the study's findings and outlined directions for future research. Through this work, a contribution was made toward developing more efficient and scalable wind energy forecasting solutions, thereby supporting the broader adoption of renewable energy.

LITERATURE REVIEW

Wind energy forecasting has been extensively studied using various parameters and methodologies, as highlighted in recent research. A comparative analysis of these studies reveals the diversity of approaches and the critical parameters influencing wind power prediction. (Alkesaiberi et al., 2022) employed a comprehensive set of parameters, including weather conditions, turbine specifications (e.g., rotor features, gearbox ratios), ambient temperature, blade pitch angle, nacelle position, hub height, and wind characteristics (direction, speed, and intensity), alongside temporal and geographic factors. (Kosovic et al., 2020) expanded the scope by incorporating additional parameters like ice accumulation and horizontal grid spacing, emphasizing the importance of localized atmospheric conditions. (Tarek et al., 2023) focused on wind speed, wind direction, air density, active power, theoretical power, and time-series data, while (Lipu et al., 2021) and (Ju et al., 2019) utilized time-series wind power values to develop forecasting models. (Hu et al., 2021) integrated meteorological data, including wind speed, wind direction, air pressure, temperature, and relative humidity. (Pang et al., 2021) further extended the parameter set by including humidity, air pressure, air density, rainfall, and surface sensible heat, employing machine learning to analyze correlations. (Franke et al., 2021) highlighted the role of air density, rotor area, Betz coefficient, and hub height in wind power calculations. Similarly, (Badran & Abdulhadi, n.d.) examined turbinespecific parameters such as swept area, power coefficient, and air velocity, while (Van Der Merwe et al., n.d.) explored the interplay between wind speed, direction, air density, pressure, and temperature.

The forecasting of wind energy has garnered significant attention, leading to a variety of studies that leverage artificial intelligence (AI) models. These studies explore diverse parameters and employ different AI techniques to enhance the accuracy of wind energy predictions. One notable study (Alkesaiberi et al., 2022) compared multiple machines learning approaches, focusing on kernel-driven models such as Support Vector Regression (SVR) and Gaussian Process Regression (GPR), alongside ensemble learning models. The findings indicated that both GPR and ensemble models achieved satisfactory performance in predicting wind power. The parameters utilized in this research included weather conditions, turbine specifications (e.g., rotor features and gearbox ratios), ambient temperature, blade pitch angle, nacelle position, hub height, turbine operational status, and wind characteristics (direction, speed, and intensity), as well as temporal factors and geographic location.

Table 13. Summary of research and parameters

Research	Parameters	AI method
(Alkesaiberi, Harrou and Sun, 2022)	Weather, type of turbine, rotor features, ambient temperature, blade pitch angle, gearbox gear ratio, nacelle position, hub height, turbine status, location, time, wind direction, wind speed	Gaussian Process Regression
(Kosovic et al., 2020)	wind direction, wind speed, wind intensity hub-height, ice accreted, temperature weather condition (shout down), horizontal grid spacing, location, time	AI analog ensemble
(Tarek et al., 2023)	wind direction, wind speed, Air Density, Active Power, Theoretical Power, Date/Time	Deep Learning Models
(Lipu et al., 2021)	wind power values (Time series data)	Machine learning
(Ju et al., 2019)	wind power values (Time series data)	Machine learning
(Hu et al., 2021)	wind speed, wind direction, air pressure, temperature, relative humidity weather data at certain time intervals	Extreme Learning Machine
(Pang, Yu and Liu, 2021)	wind speed, wind direction, temperature humidity, air pressure, air density, rainfall Surface sensible heat	Deep belief network-based hybrid forecasting method
(Franke et al., 2021)	air density, rotor area, Betz coefficient rated wind speed per time interval, hub height	
(Badran and Abdulhadi, no date)	wind speed, air density, temperature, air, pressure, area swept, hub height	
(Van Der Merwe et al., no date)	wind speed, wind direction, air density, air pressure, air temperature, area, power coefficient	

Another study (Kosovic et al., 2020) expanded on this by incorporating additional parameters such as ice accumulation and horizontal grid spacing, emphasizing the importance of localized atmospheric conditions. (Kosovic et al., 2020) also discussed existing forecasting methodologies, including the DICast forecasting system for weather predictions and various numerical weather prediction (NWP) models, such as the Weather Research and Forecasting model with Real-Time FourDimensional Data Assimilation (WRF RTFDDA). These methodologies leverage mathematical models to analyze atmospheric data, facilitating improved input forecasting for AI models. Furthermore, a comparative analysis of forecasting techniques was conducted in the study (Qureshi et al., 2023) This research evaluated the accuracy of two algorithms: Autoregressive Integrated Moving Average (ARIMA) from traditional statistical models and Gated Recurrent Unit (GRU) from deep learning frameworks. The results indicated that the GRU outperformed ARIMA, showcasing its potential for enhanced predictive capabilities. In a case study from Sri Lanka (Peiris et al., 2021) the authors proposed a LevenbergMarquardt (LM) based Artificial Neural Network (ANN) model for predicting potential wind power generation.

Additionally, another study highlighted the efficacy of Genetic Algorithms as a learning algorithm within ANN frameworks, demonstrating superior performance in wind power predictions. Overall, most studies utilize AI or machine learning methods to forecast wind power. Artificial Neural Networks (ANN) were chosen for developing the AI model due to their relative ease of use compared to alternative methods and their high accuracy.

Moreover, ANN is well-suited for modeling nonlinear relationships and is integrated into several simulation software platforms. The algorithm proposed in study (Fernandes & Yen, 2021) is capable of identifying three distinct solutions regarding training/test error and computational complexity. These solutions are achieved by removing convolution filters from multiple layers within the network architecture, employing a network pruning method. Similarly, study (Im & Dasari, 2022) utilized network pruning to diminish the computational complexity of deep neural networks. Furthermore, study (Freire et al., 2022) implemented a pruning strategy to achieve a reduction in the computational complexity of neural networks. Study (Hossain et al., 2023) describes various techniques for reducing the computational complexity of deep neural networks. One technique is pruning, which involves removing the least significant learning parameters of deep learning models to lower complexity without compromising performance. Another technique is quantization and model compression, which reduces the precision of the weights and activations of the neural network.

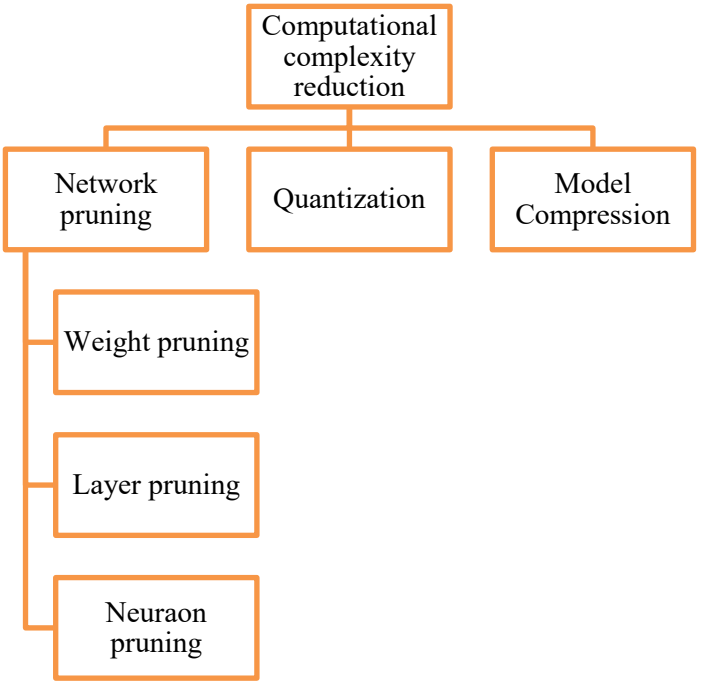


Figure 12. Computational complexity reduction techniques Methodology

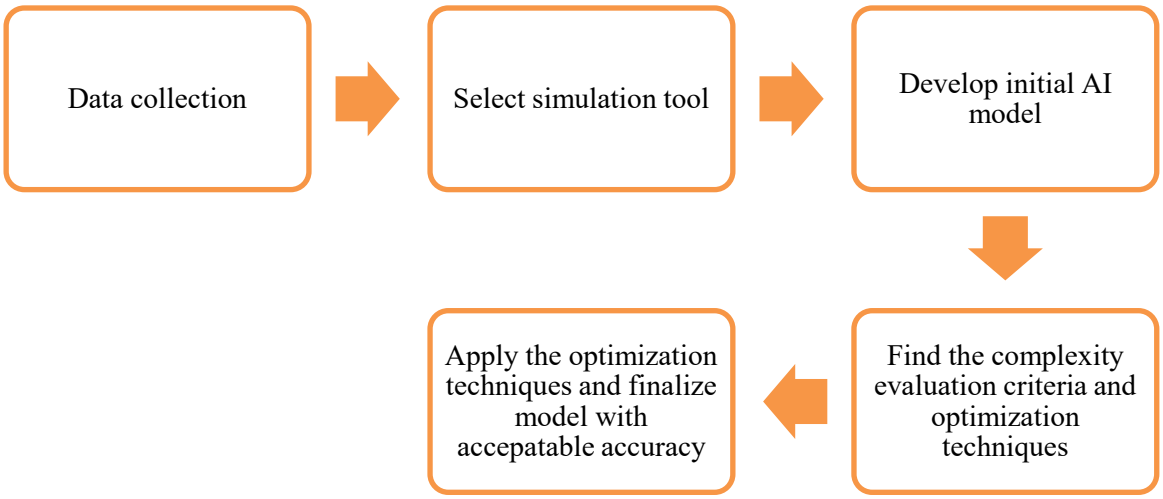


Figure 13. Methodology

DATA COLLECTION

The factors affecting wind energy were classified into two categories: meteorological factors and physical factors. Following the installation of the wind power plant, physical parameters such as rotor diameter and hub height were held constant over time. Therefore, any variations in wind power output were attributed solely to changes in meteorological parameters. As a result, physical parameters were excluded from the model. Only meteorological parameters—such as wind speed, wind direction, relative humidity, rainfall, snowfall, surface pressure, temperature, and air density—were used to train the initial model. All meteorological parameters identified in the literature were utilized in the development of the initial model.

The wind energy generation dataset was referenced in study (Effenberger & Ludwig, 2022) and was downloaded from the website cited in (*Wind Turbine Scada Dataset*, n.d.). This dataset included wind energy, wind speed, and wind direction. The actual location of the wind power plant was identified as HXPR+54J Şenköy, Çınarcık/Yalova, Türkiye. The remaining meteorological data were obtained from a meteorological data collection website (🏛 *Historical Weather API | Open-Meteo.Com*, n.d.). Subsequently, these datasets were properly combined to create the finalized dataset, which was used to develop the initial model.

Simulation tools

The simulation software gives various machine learning and deep learning techniques to develop AI model. Artificial Neural Networks (ANNs) have proven to be superior to other AI techniques for wind power forecasting due to their ability to effectively model complex non-linear relationships inherent in wind patterns (Peiris et al., 2021). Additionally, ANNs excel in processing large datasets, which is critical for wind forecasting where historical and real-time data from multiple parameters must be analyzed simultaneously. This capability improves the better generalization with large dataset (Hu et al., 2021). Another significant advantage of ANNs is their adaptability to real-time forecasting (Ju et al., 2019). Furthermore, ANNs exhibit robustness when handling noisy or incomplete data, a common challenge in wind power datasets. Techniques such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs) within ANN frameworks have shown greater resilience to data inconsistencies (Qureshi et al., 2023).

INITIAL AI MODEL STRUCTURE

The AI model was trained using the finalized dataset. The software tool was provided with a feature to change the hidden layer size. In the initial model, the hidden layer size was set to 10, which was the default value given by the software. A coefficient of determination of 93.5% was achieved with a hidden layer size of 10. The hidden layer size was then manually adjusted, and the accuracy was evaluated for each configuration. Although the maximum coefficient of determination, 95.1%, was obtained with a hidden layer size of 200.

The coefficient of determination (R^2) (Pham et al., 2019) was used as a key indicator of model accuracy, with a value approaching 1 (or 100%) indicating an excellent fit. The accuracy of the initial model was illustrated in Figure 3.

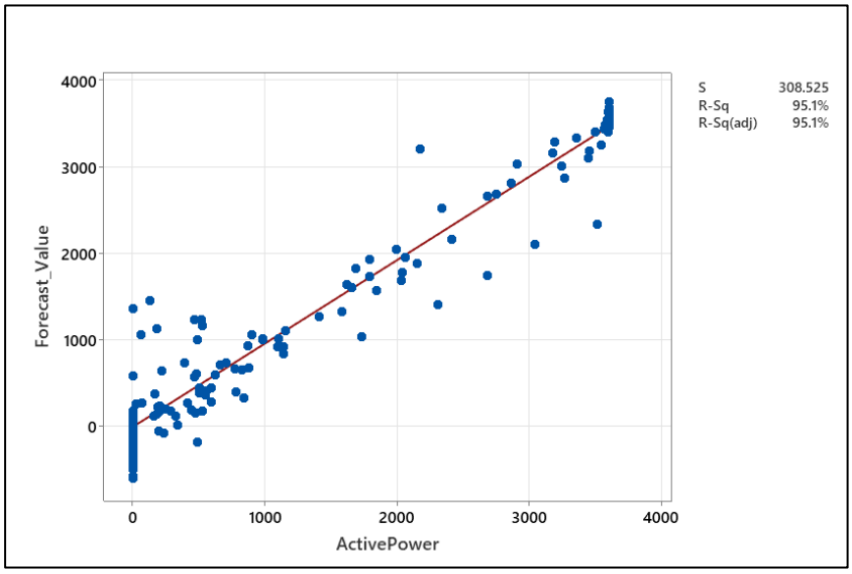


Figure 14. Graph actual power vs model output

Complexity Evaluation Criteria

According to study (Cheng et al., 2020), a higher number of parameters was associated with a more

complex model; therefore, the number of parameters was used to determine computational complexity. Study (Han et al., 2016) indicated that larger models consumed more storage and memory, allowing model size to be used as an indicator of efforts to reduce computational complexity. Additionally, if the inference time was high, it resulted in output delays and required greater computational power. Hence, according to study (Wang et al., 2020), inference time was also considered a measure of computational complexity. Study (He, n.d.) reported that models with higher computational complexity typically required longer training times; thus, training time was used as another indicator. Furthermore, according to study (Pham et al., 2019), R^2 and RMSE were applied as accuracy measures of AI models.

Optimization techniques

The computational complexity of an AI model was significantly influenced by its parameters, as these determined the amount of computation required for both training and inference. Therefore, the number of parameters was reduced in order to decrease computational complexity. To support this reduction, the linear relationship between output power and the model's parameters was required to be evaluated, as it could guide effective parameter selection.

In (Ratner, 2009), correlation analysis was employed to assess the linear relationship between two variables. A correlation value close to 1 was interpreted as indicating a perfect positive linear relationship, while a value close to -1 was considered to represent a perfect negative linear relationship. The evaluation criteria, as defined by (Ratner, 2009), were as follows:

Values between 0 and 0.3 (or 0 and −0.3) were interpreted as indicating a weak positive (or negative) linear relationship.

Values between 0.3 and 0.7 (or −0.3 and −0.7) were considered to reflect a moderate positive (or negative) linear relationship.

Values between 0.7 and 1.0 (or −0.7 and −1.0) were taken to indicate a strong positive (or negative) linear relationship.

Table 14. Summary of correlations.

Parameter	**Correlation**
Wind speed (m/s)	0.935
Wind direction (°)	0.907
Relative humidity (%)	0.305
Rain (mm)	0.1
Surface pressure (Pa)	-0.019
Air density (kg/m³)	0.115
Snowfall (mm)	-0.038
Temperature (K)	0.274

A significant positive linear relationship was identified between wind speed and wind direction; consequently, these variables were considered essential for the model. Relative humidity and temperature were found to exhibit a moderately positive linear relationship. Rainfall, surface pressure, air density, and snowfall demonstrated weak relationships and were therefore excluded from the original dataset. Subsequently, the model was trained using only wind speed and wind direction, which resulted in low accuracy (51.8%). As a result, relative humidity and temperature were reintroduced into the model. The improvement of the model after removing parameters with weak correlations is demonstrated in Table 3.

Table 15. Improvement of the model

Evaluation Parameters	**Initial model**	**Parameter removed model**
Number of parameters	8	4
Model size (KB)	5	5
Inference time (S)	0.019078	0.018402
Training time (S)	10.81253	9.265138
RMSE	308.525	305.882
R^2	95.1%	94.9%

The literature review outlined various methods for reducing computational complexity and provided a brief analysis of these approaches. Network pruning techniques were examined, and weight pruning, layer

pruning, and neuron pruning techniques were applied for further modifications. According to (Molchanov et al., n.d.), a pruning ratio of 30% (0.3) was used, meaning that the smallest 30% of the weights (in terms of absolute value) were set to zero.

In weight pruning, if individual weights were less than or equal to 0.3, those weights were removed from the neural network. In layer pruning, if the weights of a layer were less than or equal to 0.3, the entire layer was removed. In neuron pruning, if the weights associated with a neuron were less than or equal to 0.3, the entire neuron was removed from the neural network. The improvement of the model after pruning is demonstrated in Table 4.

Table 16. Improvement of the model after pruning

Evaluation Parameters	**Initial model**	**Parameter removed +pruning**
Number of parameters	8	4
Model size (KB)	5	5
Inference time (S)	0.019078	0.0072
Training time (S)	10.81253	3.11
RMSE	308.525	353.025
R^2	95.1%	93.3%

The literature review outlined various methods for reducing computational complexity and provided a brief analysis of these approaches. Network quantization techniques were examined. Quantization was presented as a model compression technique used in machine learning and deep learning to reduce the memory footprint and computational requirements of neural networks. The improvement of the model after quantization is demonstrated in Table 5.

Table 17. Improvement of the model after quantization

Evaluation Parameters	**Initial model**	**Parameter removed +pruning+ quantization**
Number of parameters	8	4
Model size (KB)	5	4
Inference time (S)	0.019078	0.0069
Training time (S)	10.81253	0.79
RMSE	308.525	379.267
R^2	95.1%	91.3%

RESULTS

The artificial neural network (ANN) model developed for wind power forecasting was shown to have achieved substantial improvements in predictive accuracy as the network architecture was enhanced. The evaluation process was conducted using a systematic approach, in which the model's performance was assessed across various levels of complexity. Table 6 presents the performance comparison between the initial and optimized models.

Table 18. Performance comparison between the initial and optimized models

Evaluation Parameters	**Initial model**	**Parameter removed +pruning+ quantization**	**Percentage improvements**
Number of parameters	8	4	50%
Model size (KB)	5	4	33.3%
Inference time (S)	0.019078	0.0069	63.83%
Training time (S)	10.81253	0.79	92.69%
RMSE	308.525	379.267	-22.94%
R^2	95.1%	91.3%	-3.99%

The initial model was reported to have achieved a final accuracy of 95.1%. A coefficient between 0.7 and 0.9 was interpreted as indicating moderate to strong predictive capability, while an R-squared value exceeding

0.9 was considered indicative of excellent performance. The final model achieved an accuracy of 91.3%, which was regarded as acceptable. Figure 3 illustrates the accuracy of the initial and final models.

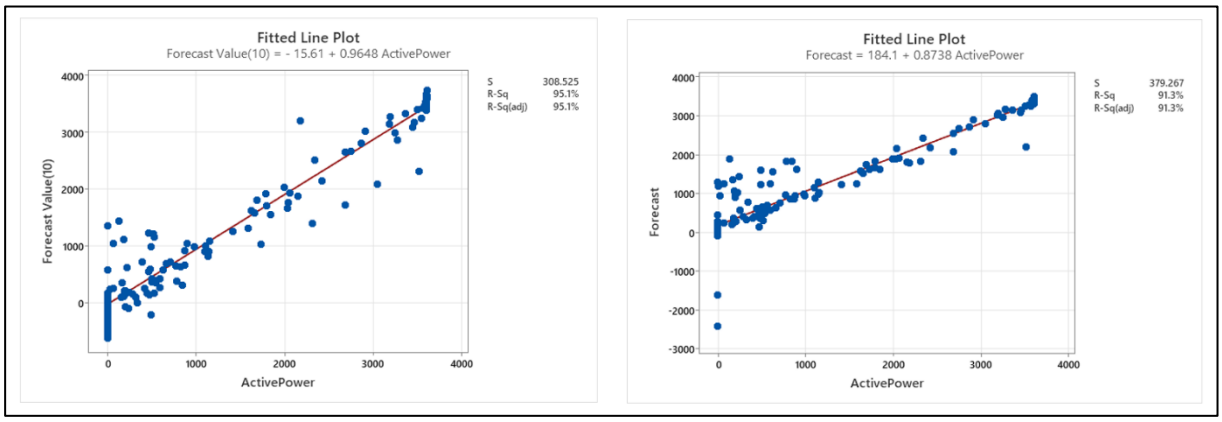


Figure 15. Accuracy of the initial and final models

DISCUSSION

The initial ANN model was successfully developed using eight meteorological parameters and achieved high predictive accuracy, with an R^2 value of 95.1%. Simulation tools were utilized to enable efficient training and validation. Weakly correlated parameters—such as rainfall, surface pressure, and others—were removed to reduce input dimensionality. The application of pruning and quantization techniques further streamlined the model, resulting in a 92.7% reduction in training time and a 63.8% reduction in inference time. Despite these reductions, the optimized model maintained an acceptable level of accuracy, with an R^2 value of 91.3%, demonstrating a practical balance between computational efficiency and predictive capability. The accuracy of an AI model can be reduced by pruning techniques, which work by removing less significant weights, neurons, or entire connections from the network to make it smaller and more efficient.

Training time was reduced from 10.81 seconds to 0.79 seconds, facilitating faster model updates. Additionally, model size was decreased by 33.3%, enhancing its suitability for deployment on edge devices within power grid systems. The reduction in inference time (63.8%) supports real-time implementation in time-sensitive grid management scenarios. While there was a 3.99% decrease in R^2 and a 22.94% increase in RMSE, these changes were considered acceptable given the substantial improvements in speed and resource efficiency. This trade-off aligns with real-world requirements, where lightweight and scalable models are prioritized.

However, it should be noted that the data used—sourced from Türkiye—may not fully capture seasonal variability present in other regions. Moreover, the model's reliance on wind speed wind direction, relative humidity and temperature may limit its accuracy in locations where other meteorological factors, such as turbulence, have a greater impact. For applications in other regions, this model can be further enhanced by incorporating additional parameters to improve its generalization capabilities, suggesting a potential direction for future work.

CONCLUSION

This research developed a short-term ANN-based wind power forecasting model and optimized it through parameter reduction, pruning, and quantization. The final model achieved a 92.7% reduction in training time and a 63.8% improvement in inference speed, while maintaining an accuracy of 91.3% (R^2), demonstrating its suitability for real-time grid applications. The resulting lightweight model enabled faster and decentralized forecasting, thereby enhancing the integration of renewable energy sources. Additionally, the reduced computational requirements allowed deployment on low-power devices, increasing the model's accessibility and practicality.

Future work will aim to develop a model tailored to diverse wind generation systems. Efforts will also be directed toward reintroducing physical parameters—such as rotor diameter and hub height—to further generalize the model for broader applicability.

ACKNOWLEDGEMENTS

I would also like to express my sincere appreciation to the Faculty of Technology at Wayamba University of Sri Lanka and the Department of Electrotechnology for providing me with the opportunity to undertake this research. The resources, support, and academically enriching environment offered by the faculty were essential

to the successful completion of this study. I am sincerely thankful for the trust and opportunity extended to me.

REFERENCES

[1] *Historical Weather API | Open-Meteo.com*. (n.d.). Retrieved June 29, 2025, from https://open-meteo.com/en/docs/historical-weather-api?utm_source=chatgpt.com

[2] Alkesaiberi, A., Harrou, F., & Sun, Y. (2022). Efficient Wind Power Prediction Using Machine Learning Methods: A Comparative Study. *Energies*, *15*(7). https://doi.org/10.3390/en15072327

[3] Badran, O., & Abdulhadi, E. (n.d.). *EVALUATION OF FACTORS AFFECTING WIND POWER GENERATION IN JORDAN*.

[4] Cheng, Y., Wang, D., Zhou, P., & Zhang, T. (2020). *A Survey of Model Compression and Acceleration for Deep Neural Networks*. http://arxiv.org/abs/1710.09282

[5] Effenberger, N., & Ludwig, N. (2022). A collection and categorization of open-source wind and wind power datasets. In *Wind Energy* (Vol. 25, Issue 10, pp. 1659–1683). John Wiley and Sons Ltd. https://doi.org/10.1002/we.2766

[6] Fernandes, F. E., & Yen, G. G. (2021). Pruning Deep Convolutional Neural Networks Architectures with Evolution Strategy. *Information Sciences*, *552*, 29–47. https://doi.org/10.1016/j.ins.2020.11.009

[7] Franke, K., Sensfuß, F., Deac, G., Kleinschmitt, C., & Ragwitz, M. (2021). Factors affecting the calculation of wind power potentials: A case study of China. *Renewable and Sustainable Energy Reviews*, *149*. https://doi.org/10.1016/j.rser.2021.111351

[8] Freire, P. J., Napoli, A., Ron, D. A., Spinnler, B., Anderson, M., Schairer, W., Bex, T., Costa, N., Turitsyn, S. K., & Prilepsky, J. E. (2022). *Reducing Computational Complexity of Neural Networks in Optical Channel Equalization: From Concepts to Implementation*. https://doi.org/10.1109/JLT.2023.3234327

[9] Han, S., Mao, H., & Dally, W. J. (2016). *Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding*. http://arxiv.org/abs/1510.00149

[10] He, Y. (n.d.). *Channel Pruning for Accelerating Very Deep Neural Networks*.

[11] Hossain, M. B., Gong, N., & Shaban, M. (2023). Computational Complexity Reduction Techniques for Deep Neural Networks: A Survey. *2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things, AIBThings 2023 - Proceedings*. https://doi.org/10.1109/AIBThings58340.2023.10292477

[12] Hu, S., Xiang, Y., Huo, D., Jawad, S., & Liu, J. (2021). An improved deep belief network based hybrid forecasting method for wind power. *Energy*, *224*. https://doi.org/10.1016/j.energy.2021.120185

[13] Im, M. S., & Dasari, V. R. (2022). *Computational complexity reduction of deep neural networks*. http://arxiv.org/abs/2207.14620

[14] Ju, Y., Sun, G., Chen, Q., Zhang, M., Zhu, H., & Rehman, M. U. (2019). A model combining convolutional neural network and lightgbm algorithm for ultra-short-term wind power forecasting. *IEEE Access*, *7*, 28309–28318. https://doi.org/10.1109/ACCESS.2019.2901920

[15] Kosovic, B., Haupt, S. E., Adriaansen, D., Alessandrini, S., Wiener, G., Monache, L. D., Liu, Y., Linden, S., Jensen, T., Cheng, W., Politovich, M., & Prestopnik, P. (2020). A comprehensive wind power forecasting system integrating artificial intelligence and numerical weather prediction. *Energies*, *16*(3). https://doi.org/10.3390/en13061372

[16] Li, N., He, F., & Ma, W. (2019). Wind power prediction based on extreme learning machine with kernel mean p-power error loss. *Energies*, *12*(4). https://doi.org/10.3390/en12040673

[17] Lipu, M. S. H., Miah, M. S., Hannan, M. A., Hussain, A., Sarker, M. R., Ayob, A., Saad, M. H. M., & Mahmud, M. S. (2021). Artificial Intelligence Based Hybrid Forecasting Approaches for Wind Power Generation: Progress, Challenges and Prospects. *IEEE Access*, *9*, 102460–102489. https://doi.org/10.1109/ACCESS.2021.3097102

[18] Molchanov, P., Mallya, A., Tyree, S., Frosio, I., & Nvidia, J. K. (n.d.). *Importance Estimation for Neural Network Pruning*. https://github.com/NVlabs/Taylor_pruning.

[19] Pang, C., Yu, J., & Liu, Y. (2021). Correlation analysis of factors affecting wind power based on machine learning and Shapley value. *IET Energy Systems Integration*, *3*(3), 227–237. https://doi.org/10.1049/esi2.12022

[20] Peiris, A. T., Jayasinghe, J., & Rathnayake, U. (2021). Forecasting wind power generation using artificial neural network: “Pawan danawi”-A case study from Sri Lanka. *Journal of Electrical and*

Computer Engineering, *2021*. https://doi.org/10.1155/2021/5577547

[21] Pham, B. T., Nguyen, M. D., Dao, D. Van, Prakash, I., Ly, H. B., Le, T. T., Ho, L. S., Nguyen, K. T., Ngo, T. Q., Hoang, V., Son, L. H., Ngo, H. T. T., Tran, H. T., Do, N. M., Van Le, H., Ho, H. L., & Tien Bui, D. (2019). Development of artificial intelligence models for the prediction of Compression Coefficient of soil: An application of Monte Carlo sensitivity analysis. *Science of the Total Environment*, *679*, 172–184. https://doi.org/10.1016/j.scitotenv.2019.05.061

[22] Qureshi, S., Shaikh, F., Kumar, L., Ali, F., Awais, M., & Gürel, A. E. (2023). Short-term forecasting of wind power generation using artificial intelligence. *Environmental Challenges*, *11*. https://doi.org/10.1016/j.envc.2023.100722

[23] Ratner, B. (2009). The correlation coefficient: Its values range between 1/1, or do they. *Journal of Targeting, Measurement and Analysis for Marketing*, *17*(2), 139–142. https://doi.org/10.1057/jt.2009.5

[24] Tarek, Z., Shams, M. Y., Elshewey, A. M., El-Kenawy, E. S. M., Ibrahim, A., Abdelhamid, A. A., & El-Dosuky, M. A. (2023). Wind Power Prediction Based on Machine Learning and Deep Learning Models. *Computers, Materials and Continua*, *74*(1), 715–732. https://doi.org/10.32604/cmc.2023.032533

[25] Van Der Merwe, S., Du Plessis, A., Rix, A. J., Van Der Merwe, S. W., Du Plessis, A. A., & Rix, A. J. (n.d.). *WIND POWER FORECASTING: A REVIEW*. https://www.researchgate.net/publication/380106361

[26] Wang, X., Han, Y., Leung, V. C. M., Niyato, D., Yan, X., & Chen, X. (2020). *Convergence of Edge Computing and Deep Learning: A Comprehensive Survey*. https://doi.org/10.1109/COMST.2020.2970550

[27] *Wind Turbine Scada Dataset*. (n.d.). Retrieved June 29, 2025, from https://www.kaggle.com/datasets/berkerisen/wind-turbine-scada-dataset/data

Enhancement of Quality Management Through Lean in Sri Lankan Construction Industry

Jayanetti J.K.D.D.T.[a*], Perera B.A.K.S.[b], Ranadewa K.A.T.O[c]

[a]Sri Lanka Institute of Information Technology, New Kandy Road, Malabe, 10115, Sri Lanka
[b]University of Moratuwa, Bandaranayake Mawatha, Moratuwa, 10400, Sri Lanka
thamasha.j@sliit.lk*, kanchana@uom.lk, tharushar@uom.lk

ABSTRACT

Quality management remains a critical concern in the construction industry of developing countries, where inefficiencies, rework, and inconsistencies in quality practices negatively impact project outcomes. Although lean construction features prominently in the global literature, the rigorous integration of lean principles into quality management frameworks remains underexplored. While lean construction is widely recognised for enhancing process efficiency and value delivery, its integration with quality management, particularly through structured frameworks, has received limited attention in the Sri Lankan context. This study addresses this gap by investigating how lean can be applied to enhance quality management in Sri Lankan construction organisations. Guided by a pragmatic research philosophy, this study employed the Delphi technique involving experts with expertise in lean construction and quality management. Data was analysed using NVivo through directed content analysis.

The study identified lean-related quality process areas and performance indicators based on established literature and expert judgement. The validated indicators were organised into four core process areas: continuous improvement, benchmarking, standardisation, and error detection and prevention. A total of 22 performance indicators corresponding to these areas were confirmed through expert consensus. The findings show strong alignment with established lean concepts such as Kaizen, the Plan Do Check Act cycle, standard work, and quality at source. The results also reflect local priorities such as proactive error management and regulatory alignment, emphasising the need for contextual adaptation. The study extends lean quality theory to a new geographic setting, offers a practical framework for Sri Lankan construction organisations, demonstrates the methodological value of the Delphi approach in data-limited contexts, and supports societal goals by promoting more reliable and accountable construction practices. These contributions advance understanding and implementation of lean-based quality management in emerging construction sectors.\

KEYWORDS: *Construction Quality Management, Construction, Lean Construction, Construction Dynamics*

INTRODUCTION

Lean Construction (LC), adapted from lean manufacturing, represents a paradigm shift in the construction industry aimed at minimizing waste, enhancing value, and improving overall process efficiency through continuous improvement and stakeholder collaboration (Khan et al., 2024; Koskela, 2000). As construction projects face increasing demands for timely delivery, cost-effectiveness, and consistent quality, the integration of lean principles has gained global traction. Tools and techniques such as the Last Planner System, 5S, visual management, and value stream mapping have been widely recognized for their potential to improve process flow, reduce rework, and enhance construction quality (Salem et al., 2005; Aziz & Hafez, 2013).

Quality management in construction is a critical dimension that ensures project outputs meet client expectations, regulatory requirements, and performance standards (Alsharef et al., 2024). However, quality failures remain a recurring issue, particularly in developing economies where project environments are often characterized by fragmented processes, poor communication, and reactive management strategies (Love et al., 2018). Within this context, the degree to which lean principles are embedded in an organization's culture and operational systems plays a vital role in determining the effectiveness of quality-related outcomes. High levels of lean integration are typically associated with a proactive quality culture, emphasizing defect prevention, continuous monitoring, and systematic improvement rather than corrective actions after problems arise (Nesensohn et al., 2015; Marhani et al., 2012).

In Sri Lanka, the construction industry has begun exploring lean practices, although adoption remains at a relatively early stage (Ranadewa et al., 2021). Empirical studies suggest a growing awareness of lean principles but also highlight gaps in practical implementation and strategic integration with quality management systems (Senaratne & Wijesiri, 2008; Jayanetti et al., 2024). The limited application of lean construction tailored to the Sri Lankan context emphasise a critical research gap, particularly in understanding how lean can be incorporated to address persistent quality challenges. Despite growing global evidence on the role of lean practices in improving construction quality, there is limited empirical research examining how lean, contributes to construction quality management in the Sri Lankan context. Existing literature seldom addresses the operational realities, fragmented processes, and quality-related challenges specific to

developing construction sector. This study addresses this gap by exploring how lean can be effectively used to enhance quality management in Sri Lankan construction organizations through the identification of relevant process areas and performance indicators. The present study aims to investigate how quality management can be enhanced through lean within Sri Lankan construction organizations. Specifically, it seeks to identify key process areas (KPAs) that reflect lean attributes relevant to quality, define performance indicators (PIs) associated with these KPAs, and explore how these can be operationalized to drive quality improvements in practice. A qualitative, three-round Delphi method is employed to achieve expert consensus, supported by NVivo-based content analysis for thematic synthesis. The findings are expected to contribute to the development of a structured, context-sensitive framework that integrates lean and quality management, offering both theoretical insight and practical value to the Sri Lankan construction sector.

The paper begins with an introduction and literature review, followed by a Delphi-based methodology. Key findings on process areas and performance indicators are presented, discussed in relation to existing literature, and summarized through theoretical, practical, methodological, and societal contributions, concluding with implications for future research.

LITERATURE REVIEW

Lean Construction and Quality Management

Lean Construction (LC) originates from the principles of lean manufacturing, focusing on maximizing value for the client while minimizing waste throughout the construction process (Aslam et al., 2024; Koskela et al., 2002; Ballard & Howell, 2003). By applying tools such as the Last Planner® System, 5S, visual management, and value stream mapping, LC enables improved planning, flow, and coordination across project phases. These tools foster collaboration, reduce variability, and support a culture of continuous improvement, which are all foundational to high-quality project delivery (Salem et al., 2005; Aziz & Hafez, 2013). Globally, LC has been credited with improving cost performance, reducing delays, and enhancing quality outcomes on complex construction projects.

Quality management in construction encompasses planning, assurance, and control measures to ensure projects meet predefined standards, client requirements, and performance expectations (Alsharef et al., 2024). Traditional quality control approaches in construction often emphasize inspection and correction rather than prevention and system optimization. Lean contributes to reversing this trend by embedding a proactive, process-driven approach to quality (Hamzo Khan et al., 2024). Organizations with higher lean maturity levels often integrate quality checks within standardized workflows, enabling early detection of issues and continuous refinement (Nesensohn et al., 2015; Salem et al., 2005). As lean implementation matures, firms tend to exhibit improved quality metrics, including reduced defect rates, enhanced documentation, and better stakeholder satisfaction.

Quality Management in Construction Through Lean Integration

Lean construction supports quality improvement by fostering structured problem-solving, standardization, and real-time monitoring of performance (Gartoumi et al., 2024). Techniques such as root cause analysis, visual management, A3 thinking, and kaizen promote transparency and accountability in day-to-day operations (Salem et al., 2005). As organizations mature, their quality systems shift from reactive to predictive models, with built-in quality controls embedded into production processes. This enables the prevention of errors before they occur and promotes a continuous feedback loop, enhancing product and process quality (Nesensohn et al., 2015; Love et al., 2018). Thus, lean acts as both a framework and an enabler for improving construction quality outcomes over time.

Key Process Areas (KPAs) in lean reflect the foundational domains that must be developed to achieve sustained lean performance. These include continuous improvement, customer orientation, leadership involvement, team development, waste elimination, and knowledge sharing (Aziz & Hafez, 2013). Each KPA is associated with measurable outcomes, typically monitored through Key Performance Indicators (KPIs). In the context of quality management, relevant KPIs include defect frequency, rework rates, quality compliance ratios, customer satisfaction scores, and cost of quality. These metrics help organizations track performance trends, identify bottlenecks, and drive improvements systematically. However, the integration of KPAs and KPIs tailored to quality enhancement in lean remains under-researched, particularly in developing construction environments.

In the Sri Lankan construction industry, quality issues such as excessive rework, defects, and client dissatisfaction are persistent challenges. Despite growing awareness of lean principles, adoption remains limited and often isolated to individual projects or tools (Senaratne & Wijesiri, 2008). Cultural resistance to change, lack of training, and absence of a systematic approach to lean implementation hinder sustained impact. Moreover, quality management practices are often reactive and inspection-driven rather than prevention-focused. Introducing a lean perspective provides an opportunity to institutionalize continuous improvement, enhance process consistency, and build a proactive quality culture across the industry. Recent local research by Jayanetti et al. (2024) highlights a set of lean elements suited to Sri Lankan conditions, indicating the growing relevance of this approach for long-term quality enhancement.

RESEARCH GAP

Theoretical Gap

Although Lean Construction has been extensively studied across various international contexts, particularly in developed economies, its integration with quality management practices remains underexplored in developing countries. Prior research has demonstrated that lean implementation can significantly improve quality performance by minimizing waste, reducing variability, and fostering a culture of continuous improvement (Gartoumi et al., 2024; Aziz & Hafez, 2013; Love et al., 2018). However, much of this evidence is derived from projects in countries with mature construction industries, robust institutional frameworks, and strong regulatory systems (Nesensohn et al., 2015). As a result, there is a limited understanding of how lean, specifically contributes to quality enhancement in emerging markets such as Sri Lanka. Furthermore, the majority of existing lean models have been developed in Western contexts and are not tailored to the socio-economic, regulatory, and cultural nuances of South Asian construction industries. These models often lack adaptability to local conditions where resource constraints, skill shortages, and inconsistent project governance are common (Ranadewa et al., 2021). While some recent attempts have been made to define localized lean indicators (Jayanetti et al., 2024), they do not yet provide a comprehensive framework for linking lean construction with quality performance outcomes. This constitutes a critical literature gap, as it limits the theoretical understanding of how lean can be effectively integrated to improve quality management processes in these contexts.

Industry Gap

From an industry perspective, Sri Lanka's construction sector continues to struggle with recurring quality issues such as high defect rates, rework, client dissatisfaction, and poor documentation practices. These problems are often exacerbated by fragmented supply chains, inconsistent monitoring, and a lack of integrated quality management systems (Ranadewa et al., 2021). Although awareness of lean tools and techniques is gradually increasing, their practical application remains irregular and largely limited to isolated initiatives. There is no widely adopted mechanism for construction firms to assess their lean maturity levels in relation to quality performance metrics. This practical gap emphasize the absence of a validated, context-specific framework that construction organizations can use to systematically improve quality through lean practices.

METHODOLOGY

This study adopts a pragmatic research philosophy, which is well suited to investigations aiming to generate practical solutions grounded in real-world contexts. Pragmatism allows researchers to integrate both objective and subjective perspectives, thereby supporting the development of context-specific knowledge which is an essential requirement for addressing complex problems in the Sri Lankan construction industry (Saunders et al., 2019). The research follows an abductive approach, enabling iterative interaction between empirical data and theoretical insights. This approach is particularly appropriate for studies seeking to build new conceptual understanding while remaining responsive to the realities of practice (Dubois & Gadde, 2002). In line with the research aim, a mono-method qualitative choice was adopted to facilitate an in-depth exploration of expert perspectives and nuanced themes that are not easily quantifiable (Creswell & Plano Clark, 2017). A survey strategy was employed through the use of the Delphi technique, The Delphi technique was selected for its strength in capturing expert consensus in complex and under-researched domains, making it particularly suited for exploring context-specific process indicators within the Sri Lankan construction industry (Hasson et al., 2000; Okoli & Pawlowski, 2004). As a qualitative approach, Delphi technique enables structured and anonymous expert engagement, supports inbuilt validation, and allows iterative refinement of findings, aligning well with the pragmatic philosophy underpinning this study (Skulmoski et al., 2007). A cross-sectional time horizon was applied, as data were collected within a defined period to reflect current industry conditions. Data collection involved three rounds of semi-structured interviews with 25 purposively selected experts from academia and industry. These interviews, conducted either in person or via online platforms, ranged from 60 to 75 minutes in duration. In Delphi studies, various thresholds are commonly employed to determine consensus (Schifano and Niederberger 2025). Considering the depth of expertise within the selected panel and the scope of the research, a consensus threshold of 75% was deemed appropriate for this study to ensure the credibility and reliability of emerging findings across Delphi rounds (Linstone & Turoff, 2002). Data were analysed using thematic content analysis supported by NVivo software, which enabled systematic coding, theme development, and traceability of expert responses (Bazeley & Jackson, 2013). This methodological design ensured that the findings were both rigorously derived and contextually grounded, thereby enhancing their relevance to quality management practices in the Sri Lankan construction sector.

The selection of Delphi experts was carried out through a rigorous screening process to ensure that participants possessed substantial and relevant expertise in lean construction and construction quality management. The studies require the expertise of the panel in both areas to achieve accuracy and credibility of the data. The panel comprised a balanced mix of industry experts and academics who met the defined eligibility thresholds based on their experience, qualifications, and professional engagement. This ensured the inclusion of highly competent individuals with both practical insights and scholarly depth. A total of 25 experts participated in Round 1, with 24 continuing through Rounds 2 and 3, demonstrating a high level of engagement and commitment throughout the process. Initially, a panel of 30 experts was considered for the study. However, by the 23rd interview, data saturation had been reached, as no new themes or insights were emerging. To ensure robustness and confirm thematic stability, two additional expert interviews were conducted, bringing the total to 25 (Diamond et al., 2014). The strength of this panel enhanced the credibility and reliability of the consensus findings, ensuring that the outcomes reflect informed judgment rooted in both theory and practice.

Table 1 summarises the objectives of each Delphi round and corresponding phase. In Round 1–Phase I, experts identified process areas related to quality management in lean, while Phase II focused on identifying relevant performance indicators. Round 2–Phase I refined and contextualised the key process areas for Sri Lankan construction organisations, and

Phase II validated the corresponding performance indicators. In Round 3–Phase I, notable indicators were assessed for each key process area, followed by Phase II, which explored their practical application in enhancing quality management through lean implementation.

Table 19.Expert Profile and Delphi Rounds

Expert Code	Compulsory (Must satisfy at least one from each)				Additional Qualifications (Must satisfy at least 3)					Participation		
	A		B									
	5+ experience in LC job role	3+ academic experience in LC	10+ experience Construction quality management	10+ experience in academia	Graduated in built environment discipline	Professional affiliation to Lean related institution	Member of professional institution in built environment	Master's qualification in built-environment or lean	2+ indexed publications	Delphi R1 25	Delphi R2 24	Delphi R3 24
E1	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓
E2	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓
E3	✓		✓		✓	✓	✓	✓		✓	✓	✓
E4	✓		✓		✓		✓	✓		✓	✓	✓
E5	✓		✓		✓		✓	✓		✓	✓	✓
E6	✓		✓		✓		✓	✓		✓	✓	✓
E7	✓		✓	✓			✓	✓		✓	✓	✓
E8	✓		✓		✓	✓	✓			✓	✓	✓
E9	✓		✓		✓		✓			✓	✓	✓
E10	✓		✓	✓	✓		✓			✓	✓	✓
E11	✓		✓	✓	✓		✓			✓	✓	✓
E12	✓		✓	✓	✓		✓			✓	✓	✓
E13	✓		✓	✓	✓		✓			✓	✓	✓
E14	✓		✓	✓	✓		✓			✓	✓	✓
E15		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
E16		✓		✓	✓	✓	✓	✓		✓	✓	✓
E17		✓		✓	✓			✓	✓	✓	✓	✓
E18		✓		✓	✓			✓	✓	✓	✓	✓
E19		✓		✓	✓			✓	✓	✓	✓	✓
E20		✓		✓	✓			✓	✓	✓	✓	✓
E21		✓		✓	✓			✓	✓	✓	✓	✓
E22		✓		✓	✓		✓	✓		✓	✓	✓
E23		✓		✓	✓		✓	✓		✓	✓	✓
E24		✓		✓	✓		✓	✓		✓	✓	✓
E25		✓		✓	✓		✓	✓		✓		

Delphi Round	**Phase**	**Objective**
Round 1	Phase I	Identify process areas related to quality management in lean construction
	Phase II	Identify performance indicators related to quality management in lean construction
Round 2	Phase I	Investigate key process areas related to quality management in lean construction in the Sri Lankan construction organisations
	Phase II	Identify performance indicators related to quality management in lean construction in the Sri Lankan construction organisations
Round 3	Phase I	Assess performance indicators for each key process area related to quality management in lean construction in the Sri Lankan construction organisations
	Phase II	Investigate how to use the key process areas and performance indicators in lean construction to enhance quality management in the Sri Lankan construction organisations

FINDINGS

This section presents the findings of the Delphi rounds.

KEY PROCESS AREAS RELATED TO QUALITY MANAGEMENT IN LEAN IN THE SRI LANKAN CONSTRUCTION ORGANISATIONS (DELPHI ROUND 1 PHASE1, DELPHI ROUND 2 PHASE 1)

Initially, three model process areas were identified through a review of existing literature. In Delphi Round 1 Phase 1 (DR1P1), experts were invited to evaluate and determine which of these areas were relevant to quality management in the construction sector. Subsequently, in Delphi Round 2 Phase 1 (DR2P1), the experts contextualised these process areas to reflect the specific practices and conditions of the Sri Lankan construction industry. Table 2 presents the key findings of the study. The italicised text presents items identified from the literature and modified for the Sri Lankan context, while the bold text presents new suggestions by experts that received a consensus of more than 75%.

Table 2: Key findings of the research

Model process areas	75% or more	Model Indicators	75% or more	How to use model indicators	75% or more
Continuous improvement	✓	*Progress monitoring and reporting*	✓	Establish regular tracking systems to monitor quality performance trends	✓
	✓	Stretch targets	✓	Set challenging but achievable goals to drive continuous performance enhancement	✓
	✓	Employee empowerment	✓	Empower employees by involving them in decision-making and quality initiatives	✓
	✓	**Encouraging innovative ideas**	✓	Create structured platforms to capture and implement innovative ideas	✓
	✓	**Continuous improvement cycles**	✓	Implement structured continuous improvement cycles like PDCA or Kaizen	✓
	✓	**Learning by doing**	✓	Encourage reflective practice through experiential learning and feedback	✓
Benchmarking	✓	Benchmark to high industry standards	✓	Benchmark organisational performance against leading industry standards	✓
	✓	Industry analysis	✓	Conduct systematic industry analyses to inform quality benchmarks	✓
	✓	*Best practices incorporated to sub-processes*	✓	Integrate established best practices into operational subprocesses	✓
	✓	**Up to date benchmark criteria**	✓	Continuously update benchmarking criteria to reflect current standards	✓
	✓	**Performance Metrics**	✓	Use well-defined performance metrics to monitor and improve outcomes	✓
	✓	**Shared internal best practices**	✓	Share effective internal practices across teams and projects	✓
Standardisation	✓	Defined work processes	✓	Define and institutionalise standard workflows to ensure consistency	✓
	✓	Maintenance of quality standards	✓	Implement robust systems to uphold project-specific quality standards	✓
	✓	Process controlling and quality assurance	✓	Apply structured quality control and assurance procedures	✓
	✓	*Proper documentation*	✓	Maintain thorough and accessible project documentation for traceability	✓
	✓	**Standard Operating procedures for site operations**	✓	Develop and enforce SOPs for site-level operations	✓
	✓	**Matching building standards to performance metrics**	✓	Align construction standards with performance requirements and objectives	✓
Error detection and prevention	✓	Prompt defect response & Error prevention	✓	Establish rapid-response systems for defect identification and resolution	✓
	✓	Establishing teams for bottlenecks	✓	Create multidisciplinary teams to address recurring process bottlenecks	✓
	✓	**First-Time Right (FTR) Rate**	✓	Promote 'First-Time Right' practices to minimise rework and defects	✓
	✓	**Independent quality assurance entity**	✓	Engage independent quality assurance bodies for objective evaluations	✓

Three key process areas, namely continuous improvement, benchmarking, and standardisation, were identified through a comprehensive review of existing literature as foundational elements of lean construction that contribute to enhanced quality management. These process areas were presented to the expert panel for validation during Delphi Round 1, Phase I.

The experts endorsed all three areas, each achieving the required consensus threshold of 75 percent or higher, thereby confirming their relevance and applicability within the context of the Sri Lankan construction industry.

In addition to validating the literature-derived areas, the expert panel proposed an additional process area based on contextual insights and professional experience. This new area, identified as error detection and prevention, was not present in the initial literature but was considered critical for addressing practical quality issues in the local construction environment. It was also validated during Delphi Round 1, Phase I, and received more than 75 percent consensus, thereby confirming its significance. As presented in Table 2, this expert-driven process area is highlighted in bold to distinguish it from those established through the literature review.

Key performance indicators for each key process area related to quality management in lean construction in the Sri Lankan construction organisations (Delphi round 1 phase 1, Delphi round 1 phase II, Delphi round 2 phase I, Delphi round 3 phase I)

Experts were presented with 12 indicators derived from the literature for identifying relevant performance indicators. In Delphi Round 1 Phase 2 (DR1P2), all 12 indicators were validated by the expert panel, each achieving a consensus level of 75% or above, confirming their relevance to quality management in lean construction.

In Delphi Round 2 Phase 2 (DR2P2), experts proposed 10 additional performance indicators based on their professional insights and contextual knowledge of the Sri Lankan construction industry. These newly suggested indicators were also validated with the required consensus level, bringing the total number of validated indicators to 22. This outcome reflects the value of expert involvement in capturing nuanced, context-sensitive dimensions of quality management that are not always evident in global literature. A detailed summary of the validated indicators is provided in Table 2. In Table 2, indicators shown in plain text represent those accepted from literature without modification, italics indicate indicators identified from literature but refined by experts, and bold denotes indicators that were newly suggested by the experts.

Subsequently, in Delphi Round 3 Phase 1 (DR3P1), all validated indicators were systematically categorised under the four key process areas of quality management: continuous improvement, benchmarking, standardisation, and error detection and prevention. This categorisation established a clear structure linking each performance indicator to its respective process area, thereby supporting a comprehensive framework for evaluating lean integration for quality management within Sri Lankan construction organisations. This structured classification enhances practical applicability by enabling targeted assessment, monitoring, and improvement initiatives aligned with each process area, while also contributing to theoretical development by contextualising lean-based quality indicators within a developing country setting.

How to use the key process areas and performance indicators in lean construction to enhance the quality management in the Sri Lankan construction organisations (Round 3 phase I, Round 3 phase II)

In Delphi Round 3 Phase 2, experts were specifically tasked with investigating how the validated key process areas and performance indicators could be utilised to enhance quality management within Sri Lankan construction organisations as shown in Table 2. This phase moved beyond identification and validation, focusing instead on the practical integration of these elements into organisational processes and quality management systems.

In the area of benchmarking, experts underscored the dual importance of external alignment and internal knowledge dissemination. *As Expert E4 commented, "Benchmarking against industry standards helps align with global practices, but sharing internal best practices across sites can lead to faster diffusion of quality-enhancing methods."* Indicators such as *benchmark to high industry standards*, *updated benchmark criteria*, *performance metrics*, and *shared internal best practices* were viewed as strategic tools for elevating organisational quality. The integration of *industry analysis* and *sub-process benchmarking* was further seen as enabling data-driven alignment with evolving quality standards.

The standardisation process area was consistently recognised as fundamental to ensuring consistency and compliance. Experts advocated for clearly defined procedures, accessible documentation, and strict adherence to established protocols. *As Expert E1 observed, "A well-documented process is a quality assurance system in itself."* Indicators such as *defined work processes*, *standard operating procedures for site operations*, and *maintenance of quality standards* were viewed as essential for minimising process variability. The emphasis on *process control*, *quality assurance*, and *alignment with building performance metrics* further supports the establishment of a systematic and repeatable approach to quality management.

Finally, the inclusion of error detection and prevention as a distinct process area was acknowledged as a valuable contextual contribution, especially for addressing quality issues at source. Indicators such as *First-Time Right (FTR) rate*, *prompt defect response and error prevention*, and *independent quality assurance entities* were considered vital for embedding preventive quality practices. *Expert E3 stated, "We must move beyond post-completion inspections. Early error detection can reduce waste, time, and cost significantly."* The use of *teams for bottleneck resolution* was also proposed to enable more agile, cross-functional responses to emerging quality issues. Together, these indicators represent a shift from reactive quality control to a more anticipatory and system-wide preventive approach tailored to the Sri Lankan construction environment.

DISCUSSION

This study aimed to explore how lean construction can enhance quality management in Sri Lankan construction organisations by identifying key process areas and associated performance indicators. The Delphi method confirmed four core process areas as continuous improvement, benchmarking, standardisation, and error detection and prevention, mapped to 22 validated indicators.

The prominence of continuous improvement aligns with Kaizen and PDCA-based models (Liker, 2004; Womack & Jones, 2003), which stress iterative feedback and worker involvement. Indicators such as progress monitoring and employee empowerment reflect widely accepted lean practices (Ballard & Howell, 2003). Expert-suggested indicators like learning by doing and continuous improvement cycles reinforce the shift from top-down to collaborative learning processes, supporting findings by Salem et al. (2006) and Sarhan & Fox (2013). Benchmarking indicators confirmed, such as industry analysis and best practice, correspond with value stream mapping and performance gap identification (Sacks et al., 2010). Additions like shared internal benchmarks and real-time metrics mirror practices promoted by Anvari et al. (2011), indicating that both external and internal benchmarking are necessary for lean construction in fragmented construction environments. The role of standardisation is consistent with lean principles that require process stability to support flow (Modig & Åhlström, 2012). Indicators such as defined work processes and quality standards match the foundation of 5S and standard work (Formoso et al., 2002). The expert-validated inclusion of site-specific SOPs and performance-matched building standards highlights the need for regulatory and operational contextualisation, as also noted in Ranadewa et al. (2020). Notably, error detection and prevention emerged as a locally emphasised area. While not prominent in initial lean models, the validated indicators, including First-Time Right (FTR) rate and independent QA entities, reflect modern interpretations of jidoka and poka-yoke for quality at source (Ohno, 1988; Shingo, 1986). These findings align with Tezel et al. (2016), who argue that lean systems require built-in quality control mechanisms beyond inspection. Overall, the study findings confirm key process areas and indicators in lean construction provide tangible mechanisms and monitors to enhance quality management in construction. This supports ideas by Marodin et al. (2019) and Ochieng et al. (2020) for lean integration for successful quality management in construction projects.

RESEARCH CONTRIBUTIONS

Theoretical Contribution

This research advances the theoretical understanding of lean integration for quality management in construction by extending existing frameworks to the Sri Lankan context, where such integration has been limited. While global literature recognises lean maturity as a driver of performance improvement (Sacks et al., 2010; Marodin et al., 2019), its adaptation to quality-specific process areas within developing economies remains underexplored. This study fills that gap by identifying four key process areas, namely continuous improvement, benchmarking, standardisation, and error detection and prevention, and validating 22 associated performance indicators through expert consensus. The identification of these indicators provides a structured and contextually grounded extension to lean practices, which have traditionally emphasised tool adoption over measurable quality performance dimensions.

Practical Contribution

The study offers a practical framework that construction practitioners in Sri Lanka can use to assess and improve organisational quality management. The performance indicators provide measurable criteria for evaluating current practices, setting improvement targets, and benchmarking progress. This operational clarity enables both detailed process improvements at the project level and broader strategic alignment with organisational quality objectives. Furthermore, the framework helps construction firms address persistent challenges within the Sri Lankan industry, including low productivity, high levels of rework, and inconsistent quality control, by embedding lean principles into their quality assurance systems. Industry regulators and professional bodies may also utilise the indicators to develop lean-oriented guidelines and training content specifically adapted to the Sri Lankan construction context.

Methodological Contribution

This study demonstrates the structured application of the Delphi technique to develop a locally validated lean integrated quality management framework. Although Delphi is commonly used for consensus building, its phased use in this study across indicator identification, validation, contextualisation, and categorisation enabled an iterative yet systematic development process. This approach provides a replicable model for similar research in contexts where empirical data are limited, and expert input is essential.

Societal Contribution

This research promotes more efficient, reliable, and quality-focused construction practices in Sri Lanka. By embedding lean thinking into daily operations, the framework encourages a culture of continuous improvement and accountability. It supports long-term objectives related to capacity building and professional development, while also delivering benefits to clients, communities, and end users through more sustainable and socially responsive construction outcomes.

CONCLUSION

This study explored how lean construction practices can be strategically applied to enhance quality management within Sri Lankan construction organisations. Using a structured Delphi technique, four core process areas were identified and validated. These include continuous improvement, benchmarking, standardisation, and error detection and prevention. The process areas were supported by a set of twenty-two performance indicators, derived through a synthesis of established literature and expert judgement. This approach ensured both theoretical rigour and contextual relevance. The integration of these

components provides a coherent and practical framework for improving quality outcomes within lean-oriented construction environments.

The findings confirm that the validated process areas and indicators offer a reliable foundation for evaluating and enhancing quality management practices in the Sri Lankan construction industry. The final conclusion of the study is clearly established. The framework developed through this research can now be applied by industry practitioners and decision-makers to support systematic improvements in quality management. Future research may extend this work by testing the framework across a range of project types, organisational structures, and regional contexts, contributing further to the advancement of lean-based approaches to quality management.

REFERENCES.

[1] Anvari, A., Zulkifli, N., & Yusuff, R. M. (2013). A dynamic modelling approach to measure lean performance within lean attributes. *The International Journal of Advanced Manufacturing Technology, 67*(9–12), 2413–2430. https://doi.org/10.1007/s00170-012-4356-0

[2] Alsharef, A., Ovid, A., Jamil Uddin, S. M., & Albert, A. (2024). Biggest challenges facing the construction industry. *Construction Research Congress 2024, CRC 2024, 3*, 652–660. https://doi.org/10.1061/9780784485286.065

[3] Aslam, N., Yang, W., Saeed, R., & Ullah, F. (2024). Energy transition as a solution for energy security risk: Empirical evidence from BRI countries. Energy, 290, 130090. https://doi.org/10.1016/j.energy.2023.130090

[4] Aziz, R. F., & Hafez, S. M. (2013). Applying lean thinking in construction and performance improvement. *Alexandria Engineering Journal, 52*(4), 679–695.

[5] Ballard, G., & Howell, G. (2003). Lean project management. *Building Research & Information, 31*(2), 119–133. https://doi.org/10.1080/09613210301997

[6] Bazeley, P., & Jackson, K. (2013). *Qualitative data analysis with NVivo* (2nd ed.). SAGE Publications.

[7] Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

[8] Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining and measuring data saturation in qualitative research: A systematic review. *Journal of Clinical Epidemiology, 67*(8), 901–907. https://doi.org/10.1016/j.jclinepi.2014.01.017

[9] Dubois, A., & Gadde, L.-E. (2002). Systematic combining: An abductive approach to case research. *Journal of Business Research, 55*(7), 553–560. https://doi.org/10.1016/S0148-2963(00)00195-8

[10] Gartoumi, K. I., Aboussaleh, M., & Zaki, S. (2024). Implementing lean construction to improve quality and megaproject construction: A case study. *Journal of Financial Management of Property and Construction, 29*(1), 1–22. https://doi.org/10.1108/JFMPC-12-2022-0063

[11] Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*(4), 1008–1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

[12] Jayanetti, T., Perera, S., & Ranadewa, K. A. T. O. (2024). Development of lean maturity parameters for the Sri Lankan construction industry: A Delphi-based study. *International Journal of Construction Management*. (In Press)

[13] Koskela, L. (2002). *The theory of project management: Explanation to novel methods*. VTT Technical Research Centre of Finland.

[14] Liker, J. K. (2004). *The Toyota way: 14 management principles from the world's greatest manufacturer*. McGraw-Hill.

[15] Linstone, H. A., & Turoff, M. (Eds.). (2002). *The Delphi method: Techniques and applications*. Addison-Wesley.

[16] Love, P. E. D., Smith, J., & Davis, P. R. (2018). Rework in construction: A litmus test for quality. *Journal of Construction Engineering and Management, 144*(2), 04017100.

[17] Marhani, M. A., Jaapar, A., Bari, N. A. A., & Zawawi, M. (2012). Lean construction: Towards enhancing sustainable

construction in Malaysia. *Procedia - Social and Behavioral Sciences, 68*, 87–98.

[18] Marodin, G. A., Frank, A. G., Tortorella, G. L., & Saurin, T. A. (2016). Contextual factors and lean production implementation in the Brazilian automotive supply chain. *Supply Chain Management: An International Journal, 21*(4), 417–432. https://doi.org/10.1108/SCM-01-2016-0015

[19] Modig, N., & Åhlström, P. (2012). *This is lean: Resolving the efficiency paradox*. Rheologica Publishing.

[20] Nesensohn, C., Bryde, D., & Ochieng, E. (2015). Maturity and maturity models in lean construction. *Engineering, Construction and Architectural Management, 22*(1), 73–93.

[21] Ohno, T. (1988). *Toyota production system: Beyond large-scale production*. Productivity Press.

[22] Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management, 42*(1), 15–29. https://doi.org/10.1016/j.im.2003.11.002

[23] Parameswaran, A., Ranadewa, K., Rajini, P. A. D., & Jayanetti, J. (2024). State-of-the-art lean learning practices in construction: A Sri Lankan case study. In *Proceedings of the 12th World Construction Symposium* (pp. xx–xx). [Conference paper; in press].

[24] Ranadewa, K. A. T. O., Sandanayake, Y. G., & Siriwardena, M. (2018). Enabling lean among small and medium enterprise (SME) contractors in Sri Lanka. In *Proceedings of the 26th Annual Conference of the International Group for Lean Construction* (pp. 392–401). https://doi.org/10.24928/2018/0428

[25] Sacks, R., Koskela, L., Dave, B. A., & Owen, R. (2010). Interaction of lean and building information modeling in construction. *Journal of Construction Engineering and Management, 136*(9), 968–980. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000203

[26] Salem, O., Solomon, J., Genaidy, A., & Luegring, M. (2005). Site implementation and assessment of lean construction techniques. *Lean Construction Journal, 2*, 1–21.

[27] Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research methods for business students* (8th ed.). Pearson Education.

[28] Senaratne, S., & Wijesiri, D. (2008). Lean construction as a strategic option: Testing its suitability and acceptability in Sri Lanka. *Lean Construction Journal, 2008*(August), 34–48.

[29] Shingo, S. (1986). *Zero quality control: Source inspection and the poka-yoke system*. Productivity Press.

[30] Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research, 6*(1), 1–21. https://doi.org/10.28945/199

[31] Tezel, B. A., Koskela, L. J., & Aziz, Z. (2018). Lean thinking in the highways construction sector: Motivation, implementation and barriers. *Production Planning & Control, 29*(3), 247–269. https://doi.org/10.1080/09537287.2017.1412522

Multifactorial Drivers of Chronic Kidney Disease of Unknown Etiology (CKDu): A Review of Cadmium Exposure, Ultraviolet B B Radiation, and the Potential Role of Vitamin D Toxicity

R.T.K. Ariyawansha[1,a*], B.F.A. Basnayake[2,b], D.A.N. Dharmasena[2,c], S. Siribaddana[3,d]
[1]Department of Environmental Technology, Faculty of Technology, Sri Lanka Technology Campus, Meepe, Padukka 10500, Sri Lanka
[2]Department of Agricultural Engineering, Faculty of Agriculture, University of Peradeniya, Peradeniya 20400, Sri Lanka
[3]Department of Medicine, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka, Mihintale 50300, Sri Lanka
renukaa@sltc.ac.lk[a*], benb@pdn.ac.lk[b], dand@agri.pdn.ac.lk[c], sisira@med.rjt.ac.lk[d]

ABSTRACT

There are many research studies conducted to determine the cause/origin of CKDu for preventing this increasingly occurring disease, particularly among poor farmers. The disease is known to be multifactorial with heat stress enduring period > 3 months, but there is difficulty in distinguishing between harsher environments than endemic locations. This study identified that the likely cause is UVB actuating Vitamin D Toxicity (VDT), thus affecting the kidneys. Hence, a focused literature review was undertaken to find the links between cadmium (Cd), calcium, fluoride, enzymes, inhibitions, and the like. Moreover, 2nd law of thermodynamics was applied to determine the entropy differences between cold and hot source. The mean values of climate models were obtained from one publication on radiative forcing (RF) in the tropopause of 1.28 Wm^{-2} and climate feedback (CF) 0.25 $Wm^{-2}K^{-1}$. The energy of RF was used to determine entropy value $S_{RF(UVB)}$ at $T_2 = 230K$ as mean atmospheric temperature and the maximum temperature, T_1 at locations. It was then equated to the energy value of UVB, $Q_{CF_{(UVB)}}$ to be found between 1/273 and $1/T_1$. It was also validated using CF. The endemic location resulted 3.697 Wm^{-2} at 303K, and low RH compared to 3.239 Wm^{-2} at 311 K high RH. Although there is much comfort in endemic location, the chances of VDT or heat stress are higher more so with Cd inhibition of enzyme 7-dehydrocholesterol reductase (DHCR7), which is crucial for cholesterol synthesis. Instead, 7-dehydrocholesterol in excess switches more to form VDT, causing symptomatic hypercalcemia. Cadmium can disrupt vitamin D metabolism, contributing to osteomalacia and osteoporosis, actuating hypercalciuria, an indirect marker of low-level cadmium exposure. The kidneys, already compromised due to cadmium (Cd) accumulation and reabsorption during systemic distribution, ultimately eliminate Cd via the urine. Notably, no significant Cd accumulation is observed in end-stage renal tissues. Further basic research is required to elucidate the VDT in response to UVB exposure.

KEYWORDS: *Cadmium, Ckdu, Climate Feedback, Hypercalcemia, Vitamin D Toxicity (Vdt), Uvb, Radiative Forcing, 7-Dehydrocholesterol Reductase (Dhcr7)*

INTRODUCTION

Chronic kidney disease (CKD) is a global health problem with an increasing number of cases. CKD is defined as abnormalities of kidney structure or function, present for >3 months with health implications [1]. Hewageegana et al. (2024) [2] did undertake an extensive review on the unknown etiology of CKDu. They have reviewed papers on most aspects of possible causes and narrowed down to heat stress with multifactorial illnesses, but not conclusive because the neighbouring Northern Province in Sri Lanka is much harsher than the affected locations. Multifactorial illnesses may seem plausible with the number of publications on heavy metal contaminants, medical health issues like non-communicable diseases, social behaviour, and poor food intake [2-7].

Cadmium is one of the key toxicants causing CKDu as reported by [8]. According to the findings, cadmium is responsible for the mobility of calcium, thus causing hypercalcemia and hypercalciuria with excessive exposure to ultraviolet B radiation (UVB). Vitamin D Toxicity (VDT) is very rare occurrence [9], but it could be one of the triggering factors. It is a hypothesis, but it could very well be valid because there is excess UVB, particularly in the tropical belt with increased radiative forcing. The other logical argument is UVB, if in excess, will heat the skin surface very quickly, leading to accelerated perspiration and dehydration, which is a known factor affecting CKDu [10].The researchers in this study have been involved in climate modelling, and they are able to distinguish between a high incidence of UVB in the affected locations as opposed to harsher environments with less intensity in the neighbouring Northern Province. This study was conducted to conduct a targeted review of publications on cadmium to derive adequate evidence for in-depth research on the effect of UVB exposure on VDT or heat stress.

MATERIALS AND METHODS

A focused literature review was undertaken to determine the influence of gastrointestinal absorption of heavy metals (HM) on organ failure, specifically examining the interaction of cadmium and metallothioneins, cadmium, and calcium, fluoride, heat stress, VDT, switching between vitamin D production and cholesterol synthesis, and the mechanism of bone damage. Studies directly related to temperature were reviewed to determine the relationship to energy received from UVB radiation. Such that, the second law of thermodynamics was used to explain the energy levels received at different temperatures in dry and humid conditions.

By the Clausius definition, if an amount of heat (Q) flows into a large heat reservoir at temperature (T) above absolute zero, then the entropy (S) increase is:

$$\Delta S = Q/T. \quad (1)$$

Whenever, ΔS is the same between energies, Eq. (1) can be written as:

$$\frac{Q_1}{T_1} = \frac{Q_2}{T_2} \quad (1a)$$

And between two reservoirs of hot and cold, Eq. (1) can be expressed as:

$$\Delta S = Q\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \quad (2)$$

Where $T_1 > T_2$, [11, 12].

The radiative forcing (RF) values and climate feedback were obtained from literature [13], as given in Table 1.

Table 1. A comparison of radiative perturbations (Wm^{-2}) at the tropopause and the associated climate feedback ($Wm^{-2}K^{-1}$) [13]

Values of Climate Models	**Radiative Forcing Tropopause (Wm^{-2})**	**Climate Feedback ($Wm^{-2}K^{-1}$)**	**Climate Feedback * ($Wm^{-2}K^{-1}$)**
Mean	1.28	0.25	0.17
Std. dev	0.40	0.07	0.05
Min	0.68	0.16	0.12
Max	2.24	0.40	0.28

*Adjusted value given by [13]

Not all the radiative forcing is UVB because most of the entropy buildup is from the climate feedback of long-wave radiation. The direct reflectance from stratospheric water vapour in the lower stratosphere is responsible for UVB, and it can be calculated considering the average atmospheric temperature. According to [14], the average atmospheric temperature over time must have reduced from 254.6 K, but that reduction is not known. This phenomenon is reflected in several of the models developed over the years, since [14] tabulated NASA as 254.6, the Heat Transfer Model as 251.7 K, and the American Chemical Society Layer Model as 242.3 K. This reduction explains the movement of gas molecules upwards into the lower

stratosphere from the top of the troposphere within the ozone layer. We decided to use a value of 230 K that was obtained from unpublished works of the authors to calculate the plausible UVB received at 30 ºC and 38 ºC, representing Giradurukotte (endemic CKDu) and Northern Province. It was compared with 254.6 K, representing 1990 values. The minimum temperatures can be determined by considering Eq. (2). Such that the entropy of radiative forcing (S_{RF}) is:

$$S_{RF} = Q_{CF}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \text{ in } Wm^{-2}K^{-1}. \quad (3)$$

And Q_{CF} is the energy of climate feedback, since the entropy of climate feedback (S_{CF}) is:

$$S_{CF} = \frac{Q_{CF}}{T_1} \quad Wm^{-2}K^{-1}. \quad (3a)$$

T_2 is the minimum temperature and T_1 is the maximum temperature of a location at a given time. It can be postulated that Q_{RF} (energy of radiative forcing) being positioned in the Stratospheric Water Vapor (SWV) region, it will be at the assumed temperature of 230 K. The energy of S_{RF} is within S_{CF}, thus considering Eq. (3a), we can write that:

$$S_{RF} = \frac{Q_{RF}}{T_1} \text{ in } Wm^{-2}K^{-1}. \quad (4)$$

Combining Eq. (3) with Eq. (4), we get:

$$T_2 = \frac{Q_{CF}.T_1}{(Q_{RF}+Q_{CF})} \quad \text{in K} \quad (5)$$

In an activated state due to UVB radiation, the temperature T_1 will increase, depending on how cold the source temperature would be. In this instance, it is the assumed average atmospheric temperature (T_A),

T_A = 230K. Such that the average entropy of radiative forcing (S'_{RF}):

$$S'_{RF} = \frac{Q_{RF}}{T_A} \text{ in } Wm^{-2}K^{-1}. \quad (6)$$

It is at a higher entropy value since $S'_{RF} > S_{RF}$. We can write Eq. (3) at an elevated temperature, T'_1:

$$S'_{RF} = Q_{CF}\left(\frac{1}{T_2} - \frac{1}{T'_1}\right) \text{ in } Wm^{-2}K^{-1}. \quad (7)$$

Since T_2 was found from Eq. (5), $T'_1 = \frac{Q_{CF}}{\left(\frac{Q_{CF}}{T_2} - S'_{RF}\right)}$ in K. (8)

Therefore, the entropy of UVB can be derived from the difference between the activated maximum temperature and the maximum temperature at the location.

$$S_{RF(UVB)} = Q_{CF}\left(\frac{1}{T_1} - \frac{1}{T'_1}\right) \text{ in } Wm^{-2}K^{-1}. \quad (9)$$

It can be validated by considering only Q_{RF} with the low temperature of the average atmospheric temperature and the maximum temperature at the location. Hence, the entropy of radiative forcing in terms of UVB ($S_{RF(UVB)}$) can be written as:

$$S_{RF(UVB)} = Q_{RF}\left(\frac{1}{T_A} - \frac{1}{T_1}\right) \text{ in } Wm^{-2}K^{-1}. \quad (10)$$

Where, T_A is the average atmospheric temperature and T_1 is the ambient temperature at locations.

To determine the energy of UVB, $Q_{CF_{(UVB)}}$ at the location, the entropy of UVB should then be expressed in terms of the entropy of non-condensable climate feedback UVB component, stated as:

$$S_{RF(UVB)} = Q_{CF_{(UVB)}}\left(\frac{1}{273} - \frac{1}{T_1}\right) \text{ in } Wm^{-2}K^{-1}. \quad (11)$$

It leads to $Q_{CF_{(UVB)}} = \frac{S_{RF(UVB)}}{\left(\frac{1}{273} - \frac{1}{T_1}\right)}$ in Wm^{-2} (11a)

There are basically four cases, where $T_2 = 230K$ and $T_1 = 311$K and 303K.

And $T_2 = 254.6\ K$ and $T_1 = 311\ K\ and\ 303\ K$

Both Q_{RF} and Q_{CF} were modified to resemble radiative forcing and climate feedback values in 1990 by considering that,

$$Q_{RF2} = Q_{RF}.\frac{230}{254.6} \quad (12)$$

$$S_{CF2} = \frac{Q_{RF2}}{Q_{RF}} . S_{CF} \quad (13)$$

The implications of increasing energy levels over time are discussed, as well as high microbial activity. In lieu of the complexities surrounding the disease, the paper focuses on precautionary measures to reduce and prevent CKDu cases. Also, going beyond preventive steps, a research study is proposed to quantify the findings of this review.

Results anD Discussion

3.1 Bioelement Interactions in Cd Toxicity

Organometallic compounds can cross biological membranes more easily than inorganic heavy metal compounds because inorganics entry into living systems is often more limited by factors like solubility and bioavailability. Organometallics pose a higher risk of toxicity because of bioaccumulation due to their increased lipophilicity and potential for metabolic transformation into more toxic forms [15]. Trihalomethanes (THM) interfere with essential biometals like zinc, magnesium, selenium, calcium, and iron, altering their homeostasis and also impairing their biological functions [16-19], paving the way for heavy metal entry. Gastrointestinal absorption of Cd metal is determined in the diet by the content of these essential elements, along with vitamins, polyphenols, antioxidants, and other active biomolecules. The enhanced intake of some bioelements may prevent the absorption and the toxic effects of Cd, while the deficiency of some of these biologically active substances can increase gastrointestinal absorption and the accumulation of Cd in the body [17,19, 20 -22].

3.1.1 Cadmium and Metallothioneins

Genchi et al. (2020) [17] clearly explain the dysfunction of the kidney by cadmium. Metallothioneins (MTs) are ubiquitous low molecular weight proteins (MW 7–8 KDa). Their name is related to the occurrence of several divalent metal ions (Fe, Co, Ni, Cu, Zn, Ag, Cd, Hg, Pb) in the isolated material (metallo-) and the high –SH (cysteine) content of the protein (-thionein). Metallothionein was first identified in the kidney cortex of equines as a Cd-binding protein that is responsible for the accumulation of Cd in the tissues. The mammalian MT is made up of 61–68 amino acid residues, 18–23 of which are cysteines. MT is a unique protein that is very rich in thiol groups.

In the human body, high levels of this protein are synthesized in the liver and kidney, and their production depends on the presence of dietary minerals, such as Zn, Cu, and Se, as well as the amino acids histidine and cysteine. After absorption, Cd is delivered to the liver via endogenous intestinal MTs. Then hepatic Cd-MT gradually redistributes this metal to the kidneys, which are the main target organ for Cd toxicity. After being released from the liver, Cd bound to MTs is distributed to other organs. In the kidney, Cd undergoes glomerular filtration, from where it is reabsorbed intracellularly from renal tubular cells. In the latter, Cd is cleaved from MTs and Cd^{2+} ions are re-excreted into the tubular fluid; finally, Cd is eliminated in the urine. An acidic medium, such as that present in lysosomes (pH 4.5–5.5), promotes the dissociation of Cd from a Cd-MT complex. Cd-MT shows very low toxicity to fetuses and the central nervous system compared with other heavy metals because it does not easily cross the placental and hemato-encephalic barriers. Cd is mainly eliminated through the urine, even if the amount of Cd excreted daily in urine is very low. This low excretion corresponds to the Cd biological half-life of more than 25–30 years [17, 23-24].

High levels of cadmium and fluoride, and complex interactions between the ratio of dissolved sodium and calcium in the presence of fluoride have been hypothesized to cause CKDu in endemic areas of Sri Lanka [25]. Cadmium may also act directly on bone. Animal studies have shown cadmium to stimulate the formation and activity of osteoclasts, breaking down the collagen matrix in bone. Osteoporosis is the main cause of fractures in post-menopausal women, a common occurrence worldwide, giving rise to disability and a high cost to health services. In bones, cadmium can disrupt vitamin D metabolism, contributing to osteomalacia (softening of the bones) and osteoporosis (weakening of the bones) [26]. The identification of

cadmium, an environmental pollutant, as one causal factor is highly significant in helping to control the incidence of this complex condition [27].

3.2 UVB and VDT

Hewageegana et al. (2024) [2] have touched on an important issue: heat stress. In the text, they state that chronic occupational heat stress and dehydration are now thought to be the key etiologic factors for CKDu [10]. Chronic dehydration is a promoting factor for the development of crystal nephropathy [28]. In Central America, the main focus of attention for research in CKDu is dehydration and heat stress, but it has been paid less attention in Asia [29], including Sri Lanka. The possible pathophysiological mechanism of CKD in chronic dehydration includes sub-clinical rhabdomyolysis, hyperuricemia and hyperuricosuria, hyperosmolarity, and vasopressin effects [2]. Hard water in the CKDu prevailing area makes the water unpalatable [28], and as a result, people drink less, which causes dehydration in a vicious cycle. The dehydration hypothesis, however, cannot fully elucidate the low disease prevalence in the neighboring Northern province of Sri Lanka, where the conditions are warmer and drier than the endemic areas [30]. It is hypothesized in this study that UVB levels tend to be higher under low humidity like in the affected areas, than humid high-temperature locations referred to in the Northern Province. However, the relationship is complex; lower humidity generally allows more UVB radiation to reach the Earth's surface [31]. In Central America, the CKDu endemic is called Mesoamerican Nephropathy (MeN), and men working with heavy physical labor in warm areas near the Pacific Ocean often develop the disease.

Vitamin D is produced in the skin when 7-dehydrocholesterol is converted to cholecalciferol (vitamin D_3) by UVB radiation. Overexposure to vitamin D produces symptomatic hypercalcemia, an abnormally high level of calcium in the blood with possible weakness, fatigue, depression, confusion, stupor or coma, polyuria, nephrolithiasis, renal failure, ectopic calcification, conjunctivitis, fever, chills, anorexia, nausea, vomiting, and constipation [32]. Vitamin D Toxicity (VDT) normally encountered by excessive vitamin D intake. Overexposure to ultraviolet B (UVB) radiation is less common than toxicity from supplements [33].

Becklund et al. (2009) [33] explain that vitamin D toxicity and hypercalcemia do not typically occur upon exposure to sunlight because of several factors that limit the endogenous production of vitamin D. These factors include the photochemical conversion of previtamin D_3 into biologically inert compounds, skin pigmentation, and latitude. Thus, the levels of $1,25(OH)_2D_3$ required to suppress experimental autoimmune encephalomyelitis (EAE) are well above those that can be produced naturally upon exposure to sunlight. Furthermore, results from their laboratory suggest that hypercalcemia is more than simply an unfortunate consequence of $1,25(OH)_2D_3$ treatment and may play an essential role in the immunosuppressive effects of $1,25(OH)_2D_3$. It is notable that the conversion of 7-dehydrocholesterol to cholesterol, the final step of cholesterol synthesis in the Kandutsch-Russell pathway, is catalysed by the enzyme 7-dehydrocholesterol reductase (DHCR7). 7-dehydrocholesterol is also a precursor of vitamin D, switching between cholesterol synthesis [34]. Although not explicitly detailed, cadmium can inhibit the enzyme 7-dehydrocholesterol reductase (DHCR7), which is crucial for cholesterol synthesis. According to [35], inhibition leads to an accumulation of 7-dehydrocholesterol (7-DHC) and a reduction in cholesterol production. The altered sterol profile can have various consequences, including potential effects on cellular signalling, membrane structure, and susceptibility to oxidative stress. We can postulate that excess of 7-dehydrocholesterol will increase vitamin D production, leading to toxic levels. It is reported that in CKD patients, cholesterol synthesis can be affected, leading to dyslipidemia [36].

3.3 Mechanism of Bone Damage

Cadmium disrupts bone metabolism in many ways, with the worst appearing to be from impairing key kidney functions and directly damaging the osteoblasts. Vitamin D, whether from diet or production in the skin, must be first hydroxylated in the

liver and then undergo a second hydroxylation in the kidneys to produce its most active form, calcitriol ($1,25(OH)_2D_3$). Cadmium poisons the kidney enzyme that converts 25(OH)D to $1,25(OH)_2D_3$. Cadmium also damages the kidneys' ability to excrete toxins and to reabsorb important molecules and minerals. This results in increased loss of calcium in the urine. In fact, hypercalciuria is an indirect marker of low-level cadmium exposure. Cadmium inhibits the activity of alkaline phosphatase, which is produced by osteoblasts and is needed to deposit calcium into newly formed bone. Cadmium also decreases bone's collagen content by stimulating the formation and activity of osteoclasts, which break down the collagen matrix in bone as part of normal remodelling. Finally, cadmium activates gene expression of "toxic response" pathways in bone cells, which further stimulates osteoclast bone resorption [37]. Bone damage, specifically renal osteodystrophy, is a common complication in CKDu patients in Sri Lanka. This condition arises because the kidneys, when damaged, struggle to regulate mineral and bone metabolism, leading to various bone problems. In CKDu patients, bone pain and fractures are common symptoms, and the severity of bone damage can impact quality of life and mortality. A common symptom, often associated with the weakening of bones [38].

In a mathematical and biochemical point of view, both time and rate are involved in the progress of the multifactorial illness. Perhaps VDT triggers further malfunctioning of the kidney in a vicious cycle by hypercalcemia condition to activate even low levels of cadmium stored in the bones to release more calcium and even phosphorus. At latter stages of the patients, vitamin D levels become low but the effectiveness may become redundant. Evidently, Vitamin D treatment in CKDu patients involves supplementing with active vitamin D analogues or calcitriol, especially in later stages of CKD, to address secondary hyperparathyroidism. These analogues help regulate calcium and phosphate metabolism, reducing the risk of bone complications and cardiovascular issues. Nutritional vitamin D supplements are also used to address vitamin D deficiency and insufficiency in the general population with CKD [39].

CLIMATE CHANGE

A rise of the incidence of CKD was reported from the dry zone of the country in the 1990s [2] and the number of CKDu patients in Sri Lanka is increasing, particularly in the North Central Province. According to data from 2020, a total of 164,000 individuals were diagnosed with CKD. Furthermore, a comprehensive study indicates that approximately 10% of the population in Sri Lanka is afflicted by kidney-related ailments. There has been a notable increase in the burden of CKD in the country in the last two decades, with the identification of CKD cases among a population residing in the North Central Province, specifically in the Anuradhapura and Polonnaruwa districts. The occurrence of CKD in these regions varies between 5% and 15% [40].

Such an increase could be related to climate change because UVB radiation is increasing directly from Stratospheric Water Vapor in the ozone layer. The increase in activated water vapour is expressed in terms of radiative forcing of the atmospheric perturbations and radiated to the Earth in the form of climate feedbacks (CF). The increase in water vapour in the lower stratosphere causes elevated levels of UV radiation to the Earth otherwise be reflected from the ozone layer. A study conducted to compare atmospheric modelling of atmospheric perturbations resulting in climate feedback shows vast variations, but the mean values given in Table 2 by this research could be used to comprehend the differences between unaffected high-humidity locations to those of multifactorial regions with high incidence of CKDu.

The results of the study analysis given in Table 2 show that there is a difference in the lower maximum temperature location, like Giradurukotte to that of a higher temperature location in the Northern Province. The energy of UVB received at the lower maximum temperature is 14.16% more than in the higher maximum temperature and 8.70% higher in 1990. 17.27% increase can be expected if the maximum temperature rises to 40 ^oC, 313K. The values can be validated by applying Eq. (10).

Table 2. show the radiative force and climate feedback obtained from [13] and the modified values to simulate 1990 conditions at assumed average atmospheric temperatures and likely prevailed value in 1990 for location temperatures of

311K and 303K and the derived minimum and maximum temperatures, UVB entropy component of radiative forcing and UVB energy component of climate feedback at the two locations and hypothetical values of them in 1990

Q_{RF} (Wm^{-2})	S_{CF} ($Wm^{-2}K^{-1}$)	Q_{CF} (Wm^{-2})	T_A (K)	T_1 (K)	T_2 (K)	T_1' (K)	$S_{RF(UVB)}$ ($Wm^{-2}K^{-1}$)	$Q_{CF(UVB)}$ (Wm^{-2})
1.28	0.25	77.75	230	311	305.96	312.81	0.00145	3.239
1.28	0.25	75.75	230	303	297.97	304.63	0.00134	3.697
1.16	0.23	70.24	254.6	311	305.96	312.14	0.00082	1.840
1.16	0.23	68.43	254.6	303	297.97	303.98	0.00073	2.000

The climate feedback values for UVB were derived from the average model predictions; however, specific values for tropical locations can be much more than 3.697 Wm^{-2}. The extent of endemic locations is becoming more as reported by [2], because of the increasing radiative forcing and climate feedback year after year. The relationship between UVB radiation, humidity, and wind is complex. Generally, UVB levels are higher when humidity is low and wind is high. This is because low humidity allows more solar radiation, including UVB, to reach the Earth's surface, and high wind speeds can reduce the amount of water vapor in the atmosphere, further contributing to higher UVB levels. However, monsoon rains with high winds increase humidity in coastal locations compared to inland landmasses. UVB radiation intensity is generally higher in conditions of low humidity. This is because high humidity can lead to the formation of water droplets in the atmosphere, which scatter and reduce the amount of direct solar radiation reaching the surface, including UVB. In other words, the energy level of 3.239 Wm^{-2} given in Table 2 will reduce further with high humidity, giving normal values of UVB for absorption by the skin. Low humidity, therefore, allows more UVB radiation to reach the Earth's surface [31]. Farmers tend to work long hours under low humidity and windy conditions rather than high humidity and low wind conditions. Farmers' perceptions of climate change should be viewed to find better solutions, like conservation farming.

3.5 Synthesis in Developing Hypothesis

Metallothioneins play a protective role in neutralizing the toxicity of heavy metals. They prevent cellular damage. It is not an enzyme but can conjugate HMs. They transport HM into different parts of the body, thus leaving behind HMs in the dissociation reactions. Vitamin D is a nutrient the body needs, along with calcium, to build bones and keep them healthy. The body can absorb calcium only if it has enough vitamin D. Calcium is a major part of the bones. Cadmium can disrupt vitamin D metabolism in bones, thus causing hypercalcemia. Vitamin D also has many other uses in the body. It supports immune health and helps keep muscles and brain cells working.

In cholesterol synthesis, 7-dehydrocholesterol is indeed a key intermediate. When 7-dehydrocholesterol levels are high, the enzyme 7-dehydrocholesterol reductase (DHCR7) facilitates its conversion into cholesterol, thus continuing the cholesterol synthesis pathway [41] . Boland and Tatonetti (2016) [41] found that Cd and including a metabolite of arsenic, As inhibits enzyme (DHCR7). When there is inhibition of the enzyme (DHCR7), there will be a build-up of high level of 7-dehydrocholesterol, and in the presence of high UVB, there will be an excess of vitamin D. It is postulated that VDT conditions are of short durations because 7-dehydrocholesterol productions are suppressed over time. Individuals with inhibited or deficient enzyme DHCR7 will likely indicate low vitamin D levels as reported in Colombo and Kandy in Sri Lanka. It is expected that individuals with low vitamin D will get hydrated rapidly, causing oxidative stress. Under excessive exposure to UVB, individuals become feeble, experiencing both heat stress and VDT. Farmers become even more susceptible

to excess calcium in water, with catalytic reactions of fluoride causing rapid kidney failure.

CONCLUSION

The study distinguished the endemic region with higher UVB levels at lower temperatures than the adjoining region, signaling the possibility of VDT. It may be an additional influence in the multifactorial clinical conditions causing CKDu. The key factor being cadmium accumulation in the kidney and bones, which dysfunction the kidney and creates other health issues. Cadmium found in the urine of CKDu patients is a sign of excess levels in the body. In the presence of calcium or fluoride or both, cadmium in excess causes hypercalcemia and hypercalciuria in the affected patients. VDT also shows the same symptoms because VDT productions are more with less cholesterol synthesis. The worst-case scenario is when the manifestation of VDT with excess cadmium and calcium, catalyzed by fluoride, can lead to various adverse effects, including cellular damage and dysfunction of the kidneys. Most times, the poor health conditions of the farmers with weakened kidneys due to cadmium (Cd) exposure, exacerbated by UVB radiation, can lead to hypercalcemia and impaired cholesterol synthesis. Cd, even at low levels, can form complexes with calcium in the bones, disrupting calcium balance and potentially causing hypercalcemia. This, coupled with impaired kidney function, can lead to bone damage and reduced calcium levels, which are further compounded by a condition of low cholesterol synthesis. Basic research on UVB triggering VDT should be undertaken to prevent CKDu and find new management practices for affected patients.

REFERENCES

[1] S. Imbulana, K. Oguma, Groundwater as a potential cause of chronic kidney disease of unknown etiology (CKDu) in Sri Lanka: A review. Journal of Water and Health, 19(3) (2021) 393-410. DOI: 10.2166/wh.2021.079.

[2] T. Hewageegana, N. Rajakaruna, C. Jayasinghe, Chronic Kidney Disease of Unknown Origin in Sri Lanka: A Literature Review. IntechOpen. (2024). http://dx.doi.org/10.5772/intechopen.1004707.

[3] Y. Gelaye, Public health and economic burden of heavy metals in Ethiopia: Review. Heliyon. 10(19) (2024) e39022. doi: 10.1016/j.heliyon.2024.e39022.

[4] Z. Haidar, K. Fatema, S.S. Shoily, A.A. Sajib, Disease-associated metabolic pathways affected by heavy metals and metalloid. Toxicology Reports 10 (2023). 554-570. DOI: 10.1016/j.toxrep.2023.04.010.

[5] M.R.D.L. Kulathunga, M.A.A. Wijayawardena, R. Naidu, Dietary heavy metal(loid)s exposure and prevalence of chronic kidney disease of unknown aetiology (CKDu) in Sri Lanka. Environ Geochem Health 44(11)(2022) 3863-3874. doi: 10.1007/s10653-021-01144-1.

[6] R.A. Perera, R.T. Perera, U.P. Liyanage, J. Premaratne, J.A. Liyanage, Chemometric assessment of bioaccumulation and contamination pathways for toxic metals in diet and environment: implications for chronic kidney disease of unknown etiology (CKDu) in Sri Lankan agricultural regions. Environ Monit Assess 196(12)(2024) 1179. doi: 10.1007/s10661-024-13316-4.

[7] K.P. Wanigasuriya, R.J. Peiris-John, R. Wickremasinghe, Chronic kidney disease of unknown aetiology in Sri Lanka: Is cadmium a likely cause? BMC Nephrology 12(32)(2011). https://doi.org/10.1186/1471-2369-12-32.

[8] N. Jayatilake, S. Mendis, P. Maheepala, M.F.R. Jayatilake, Chronic kidney disease of uncertain etiology: prevalence and causative factors in a developing country. BMC Nephrol 14(180)(2013).

[9] S. Ruwanika, N. Gunawardena, C. Arambepola, Prevalence of low vitamin D status in an urban district in Sri Lanka: a population-based study. BMC Nutr. 10(115)(2024). https://doi.org/10.1186/s40795-024-00923-0.

[10] C. Jayasumana, C. Orantes, R. Herrera, M. Almaguer, L. Lopez, L.C. Silva, P. Ordunez, S. Siribaddana, S. Gunatilake, M.E. De Broe Chronic interstitial nephritis in agricultural communities: a worldwide epidemic with social, occupational and environmental determinants. Nephrol Dial Transplant. 32(2)(2017) 234-241. doi: 10.1093/ndt/gfw346.

[11] Y. Cengel, M. Boles, Thermodynamics: An Engineering Approach, 8th Edition, McGraw-Hill, 2015.

[12] G.W.F. Drake, "entropy". Encyclopedia Britannica (15, Feb. 2024). https://www.britannica.com/science/entropy-physics.

[13] A. Banerjee, G. Chiodo, M. Previdi, M. Ponater, A. J. Conley, L. M. Polvani, Stratospheric water vapor: an important climate feedback. Clim. Dyn. 53 (2019) 1697–1710.

[14] E. W. Beans, A heat transfer model for global warming. Case Studies in Thermal Engineering 43 (2023) 102677, https://doi.org/10.1016/j.csite.2022.102677

[15] ILO, Metals and organometallic compounds. Encyclopaedia of Occupational Health and Safety (2011). https://www.iloencyclopaedia.org/part-iv-66769/biological-monitoring-65407/item/404-metals-and-organometallic-compounds#:~:text=Following%20occupational%20exposure%20to%20inorganic,for%20the%20most%20part%20unchanged.

[16] M.E. El-Boshy , E.F. Risha, F.M. Abdelhamid, M.S. Mubarak, T.B. Hadda, Protective effects of selenium against cadmium-induced hematological disturbances, immunosuppressive, oxidative stress and hepatorenal damage in rats. Trace Elem. Med. Biol. 29 (2015) 104–110. doi: 10.1016/j.jtemb.2014.05.009.

[17] G. Genchi, M.S. Sinicropi, G. Lauria, A. Carocci, A. Catalano, The Effects of Cadmium Toxicity. Int J Environ Res Public Health. 17(11)(2020) 3782. doi: 10.3390/ijerph17113782.

[18] O. Jamakala, U. Rani, Mitigating role of zinc and iron against cadmium induced toxicity in liver and kidney of male albino rat: A study with reference to metallothionein quantification. Internat. Pharm Sci. 6 (2014) 411-417.

[19] Y. Wang, Y. Wu, K. Luo, Y. Liu, M. Zhou, S. Yans, H. Shi, Y. Cai, The protective effects of selenium on cadmium-induced oxidative stress and apoptosis via mitochondria pathway in mice Sidney. Food Chem. Toxicol. 58 (2013) 61-67. doi: 10.1016/j.fct.2013.04.013.

[20] M.M. Brzóska, S. Borowska, M. Tomczyk, Antioxidants as a potential preventive and therapeutic strategy for cadmium. Curr. Drug Targets, 17 (2016) 1350-1384. doi: 10.2174/1389450116666150506114336.

[21] Y.S. Lin, W.C. Ho, L. Caffrey, B. Sonawane, Low serum zinc is associated with elevated risk of cadmium nephrotoxicity. Environ. Res. 134 (2014) 33-38. doi: 10.1016/j.envres.2014.06.013.

[22] S. Satarug, M.R. Moore, Adverse health effects of chronic exposure to low-level cadmium in foodstuffs and cigarette smoke. Environ. Health Perspect. 112 (2004) 1099 - 1103. doi: 10.1289/ehp.6751.

[23] A. Bernard, Cadmium and its adverse effects on human health. Indian Med. Res. 128(2008) 557-564.

[24] G. Nordberg, K. Nogawa, M. Nordberg, L. Friberg, Cadmium. In: Nordberg G., Fowler B., Nordberb M., Friberg L., editors. Handbook of Toxicology of Metals. Academic Press; Cambridge, MA, USA, (2007) 445–486.

[25] S. Rajapakse, M.C. Shivanthan, M. Selvarajah, Chronic kidney disease of unknown etiology in Sri Lanka. Int. J. Occup. Environ. Health 22(3)(2016) 259-264. doi: 10.1080/10773525.2016.1203097.

[26] Y. Yang, M.F. Hassan, W. Ali, H. Zou, Z. Liu, Y. Ma, Effects of Cadmium Pollution on Human Health: A Narrative Review. Atmosphere 16 (2004) 225. https://doi.org/10.3390/atmos16020225.

[27] G. Kazantzis, Cadmium, osteoporosis and calcium metabolism. Biometals 17(5)(2004) 493-8. doi: 10.1023/b:biom.0000045727.76054.f3.

[28] S.J. Wimalawansa, C.B. Dissanayake Nanocrystal-induced chronic tubular-nephropathy in tropical countries: Diagnosis, mitigation, and eradication. European Journal of Medical Research 28 (2023) 221. DOI: 10.1186/s40001-023-01162-yut B

[29] G. Paidi, I. A.I. Jayarathna, D.B.A.M.R. Salibindla, J. Amirthalingam, K. Karpinska-Leydier, K. Alshowaikh, Chronic kidney disease of unknown origin: A mysterious epidemic. Cureus13(8) (2021) e17132. DOI: 10.7759/cureus.17132

[30] D.G.A. Paranagama, M.A. Bhuiyan, N. Jayasuriya, Factors associated with chronic kidney disease of unknown aetiology (CKDu) in north central province of Sri Lanka: A comparative analysis of drinking water samples. Applied Water Science 8 (2018) 151. DOI: 10.1007/s13201-018-0792-

[31] J. Liu, W. Zhang, The Influence of the Environment and Clothing on Human Exposure to Ultraviolet Light. PLoS ONE, 10(4) (2015) e0124758. doi:10.1371/journal.pone.012475.

[32] J.F. Marcus, S.M. Shalev, C.A. Harris, D.S. Goodin, S.A. Josephson, Severe Hypercalcemia Following Vitamin D Supplementation in a Patient With Multiple Sclerosis: A Note of Caution. Arch Neurol. 69(1) (2012) 129–132. doi:10.1001/archneurol.2011.1199.

[33] B.R. Becklund, K.S. Severson, S.V. Vang, H.F. DeLuca, UV radiation suppresses experimental autoimmune encephalomyelitis independent of vitamin D production. Proc. Natl. Acad. Sci. U.S.A. 107 (14) (2010) 6418-6423. https://doi.org/10.1073/pnas.1001119107.

[34] A.V. Prabhu, W. Luu, D. Li, L.J. Sharpe, A.J. Brown, DHCR7: A vital enzyme switch between cholesterol and vitamin D production, Progress in Lipid Research 64 (2016) 138-151. https://doi.org/10.1016/j.plipres.2016.09.003.

[35] N. Yamada, T. Karasawa, J. Ito, et al. Inhibition of 7-dehydrocholesterol reductase prevents hepatic ferroptosis under an active state of sterol synthesis. Nat. Commun. 15 (2024) 2195. https://doi.org/10.1038/s41467-024-46386-6.

[36] S.D.O. Lacount, L. R. Tannock, Dyslipidemia in Chronic Kidney Disease. Endotext [Internet]. National Library of Medicine**.** National Centre for Biotechnological Information (2025) https://www.ncbi.nlm.nih.gov/books/NBK305899/ .

[37] J. Pizzorno, L. Pizzorno, Environmental Toxins Are a Major Cause of Bone Loss. Integr. Med. (Encinitas). 20(1)(2021) 10-17.

[38] H.M. Abeywickrama, S. Wimalasiri, Y. Koyama, M. Uchiyama, U. Shimizu, N. Kakihara, R. Chandrajith, N. Nanayakkara, Quality of Life and Symptom Burden among Chronic Kidney Disease of Uncertain Etiology (CKDu) Patients in Girandurukotte, Sri Lanka. Int. J. Environ. Res. Public Health. 17(11) (2020) 4041. doi: 10.3390/ijerph17114041.

[39] M.L. Melamed, R.I. Thadhani Vitamin D therapy in chronic kidney disease and end stage renal disease. Clin. J. Am. Soc. Nephrol. 7(2)(2012) 358-65. doi: 10.2215/CJN.04040411.

[40] D.C.R. Weerakoon, E.P.E.D.Z. Siriwardana, J.M.K.B. Jayasekara, H.D.W.T. Damayanthi, T. Dorji, D.E. 3rd. Lucero-Prisno, Chronic kidney disease in Sri Lanka: Health systems challenges of patients on hemodialysis. Public Health Chall. 3(1)(2024) e155. doi: 10.1002/puh2.155.

[41] M. Boland, N. Tatonetti, Investigation of 7-dehydrocholesterol reductase pathway to elucidate off-target prenatal effects of pharmaceuticals: a systematic review. Pharmacogenomics J 16 (2016) 411–429 (2016). https://doi.org/10.1038/tpj.2016.48.

Developing A Web-Based Augmented Reality Tool For Promoting Sustainable Fashion Consumption

Fernando D.T. [a], Methma S.L.K.K. [a], De Silva R.K.J.[a*]

[a]*Department of Textile and Apparel Engineering, University of Moratuwa, Katubedda, 10390, Sri Lanka*

franandodt.21@uom.lk,methmaslkk.21@uom.lk, jayamalide@uom.lk*

ABSTRACT

This study explores the use of Augmented Reality (AR) as an intervention to promote sustainable fashion consumption among Sri Lankan consumers. It also involves developing a web-based AR tool designed to educate users about sustainable fashion and evaluating its acceptance within the target audience. This research identified that young consumers in Sri Lanka demonstrate a low level of awareness regarding sustainable fashion consumption. Therefore, this study addresses an important gap by analyzing how interactive digital tools can influence consumer education and promote green purchase behaviors. To gather requirements for developing the AR tool, a qualitative research method was employed through focus group discussions with 8 Gen Z participants representing diverse fashion preferences. For validation, the tool was further tested with 30 participants to evaluate usability, engagement, and effectiveness. Nine themes relevant to AR tool development were identified through thematic analysis, which highlighted the awareness about sustainable fashion, patterns of digital learning, and sensitivity to AR capabilities. This study revealed a strong consumer intention to engage with visually dense, socially sharable, and mobile-optimized AR applications. To address these points, a prototype web-based AR platform was designed using MyWebAR platform, alongside real time information on garment sustainability through the scanning of QR codes. The results demonstrate that AR can successfully engage consumers, increase awareness of the environmental footprint of fashion, and enable behavioral change when deployed on familiar social media platforms. The research concludes that interactive, culturally relevant AR tool experiences have strong potential to influence sustainable fashion practice among consumers.

KEYWORDS: *Augmented Reality Consumer Behavior; Digital Technologies; Fashion Consumption; Sustainable Fashion*

INTRODUCTION

Fashion is the second-largest polluting industry globally, producing about 92 million tonnes of textile waste each year (Alfie, 2025), most of which ends up in landfills. Consumers always seek new clothing, often influenced by the visual appearance of how garments are displayed on hangers (De Silva & Rupasinghe, 2018). Therefore, it is challenging to change the mindset of fast fashion consumers, as they have become accustomed to constantly pursuing the newest and most disposable garments. Moreover, the fast-growing fashion retail industry failed to achieve corresponding growth regarding sustainability due to a lack of awareness and engagement. Sustainable fashion consumption was described as less purchasing, choosing ethically manufactured clothing, and extending the life of clothing (Schiaroli et al., 2024). Innovation in educating and influencing the consumption pattern for sustainable fashion is thus needed for the tech-savvy consumer. Augmented Reality (AR) is used in retail to enhance product experience and provide more individualized consumer experience for prospective consumers. According to Rumokoy & Frank (2024), that AR is used in retail to improve product experience and enable individualize consumer experiences by aligning technological features with consumer needs with

increasing perceived value and engagement. However, its application within the fashion industry remains limited. Despite its potential, there is a dearth of studies exploring the opportunities that AR can offer for promoting sustainable fashion consumption. As sustainable fashion is still an emerging concept, the use of AR applications to drive behavioral change toward sustainability remains insufficiently explored.

Sri Lankan fashion consumers were chosen for this study due to accessibility and the rising popularity of AR among the younger generation. The identified gap in research on sustainable fashion consumption in Sri Lanka indicates a number of Sri Lankan consumers have only a limited understanding of the concepts involved in sustainability and are able to recognize major terms such as "green" and "environmentally friendly" but lack any deeper understanding of the sustainable consumption. This represents a knowledge gap that interferes with informed decision-making, since a large part of the population does not feel empowered to act on its positive attitudes about environmental protection. Furthermore, price and product quality factors turn out to be the dominating issues in purchase decisions when compared with sustainability attributes. According to Alhumaid (2020) digital technologies improve the teaching and learning process by offering dynamic, interactive, and student-centered learning experiences that typically lack in traditional methods. Digital technologies improving accessibility and efficacy in developing critical thinking and active engagement.

This research aims to explore the use of AR in stimulating customer awareness in sustainable fashion consumption, and subsequently, developed an AR tool and received responses to confirm the impact. There are three research objectives were addressed, namely, 1) to identify consumers' awareness of sustainable fashion and the use of AR technologies, 2) analyze the requirements for the AR tool, and formulate key constructs in developing the tool, 3) develop and measure the effectiveness of the proposed AR tool in establishing the awareness for sustainable fashion consumption.

LITERATURE REVIEW

Consumer Awareness in Sustainable Fashion Consumption

Connecting consumer awareness and engagement in sustainable fashion presents a major opportunity to encourage more responsible fashion consumption. Unfortunately, Sri Lankan consumers are yet unaware of sustainable fashion practices in the industry (Jimenez-Fernandez et al., 2023). This knowledge gap is critical because it creates limitations for the consumer to make sustainable fashion choices. However, Thilakarathne and Wijesundara (2017) the need for online platforms in shaping how consumers interact with the fashion industry, particularly in influencing their purchasing behavior toward sustainability. These platforms create opportunities for promoting sustainable choices and filling the most important awareness gap in relation to the environmental impacts associated with fast fashion.

The role of technology is increasingly important in enabling sustainable fashion consumption. Mobile applications and other digital tools have proven their value as effective dissemination channels. According to Muralikrishnan et al. (2024), there is a need to discuss the use of smart phone applications as facilitators in sustainable fashion; such applications may offer instant and accessible data about the sustainability aspects of fashion products and thus enable better choices by the consumer. This is particularly relevant in the Sri Lankan context, where ease of access translates into significant potential for awareness, particularly among young, digitally literate consumer.

To foster sustainable fashion, the suitable stimuli that shape consumer behavior and awareness must be understood. According to the published literature, it is evident that a combination of emotional and contextual stimuli is highly effective for sustainable decision making. Chen et al. (2024) recommends green information that is brief, engaging and educates consumers but does not overwhelm them. When such data are offered in interactive or graphic forms, it can increase affective

engagement and perceived value. Zollo (2024) also explains further those consumers usually experience cognitive dissonance if their consumption act is not congruent with their values of sustainability, and tension towards more mindful consumption is created when proper knowledge is offered.

Previous research has highlighted value-based communication, transparency in production processes, and storytelling elements such as ecological footprint narratives, as powerful psychological levers. Gee Gamage et al. (2024) found that secondhand clothing is more fashionable in Sri Lanka because it is not only more affordable but also aligns with eco-friendly values, providing both economic and ethical incentives that serve as potential awareness triggers. In addition, digital media platforms such as Instagram and TikTok support the stimuli based on peer influence and popular narratives. After Kurukula, 2022; Sandunima & Jayasuriya, 2024, influencer activities and user posts are facilitators towards making sustainability familiar and appealing. Thus, leveraging and incorporating those driving

Existing AR tools and their application in the Fashion Industry

Though Augmented reality (AR) technology is an emerging technology, it has been shown to enhance consumer engagement and interest in sustainable fashion worldwide. According to Fan et al. (2019), AR influences online retailing since shoppers' experience interactivity and immersive shopping that influences their attitude toward products. By enabling customers to visualize the product in the relevant environment, AR takes a more experiential format that can engage the attention of Sri Lankan consumers effectively. Besides, according to Gupta & Bhavsar (2024c), AR could lead to a better comprehension of environmental information, somewhat complex for consumers, and develop high intent for sustainable choices. AR's potential in this context is particularly appropriate to bridge the information gap that exists on sustainable fashion in Sri Lanka.

Thus, AR has been effective in consumer behavior change through improvement in the interactive experiences of consumers. This improves the consumer's comprehension and dedication to sustainability. Attri et al. (2024) highlighted that in-store AR builds positive behavior and improves consumer perceptions because of immersive experiences which attach consumers with the product on a deeper level. The more salient and thus relatable the benefits of sustainable fashion become with increased transparency, the more it may foster pro-environmental behavior. Interactivity in AR is crucial to drive buying behavior, specifically in online shopping, because it contains no chance for the consumer to directly interact with the product. Referencing interactivity and visual aesthetics of AR, Konstantoulaki et al. (2024) noted that both led to higher purchase intentions due to a more realistic and interactive product experience. Furthermore, Chen et al. (2024) mentioned "green information" and how it could affect the engagement of people in sustainability. For instance, an AR tool showing precisely how much carbon footprint a choice for a particular sustainable product has reduced will possibly motivate Sri Lankan consumers toward making purchases in an environmentally friendly manner.

The possibility of AR in visualizing sustainable practice and giving real-time feedback on consumer choices makes it an effective tool in the promotion of sustainable fashion consumption. According to Thamoda et al. (2024), Sri Lankan consumers with a preference for secondhand fashion may be more willing to buy sustainable products when clear information about their positive environmental impact is presented visually. Therefore, AR will effectively stimulate consumer behavior for sustainable fashion, however, an effective tool still needs to be developed. The application of AR in sustainable fashion education is one novel method of addressing the awareness gap identified above in this review. As indicated, Sri Lankan consumers are increasingly digitally connected, especially the Gen Z consumers, therefore there is a higher potential to adopt AR in their life style. According to Karagozlu (2021), AR improve learning environments by giving interactive experiences which help users to understand complex topics effectively. When applied this to sustainable fashion, users can get a better knowledge about the issues such as overconsumption, textile wastage, and the environmental impacts of fast fashion.

Abad-Segura et al. (2020) further observe that AR-based learning technologies are sustainable learning environments

in so far as they facilitate active participation and long-term memory retention. These technologies are scalable but flexible and thus adaptable to incorporation in digital platforms Sri Lankan consumers use on a regular basis. As discussed earlier, the use of AR in commercial environments can be directly applied to learning contexts, particularly where the goal is to drive behavior change through increased awareness.

Additionally, Ayer, Messner, and Anumba (2016) observed that AR can facilitate the development of thinking through allowing the user to experience virtual environments where users are directly eyewitness to consequences of design and consumption decisions. Such experience-based, gamified formats of learning are effective in creating sustainable behavior in young consumers and students. Adding AR to educational curricula can thus contribute a great deal towards enabling the pursuit of sustainable fashion in Sri Lanka through the visualization and sense-making of abstract impacts on the environment.

METHODOLOGY

A qualitative research method was adopted in the research design, primarily, Focus Group Discussions (FGDs) were conducted with 8 participants during the requirements gathering stage. A purposive sampling method was adopted to select participants within the age between 18-27 years, who are highly digitally literate and fashion-conscious. They are tech-savvy generation and, hence, are used to tools such as AR and are strongly responsive towards eco-friendly and sustainability measures. The selected participants represented a new generation of digital technology users, and thus exemplify AR's potential contribution to sustainable fashion (Kurukula, 2022; Sandunima & Jayasuriya, 2024). Participant's demographics were collected to tests AR tool suitability for relevance, inclusivity, and flexibility in different socio-cultural contexts and common usage contexts.

A semi-structured questionnaire was employed with stimulating of a discussion on areas of interest such as awareness and knowledge towards sustainable fashion by the participants, present consumption attitude towards fashion, and exposure to AR technology More precisely, this study examines the consumers' attitudes, perceptions, and motivations towards sustainable fashion and AR adoption in Sri Lanka and gathered constructs to develop an effective AR tool. Two rounds of FGDs were conducted during the month of May-June 2025, each session conducted approximately 60 to 75 minutes. All the sessions were audio-recorded and transcribed after the informed participant's consent to ensure accuracy and ethical handling of the data. FGDs are among the most effective research methods for behavioral and educational studies (G. Adekola & E.S. Olumati, 2023) enabling observations on perceptions of sustainable versus fast fashion, preferences for AR feature usage, and the psychological and social norms that influence fashion choices.

All the participants in the study gave their informed consent prior to their involvement. Their identities were totally kept confidential, and any personal details were anonymized to maintain privacy. The data collected was used solely for academic purposes. Participants were made aware of their right to withdraw from the study at any point.

DATA ANALYSIS AND RESULTS

Thematic analysis was used in data analysis to obtain main themes and sub-themes which allowed for the identification of key constructs for the development of the AR tool. Through this approach, key insights were identified, including expectations regarding sustainability-related content. Furthermore, preferences for presenting this information were recognized, including visually compelling, interactive, and easily accessible formats.

Participant 3 and 5 noted a strong inclination for desired technology to be presented in a digital format, delivering visual and interactive learning experiences. Many participants acknowledged that they lacked time to read in-depth articles but were willing to engage with digital content if it was engaging and bite sized. They also suggested gamified learning and AR overlays as attractive features. This shows that the learning process should be quick, intuitive, and visually driven matching current digital content consumption habits. Even though most of the participants did not have previous AR experience regarding fashion related platform, they were interested to experience them. Participants have preference to explore AR features that offered educational, stylistic, and informative value. This suggests that AR's potential goes beyond aesthetics and act as a powerful educational interface. Apart from that, irrespective of the knowledge inputs, participants wanted to engage with AR in meaningful ways. This is positive feedback of interactive visualizations and simulations that guide the user through pathways of decision. For instance, users could clearly visible the resource savings while using these tools. Most participants showed a strong willingness to try a web-based AR tool if it was accessible, easy to use, and relevant. Participant 4 remarked that, it needs to be mobile friendly and not too complicated. These answers point to the importance of developing a light and privacy solution. Participants also expressed app permission and technical performance issues. The participants mainly aged 18-24 gave a positive response to the AR tool. There were 7 out of 9 participants were willing to use similar tools in future. This indicates that minimalization and convenience are key factors in the acceptance of technology. Table 1 depicts the key findings and the expected features of the AR tool, which were extracted from the interview transcripts.

Table 1: Key Findings of the Focus Group Discussions

Key Findings	Expected Features in The AR Tool	Participant(s)
Awareness and Attitudes gap in Sustainable Fashion. Lack of awareness of what sustainable fashion entails; following global trends without understanding their long-term environmental effects.	AR storytelling of garment lifecycle (e.g., visual storytelling) Infographic overlays of environmental information on product scan.	1 & 2
Challenges to Sustainable Practices. Consumers value affordability and convenience over ethical fashion, citing price and accessibility as main barriers.	Cost vs. impact comparison feature (eco vs. fast fashion) Incentives for sustainable choices (points, badges, discounts)	1,2,3,4 & 5
Education Through Interactivity. Participants showed strong interest in interactive technology, especially when learning is visual and immersive.	Interactive educational modules ("Swipe to learn" quizzes) Clickable hotspot layers on garments (material info, ethical sourcing)	3, 5 & 7
Interest in AR Features. They tend to go to social media for fashion inspiration and selection, and it has a big influence on their shopping habits. The idea of refining fashion decisions through time was appealing to them, but they stressed the tool had to be natural and interactive.	Shareable AR experiences or "Style with Sustainability" filters Instagram/TikTok integration for social challenges or badges Progress dashboard (points, badges, tips for improvement) Weekly sustainability tips via notifications or pop-ups	1,2,3,5,6 & 8
Willingness to Use AR for Fashion. Others are keen on using new technology like AR, especially if it is fun and convenient.	Simple AR scan-to-learn functionality via QR code or clothing label Try-on simulations in real-time with sustainability scoring	4,5,6 & 7
Digital Engagement and learning Preferences. Youth in particular think that digital media can be used to modify behavior if combined with incentives and attractive content. People want good information at the point of purchase but don't wish to read lengthy pieces.	Level-up system for completing eco-learning pathways Animated AR avatars to guide users through lessons. Pop-up cards with bite-sized facts on origin, material, recyclability 5-second animated explainers triggered once a product is scanned	1,2,3,6 & 8

Development of the AR tool and Its Key Features

The participants in the focus group have expressed precisely that they prefer interactive, shorter and visual content instead of traditional educational methods as their main learning method. From gathered information about favorite learning formats, participants in the focus group have expressed high interest in AR simulation, short interactive videos, gamified quizzes formats. Also, moderate interest for infographics and low interest for long form text articles. According to the participants' opinions and the reference to credible academic research, the AR tool suggested can be executed as a website, which will incur no download for the users. Users will be able to make use of a product that is QR code scanned and, upon doing so, an AR real-time multimedia that further adds sustainability information of the product would appear to them. The features directly address participant preferences for visual, quick, and mobile learning. As Participant 3 stated, “If it’s on my phone and easy to use, I’d probably check it while shopping." Similarly, Participant 6 emphasized, "I didn’t know fashion

could use AR previously, but the sound is good, especially for trying things on virtually. “Social validation and peer influence thus become very important. The AR tool's social integration module would allow users to share their "impact stories" or completed sustainable challenges on popular social media platforms like Instagram or TikTok, which have the capacity to initiate a ripple effect within their digital communities, precipitating more sustainable behavior. The key proposed features, as affirmed by participants and literature, are listed below.

AR Prototype Development

Based on the findings obtained from the FGDs, AR prototype was developed using **MyWebAR**, including the user interface and configuration options. It reflects the environment in which the tool’s core structure was created.

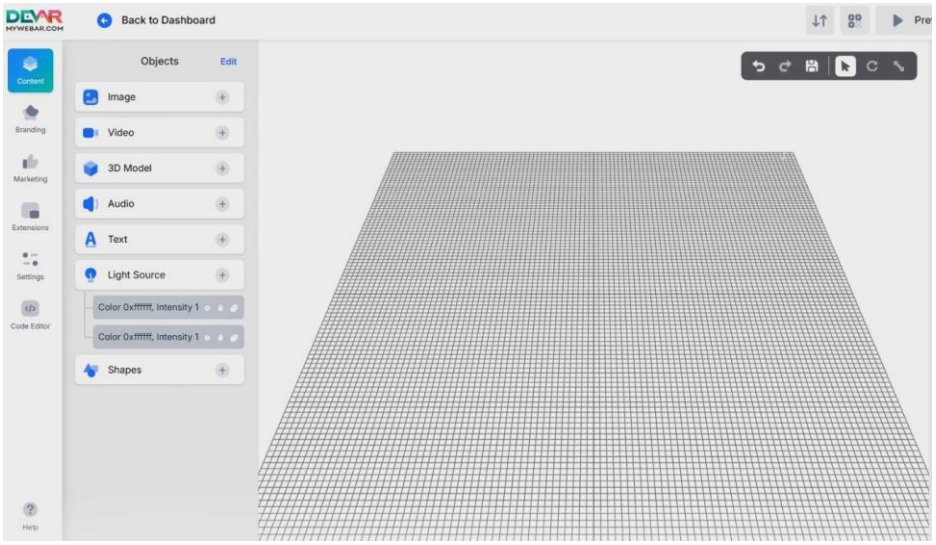


Figure 1: Initial Setup and Configuration in MyWebAR

The following image illustrates the second stage of the AR experience, where interactive features such as 3D models and other components are added.

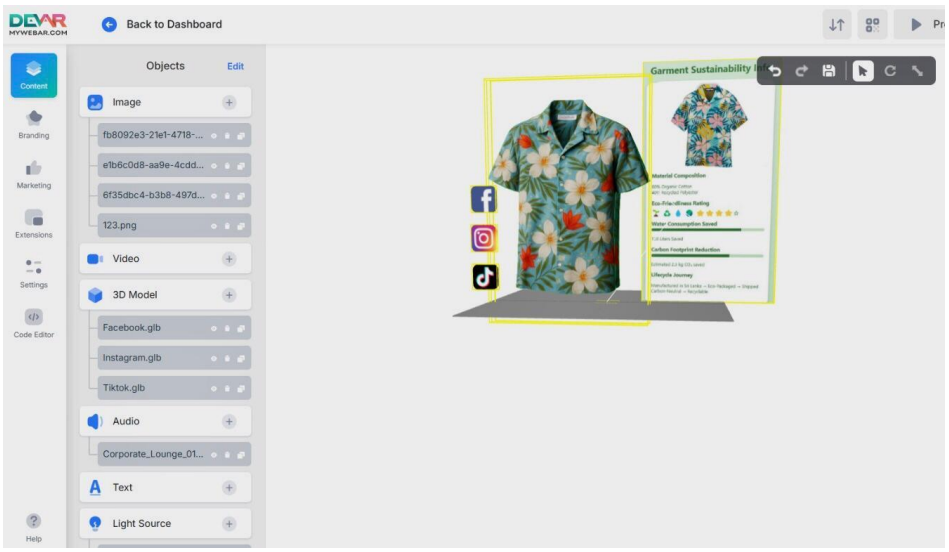


Figure 2: Adding Interactive Elements to the AR Tool

At this stage, the process of adding interactive features to the AR experience, such as a 3D realistic view of the garment and other components, was carried out. The interface is shown with the toolset for integrating these elements. Figure 3 shows the testing phase as the third step, where real-world interactions are used to test how well the AR tool functions in different scenarios.

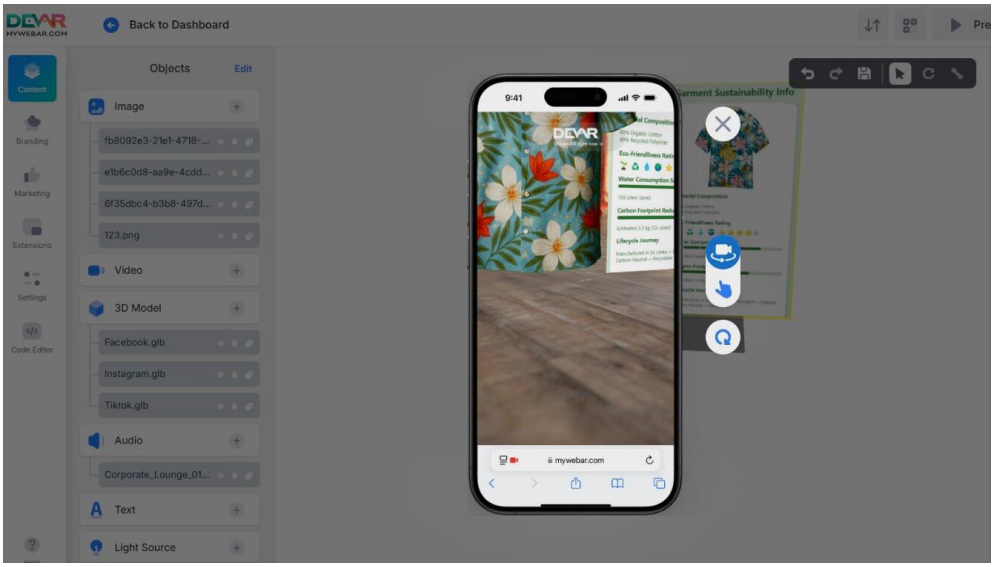


Figure 3: AR Tool Testing with User Input

These features are to fill gaps in transparency and awareness, facilitate aware and knowledge-based decision-making, and enable users for better understanding of sustainable and fast fashion ethics. They are to facilitate engagement through knowledge gamification, leverage peer influence through social sharing, and facilitate long-term use of the platform. The feedback loop also enables evidence-based refinement together with tracking long-term behavioral impact, thus making the tool not only educative but also effective and sustainable change-facilitating.

In the final stage, when the prototype was ready, the QR code and a questionnaire was distributed and feedback was collected from the 30 participants. Notably, 83% of respondents consider that the tool has the potential to impact sustainable

fashion purchase habits, although price, availability, and awareness are issues. The overall experience and likelihood to recommend the tool scored 3.7 and 3.9 out of 5, respectively, reflecting satisfactory user experience and peer recommendation potential. Suggestions for improvement focused on increasing local relevance, applying gamification, provision of Sinhala and Tamil languages, facilities for offline access, and expansion of facilities of social sharing. Although users enjoyed interactive learning and visualize facilities of the AR tool, internet speed, coverage of networks in the rural region, and more locally-oriented and peer-to-peer collaboration deficiencies suggest lines of further improvement and localization.

Prototype Display: Web-Based AR Interface

Figure 4: QR code for AR prototype

The AR prototype was developed for deployment via mobile and desktop web browsers, and consumers were able to use scan QR code scanner to view clothing in real space. While developing this initial prototype, it was focused on creating a clean, natural AR interface that can accurately simulate an in-store retail setting. The prototype developed using key elements such as the sustainability card, and social media logos around the digital apparel.

Using a functioning QR code tagged with the garment, users were allowed to start the web-based AR experience instantly and view product sustainable information in real time. This first prototype was meant to test spatial organization, visual presentation, and user interaction as a preliminary step toward building on to create a more intricate and scalable platform for learning about sustainable fashion.

DISCUSSION

A clear disconnect between consumer preference and actual behavior regarding sustainable fashion in Sri Lanka was identified through the key findings of this study. While participants discussed an awareness on topic such as eco-friendly practices and secondhand clothing, limited understanding was observed in areas like the environmental impact of the garment lifecycle. This reflects a wide global pattern noted in previous studies (Jimenez-Fernandez et al., 2023; Zollo, 2024), where individuals express a preference to make ethical choices but struggle to implement them effectively. To identify these issues, a web-based AR tool has been proposed an educational solution with offering an engaging accessible means of delivering sustainability knowledge. By combine the AR experience into shopping behavior and eliminating the consumer need for separate app installation, the barrier to user engagement is significantly reduced.

Respondents exhibited a fairly positive attitude towards sustainable fashion with understanding it as an ethical necessity and not a trend. Participants have quite understood about facing consequences due to climate change, and how fashion contributes to it. However, some participants acknowledged purchasing attractive and inexpensive clothing, despite experiencing guilt afterward. This suggests that while consumers may align with sustainability values, their ability to make ethical choices is often hindered by external factors such as price sensitivity, limited availability, and inadequate access to relevant information.

Participants identified several barriers that limit their engagement in sustainable fashion choices, with cost emerging as the most commonly cited obstacle. Some participants expressed a desire to purchase sustainable clothing but noted that such items are often expensive and difficult to find. Beyond affordability, concerns were also raised about limited product availability and a narrow range of styles. For instance, constraint affecting the practicality of making sustainable fashion decisions.

When the effectiveness of key AR features; the AR prototype that highlighted the garment's sustainability values was identified as one of the strongest features. The breakdown of material composition and the eco-friendliness star rating were liked by the participants. In addition, the participants were enabled to intuitively trace the flow of information through the quantitative environmental data provided. The social media logos (Facebook, Instagram, TikTok), which were displayed on the left side of the clothing, were seen as a smart touch. This feature was to be appealing to the behaviors of Gen Z and Millennials. It was suggested by the models that peer-to-peer sharing the visibility of sustainable products would be increased, as screenshots of the AR view were likely to be captured and shared by buyers on their own platforms. However, it was proposed by some participants that social media links should be shown only after user interaction (e.g. clicking a "Share" button) rather than being always displayed, in order to minimize digital clutter. When the technological and cultural suitability in Sri Lanka was considered, it was explained by respondents that the AR experience was culturally appropriate for Sri Lanka's tech environment. Although smartphones were found to be widely used, concerns were raised about data availability in rural areas. Therefore, a light, web-based AR (as opposed to app- based platforms) was recommended as the most suitable option for this market.

Participants, especially those from Generation Z, mentioned that sharing on social media could influence and encourage sustainable behaviors in their peer groups. The power of AR technology has been recognized for its potential to turn people into micro-ambassadors of sustainability through aesthetically appealing and accessible content. As a result, an impact towards ethical consumption was considered possible. As a next step, user-generated content (UGC) has been proposed, where people can post their reviews or stories about sustainability. This was done to add an increased level of interaction and authenticity.

CONCLUSION

A web-based AR tool was developed to raise awareness about sustainable fashion as a means of promoting behavioral change and serving as an educational tool. The requirements of the tool have been gathered and analyzed using focus group discussions, thematic analysis and iterative AR prototype development. The findings indicated that Sri Lankan youth are increasingly receptive to technology facilitated experiences, particularly those that are interactive, visual, and socially shareable. The AR prototype met these needs by allowing users to scan QR code in garment labels and access real time sustainability information, effectively embedding learning at the point of purchase. This user-friendly tool demonstrated strong potential to positively influence purchasing attitudes and help consumers make more informed choices. This research demonstrated that AR technology has the potential to enhance the fashion industry, emphasizing the need for meaningful content creation to maximize its effectiveness.

REFERENCES

[1] Alfie. (2025c, January 14). *Fashion Waste – 2025 Facts & Statistics*. Waste Management Services | Recycling | WasteManaged. https://www.wastemanaged.co.uk/our-news/retail/fashion-waste-facts-and-statistics/

[2] De Silva, R. K. J., Rupasinghe, T. D., & Apeagyei, P. (2018b). A collaborative apparel new product development process model using virtual reality and augmented reality technologies as enablers. *International Journal of Fashion Design Technology and Education*, *12*(1), 1–11. https://doi.org/10.1080/17543266.2018.1462858

[3] Schiaroli, V., Fraccascia, L., & Dangelico, R. M. (2024b). How can consumers behave sustainably in the fashion industry? A systematic literature review of determinants, drivers, and barriers across the consumption phases. *Journal of Cleaner Production*, *483*, 144232. https://doi.org/10.1016/j.jclepro.2024.144232

[4] Rumokoy, F. S., & Frank, B. (2024b). Retail value creation through augmented reality: The role of task-technology fit, consumer knowledge, and personality. *Journal of Retailing and Consumer Services*, *84*, 104173. https://doi.org/10.1016/j.jretconser.2024.104173

[5] Alhumaid, K. F. (2020b). Qualitative evaluation: Effectiveness of utilizing digital and social media in education. *Utopía Y Praxis Latinoamericana: Revista Internacional De Filosofía Iberoamericana Y Teoría Social*, *25*(6), 466–476. https://doi.org/10.5281/zenodo.3987663

[6] Jimenez-Fernandez, A., Aramendia-Muneta, M. E., & Alzate, M. (2023b). Consumers' awareness and attitudes in circular fashion. *Cleaner and Responsible Consumption*, *11*, 100144. https://doi.org/10.1016/j.clrc.2023.100144 [7] Wanniarachchi, T. M., Thilakarathne, B. L. S., & Wijesundara, B. (2017). Factors influencing purchasing intention of fashion clothing through online platforms. *Sri Lankan Journal of Marketing*, 3(2), 16–36. https://fcms.kln.ac.lk/dep/dmm/images/SLJM/3-2/T.M-Wanniarachchi-B.L.Sanjaya-Thilakarathne-Bimali-Wijesundara----Online-Publish-Article.pdf

[7] Muralikrishnan, N., Sehgal, N., & Kumar, T. (2024b). A smartphone application facilitating sustainable fashion. *Procedia CIRP*, *128*, 55–60. https://doi.org/10.1016/j.procir.2024.05.093

[8] Chen, L., Haider, M. J., & He, J. (2024b). Should "green information" be interactive? The influence of green information presentation on consumers' green participation behavior for driving sustainable consumption of fashion brands. *Journal of Cleaner Production*, *470*, 143329. https://doi.org/10.1016/j.jclepro.2024.143329

[9] Zollo, L. (2024b). Ethical identity, social image and sustainable fashion: still an impossible deal? A sociopsychological framework of ethical consumers' attitude-behavior gaps. *Journal of Consumer Marketing*, *41*(5), 564–582. https://doi.org/10.1108/jcm-08-2023-6213

[10] Geegamage, P. H. T., Ranaweera, H. R. A., & Halwatura, R. U. (2024). As Good as New: A Study on Sri Lanka Consumers' Secondhand Fashion Consumption. In *A study on Sri Lanka consumers' secondhand fashion consumption* (pp. 143–165). https://doi.org/10.1007/978-981-97-0569-6_6

[11] [12] Kurukula, U. (2022b). *Impact of social media influencers on sustainable fashion purchasing a study on buying behavior of gen Z in Sri Lanka*. http://dlib.iit.ac.lk/xmlui/handle/123456789/1526

[12] [13] Sandunima, K. C., & Jayasuriya, N. (2024b). Impact of firm-created and user-generated social media marketing on customers' purchase intention in the fashionwear industry: evidence from Sri Lanka. *South Asian Journal of Marketing*, *5*(1), 61–73. https://doi.org/10.1108/sajm-04-2023-0029

[13] [14] Fan, X., Chai, Z., Deng, N., & Dong, X. (2019b). Adoption of augmented reality in online retailing and consumers' product attitude: A cognitive perspective. *Journal of Retailing and Consumer Services*, *53*, 101986. https://doi.org/10.1016/j.jretconser.2019.101986

[14] Söderström, C., Mikalef, P., Landmark, A. D., & Gupta, S. (2024b). Augmented reality (AR) marketing and consumer responses: A study of cue-utilization and habituation. *Journal of Business Research*, *182*, 114813. https://doi.org/10.1016/j.jbusres.2024.114813

[15] Attri, R., Roy, S., & Choudhary, S. (2024b). In-store augmented reality experiences and its effect on consumer perceptions and behaviour. *Journal of Services Marketing*, *38*(7), 892–910. https://doi.org/10.1108/jsm-01-2024-0005

[16] Konstantoulaki, K., Rizomyliotis, I., Ang, E., & Quynh, N. T. (2024b). Augmented reality characteristics as drivers of consumers' purchase intention. *EuroMed Journal of Business*. https://doi.org/10.1108/emjb-04-2023-0114

[17] Karagozlu, D. (2021b). Creating a Sustainable Education Environment with Augmented Reality Technology. *Sustainability*, *13*(11), 5851. https://doi.org/10.3390/su13115851

[18] Abad-Segura, E., González-Zamar, M., La Luque-De La Rosa, A. L., & Cevallos, M. B. M. (2020b). Sustainability of Educational Technologies: An approach to Augmented Reality research. *Sustainability*, *12*(10), 4091. https://doi.org/10.3390/su12104091

[19] Ayer, S. K., Messner, J. I., & Anumba, C. J. (2016b). Augmented reality gaming in sustainable design education. *Journal of Architectural Engineering*, *22*(1). https://doi.org/10.1061/(asce)ae.1943-5568.0000195

[20] Adekola, G., & Olumati, E. (2023b). Focus group discussion: A Research Method in Community Development. *International Journal of Research and Innovation in Social Science*, *VII*(V), 392–399

Piezoelectric Energy Harvesting in Interactive Environments: A Case Study Using a Playpen-Based Prototype

H. M. U. J. Bandara[1], M. N. V. Fernando[2], H. D. K. Samaranayake[3]

[1,2,3]Faculty of Engineering, NSBM Green University, Sri Lanka
[1] hmujbandara@students.nsbm.ac.lk
[2] nipuni.f@nsbm.ac.lk
[3] hdksamaranayake@students.nsbm.ac.lk

ABSTRACT

As modern cities evolve towards sustainable and efficient infrastructure, implementation of micro-energy harnessing technologies caught the attention of researchers. This is mostly because non-renewable energy sources are being depleted, and the consequences of them are also marginal to the world. Piezoelectric materials are known for their conversion ability of mechanical stress, such as motion, vibrations, pressure, to electrical energy. This study provides a piezoelectric mechanism-based prototype that demonstrates the suitability of converting mechanical energy to electricity, especially footsteps from people, while being low maintenance and cleaner. In this design, 70 PZT piezoelectric disks are implemented beneath a playpen floor, suspended by a spring mechanism and further suspension components.

Output voltages of 9.2V to 11.4V could be observed from this demonstration, with the help of 2-5kg loads applied. 76 LEDs were used to visualize this power in real-time, which also offers interactive feedback for children. The output analysis shows a low-current and somewhat non-linear power output.

These results validate the possible outcome of deploying a piezoelectric system in high footfall infrastructure and recreational spaces. Even though it is low powered, such systems can enhance built environment functionality, and further improvements in the design can increase the output of such designs. This project represents a small-scale demonstration of this unpopular yet effective technology.

KEYWORDS: *Footstep Energy, Interactive Environments, Kinetic-To-Electrical Conversion, Low-Power Applications, Piezoelectric Energy Harnessing, PZT (Lead Zirconate Titanate), Smart Cities, Sustainable Infrastructure*

INTRODUCTION

The desire for renewable and sustainable energy sources has grown substantially in the 21st century, especially for smart cities. Due to the astonishing speeds of non-renewable energy, such as fossil fuel supplies are being depleted, the energy security in the world is considered in danger. According to the current forecasts, current fossil fuel reserves may significantly run out in the years ahead, if the current rates of consumption continue. Some estimate that Coal may run out in 139 years, Oil in 54 years and Petroleum gas in 49 years, from a 2020 study (METGroup Countries, 2021). This shows the importance of eco-friendly and alternative energy sources in the upcoming future. (Nations, n.d.) (Sorrell, Speirs, Bently, Brandt & Miller, 2010)

Energy extraction from non-renewable energy sources is relatively straightforward, but the negative effects and the inefficiency outweigh the advantages (United Nations, 2015). Therefore, it is crucial for engineers to find, study, develop and implement renewable and sustainable technologies. Even though rarely discussed, piezoelectric energy harnessing stands out among the recent technologies, as a possible solution to generate clean energy (Ghazanfarian, Mohammadi & Uchino, 2021). Piezoelectric materials are known for converting mechanical stress, such as vibrations, pressure, motion, into electrical energy (Rajapaksha, Senevirathne & Fernando, 2025) (Robeldo, 2023). This property is used in piezoelectric technology. The energy harnessed is ideal for low-power applications, such as sensors, LEDs, small electronic devices etc.

Continuous research and development would help to improve the efficiency of this technology by a considerable margin.

The goal of this study is to investigate the usefulness of piezoelectric energy harnessing. Frequent movement can be converted to electrical energy, by integrating this technology into dynamic spaces, such as playgrounds, walkways, roads etc. The aim is to evaluate the effectiveness and the potential of applying this technology to help the wider search for renewable and sustainable energy. (Rajapaksha et al, 2025) Our project utilizes a small-scale implementation example to demonstrate this idea.

The diagram below represents how mechanical stress applied on a piezoelectric material converts to electrical energy, and the losses involved with it (Ghazanfarian et al, 2021).

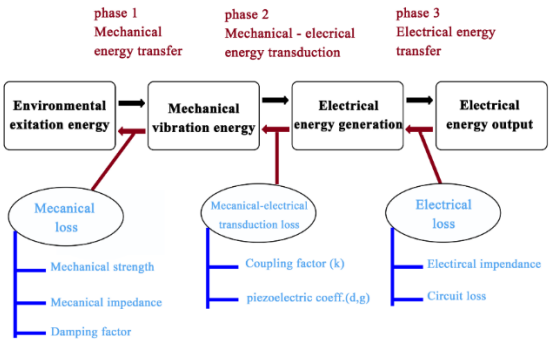


Figure 16 - Three major phases associated with piezoelectric energy harvesting. Retrieved from https://www.mdpi.com/2076-0825/10/12/312

In this demonstration, mechanical energy from footsteps is converted to electricity, with the use of piezoelectric disks beneath a rigid floor construction. When pressure is applied onto the disk, the crystal lattice of the piezoelectric material distorts, which creates an electrical imbalance, thus generating an alternating current between the two terminals.

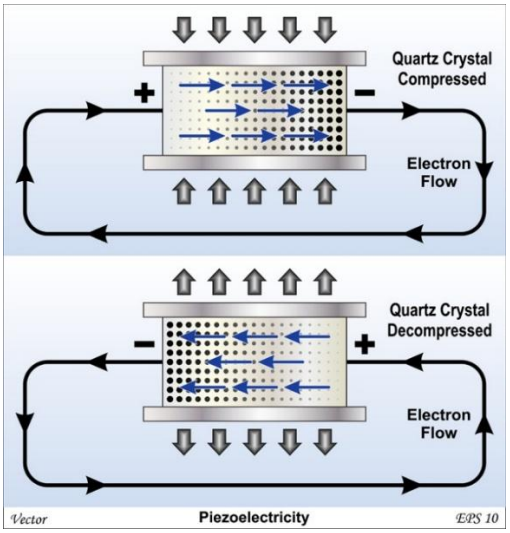


Figure 17 - How piezoelectricity works. Retrieved from www.autodesk.com/products/fusion-360/blog/piezoelectricity/

This current generated can be directly used to power small loads. Here, small LEDs are used as the load, which are then implemented to a play pen (Figure 3). The idea is to create an interactive and attractive environment for children, which also represents a small-scale and practical example of energy generation of this technology with human motion.



Figure 18 – Completed piezoelectric energy harvesting playpen prototype

MATERIALS AND METHODOLOGY

Materials

The main energy harnessing elements were 70 piezoelectric disks, which is constructed from PZT – Lead Zirconate Titanate (Rajapaksha et al, 2025) (Ghazanfarian et al, 2021) (Anton & Safaei, 2022) material, 3.5cm diameter. A metal washer served as a spacer under each disk, to allow some flexibility, and minimize the possibility of cracking under load. As the visual indicators of the generated current, 76 LEDs and 76 resistors (1kΩ) were used. The resistors act as current limiters, to avoid damaging the LEDs.

These piezoelectric disks were installed onto a 0.60m x 0.55m aluminium composite board, which served as the base structure. To spread the force applied evenly, a 0.53m x 0.50m plywood board was used, which acts as the upper flooring layer. PVC tubing, and 90-degree connectors were used, for the construction of the playpen frame, and these frames were covered by a net.

To provide suspension, to cushion the impact from the force applied and secure the floor and frame to the base, springs, hot glue stick pieces, cut to 1cm length and mounting hardware were used. Wiring was used according to the needs for the connection of the circuit. Additionally, for assembly and testing, soldering equipment and basic testing tools, such as a multimeter and an oscilloscope, were used respectively.

METHODOLOGY

The construction of the playpen was divided into three main stages: building the circuit, assembling the mechanical frame, and testing the system. One terminal of the 70 piezoelectric disks were joined into parallel groups and placed on the washers in the first stage. The positive terminal of every LED was soldered with the 1kΩ ohm and connected to the appropriate piezoelectric disk, while the negative terminals of the LEDs and the paralleled piezoelectric terminals are connected to a common ground. This arrangement allows the LEDs to be directly powered by the voltage generated from each footstep. Also, avoiding rectification allows the circuit to be further optimized for this setup, as LEDs do not suffer much from weak alternating current.

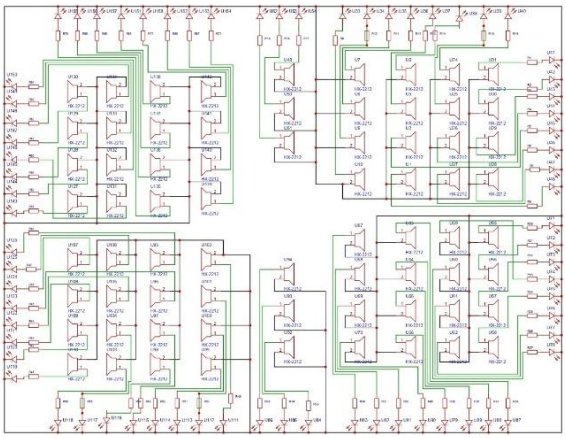


Figure 19 - Connections of piezoelectric disk to LEDs

The plywood board was placed over the disks, which were mounted on the aluminium composite base. The glue stick pieces, 1cm length, were positioned between the disks and the flooring. Springs were inserted between the layers, which serve as the suspension mechanism, further reducing the risk of damage (Figure 5) (Figure 6).

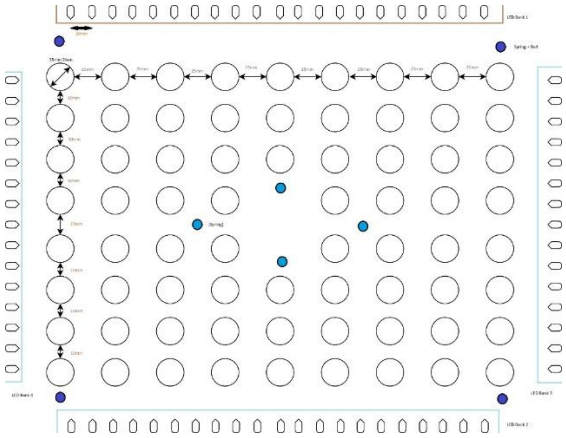


Figure 20 - Piezoelectric disk mounting layout on the composite board

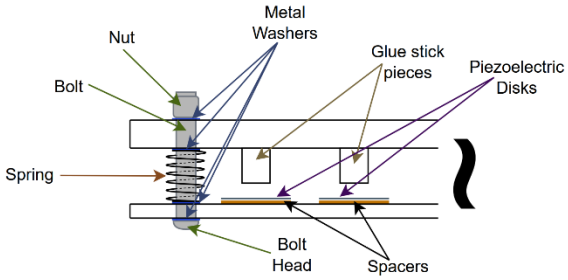


Figure 21 - Diagram of piezoelectric disk mounting (Side view)

PVC tubing was used to build the frame, which was mounted to the base with bolts and 90-degree PVC connectors (Figure 7).

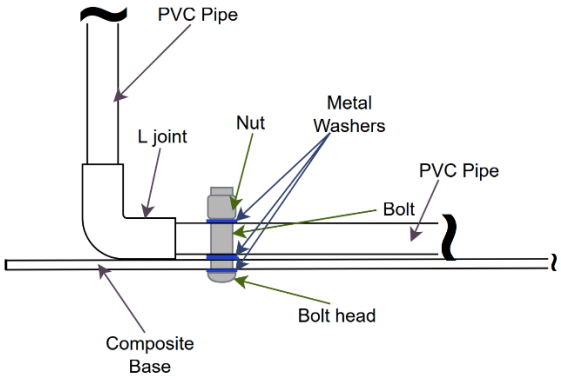


Figure 22 - Playpen frame diagram (Side view)

To simulate footstep pressure, flat objects weighing from 2kg to 5kg were implemented in the final stage. This is done by dropping the objects onto the plywood floor, from a height of ~0.15m. To analyze the output of the project, LED response and oscilloscope output were recorded. The results show reliable voltage generation and visual feedback from the interactive system. The complete project can be explained by the flow chart below (Figure 8).

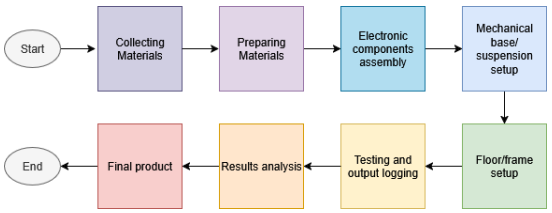


Figure 23 – Flowchart showing the step-by-step methodology followed in constructing and testing the piezoelectric playpen prototype

RESULTS AND DISCUSSION

Sample Collection

After assembling, the project was tested by simulating footsteps, by applying varying loads. Several piezoelectric disks were activated by physically applying load across various sections of the floor, weighing from 2kg to 5kg. Using a digital oscilloscope, the output voltages were recorded, and the LED behavior was recorded simultaneously.

Sample Analysis

Even with relatively smaller loads (~2kg), individual piezoelectric disks presented 9.2V peak-to-peak output on average (Figure 9).

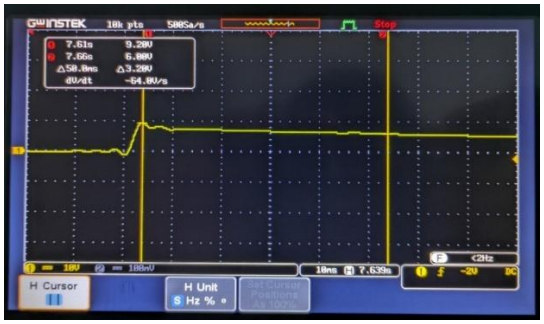


Figure 24 – Voltage spikes generated by piezoelectric discs during simulated footstep testing (oscilloscope output) - 2kg

This shows the mechanical coupling efficiency of this project prototype. High voltage spikes from the piezoelectric disks were verified by the oscilloscope graph, which suggested an effective electrical output from

the PZT disks. Though the mechanical dampening caused output gain to become less noticeable with higher loads, the voltage output remained somewhat consistent. Even though the output voltage was considerably large, the current remained low, which proves that powering low-power devices are possible with this mechanism.

RESULTS

Figure below represents the collected output data (Figure 10):

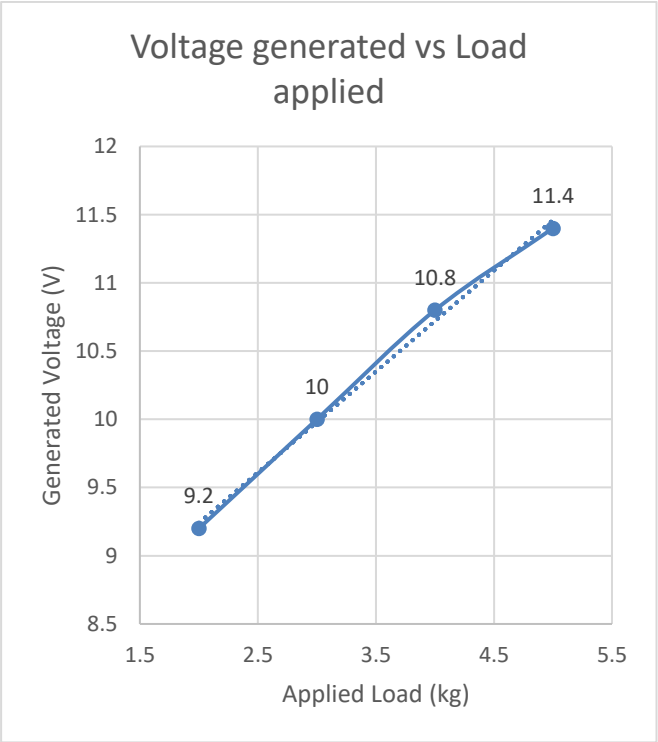


Figure 25 – Graph for Voltage Generated vs Load Applied

Below is an example of LED illumination under a simulated load (Figure 11):



Figure 26 – LED illumination under (simulated) applied footstep load

DISCUSSION

The experimental test findings verify that the piezoelectric energy harnessing mechanism operated with great efficiency, even under light loads. Notably, with 2-3kg load, the peak-to-peak output was ~10V.

The voltage spike pattern occurred during the compression and release of the piezoelectric disks is clearly visible in the oscilloscope output (Figure 9). From these voltage spikes, LEDs can be powered in real time, which also shows the outcome of the energy harvesting of this prototype. Due to energy distribution across multiple contact

points and mechanical saturation, the results showed minimal gains, even when higher loads were applied. Even then, the visible light intensity gradually increased with the increase of the load applied.

These results show that this design can provide useful electrical energy, with little mechanical interactions, such as a child's footsteps. This shows the potential for piezoelectric energy harnessing in crowded or low-impact environments, even though this system is limited to low-power applications since piezoelectric elements have inherently low current output.

In future, we hope to improve the mechanical design, such as suspension, coupling, and mounting, to accommodate more load efficiently, while providing exceptional dampening and minimal saturation to the piezoelectric disks. Also, implementing efficient rectification circuits and storage will help this technology to be used in a wide range of applications, such as energy supply for autonomous devices and public lighting. Further long-term testing and improvements will also help this technology to be used in day-to-day life, such as dancing floors, clothing, footwear, road infrastructure etc.

CONCLUSION

This study successfully demonstrates the feasibility of harnessing low-intensity mechanical motion, by piezoelectric mechanisms, and demonstrated using a playpen setup. The LED illumination gave real-time feedback, which effectively visualizes the energy conversion process for demonstrative and educational purposes. The results indicated a considerably reliable mechanical coupling between the floor and the piezoelectric disks and good efficiency under small loads. While the power output is enough only for low-power devices, due to the limited current generated by the PZT elements, the prototype promotes the potential for applying piezoelectric harnessing technology into interactive or public environments, specifically in areas with frequent human movement.

ACKNOWLEDGEMENTS

The authors would like to present their appreciation to everyone who supported this research. Specifically, our parents, lecturers, colleagues, and everyone who supported this research even with a little help. Also, special heartfelt gratitude to Ms. Nipuni Fernando, who guided us on the right path for this research, correcting us, always checking our progress and motivating us!

REFERENCES

[1] Countries, M. (2021, 01 18). *When will fossil fuels run out?* Retrieved from group.met.com: https://group.met.com/en/mind-the-fyouture/mindthefyouture/when-will-fossil-fuels-run-out/

[2] Nations, U. (n.d.). *Renewable energy – powering a safer future*. (United Nations) Retrieved 06 19, 2025, from un.org: https://www.un.org/en/climatechange/raising-ambition/renewable-energy

[3] Steve Sorrell, J. S. (2010, 09). *Global oil depletion: A review of the evidence.* Retrieved from sciencedirect.com: https://www.sciencedirect.com/science/article/abs/pii/S0301421510003204

[4] Nations, U. (2015, 12 23). *The Role of Fossil Fuels in a Sustainable Energy System*. Retrieved from un.org: https://www.un.org/en/chronicle/article/role-fossil-fuels-sustainable-energy-system

[5] Ghazanfarian J., Mohammadi M. M., Uchino K. (2021, 11 23). *Piezoelectric Energy Harvesting: A Systematic Review of Reviews.* Retrieved from mdpi.com: https://www.mdpi.com/2076-0825/10/12/312

[6] Rajapaksha R. M. A. D., Senevirathne P. K. V. M., Fernando M. N. V. (2025, 04 10). *Design and Optimization of Piezoelectric Energy Harvesting Systems for Efficient.* Retrieved from jrte.org: https://www.jrte.org/wp-content/uploads/2025/04/Piezoelectric_Energy_Harvesting_in_Smart_Pavements.pdf

[7] Robeldo, E. (2023, 02 12). *How Piezoelectricity Works to Make Crystals Conduct Electric Current* . Retrieved from autodesk.com: https://www.autodesk.com/products/fusion-360/blog/piezoelectricity/

[8] Anton, S. R., Safaei M. (2022). *Piezoelectric Energy Harvesting.* Retrieved from sciencedirect.com: https://www.sciencedirect.com/science/article/abs/pii/B9780128157329000449

Development Of An Ai-Based Model With Low Computational Complexity For Accurate Solar Energy Forecasting

Sudesh Chandrasinghe , Nimesha Fernando

Wayamba University of Sri Lanka
isurusudesh2000@gmail.com, nimeshaf@wyb.ac.lk

ABSTRACT

This paper introduces a short-term solar energy forecasting model that is designed with a focus on low computational complexity and addresses the challenges posed by fluctuations in solar energy generation, which are significantly influenced by environmental factors. These fluctuations can lead to instability when solar power generation systems are integrated into national energy grids, creating difficulties in maintaining a balanced supply and demand. If solar energy generation can be accurately forecasted before fluctuations occur, potential issues can be identified in advance, allowing for better management of the energy system, including optimizing storage facilities when energy generation is high. Current solar energy forecasting systems face significant challenges due to their high computational complexity, which results in increased power consumption and lower accuracy. To address these issues, this study focuses on the development of an artificial intelligence (AI)-based forecasting model using an Artificial Neural Network (ANN). The goal is to reduce the computational complexity of the model while maintaining high accuracy. To achieve this, various data analysis and complexity reduction techniques, such as variable reduction, pruning, and quantization, were applied. The performance of the optimized AI model was evaluated by comparing the forecasted values to actual solar energy generation data. The results demonstrate that the proposed model successfully reduces computational complexity while maintaining a satisfactory level of accuracy. This optimization makes the model more suitable for real-time forecasting, particularly in resource-constrained environments, and provides a more efficient approach to solar energy management. The findings of this study suggest that AI-based forecasting models can play a critical role in enhancing the integration of solar energy into national grids, ensuring a more reliable and sustainable energy supply. Further research could explore additional optimization techniques and the introduction of generalization techniques to improve transferability of the model and applicability across diverse geographical regions. Additionally, focus on utilizing AI techniques that minimize computational complexity without compromising the accuracy of the model, aiming to maintain high forecasting precision while optimizing the efficiency of the system.

KEYWORDS: *Solar forecasting, artificial intelligence, computational complexity, simulation, optimization, machine learning.*

INTRODUCTION

A significant amount of energy is utilized in daily lives, with conventional methods of energy generation, such as hydro power generation and the burning of fossil fuels, being commonly employed. However, as the demand for energy continues to rise, there has been a growing emphasis by researchers and other responsible entities on exploring alternative sources of energy. Among these, solar energy is particularly promising due to its inherent qualities it is a clean, eco-friendly, and renewable source of power. Its potential to provide sustainable energy solutions has garnered much attention as a viable

alternative to traditional energy generation methods.

The integration of solar power generation has led to increased recognition within the community of its numerous advantages, particularly its environmental sustainability and potential for long-term energy cost reduction. As awareness of these advantages grew, both small and large-scale investors became increasingly interested in contributing to the expansion of solar power infrastructure. Consequently, a significant number of solar power generation systems have now been integrated into the national grid. However, the management of these systems has become more challenging due to the fluctuations in solar power generation, which are influenced by varying environmental factors such as weather conditions. These fluctuations can complicate efforts to maintain a stable and reliable energy supply, posing a challenge to the smooth operation of the overall energy system(Chodakowska et al., 2024).

Forecasting solar power generation is of utmost importance for maintaining a well-balanced and reliable national energy system. Accurate forecasts are vital for ensuring that supply and demand are harmonized, storage management is optimized, and the overall system remains dependable. However, traditional forecasting methods, while effective in some respects, present certain challenges. These methods tend to be resource-intensive and may not always provide the level of precision required for efficient decision-making. They often demand significant computational resources, rely on complex machine learning models, and require the collection and management of large datasets. Moreover, real-time systems can struggle to accommodate such heavy models, leading to delays in processing and a time-consuming forecasting process. As a result, the accuracy of traditional methods may fall short, potentially affecting the effectiveness of the forecasts and the smooth functioning of the energy system.

Advancements in AI have significantly improved solar energy forecasting, offering solutions to overcome the limitations of traditional methods. AI models provide more accurate and reliable forecasts of solar energy generation, which enhances overall system performance. However, despite these advancements, AI models also face certain challenges. One major issue is computational complexity, especially when scaled for large datasets or deployed in environments with limited resources. Additionally, when operating in real-time systems, these models can become difficult to manage, as they require substantial computational power and can impact the efficiency of the forecasting process. These challenges highlight the need for further innovations to make AI-driven solar forecasting more efficient and applicable in various settings.

This research aims to address the challenges posed by datasets that require substantial computational resources, as well as other factors such as algorithmic complexity, data dimensionality, and system limitations, all of which contribute to increased computational complexity. Short-term forecasting (within a 0–48-hour timeframe) facilitates efficient grid scheduling, optimize solar plant dispatch, and enhances participation in energy markets. It will examine various AI techniques designed to reduce this complexity, thereby improving efficiency. The study will provide a comprehensive overview of the behavior of the different complexity-reduction variables, analyzing each one individually to better understand their impact. Furthermore, the research will explore strategies to achieve the highest level of accuracy in a complexity-reduced Neural Network model, focusing on how to optimize performance while minimizing computational demands.

LITERATURE REVIEW

Past approaches to solar forecasting have evolved from traditional statistical methods to more advanced machine learning (ML) and deep learning techniques. Traditional forecasting relied on manual calculations, empirical formulas, and simplified assumptions based on historical data and meteorological parameters. These methods were often time-consuming, error-prone, and limited in capturing the complex, non-linear relationships inherent in solar energy production. As solar power generation

depends heavily on fluctuating meteorological factors such as solar radiation, temperature, and cloud cover, forecasting became increasingly challenging(Chodakowska et al., 2024; Sammar et al., 2024). In response, machine learning methods, including neural networks (NN), support vector regression (SVR), and time-series models, emerged to enhance prediction accuracy(Amer et al., 2023). More recently, deep learning approaches such as long short-term memory (LSTM) networks, convolutional neural networks (CNN), and hybrid methods like extreme gradient boosting trees (XGBT) have been applied to improve solar irradiance forecasting. These advanced techniques offer more precise and reliable predictions by better capturing the dynamic behavior of solar energy generation in relation to changing environmental conditions.

Several models are commonly used in solar energy forecasting, leveraging the advancements in machine learning (ML) and deep learning to improve prediction accuracy. Among these, Artificial Neural Networks (ANN) are widely used due to their ability to capture complex, non-linear relationships between weather variables and solar energy production. Support Vector Machines (SVM), particularly Support Vector Regression (SVR), are also popular, as they can efficiently predict solar energy generation by identifying optimal patterns in the data. Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), are particularly effective for time-series forecasting, as they can learn sequential dependencies in data, making them ideal for predicting solar energy generation based on historical weather and irradiance data. Additionally, other deep learning models like Convolutional Neural Networks (CNN) and hybrid approaches such as Extreme Gradient Boosting Trees (XGBT) have been introduced to improve the accuracy and efficiency of solar irradiance predictions. These models offer a more reliable and precise forecasting method compared to traditional approaches, particularly in handling the complexity and volatility of solar energy generation.

The main challenge with existing solar forecasting and real-time prediction systems is their high computational demand. These systems, which often rely on complex models such as statistical methods, machine learning (ML), and deep learning techniques, require substantial computational resources. The intricate nature of these models, combined with the vast amount of data they process, makes it challenging to deploy them in resource-constrained environments. The need for significant processing power can result in delays and inefficiencies, especially in real-time applications, where timely predictions are crucial.(Hossain et al., 2023)

If the AI model is designed with low computational complexity, several significant benefits can be achieved. Data can be processed more rapidly, enabling faster real-time predictions. The consumption of resources and power can be minimized, leading to lower operational costs. The system's scalability is enhanced, allowing it to handle larger datasets or expanded infrastructure without difficulty. Latency can be reduced, ensuring quicker response times to new inputs. Additionally, cost efficiency is attained, as less computational power is required, reducing the need for expensive hardware and cloud services. The model can also be more easily adapted to changing requirements, making updates and modifications simpler and more efficient. Furthermore, small-scale generation stations, which may have limited resources, can benefit from the model's reduced complexity, as it allows for its deployment without the need for high-end infrastructure. These advantages demonstrate how a less complex model can make advanced forecasting systems more accessible, practical, and sustainable in various applications.

METHODOLOGY

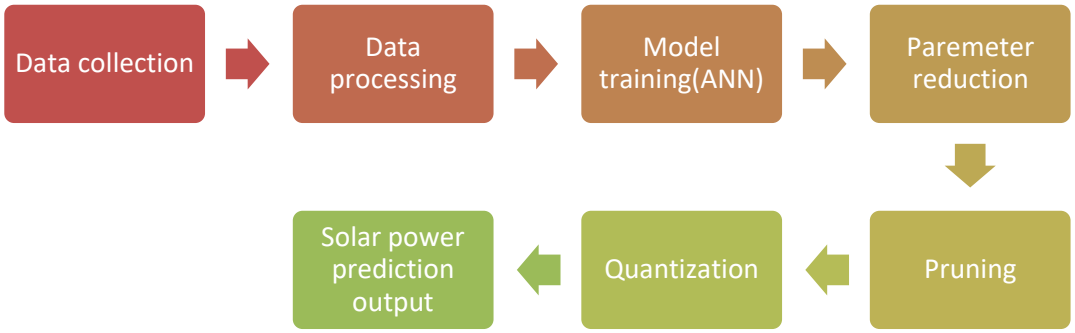


Figure 1. methodology

Data Collection

As the first stage of the development of the AI model, the data collection process is initiated. Throughout this research, the model is developed based on open-source datasets that are publicly available. For the initial dataset, a dataset containing data from the Northeast region of Brazil has been utilized. (*Solar Energy Northeast Brazil*, n.d.)This dataset includes around 4,000 data points collected over a period of five months, with a 5-minute interval between each data point.

After the collection of solar energy data, meteorological data that affects solar energy generation was identified. Meteorological factors such as weather patterns, wind speed and direction, humidity, solar radiation, ambient temperature, and module temperature, among others, have been recognized as influential through a review of past research(Amer et al., 2023). To gather the necessary weather data, efforts were made to collect this information using an open-source platform. In the initial stage, eight input variables were considered for analysis. The Historical Weather API was used to obtain the relevant data for these meteorological factors(🏛 *Historical Weather API | Open-Meteo.Com*, n.d.).

Due to the limited number of data points in the initial dataset, which could potentially affect the accuracy of the AI model and lead to improper training, a decision was made to transition to a different dataset. This new dataset was identified through research conducted on data collection in China. The dataset comprises approximately 70,000 data points, spanning two years (2019-2021), with a 15-minute interval between each data point (Chen & Xu, 2022)

The dataset effectively captures the solar power fluctuations of a 50MW solar station, providing a more comprehensive record of solar power generation. It includes six key input variables, Total Solar Irradiance (W/m²), Direct Normal Irradiance (W/m²), Global Horizontal Irradiance (W/m²), Air Temperature (°C), Atmospheric Pressure (hPa), and Relative Humidity (%). (*GitHub - Bob05757/Renewable-Energy-Generation-Input-Feature-Variables-Analysis*, n.d.)

Simulation Platform

The Artificial Neural Network (ANN) is a widely recognized and powerful method used to find solutions, particularly when other statistical techniques may not be applicable. Its advantages, such as the ability to learn from examples, operate in real-time environments, and forecast non-linear data, make it an invaluable statistical tool. One of the key strengths of ANN is its capacity to model complex, non-linear relationships between known input and output values.(Amer et al., 2023)

Additionally, ANN is well-suited for handling large datasets, making it particularly useful when working with vast amounts of data points and complex datasets.(Abuella & Chowdhury, n.d., 2015) This ability to process large quantities of data effectively sets it apart from linear models that may struggle with high-dimensional data or non-linear trends.(Amer et al., 2023)

MATLAB was chosen as the development environment for implementing the ANN model. MATLAB's Neural Network Fitting feature was employed to create and train the model, using 80% of the data for training and 20% for testing. This feature provides a comprehensive and efficient framework for building neural networks, allowing for accurate fitting of the data and effective prediction of solar energy generation. By utilizing ANN, the model is able to handle the intricacies of the data, including non-linear relationships, making it a powerful tool for accurate forecasting in this context.(*Deep Learning Toolbox - MATLAB*, n.d.)

Optimization methods

Before the first model was prepared, the dataset was arranged on an hourly basis, and data points with unexpected values were removed, provided the input variables were in optimal conditions. Eighty percent of the arranged dataset was used for training the model, while the remaining 20% was reserved for testing. However, after the model was trained on this dataset, an accuracy of only 30% was achieved. This low accuracy was attributed to the insufficient data and the inaccurate training of the model due to the limited dataset available.

To improve the dataset and accuracy, a decision was made to shift to the second dataset, which had been sourced from research conducted in China. This dataset, covering one year of data (2019-2020), was processed by removing zero output data when the input conditions were normal. The model's accuracy improved to 58% after this larger dataset was used. Subsequently, after the data was arranged on an hourly basis and the model was refined, an accuracy of 84% was achieved. This improvement in accuracy was attributed to the more comprehensive and reliable dataset, which allowed for better training and more accurate predictions.

As the next step in reducing computational complexity and improving accuracy, input variables with minimal correlation to the output were removed. The second model initially included six input variables: Total Solar Irradiance (W/m²), Direct Normal Irradiance (W/m²), Global Horizontal Irradiance (W/m²), Air Temperature (°C), Atmospheric Pressure (hPa), and Relative Humidity (%). Upon evaluating the correlation between these variables and the output, it was found that four of them exhibited a considerable correlation with the output, while two of them showed very little correlation and accuracy.

To improve the model's performance, the two variables with minimal correlation were removed, and the model was retrained. After this adjustment, the model's accuracy increased to 90.3%. This enhancement in accuracy was attributed to the removal of less relevant variables.

Table 20. Correlation ranges (Ratner, 2009)

Correlation	**Range**
Strong positive correlation	0.7 to 1
Moderate positive correlation	0.3 to 0.7
Weak or negligible correlation	-0.3 to 0.3
Moderate negative correlation	-0.7 to -0.3
Strong negative correlation	-1.0 to - 0.7

Table 21 . Variable and correlation

Input Variable	**Correlation**
Total Solar Irradiance (W/m²)	94.6%
Global Horizontal Irradiance (W/m²)	64.7%
Direct Normal Irradiance (W/m²)	57.8%
Air Temperature (°C)	22.6%
Relative Humidity (%)	18%
Atmospheric Pressure (hPa)	-0.8%

As the third method of complexity reduction, Pruning, or sparsity induction, was applied.(Fernandes & Yen, 2021) . The pruning threshold of 0.3 was selected after testing various values ranging from 0.1 to 0.5. This value provided the optimal balance between reducing network size and maintaining prediction accuracy(Molchanov et al., n.d.). This model optimization technique involves removing 15-25% of the smallest (near-zero) weights from a neural network. The aim was to enhance computational efficiency while maintaining accuracy. After applying this technique to the MATLAB model, the accuracy was slightly reduced to 89.7%. The slight decrease in accuracy was attributed to the impact of the complexity reduction technique, as some weights were removed to improve efficiency.

To measure the effectiveness of the complexity reduction, several parameters were used, Model Size (MB), Inference Time (ms), Total Time (s), RMSE (Root Mean Squared Error), and the Training Epoch. These parameters provided a comprehensive evaluation of the trade-off between reducing model complexity and preserving performance, ensuring that the model's accuracy remained acceptable while achieving improvements in computational efficiency (Pham et al., 2019).

The final method applied for computational complexity reduction was Quantization. This technique was used to reduce the computational complexity and memory usage of the model by approximating the weights and activations with lower precision. Instead of using 32-bit or 64-bit floating-point numbers to represent these values, quantization utilized fewer bits, such as 8-bit integers, to represent the model's parameters(Im & Dasari, 2022).

After applying quantization, the model's accuracy slightly decreased to 89.6% with respect to the actual values. During the research, it was observed that while the application of complexity reduction techniques consistently reduced model complexity, it also led to a reduction in accuracy. Given this trade-off, it was concluded that no further complexity reduction techniques would be applied beyond this stage, as the accuracy began to suffer with each additional reduction method. Thus, the model reached an optimal balance between reduced complexity and acceptable accuracy at this point.

COMPLEXITY EVALUATION MATRIXES

The level of complexity and the reduction in computational complexity were measured using several parameters of the model, including Model Size (MB), Inference Time (ms), Total Time (s), RMSE (Root Mean Squared Error), and the Training Epoch. When these values were high, it was understood that the computational complexity was also higher. It was important to maintain the Training Epoch at a moderate level to avoid excessive training time and overfitting. By observing the changes in these values, insights were gained into the reduction of complexity in the model and the behavior of each parameter with the application of different techniques. Each technique's impact on these values provided a clear understanding of the model's complexity was being reduced while balancing accuracy and performance.

Table 22 . Initial and final model parameters

Metrics	Initial model	Final model
Parameters	8	4
Model size (MB)	0.06	0.04
Inference Time(ms)	15.6	7.75
Total time (s)	19.718	5.23
RMSE	9.44	4.123
Training Epoch	330	18
Model accuracy(R^2)	30.3 %	89.6%

EVALUATION METRICS FOR PERFORMANCE

To measure the performance of the model, the RMSE (Root Mean Squared Error) value was calculated by comparing 20% of the actual values from the test dataset with the predicted values generated by the model. This comparison provided a quantitative assessment of how accurately the model was able to predict the data. A lower RMSE value indicated that the model's predictions were closer to the actual values, reflecting better performance, while a higher RMSE suggested a larger discrepancy between the predicted and actual values, signaling a need for improvement in the model's accuracy.(Pham et al., 2019)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (1)$$

R-squared (R^2), also referred to as the coefficient of determination, was understood to be a measure of the proportion of variance in the dependent variable that could be predicted from the independent variables in the model. After the model had generated the output, the deviation between the actual and predicted values was examined. It was expected that the R^2 value should be close to 1, indicating a high level of accuracy. When the R^2 value approached 1, it suggested that the model had explained a significant portion of the variance in the data, leading to better predictive performance. The closer this value was to 1, the more reliable the model was considered to be in terms of its accuracy in predicting the dependent variable.(Pham et al., 2019)

$$R^2 = 1 - \frac{\sum_{i-1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i-1}^{n}(y_i - \bar{y})^2} \quad (2)$$

RESULTS

Model Performance

The accuracy of the model initially began at 30% and was improved to 90.3% due to the enhancement in the number of data points and the data analysis techniques that had been discussed previously. Subsequently, the accuracy slightly decreased from 90.3% to 89.6%, which was attributed to the application of computational complexity reduction techniques, such as pruning and quantization. However, the demand for computational resources was notably reduced as a result of these techniques.

Table 23. Model Performance

Metrics	**Initial model**	**Variable Reduction**	**Pruning**	**Quantization**
Parameters	8	4	4	4
Number of Layers	3	3	3	3
Model size (MB)	0.06	0.04	0.04	0.04
Inference Time(ms)	15.6	11.34	8.16	7.75
Total time (s)	19.718	9.42	5.61	5.23
RMSE	9.44	5.42	4.20	4.123
Training Epoch	330	43	21	18
Model accuracy(R^2)	30.3%	90.3 %	89.7%	89.6%

Computational Metrics

The time required for the full model training, measured in seconds, was reduced from 19.7 to 5.25. Additionally, the time for generating a single output with a single input was also decreased from 15.6 milliseconds to 7.75 milliseconds. The model size was reduced from 0.06MB to 0.04MB, and the number of variables was halved, leading to a corresponding reduction in input variables. These improvements collectively demonstrate a significant reduction in computational complexity, showcasing the effectiveness of the optimization techniques employed.

Forecasting Accuracy Visualization

The forecasting accuracy was measured using the R-squared (R^2) value, which improved significantly from 30.3% to 89.6%, reflecting the model's enhanced predictive performance in the final result. (Abuella & Chowdhury, n.d., 2015)

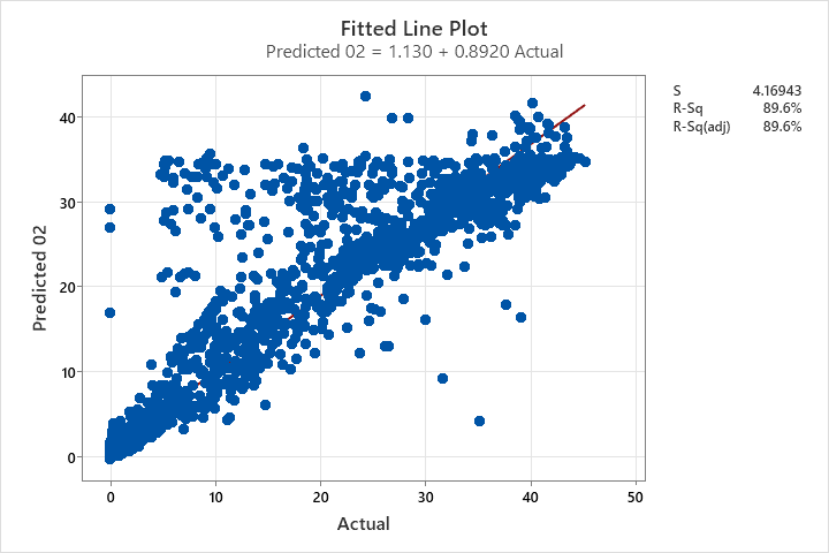


Figure 3. Final model accuracy

DISCUSSION

The results of this study demonstrate the successful development, optimization, and evaluation of an AI-based solar energy forecasting model. The initial objective of developing an AI model was achieved through the use of Artificial Neural Networks (ANN), which effectively captured the complex, non-linear relationships between meteorological data and solar

power generation, leading to improved accuracy. The optimization goal was met by applying various techniques like variable reduction, pruning, and quantization, significantly reducing the model's computational complexity while maintaining high performance.

Forecasting accuracy increasing from 30.3% to 89.6%, interference time decreasing from 15.6ms to 7.75ms, total processing time reducing from 19.71 s to 5.2 s, and RMSE improving from 9.44 to 4.12. Additionally, the training epochs decreased from 330 to 18, model size reduced from 0.06 MB to 0.04 MB, and parameter count halved from 8 to 4, resulting in faster inference and a more compact architecture suitable for real-time and embedded systems. While a slight accuracy trade-off occurred during optimization, the balance between reduced complexity and maintained high prediction accuracy was deemed acceptable. Compared to similar studies, this work stands out by achieving a computationally efficient model without significantly compromising performance, addressing the high resource demands common in many forecasting models.

The study has limitations, including the need for further testing in diverse geographical regions and over different time horizons to assess the model's generalizability. Additionally, as further computational complexity reduction techniques were applied, there was a slight decrease in accuracy. Therefore, it is crucial to apply computational complexity reduction techniques that do not compromise the model's accuracy. Real-world integration and testing in live solar power systems could further validate the model's performance and adaptability, ensuring its broader applicability and potential for sustainable solar energy solutions.

CONCLUSION

This research focused on the development of a low-complexity ANN-based short-term solar forecasting model. The model achieved an impressive 89.6% accuracy with a significant reduction in computational time, including inference time (from 15.6ms to 7.75ms) and total processing time (from 19.71s to 5.2s). The results demonstrate the model's suitability for deployment on low-power embedded devices, ensuring efficient real-time solar energy forecasting even in resource-constrained environments. Future work will involve testing the model under a wider range of weather conditions to further enhance its robustness and adaptability.

ACKNOWLEDGEMENTS

My deepest appreciation goes to all who have supported me with their invaluable guidance, expertise, and continuous encouragement throughout the course of this project. I am also grateful to the faculty and staff of the Department of Electrotechnology for providing the necessary resources and facilities, and to Wayamba University of Sri Lanka for creating a conducive academic environment that facilitated the development of this work.

I also appreciate the contributions of colleagues and advisors not listed as authors, whose insights and feedback have significantly improved the quality of this work. Lastly, I am profoundly thankful to my family and friends for their unwavering support and encouragement, which served as a constant source of motivation during the course of this research. This project would not have been possible without the collective contributions and support of all those mentioned.

REFERENCES

[1] *Historical Weather API | Open-Meteo.com*. (n.d.). Retrieved June 28, 2025, from https://open-meteo.com/en/docs/historical-weather-api

[2] Abuella, M., & Chowdhury, B. (n.d.). *Solar Power Forecasting Using Artificial Neural Networks*.

[3] Abuella, M., & Chowdhury, B. (2015, November 20). Solar power forecasting using artificial neural networks. *2015 North American Power Symposium, NAPS 2015*. https://doi.org/10.1109/NAPS.2015.7335176

[4] Amer, H. N., Dahlan, N. Y., Azmi, A. M., Latip, M. F. A., Onn, M. S., & Tumian, A. (2023). Solar power prediction based on Artificial Neural Network guided by feature selection for Large-scale Solar Photovoltaic Plant. *Energy Reports*, *9*, 262–266. https://doi.org/10.1016/j.egyr.2023.09.141

[5] Chodakowska, E., Nazarko, J., Nazarko, Ł., & Rabayah, H. S. (2024). *Solar Radiation Forecasting — State of Art and Methodological Discussion*. https://doi.org/10.20944/preprints202406.0128.v1

[6] *Deep Learning Toolbox - MATLAB*. (n.d.). Retrieved June 28, 2025, from https://www.mathworks.com/products/deep-learning.html

[7] Fernandes, F. E., & Yen, G. G. (2021). Pruning Deep Convolutional Neural Networks Architectures with Evolution Strategy. *Information Sciences*, *552*, 29–47. https://doi.org/10.1016/j.ins.2020.11.009

[8] *GitHub - Bob05757/Renewable-energy-generation-input-feature-variables-analysis*. (n.d.). Retrieved June 28, 2025, from https://github.com/Bob05757/Renewable-energy-generation-input-feature-variables-analysis

[9] Hossain, M. B., Gong, N., & Shaban, M. (2023). Computational Complexity Reduction Techniques for Deep Neural Networks: A Survey. *2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things, AIBThings 2023 - Proceedings*. https://doi.org/10.1109/AIBThings58340.2023.10292477

[10] Im, M. S., & Dasari, V. R. (2022). *Computational complexity reduction of deep neural networks*. http://arxiv.org/abs/2207.14620

[11] Molchanov, P., Mallya, A., Tyree, S., Frosio, I., & Nvidia, J. K. (n.d.). *Importance Estimation for Neural Network Pruning*. https://github.com/NVlabs/Taylor_pruning.

[12] Pham, B. T., Nguyen, M. D., Dao, D. Van, Prakash, I., Ly, H. B., Le, T. T., Ho, L. S., Nguyen, K. T., Ngo, T. Q., Hoang, V., Son, L. H., Ngo, H. T. T., Tran, H. T., Do, N. M., Van Le, H., Ho, H. L., & Tien Bui, D. (2019). Development of artificial intelligence models for the prediction of Compression Coefficient of soil: An application of Monte Carlo sensitivity analysis. *Science of the Total Environment*, *679*, 172–184. https://doi.org/10.1016/j.scitotenv.2019.05.061

[13] Ratner, B. (2009). The correlation coefficient: Its values range between 1/1, or do they. *Journal of Targeting, Measurement and Analysis for Marketing*, *17*(2), 139–142. https://doi.org/10.1057/jt.2009.5

[14] Sammar, M. J., Saeed, M. A., Mohsin, S. M., Akber, S. M. A., Bukhsh, R., Abazeed, M., & Ali, M. (2024). Illuminating the Future: A Comprehensive Review of AI-Based Solar Irradiance Prediction Models. *IEEE Access*, *12*, 114394–114415. https://doi.org/10.1109/ACCESS.2024.3402096

[15] *Solar Energy Northeast Brazil*. (n.d.). Retrieved June 28, 2025, from https://www.kaggle.com/datasets/oakthyago/solarenergynortheastbrazil

Li-ion Battery Cooling - A Computational Study of Different Phase Change Material Configurations

Sumudu Adikaram[a], Adel Nasser[b], Cristina Vallés[c], Chamil Abeykoon[a*]

[a] Northwest Composites Centre, Henry Royce Institute and National Graphene Institute, Department of Materials, Faculty of Science and Engineering, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[b] Department of Mechanical and Aerospace Engineering, Faculty of Science and Engineering, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[c] Department of Materials, National Graphene Institute and Henry Royce Institute, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
sumudu.adikaram@manchester.ac.uk, a.g.nasser@manchester.ac.uk, cristina.valles@manchester.ac.uk, chamil.abeykoon@manchester.ac.uk*

ABSTRACT

Overheating of Li-ion batteries in Electric Vehicles (EVs) degrades performance and reduces lifespan. Hence, energy-efficient and reliable Battery Thermal Management Systems (BTMS) are required. This paper investigates the use of Phase Change Materials (PCMs), a passive cooling method with high heat storage capacity, for the thermal management of prismatic Li-ion battery cells in EVs. This computational study models the influence of buoyancy-driven convective flow on the PCM cooling performance, compared against thermal conduction-only models. In addition, this study investigates how convective flow influences the cooling performance with variations in cell orientation between vertical and horizontal alignments. n-Octadecane is used as the PCM, and Computational Fluid Dynamics (CFD) simulations were conducted with the Solidification and Melting model in ANSYS Fluent. A 12 mm PCM layer placed around the cell periphery reduced the centre temperature after 1800 s by 2.7 K in the vertical orientation and 3.7 K in the horizontal orientation compared to air-cooling. The effect of natural convection was more pronounced in the horizontal orientation, providing superior cooling performance relative to the vertical case. When the same PCM volume was used to fully enclose the cell, the cooling effect was further enhanced, achieving a maximum temperature reduction of 8.3 K within the first 1800 s. The findings demonstrate that natural convection significantly enhances the PCM-based cooling effectiveness, particularly in horizontally oriented cells, while thinner PCM layers with increased heat transfer area promote faster melting and improved cooling performance.

KEYWORDS: *Battery Thermal Management (BTMS), Phase Change Materials (PCMs), Natural Convection, Computational Fluid Dynamics (CFD), Prismatic Li-ion Cells*

INTRODUCTION

Li-ion batteries are popular in EVs due to their high energy densities and long operating life (Subramanian et al., 2021). However, these batteries tend to overheat during charging and operation, while the overheating of one cell can lead to thermal runaway of the whole battery pack (Shirazi et al., 2016). Hence, it is important to maintain the temperature of the battery cells within the recommended range, usually below 40 °C for effective operation (Subramanian et al., 2021). Furthermore, monitoring the battery cells and regulating their temperature within the optimal limits helps to extend their life (Warner, 2014). Another key design criterion for BTMS is to control the temperature change of the battery pack within a 2-5 °C difference throughout its lifetime (Warner, 2014). The cooling requirement also depends on the region of operation and the nature of its performance (Warner, 2014), as the heating rates of the battery vary due to both the ambient

temperature and the charging and discharging rates.

With increased energy efficiency and fuel economy becoming rising priorities in the world, passive cooling methods are becoming an ideal solution for thermal management of EV batteries as they do not consume extra energy (Jankowski & McCluskey, 2014). Furthermore, they can offer cost-effective and easy installation solutions without complicating the design (Li et al., 2025). However, passive air cooling does not provide sufficient cooling for EV batteries; hence, alternative cooling methods are required.

Phase Change Materials (PCMs) are popular for thermal management and heat storage due to their high heat storage densities, as they absorb and release heat as latent heat while going through a phase change. They become good passive cooling options due to their high heat-absorbing and constant temperature heat storage capabilities, and hence, are currently under a lot of research. They are ideal for EV battery cooling as they are less bulky, less complex and less expensive than forced air cooling systems and liquid cooling methods (Jaguemont et al., 2018).

This work involves modelling a PCM cooling unit for prismatic Li-ion battery cells, which are widely used in EVs due to their high-density energy storage capabilities. The study is conducted to evaluate the cooling performance of different PCM unit designs with varying PCM layer thicknesses and configurations. The effect of gravity on the cooling performance of the PCM unit was also evaluated for different orientations of the battery cell. Two studies that have been conducted on the same battery model by Javani et al. (2014) and Verma et al. (2019) were used as the basis to develop the numerical model, which is detailed below.

Javani et al. (2014) studied the thermal performance of the battery with PCM unit under different heat generation rates of the cell caused by different operational conditions and ambient temperatures. The authors evaluated the cooling performance variation under different PCM thicknesses of 3, 6, 9 and 12 mm. Here, n-Octadecane was used as the PCM, and the PCM layer was only added around the periphery of the battery (refer to Figure 27) due to the orthotropic nature of the thermal conductivity of the prismatic-type battery cell (Javani et al., 2014; Li et al., 2025). Furthermore, the authors only modelled the thermal conduction, ignoring the effects of the convective flow of the liquid PCM. The temperature variation of the battery at fixed points was observed in the study to analyse the PCM cooling performance. The authors found that the maximum temperature reduction of the battery after 20 minutes only improved from 2.77 K to 3.04 K when the PCM thickness was increased from 3 mm to 12 mm. Furthermore, Javani et al. (2014) defined constant values for the specific heat (C_P), thermal conductivity and the density of the PCM for each region, solid, liquid and phase transition, as step functions. This approach is not suitable for modelling the PCM with gravity effects for tracking the flow front and to model the buoyancy-driven flow (Pan et al., 2018).

Verma et al. (2019) used the same battery model and extended the study using a different PCM, Capric acid. They observed the PCM cooling performance with the thickness of the PCM layer varying from 3, 7, 9 and 12 mm. Furthermore, the cooling performance was studied under two ambient conditions: 294 K for normal conditions and 323 K for desert conditions. It was

concluded that the PCM layer with a 3 mm thickness provides the optimum cooling. Similar to the previous study, Verma et al. (2019) have not considered the PCM flow in the model. Contrastingly, Verma et al. (2019) used constant values for the PCM properties of C_P, density and the thermal conductivity and used the Solidification and Melting model in ANSYS Fluent. However, despite the difference in the model, it was validated using the results of Javani et al. (2014), with an error below 7%.

In this study, the same battery model was used with n-Octadecane as the PCM. The model was improved to analyse the convection effects of the liquid PCM on its melting rates and the cooling performance. Effects of gravity and temperature-dependent density variations were incorporated into the model to simulate the buoyancy-driven convective flow and its effects. Furthermore, the orientation of the battery affects the PCM performance when convective flow is present. Due to the stacking of battery cells when creating the battery pack, vertical orientation is preferred for prismatic-type cells, where horizontal cell orientation is avoided to prevent upper cells from exerting force on the lower cells. Hence, vertical cell orientation is also modelled in comparison with the horizontal orientation. Finally, the study evaluates the effect of PCM configuration, i.e. PCM around the periphery of the cell (**Side Pack**) and PCM around the entire cell (**Whole Pack**), to understand how increasing the heat transfer surface area while reducing the PCM layer in the direction of high thermal conductivity influences the cooling performance.

METHODOLOGY

Initially, the model was created similarly to the models by Javani et al. (2014) and Verma et al. (2019), neglecting the convective flow for validating the initial model. Later, the gravity effects and density variation were incorporated into the model to analyse the convective flow effects.

Design

The parameters of the battery and the performance conditions were taken from Javani et al. (2014). The cell dimensions used in the models are provided in Table 24. The models and their boundary conditions are presented in

Figure 27. Five PCM configurations were analysed, and their thermal performance was compared, namely, the air-cooled battery without PCM cooling (**Cellw/oPCM**), the battery with PCM cooling without convective flow analysis (**noFlow**), and the battery with PCM cooling with convective flow analysis in the vertical orientation (**FlowY**) and the horizontal orientation (**FlowZ**). Finally, the same PCM volume was applied around the entire cell (except for the terminal face) to create the **WholePack** configuration, which includes a thin layer of PCM around the cell as shown in

Figure 27 (b). The performance of the WholePack configuration was compared to the **SidePack** configurations, i.e. noFlow, Flow Y and FlowZ, where the PCM only lies around the periphery of the cell as shown in

Figure 27 (a). Here, the thickness of the WholePack configuration was set so that the total PCM volume matched that of the SidePack configuration.

Table 24. Battery cell dimensions

Component	**Length (mm)**	**Height (mm)**	**Thickness (mm)**
Cell	146	194	5.4
Terminals	35	15	0.6
PCM (SidePack config)	3 and 12 mm around the periphery of the cell		5.4
PCM (WholePack config)	0.6 mm around the entire cell (except the face with terminals)		

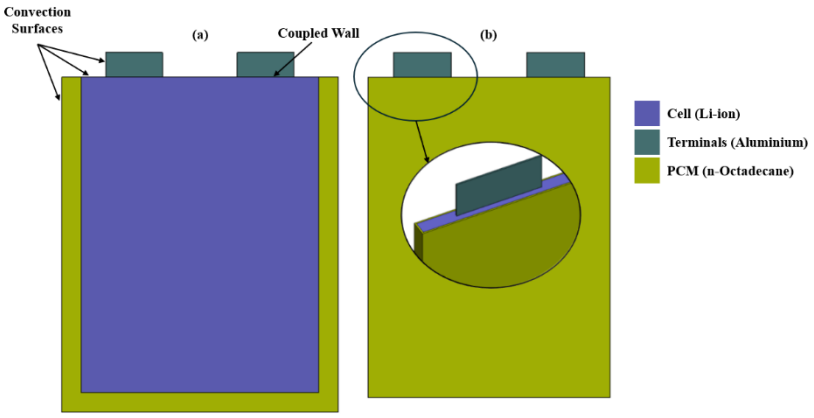


Figure 27. Geometry and boundary conditions of a) SidePack, b) WholePack configurations

Table 25. Boundary conditions

Initial cell temperature & ambient temperature	294.15 K

Uniform heat generation in the cell (when the battery is discharging at a 2C rate)	63,970 W/m^3
Outer surfaces (PCM walls, terminals & the cell top)	Convection
	Heat transfer coefficient: 7 W/m^{2}K
Coupled wall between the cell and the terminal	
Terminals: no heat generation, acts as cooling fins, material- Aluminium	
The PCM compartment and the cell are covered by a thin Aluminium layer of 0.5 mm thickness, which acts as the convection surface.	

The boundary conditions used for the simulations are provided in Table 25. The specifications of the battery, taken from Javani et al. (2014), are shown in Table 26. The heat generation rate when the battery is discharging at a 2C rate is 63,970 kW/m^3, which is 4.45 W per cell. The 2C rate signifies the battery discharges completely within half an hour. This rate was used for the simulations first to validate the results with the two previous studies, and then to compare the results of different cooling configurations with the initial results. Furthermore, at this high discharge rate, the battery tends to heat up more than the desired levels; hence, cooling is important.

Table 26. Battery Specifications

Heat generation rate (considered to be uniform throughout the cell)	At standard working conditions	6855 W/m^3
	When the cell is discharged at a 2C rate	63,970 W/m^3
	At full power	200,000 W/m^3
Thermal conductivity of the cell (Orthotropic)	In planar directions	25 W/m K
	In the direction normal to the cell	1 W/m K
Specific heat capacity (C_P)		1097 J/kg K
Density		1881 kg/ m^3

n-Octadecane was used as the PCM, and its properties are shown in Table 27. The Boussinesq model is used to approximate the density variation with temperature. Data was not available for the viscosity and thermal expansion coefficient of n-Octadecane in the paper by Javani et al. (2014); hence, it was taken from G. Zaki, E., & S. Dhmees (2022).

Table 27. Properties of n-Octadecane (Javani et al., 2014)

Solidus temperature ($T_{solidus}$)	301.15 K
Liquidus temperature ($T_{liquidus}$)	303.15 K
Thermal conductivity (k)	0.152 W/m K
Latent heat (L)	225 kJ/kg
Specific heat capacity (C_P)	2180 J/kg K
Density (ρ)	724 kg/ m^3
Viscosity	3.878×10^{-3} Pa s (G. Zaki, E., & S. Dhmees, 2022)
Thermal expansion coefficient	0.00091 1/K (G. Zaki, E., & S. Dhmees, 2022)

Mathematical Model

Initially, the noFlow configuration was modelled, neglecting the PCM convective flow, similar to the assumptions made in the previous two studies, to validate the results using those of the two papers. The density was considered constant, and only the thermal conduction was modelled. This method is called the Effective Heat Capacity (EHC) method. Here, only the energy equation is solved, where the phase change process is incorporated in the model with varying thermophysical properties defined. This model requires less computational power and permits larger mesh and time step sizes.

However, when modelling the Solidification and Melting of PCM with the convective flow, the Enthalpy Porosity (EP) method was used. The phase change behaviour was modelled using the Solidification and Melting model in ANSYS Fluent. The liquid front was modelled as an enthalpy-porosity formulation. The solid-liquid mushy zone is modelled as a porous zone with its porosity equal to the liquid fraction (a value between 0 and 1) of each cell in the domain. PCM is considered to be Newtonian and incompressible. The density variations are only considered in the buoyancy source term and are modelled using the Boussinesq model. Transient simulations were used to model the melting of the PCM over time with a **laminar viscous model.**

Enthalpy of the material (H) is calculated as the sum of sensible enthalpy (h) and the latent heat enthalpy (ΔH) and is given in Eq. (1). h is calculated by Eq. (2).

$$H = h + \Delta H \quad (1)$$

$$h = h_{ref} + \int_{T_{ref}}^{T} C_P dT \quad (2)$$

Here, h_{ref} is the reference enthalpy, T_{ref} is the reference temperature and C_P is the specific heat at constant pressure. The liquid fraction β is calculated using the PCM temperature (T) as given in the set of equations in Eq. (3).

$$\beta = 0 \text{ if } T < T_{solidus} \quad (3)$$

$$\beta = 1 \text{ if } T > T_{liquidus}$$

$$\beta = \frac{T - T_{solidus}}{T_{liquidus} - T_{solidus}} \text{ if } T_{solidus} < T < T_{liquidus}$$

The latent heat enthalpy (ΔH) can be calculated as given by Eq. (4).

$$\Delta H = \beta L \quad (4)$$

Here, L is the latent heat of the PCM, and the latent heat enthalpy changes from 0 to L as the solid PCM melts to become a liquid. The continuity equation, which ensures mass conservation, is given by Eq. (5), where the mass flux divergence is made zero. Here, density (ρ) is considered a constant, which is the reference density.

$$\nabla \cdot \rho \vec{v} = 0 \quad (5)$$

The momentum equation, which models the motion of PCM including inertia, viscous forces, and buoyancy forces due to thermal expansion, is given by Eq. (6). Here, the Boussinesq approximation is used to calculate the buoyancy force, where γ is the thermal expansion coefficient. $\vec{S}$ is the sink term, which is defined in Eq. (7).

$$\rho \frac{\partial \vec{v}}{\partial t} + \rho(\vec{v} \cdot \nabla)\vec{v} = -\nabla P + \mu(\nabla^2 \vec{v}) - \rho\gamma\left(T - T_{ref}\right)\vec{g} - \vec{S} \quad (6)$$

$$S = \frac{(1-\beta)^2}{(\beta^2 + \epsilon)} A_{mush} \cdot \vec{v} \quad (7)$$

The sink term calculates the momentum sink due to the reduced porosity in the mushy zone. Here, ε is a small number (0.001) to prevent division by zero, $\vec{v}$ is the velocity due to the buoyancy-driven natural convection and A_{mush} is the mushy zone constant. The mushy zone constant was selected as 100000. This parameter determines how quickly the velocity drops to zero as the material solidifies. For PCMs, the value is typically set to 100000 (Li et al., 2025).

The energy equation, which considers conduction and convection, is given by Eq. (8).

$$\frac{\partial}{\partial t}(\rho H) + \nabla \cdot (\rho \vec{v} H) = \nabla \cdot (k \nabla T) - S_L \quad (8)$$

Here, H is the enthalpy (from Eq. (1)), ρ is the density, $\vec{v}$ is the local velocity vector (initially in no flow condition, this was assumed to be 0), k is the thermal conductivity and S_L is the source term that accounts for the latent heat release and absorption during phase transition, which is given by Eq. (9). Here, ε is the porosity, which is equal to the liquid fraction.

$$S_L = \frac{\varepsilon \partial \rho \beta L}{\partial t} + \nabla(\rho \vec{v} \beta L) \quad (9)$$

It is important to capture convective flow as it promotes mixing, which could either even out the temperature gradients or could adversely affect the even cooling of the cell. Therefore, excluding the PCM flow can lead to inaccurate results.

Computational Setup

A structured mesh was used as presented in **Error! Reference source not found.**.

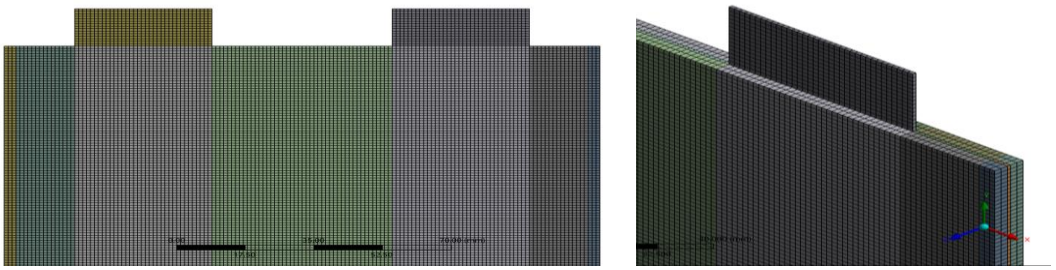


Figure 28. Structured mesh used for the cell-PCM model

All results were generated for five points along the centreline of the cell, as shown in Figure 29, which shows the plane midway across the thickness of the battery. The distances to Points 1-5 are 184, 140.5, 97, 43.5, and 10 mm, respectively, from the bottom edge of the cell.

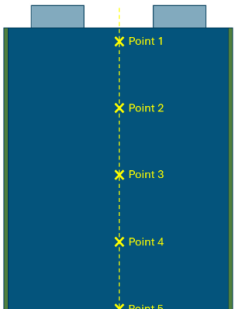
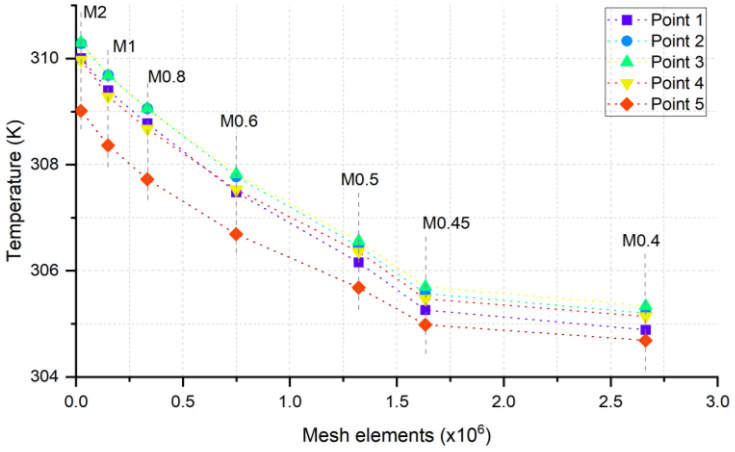


Figure 29. Locations along centreline to monitor data



A grid independence study was performed for the EP model, which is presented in **Error! Reference source not found.**, and t he mesh size of 0.6 mm was selected for the simulations. This size provided reliable results with an error of less than 0.85% while minimising computational cost. Since the study focuses on a qualitative comparison between different configurations, the same mesh element size was applied consistently across all simulations.

Validation

First, the model without PCM convection was validated using the results of the two papers, Javani et al. (2014) & Verma et al. (2019). The validation process was carried out initially for the air-cooled cell without PCM cooling (**Error! Reference s ource not found.**(a)), followed by the cell with PCM cooling with a 3 mm thick PCM layer (**Error! Reference source not found.**(b)). Since the EHC model has a reduced computational cost and permits larger mesh and time step sizes, a coarser mesh was used for validation. Initially, the mesh size and time step size were varied until the results showed correlation with the data of the two papers. The accuracy of the results also depends on the time step size; therefore, the time step was varied

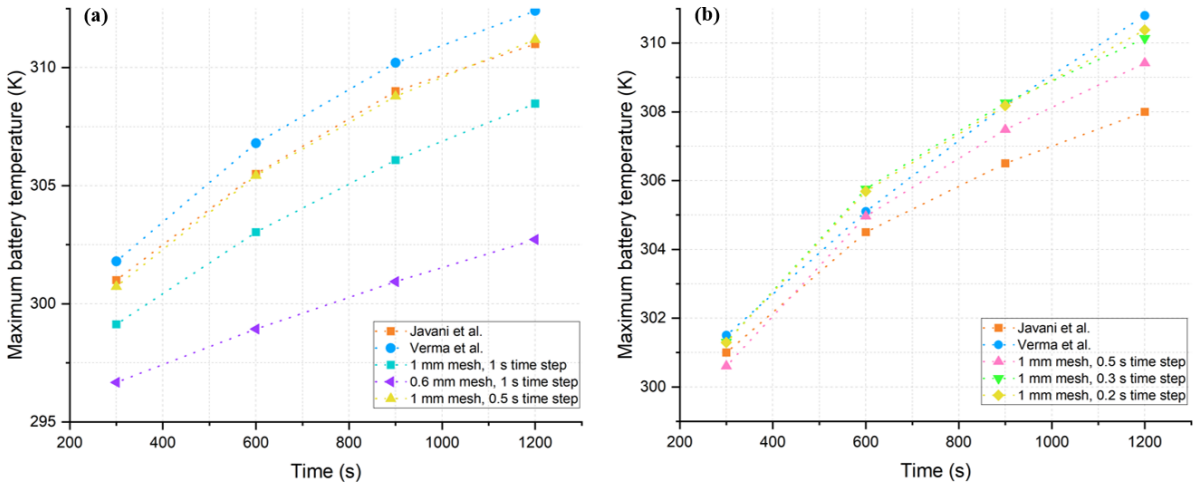


Figure 30. Validation of the battery model a) without and b) with PCM cooling

until the results converged. The maximum temperature of the cell was compared for each case.

To model the cell without PCM cooling, a 1 mm mesh with a 0.5 s time step size was sufficient, as it coincides with results from Javani et al. (2014). However, for the battery model with PCM cooling, the model with 1 mm mesh size and 0.3 s time step was chosen as it gives converged values from the time step convergence study, and the values are close to both results

by Javani et al. (2014) and Verma et al. (2019) with Root Mean Square Error (RMSE) values of 1.53 and 0.48, respectively, hence, the model can be used to obtain reliable results for the EHC model in comparison with the EP model. However, since the EP model requires smaller mesh sizes to generate reliable results, the same mesh element size of 0.6 mm was used for all models for an independent comparison of results.

RESULTS

Error! Reference source not found. presents the temperature distribution of the battery without PCM cooling and with P CM cooling under FlowZ and FlowY configurations at 1800 s. With PCM cooling, the maximum temperature was reduced from 314.12 K to 309.64 K in the FlowZ configuration and to 310.57 K in the FlowY configuration. The location of the maximum temperature shifted upward, attributed to the PCM layer at the bottom. Between the two PCM configurations, FlowZ demonstrated superior cooling performance.

Although at 2C mode, the maximum duration the battery works is for 1800 s, the cooling performance was further analysed for a higher duration to understand the PCM melting behaviours and the cooling performance associated with it.

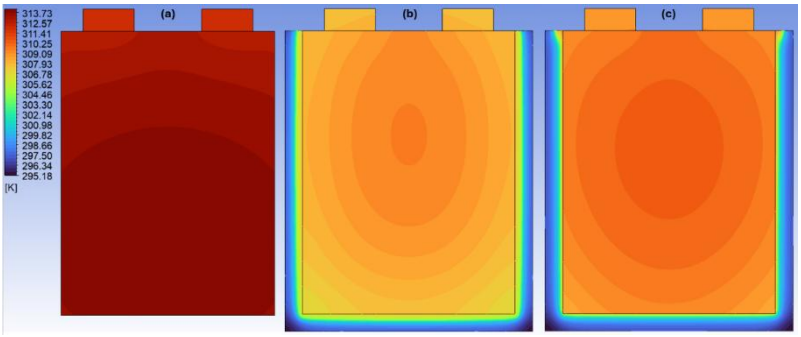


Figure 31. Temperature variation of the cell a) Cellw/oPCM, b) FlowZ config, c) FlowY configurations at 1800

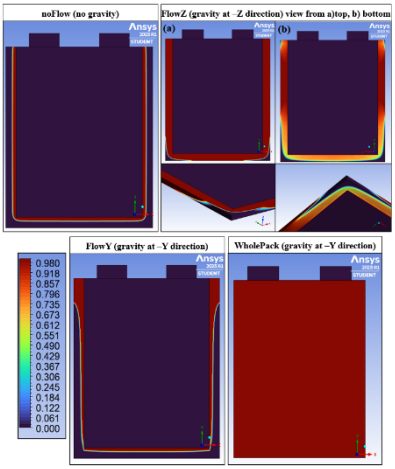


Figure 32. Liquid fraction of PCM for different configurations at 3

Error! Reference source not found. shows the liquid fraction of PCM for different configurations at 3600 s. Figure 33 (a) shows the temperature variation of Point 3 for each PCM configuration and Figure 33 (b) shows the liquid fraction variation

of PCM with time.

Cooling

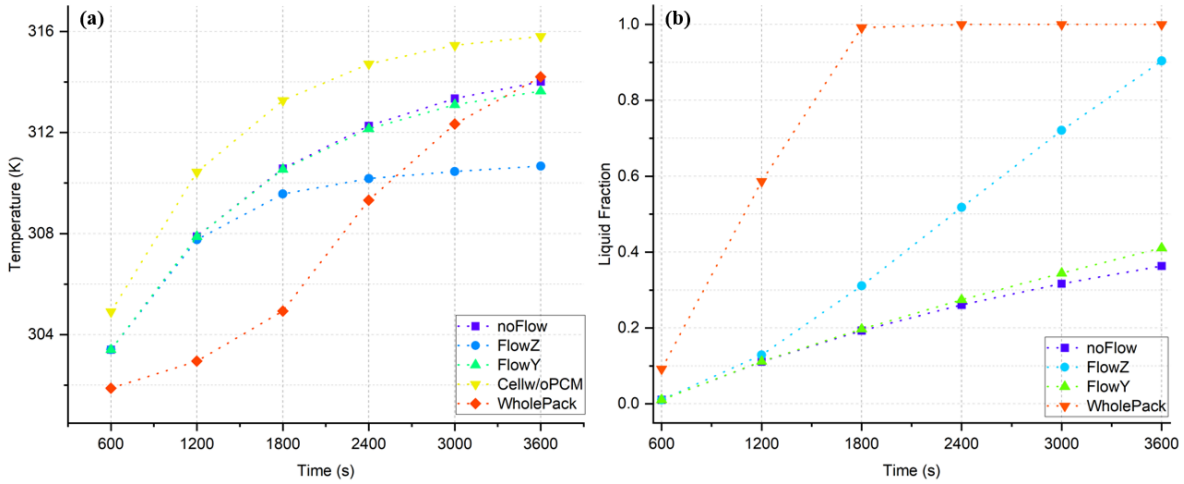


Figure 33. a) Temperature variation at Point 3, b) liquid fraction variation of PCM, for different PCM configurations

Performance Evaluation: EHC Model vs. EP Model

When comparing the temperature-time curves of noFlow, FlowZ, and FlowY in Figure 33(a), the noFlow and FlowZ curves are nearly identical at first, with noticeable differences emerging only at later time steps. The deviation in the FlowZ curve arises from distinct melting behavior driven by convective flow, as illustrated by the liquid fraction graphs in Figure 33(b). Initially, all temperature curves overlap because liquid fractions are low and heat transfer is dominated by conduction, resulting in similar performance across the noFlow, FlowZ, and FlowY models. As melting progresses, however, convective flow begins to significantly influence cooling performance.

Hence, modelling the convective flow is important when evaluating the cooling performance of PCM units. Although the EHC model can be used for small time durations during the initial melting phase to save computational efforts, as the PCM melts, incorporating the effects of convective flow using the EP model is important to accurately predict the thermal performance.

Impact of Cell Orientation on PCM Cooling Efficiency

When structural loading is considered, the vertical orientation is preferred for stacking prismatic cells in a battery pack. However, in terms of cooling performance, the horizontal orientation (FlowZ) outperforms the vertical orientation (FlowY) over the studied duration, as shown in Figure 33(a) and supported by the temperature contours in **Error! Reference source not found.**. The FlowZ configuration achieves higher liquid fractions faster, enabling greater utilization of latent heat for cooling. Convective flow in the horizontal orientation also promotes more effective heat propagation through the thin PCM layer, as illustrated by the melt flow patterns in **Error! Reference source not found.**.

Over the observed duration, the FlowY configuration showed only marginal improvement compared with the conduction-only model. This behavior arises because the melt flow is concentrated near the top of the cell, while the rest of the melting pattern remains similar to the noFlow model. Thus, although convective flow is present in the PCM unit, its impact is limited if it is not well distributed throughout the material. The influence of convection is expected to become more pronounced at higher liquid fractions over longer operating durations.

Comparison of Peripheral vs Fully Encapsulated PCM Configurations

The effect of PCM placement was evaluated, accounting for the orthotropic thermal conductivity of the battery, in two configurations: SidePack and WholePack. As shown in Figure 33 (b), the liquid fraction of the WholePack configuration reached 1 at 1800 s, while all SidePack configurations exhibited slower melting. Consequently, the WholePack initially

demonstrated superior cooling performance. However, once the PCM fully transitioned through the phase change region, the cell temperature increased sharply, emphasizing the importance of designing PCM units to operate effectively within the phase change range. At later stages, the FlowZ configuration outperformed the WholePack, maintaining better cooling performance as melting progressed.

The initial strong performance of the WholePack configuration is attributed to the large heat transfer area available, as the PCM fully encapsulates the battery with a thin layer. In contrast, with the SidePack designs, although the peripheral thermal conductivity of the cell is higher, the inherently low thermal conductivity of the PCM delays heat propagation until sufficient melting occurs to initiate convective flow and enhance thermal transfer.

Hence, the PCM unit should be designed according to the specific requirements of the application, such as operating duration and acceptable temperature range. For applications requiring long-term temperature regulation where minor increments are permissible, the SidePack configuration is more suitable. Conversely, applications demanding strict thermal control over shorter periods would benefit from designs similar to the WholePack. Therefore, a careful evaluation of design requirements is essential for developing an optimized PCM cooling unit. Furthermore, the initial low thermal performance of the SidePack configurations, caused by the low inherent thermal conductivity of the PCM, can be improved by incorporating nanoparticles to enhance conductivity. However, a critical evaluation is required, as the addition of nanoparticles also increases the viscosity of the nanocomposite PCM (NCPCM), thereby reducing thermal convection. Thus, the nanoparticle concentration should be carefully optimized by considering the combined effects of conduction and convection.

Comparison of centerline temperature evolution for all PCM configurations

Figure 34 shows the temperature variation of the cell along the centreline at the five points (Points 1–5) for the four PCM configurations over time. The WholePack configuration exhibited the best cooling performance initially, up to 2400 s, due to the rapid phase change of the PCM enabled by effective thermal propagation and the large heat transfer area. However, its performance declined once the PCM entered the sensible heat storage region.

Among the SidePack configurations, the FlowZ orientation demonstrated the highest cooling efficiency throughout the observed duration. Between 1800 s and 3600 s, FlowZ controlled the temperature rise effectively, showing only a slight increase over time. These results indicate that the horizontal orientation enhances convective flow, improving the ability of PCM to absorb and propagate thermal energy. In the FlowZ configuration, the cooling performance is higher towards the bottom of the cell, caused by the cooling effects of the PCM layer at the bottom.

Initially, at 600 s, all three SidePack configurations exhibited similar behavior, as heat transfer was dominated by conduction, which is consistent across the models. The noFlow and FlowY configurations showed nearly identical performance in the early stages, indicating that the simplified EHC model can be applied to the vertical orientation for a limited duration. However, as the PCM melted, convective flow effects became more pronounced in the FlowY configuration which could no longer be neglected. At higher time steps, the FlowY model demonstrated lower temperatures near the bottom of the cell compared to the noFlow model. This resulted from upward heat transport by the melted PCM, which left the lower region at relatively lower temperatures and thereby improved cooling performance.

Hence, these results provide insight into the impact of convective flow on the cooling performance of PCMs and highlight the importance of modeling it to accurately capture melting patterns and cooling behavior.

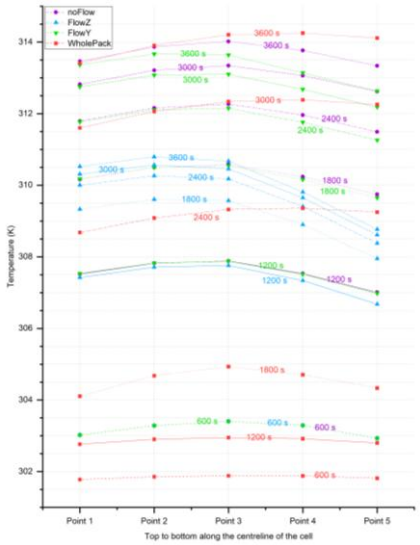


Figure 34. Temperature variation of the cell along the centreline at 600s, 1200 s, 1800 s, 2400 s, 3000 s, 3600 s

CONCLUSION

The cooling performance of different PCM configurations for the thermal regulation of prismatic battery cells was analyzed. Two modelling approaches were considered: the EHC method, which accounts only for thermal conduction in the propagation, and the EP method, which incorporates both conduction and convective flow. The cooling performance results from the two models were compared, and it was found that the simplified EHC method can be used for a limited period of about 1200 s, during which conduction is the dominant heat transfer mechanism. However, as the PCM melts, more complex models that account for convective flow, such as the EP model, are required to capture the melt dynamics and the resulting cooling performance.

Furthermore, two battery orientations, vertical (FlowY) and horizontal (FlowZ), were compared to assess the effect of convective flow on the cooling performance of the PCM unit. The FlowZ configuration exhibited superior cooling, achieving higher liquid fractions over time and resulting in a 3.9 K temperature drop at Point 5, located at the bottom of the cell, after 3600 s. Hence, the horizontal orientation enhances convective flow, enabling more effective heat propagation throughout the PCM unit. However, since the vertical orientation is preferred for prismatic cells due to cell stacking in battery packs, additional techniques are needed to enhance thermal propagation and melt rates in the FlowY configuration. Approaches such as incorporating fins or using graphite/metal foams could be explored to improve heat transfer within the PCM unit.

Finally, two PCM configurations were evaluated using the same PCM volume: a peripheral layer (SidePack) and full encapsulation of the cell (WholePack), to assess the effect of increasing the heat transfer area given the orthotropic thermal conductivity of the cell. The WholePack configuration demonstrated superior cooling performance, achieving a temperature reduction of up to 6.1 K compared with the FlowY orientation, and a maximum cell temperature reduction of 8.3 K relative to the air-cooled battery after 1800 s of operation. These results indicate that the WholePack is more effective than the SidePack in regulating cell temperatures. However, the cooling performance of the WholePack configuration substantially decreased once the PCM fully melted and entered the sensible heat storage region, leading to a gradual temperature rise after 1800 s. In contrast, the SidePack configuration, particularly in the FlowZ orientation, showed lower initial cooling, but delivered better performance with the melting of PCM, with cell temperatures dropping below those of the WholePack

configuration. Hence, the design of the PCM unit should account for the required duration of operation. The WholePack configuration is more suitable for short-term cooling, while the SidePack configuration is preferable for longer durations, as its thicker PCM layer slows heat propagation and sustains latent heat cooling.

Therefore, this study highlights the influence of different PCM cooling unit designs on cooling performance and emphasizes the need to align the design configuration with application requirements. Future work requires enhancing PCM cooling for vertically oriented cells, which are typically preferred in battery packs. Approaches to improve convective flow distribution within the PCM unit, such as optimized fin patterns or the use of metallic/graphite foams should be explored through further computational analysis.

REFERENCES

[1] G. Zaki, E., & S. Dhmees, A. (Eds. . (2022). Paraffin - Thermal Energy Storage Applications (E. G. Zaki & A. S. Dhmees (eds.)). IntechOpen. https://doi.org/10.5772/intechopen.90983

[2] Jaguemont, J., Omar, N., Van den Bossche, P., & Mierlo, J. (2018). Phase-change materials (PCM) for automotive applications: A review. Applied Thermal Engineering, 132, 308–320. https://doi.org/10.1016/j.applthermaleng.2017.12.097

[3] Jankowski, N. R., & McCluskey, F. P. (2014). A review of phase change materials for vehicle component thermal buffering. Applied Energy, 113, 1525–1561. https://doi.org/10.1016/j.apenergy.2013.08.026

[4] Javani, N., Dincer, I., Naterer, G. F., & Yilbas, B. S. (2014). Heat transfer and thermal management with PCMs in a Li-ion battery cell for electric vehicles. International Journal of Heat and Mass Transfer, 72, 690–703. https://doi.org/10.1016/j.ijheatmasstransfer.2013.12.076

[5] Li, W., Yin Yuen, A. C., Li, W., Li, W., Lin, P., Li, A., De Cachinho Cordeiro, I. M., Chen, Q., & Yuan Chen, T. B. (2025). Optimisation of PCM passive cooling efficiency on lithium-ion batteries based on coupled CFD and ANN techniques. Applied Thermal Engineering, 259(November 2024). https://doi.org/10.1016/j.applthermaleng.2024.124874

[6] Pan, C., Charles, J., Vermaak, N., Romero, C., Neti, S., Zheng, Y., Chen, C. H., & Bonner, R. (2018). Experimental, numerical and analytic study of unconstrained melting in a vertical cylinder with a focus on mushy region effects. International Journal of Heat and Mass Transfer, 124, 1015–1024. https://doi.org/10.1016/j.ijheatmasstransfer.2018.04.009

[7] Shirazi, A. H. N., Mohebbi, F., Azadi Kakavand, M. R., He, B., & Rabczuk, T. (2016). Paraffin Nanocomposites for Heat Management of Lithium-Ion Batteries: A Computational Investigation. Journal of Nanomaterials, 2016. https://doi.org/10.1155/2016/2131946

[8] Subramanian, M., Hoang, A. T., B, K., Nižetić, S., Solomon, J. M., Balasubramanian, D., C, S., G, T., Metghalchi, H., & Nguyen, X. P. (2021). A technical review on composite phase change material based secondary assisted battery thermal management system for electric vehicles. Journal of Cleaner Production, 322(September). https://doi.org/10.1016/j.jclepro.2021.129079

[9] Verma, A., Shashidhara, S., & Rakshit, D. (2019). A comparative study on battery thermal management using phase change material (PCM). Thermal Science and Engineering Progress, 11(March), 74–83. https://doi.org/10.1016/j.tsep.2019.03.003

[10] Warner, J. (2014). Lithium-Ion Battery Packs for EVs. In Lithium-Ion Batteries: Advances and Applications. Elsevier. https://doi.org/10.1016/B978-0-444-59513-3.00007-8

Public Sector Role as a Key Stakeholder Towards the Circular Economy in the Built Environment in Sri Lanka

Gunasekara G. S. A.[a], Gunarathna K. A. N.[b], Karunaratne B. C. T. M.[c*]

[a] Freelance Quantity Surveyor
[b] Lecturer, Department of Quantity Surveying, Faculty of Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
[c] Oxford College of Business, Colombo 07, Sri Lanka
savishka3314@gmail.com, nishanthi.g@sliit.lk, karunarathne7tania@gmail.com*

ABSTRACT

The circular economy (CE) concept is a more appropriate approach to meet the sustainability challenges of today's construction industry, allowing construction activities to operate in a closed loop, away from the traditional linear economy model (LE). Although some CE principles are applied in Sri Lanka, they are not fully implemented. As a policy maker, the public sector should play a significant role in implementing CE in the built environment. However, limited attention has been paid to this, and no research has identified the public sector's role in this regard. The primary aim of this research is to explore the role and potential contribution of the public sector as a key stakeholder in implementing CE in Sri Lanka's built environment. To achieve this, global and local practices were examined to understand how CE principles are applied and to provide recommendations for improving CE adoption. The benefits and challenges to the public sector and private sector views on its role in CE implementation were also explored.

A mixed approach was used for data collection. Expert interviews and questionnaires were developed after gaining a thorough understanding of CE principles through a comprehensive literature review. Ten expert interviews and thirty-eight questionnaire responses were analyzed. Although the public sector has initiated efforts such as green building standards and waste management, the study found a need for a stronger regulatory framework, institutional support, and collaboration with the private sector.

Challenges such as limited awareness, high start-up costs, and regulatory gaps were identified. The research emphasized the importance of CE policies, financial incentives, and public-private partnerships. It also stressed the need for training and awareness programs to equip stakeholders with the knowledge required to implement CE.

KEYWORDS: *Public Sector, Stakeholder, Circular Economy, Built Environment, Sri Lanka*

INTRODUCTION

Today's fast-growing world has given more prominence to sustainable development. Factors such as climate change, resource scarcity, increasing energy demand, and food security affect construction project time, cost, and quality (Lin et al., 2019). Maintaining a project's time, cost, and quality within these parameters leads to risk reduction, and a change in one of these parameters could directly affect the other two (Rezaian, 2011). According to Victar and Waidyasekara (2023), material waste is responsible for 21%–30% of cost overruns in building projects on average. Approximately 46% of all waste generated annually is produced by the construction industry in Europe, which accounts for 820 million tons of building and demolition waste. Minimizing waste and misuse of resources, as well as using recycling and renewable energy, is important to reduce construction cost overruns (Rezaian, 2011).

The relationship between construction activities and the environment is well recognized. A significant amount of waste generated by the construction industry ultimately ends up in landfills. Globally, the construction industry and its

manufactured materials consume large amounts of energy and are responsible for a significant portion of pollution caused by harmful emissions (Manewa et al., n.d.). Furthermore, the construction industry consumes a large number of resources to achieve its goals and contributes to wastage of resources (Wouterszoon et al., 2020). The industry is directly responsible for more than 30% of natural resource extraction and more than 25% of the world's solid waste (Benachio et al., 2020). About 25% of carbon dioxide is released during the building process, which also uses 40% of the world's basic materials and produces 40% of waste. For example, according to Ginga et al. (2020), in the European Union the construction industry contributed approximately 930 million tons of waste in 2016 and 2.36 billion tons in China in 2018. Waste generation has been recognized as an issue that impacts both the environment and the construction industry's efficiency (Victar and Waidyasekara, 2023). One of the main reasons for this is that the Linear Economy (LE) model is still widely used in the construction industry (Benachio et al., 2020). The degradation of the environment and resource depletion are made worse by major barriers to Sri Lanka's Construction and Demolition (C&D) Waste Management (WM) (Victar and Waidyasekara, 2024). Over the past decade, Sri Lanka's construction industry has experienced remarkable growth, which has increased waste from C&D. These pollutants are often improperly disposed of, posing health and environmental risks, or they end up in landfills (Victar & Waidyasekara, 2023).

The Circular Economy (CE) concept originated in the 1960s and has recently regained popularity in the industrial and policy arena as a method of supporting sustainable development activities (Lin et al., 2019). The construction industry has recognized the need to focus on optimizing its resource consumption and environmental impact. Recognizing the benefits of CE is considered the first step in achieving these goals (Victar & Waidyasekara, 2024). The CE concept is widely applied in the construction industry with key terms such as recycling building materials, adaptable and modular structures, and standardized designs aimed at reducing the environmental impact of construction (Victar & Waidyasekara, 2024). This can be achieved by reducing the use of primary materials, protecting material resources, and minimizing the carbon footprint through the implementation of a CE system. Through efficient and ongoing cycling of resources into goods and processes, CE seeks to create a more sustainable economy. In addition, CE could help reduce environmental issues such as resource depletion, urbanization, pollution, and climate change (Owojori and Okoro, 2022). Economic gains are also expected, such as increases in GDP, net material savings, job creation, and a decrease in the risk of supply and price instability of materials (Adams et al., 2017).

Previous studies on the CE domain highlight that the 3Rs—reduce, reuse, and recycle—as well as waste management and minimizing environmental impacts, are key areas of focus in industrialized countries such as the USA, UK, and EU. In the Asian region, Japan and South Korea place strong emphasis on educating the public about their responsibilities for material consumption and waste. China, on the other hand, embraced the CE concept to encourage urban growth and achieve balanced development in both urban and rural areas.

In the construction industry, there is a clear need to articulate the benefits of CE in a transparent and measurable way. While emphasis is placed on raising awareness of the CE concept, it is also vital to identify the challenges and enablers for implementing CE. This includes introducing relevant legislation and establishing the required institutional framework for implementation. Currently, there is no specific legislation regarding CE, which poses a challenge to developing sectors such as CE-related green procurement (Adams et al., 2017). The public sector can contribute by introducing relevant legislation and acting as a regulator and policymaker at national, regional, and local levels (Klein et al., 2022). Private and public sector collaboration is a key factor in developing CE, and the public sector must play a leading role in implementation (Bogovac et al., 2021).

The LE model is widely used in the construction industry across the world. The CE model is more environmentally friendly than the LE model as it uses resources more efficiently. However, in Sri Lanka and other developing countries there is a lack

of interest and awareness among stakeholders to transition from the LE model to CE. The main bottleneck for this transition is the minimal contribution from the public sector. Although some research has been carried out on CE worldwide, studies focusing on stakeholder contributions remain limited. Research on the contribution of the public sector as a stakeholder in CE implementation has not yet been conducted in Sri Lanka. For developing countries like Sri Lanka, CE has the potential to reduce project costs, improve resource efficiency, and strengthen the economy. Therefore, this research seeks to study the role of the public sector as a key stakeholder in CE implementation in Sri Lanka's built environment.

LITERATURE REVIEW

CE in Built Environment

The built environment, one of the industries with the highest global resource consumption, is increasingly recognizing the Circular Economy (CE) as a transformative path to sustainability (Weerakoon and Thayaparan, 2023). This is primarily because the built environment and the construction sector are closely linked, with the former driving the development and maintenance of the latter (De Valence, 2019). Organizations like the World Green Building Council (WGBC) aim for all buildings to achieve net-zero upfront carbon by 2050 and for all new constructions and renovations to reduce upfront carbon by at least 40% by 2030 (Selvaraj and Chan, 2024).

Whole life cycle thinking and life cycle costing are essential elements of CE in construction, emphasizing environmental impact from design to end-of-life. The World Business Council for Sustainable Development values CE as a $4.5 trillion opportunity that can foster new industries, create jobs, cut emissions, and improve resource efficiency. Many developed nations—such as Australia, the USA, Netherlands, France, UK, Denmark, Germany, Finland, Sweden, Norway, and China—have embraced CE, leading to innovation (Halog and Anieke, 2021). Globally, CE concepts are being integrated into construction regulations to promote sustainable practices, with cities offering tax incentives for using recycled materials or penalties for excessive waste (Adams et al., 2017). There is also growing interest in digital tools like Building Information Modeling (BIM) to support life cycle management and efficient resource use (Selvaraj and Chan, 2024).

In Sri Lanka, CE implementation in the built environment is still at an early stage. The construction industry typically follows a Linear Economy model, relies heavily on imported resources, and generates considerable waste (Victar and Waidyasekara, 2024). However, global sustainability trends, economic pressures, and environmental concerns are increasing awareness of the need for sustainable practices (Weerakoon and Thayaparan, 2023). To promote CE adoption, it is vital for the Sri Lankan government to collaborate with stakeholders to integrate CE into national policies and construction standards (Victar and Waidyasekara, 2024). Despite its benefits, CE implementation faces several challenges in Sri Lanka, including limited industry knowledge, lack of sustainable materials, weak regulatory frameworks, financial constraints, and the need for significant changes in design, construction, and operations to align with CE principles (Wanaguru et al., 2022).

Role of the Public Sector Implementing the CE

The Organization for Economic Cooperation and Development (OECD, 2019) defines public sector organizations as government-controlled entities providing public goods or services. According to Klein et al. (2020), the public sector functions as a regulator, policymaker, and major purchaser in economic development. For instance, in 2018, public spending accounted for 46.7% of the EU's GDP, with 15% allocated to public procurement. The public sector significantly influences the economy by enacting laws and providing infrastructure while also serving as a role model to other sectors. With over 55 million employees, the EU's public sector plays a major role in resource and energy consumption, with service providers consuming 13% of energy from 1990 to 2014—a growth of over 30% (Droege et al., 2021).

The public and private sectors are both crucial to the CE transition, although the public sector operates under a different structure with non-competitive missions, hierarchical management, and rule-based operations (Droege et al., 2021). It also

faces budget, governance, and legal constraints (Klein et al., 2022). Due to the sector's size and influence, its contribution to CE through institutional approaches is vital for sustainability. British Standard BS 8001:2017 offers guidelines for circularity through principles like systems thinking, innovation, and transparency, while the Ellen MacArthur Foundation's "ReSOLVE" framework—regenerate, share, optimize, loop, virtualize, exchange—is more common in the private sector (Klein et al., 2020).

The public sector can influence CE as a client, contractor, and regulator. As a client, public purchasing decisions can set market trends and promote CE by prioritizing suppliers who follow circular practices (Droege et al., 2021; Klein et al., 2020). As a contractor, the public sector can apply CE in project delivery, generating knowledge and promoting industry-wide adoption (Clifton et al., 2024; Droege et al., 2021). As a regulator, it can implement laws to mandate waste minimization and sustainable materials in construction, and support CE in urban planning, zoning, and recycling infrastructure (Clifton et al., 2024; Droege et al., 2021).

Public Procurement (PP) plays a critical role in advancing CE, as the public sector is a key contributor to the construction industry. Governments can shape market demand and industrial performance through CE-focused procurement, such as the use of recycled materials, modular designs, and deconstruction strategies (Clifton et al., 2024). Globally, PP accounts for 15–20% of GDP (Klein et al., 2022), with notable examples like Poland allocating over 12% and the EU around €2 trillion—much of it related to construction (Plebankiewicz, 2022). CE integration in the public sector often centers on PP and internal operations (Droege et al., 2021).

According to Klein et al. (2022), most CE-related public sector research focuses on Circular Public Procurement (CPP), which embeds CE principles such as the 3Rs (reduce, reuse, recycle) into Green Public Procurement (GPP) or Sustainable Public Procurement (SPP). The European Commission defines Circular Procurement (CP) as a GPP approach that promotes closed energy and material loops while minimizing negative impacts across a product's life cycle. Public procurement has proven effective in fostering innovation and technology development; for example, in Finland, nearly 48% of commercialized technologies between 1984 and 1998 were linked to public procurement or regulation (Plebankiewicz, 2022).

Challenges, barriers and enablers of implementing CE in built environment

There are lot of enablers and barriers affecting the implementation of the CE in the construction industry including the lack of certain laws and regulations (Owojori and Okoro, 2022). A key barrier is the lack of specific CE legislation and developing CE-related green buying practices requires appropriate legislative frameworks. Table 1 identifies the challenges and barriers of implementing CE in built environment.

Table 28. Challenges and barriers

Challengers and Barriers	**A**	**B**	**C**	**D**	**E**
The low number of end-of-life products	✓	✓	✓	✓	✓
The lack of market mechanisms to encourage a greater recovery	✓	✓			✓
The vagueness of the financial case	✓	✓	✓		✓
Some of the economic issues	✓	✓	✓	✓	✓
Fragmented supply chain structure		✓			✓
Lack of interest, awareness, and understanding	✓	✓	✓	✓	✓
Design considerations		✓	✓		
Particularly building complexity		✓			

Lack of end-of-life issues at the building level	✓	✓	✓		✓
No guarantee of financial gain			✓	✓	
Lack of technology and information	✓	✓	✓	✓	✓
Lack of rules and policies	✓	✓	✓	✓	✓
Limited cooperation among stakeholders		✓	✓		
Lack of standardized procedures and insufficient facilities	✓	✓	✓		✓
A-(Weerakoon and Thayaparan, 2023); B-(Adams et al., 2017); C-(Giorgi et al., 2022); D- (Lin et al., 2019); E-(Droege et al., 2021)					

Table 2 identifies the enablers of implementing CE in built environment.

Table 2. Enablers

Enablers	**A**	**B**	**C**	**D**	**E**
Crucially in breakout sessions with the requirement of understanding commercial viability		✓	✓		✓
Enabling the recovery of materials through viable (logistically and commercially) take back schemes		✓			✓
Development of markets with higher value and enabling technologies		✓	✓		✓
Knowledge, information and technology regarding the CE	✓	✓	✓	✓	✓
A-(Weerakoon and Thayaparan, 2023); B-(Adams et al., 2017); C-(Giorgi et al., 2022); D- (Lin et al., 2019); E- (Droege et al., 2021)					

RESEARCH METHODOLOGY

A literature review was conducted to explore the importance of adopting the Circular Economy (CE) model over the Linear Economy (LE) model and to identify research gaps. A mixed-method approach integrating both quantitative (questionnaire survey) and qualitative (expert interviews) data was adopted to comprehensively explore the role of the public sector in CE implementation in Sri Lanka. Quantitative data supported broad evaluation and trend analysis, while qualitative data provided deeper insights into CE principles, benefits, and stakeholder interactions. Ten expert interviews were conducted with professionals including civil engineers, quantity surveyors, and operations managers with 5–30 years of experience, selected through non-random purposive sampling. The number of interviews was determined based on the principle of thematic saturation, as no new themes emerged after the eighth interview, which is consistent with qualitative research standards (Guest et al., 2020). Additionally, 38 construction professionals from both public and private sectors participated in an online questionnaire survey to gather further data. Thematic analysis was applied to qualitative responses to identify recurring themes, and descriptive statistics (frequencies and percentages) were used to analyze quantitative data and assess awareness levels, perceived barriers, and industry support for CE. The study began with identifying the research gap and establishing five research objectives, followed by a background study using literature from books, journals, and research articles. Findings from the literature informed the design of the interview guide and questionnaire. Expert interviews helped assess CE implementation practices and stakeholder views, while the survey expanded understanding of public sector roles and challenges and supported the development of a framework for private sector engagement in CE adoption within the Sri Lankan construction industry. The profiles of experts who participated in the interviews and the industry experience of

questionnaire participants are presented in Tables 3, 4, and 5.

Table 3. Profiles of interviewees

Interviewee	**Designation**	**Experience (Years)**
I-01	Quantity surveyor	08
I-02	Planning engineer (civil engineer)	07
I-03	Senior quantity surveyor	23
I-04	Contract administrator (quantity surveyor)	21
I-05	Works engineer (Charted civil engineer)	31
I-06	Works engineer (Civil engineer)	10
I-07	Chief quantity surveyor (Charted quantity surveyor)	40
I-08	Quantity Surveyor	06
I-09	General manager (Electrical engineer)	30
I-10	Director (Charted quantity surveyor)	31

Table 4. Experience of questionnaire participants in the construction industry

Years of experience	No. of respondent	Percentages
Less than 5	12	31.6%
5-10	13	34.2%
11-15	07	18.4%
16-20	03	7.9%
Over 20	3	7.9%

Table 5. Awareness of CE among professionals

Awareness	No. of respondent	Percentages
Not	01	2.6%
Not much	12	31.6%
Some extent	22	57.9%
Widely	3	7.9%

DATA ANALYSIS

The Role and Implementation of Circular Economy (CE) Practices in Sri Lanka's Construction Industry: Public Sector Insights and Challenges

Understanding of Circular Economy (CE) principles in the Sri Lankan construction industry is limited compared to global contexts. While countries like the USA, EU, China, India, the UK, and New Zealand have integrated CE into construction through policies and green certifications like LEED and BREEAM, Sri Lanka is still at an early stage. Interviewees acknowledged that while concepts like recycling, reuse, and sustainable design are recognized, broader awareness and practical adoption remain low. Local implementation is hindered by economic issues, high material costs, and past crises like COVID-19, as noted by I-07. Some interviewees (I-01, I-04, I-10) cited existing practices such as the reuse of construction

materials, waste-to-energy initiatives, and examples like Clearpoint Residencies and solar projects as partial implementations of CE. However, I-04 emphasized that societal behavior and a lack of discipline are major barriers to effective adoption.

The public sector's involvement in CE in Sri Lanka's built environment remains limited and fragmented. While some initiatives exist—such as green building certifications, waste management systems, and solar projects—interviewees (I-02, I-06, I-08) agreed that these are isolated and not part of a cohesive national strategy. Regulations are outdated or insufficient; although organizations like the Central Environment Authority (CEA) and Urban Development Authority (UDA) play roles, there is no dedicated CE policy framework. Some public sector bodies like SLGBC and CIDA support green initiatives, but I-04 and I-06 noted issues like fraud, corruption, and a lack of commitment from public institutions hinder CE progress. Adopting CE offers multiple benefits for Sri Lanka's public sector. All interviewees agreed that reuse and recycling can reduce public spending, and sustainable energy practices like solar panels and water reuse can lower utility costs. I-01, I-06, and I-08 emphasized long-term financial efficiency and resource management. Environmental benefits include reduced carbon emissions, better public health, and conservation of resources. I-10 suggested CE could also stimulate new industries, such as waste management, contributing to GDP growth and employment. Moreover, I-04 argued that the adoption of CE could positively impact societal behavior and improve mental and physical well-being, with I-07 highlighting enhanced global recognition as another benefit.

The main barriers to CE implementation are a lack of awareness, financial limitations, inadequate regulations, and cultural resistance. Interviewees (I-01, I-03, I-05) pointed out the general unfamiliarity with CE among public officials. Financial challenges were emphasized by all, particularly the high initial costs and the need to establish new supporting industries. I-03 and I-04 cited policy and regulation gaps, while I-08 confirmed that CE-specific policies are currently absent. Corruption and inefficiencies further undermine trust and deter investment. Cultural resistance to change, low ethical standards, and reluctance to adopt new technologies were also flagged. Technological limitations and the need for foreign expertise and funding add to the burden, particularly for retrofitting existing infrastructure. To accelerate CE adoption, collaboration between public and private sectors is essential. Interviewees (I-01, I-08, I-10) emphasized the need for clear policies, regulatory frameworks, and financial incentives like tax concessions and loans to motivate the private sector. I-09 highlighted public procurement as a strategic tool to promote CE-friendly services. Training programs, workshops, and government-led forums can facilitate knowledge sharing and innovation. Interviewees (I-02, I-03, I-07) recommended mechanisms such as contractor certification and PPP (Public-Private Partnership) agreements. Through initiatives like feasibility studies and support from organizations like SLGBC, the public sector can guide and incentivize the private sector to align with national CE goals.

Benefits of adopting CE in the public sector

Table 6. Benefits of adopting CE in the public sector

Benefits	Frequency	Percentages
Improvement resources efficiency and material reuse.	33	86.8%
Reduced environmental impact.	29	76.3%
Long-term cost savings in construction and infrastructure projects	28	73.7%

Enhanced public perception and leadership in sustainability and innovation	26	68.4%
Wastage management	25	65.8%

The table 6 shows that the majority of respondents, 86.8%, indicated improving resource efficiency and material reuse as a primary benefit. Additionally, 76.3% identified reduced environmental impact as a critical benefit, 73.7% indicated potential long-term cost savings in construction and infrastructure projects, and 68.4% improved public awareness and leadership on sustainability and innovation. 65.8% indicated the need for waste management.

Additional benefits cited by respondents include CE models fostering innovation and creating new markets for recycled and reused materials. This could lead to the creation of new jobs and economic opportunities in areas such as waste management, materials recovery and green technology, which could benefit the local economy and give the public sector a reputation.

As the majority of responses were quantity surveyors, the primary reason why improving resource efficiency and material reuse was considered a significant benefit was that material costs were high and that direct cost savings could be achieved by optimizing resource availability and material use. These responses may lead to material shortages in the future. As the construction industry uses the majority of natural resources, it causes environmental pollution and global warming. It can be assumed that the latter will pay more attention to the environmental impact, as global problems such as sea level rise may arise due to this. Additionally, the emphasis on potential long-term cost savings from reduced infrastructure development suggests that respondents recognize the value of sustainable practices in addressing issues such as carbon footprints and energy consumption. Furthermore, waste management has been identified as a problem faced by the public sector in Sri Lanka and the respondents have indicated that it is appropriate to give priority to get a solution to it and the opinion of others who did not respond here may be that implementing CE is not a grant benefit as it will reduce waste production anyway.

Barriers to adopting Circular Economy (CE) principles in the public sector

4.3.1 Economic barriers to implementing CE

Table 7. Economic barriers to implementing CE

Main economic barrier	Frequency	Percentages
Lack of founding	27	71.1%
Lack of investment	26	68.4%
Limited access to adequate resources (trained workforce and modernized equipment)	23	60.5%
Low demand for CE implementation	18	47.4%

According to the table 7, 71.1% of respondents cited this issue, indicating that financial constraints were a primary barrier to adopting CE practices, while 68.4% indicated a lack of investment. In addition, 60.5% of respondents indicated limited access to adequate resources such as a trained workforce and modernized equipment, and finally, 47.4% of respondents indicated low demand as another constraint to implementing the CE.

Other economic barriers respondents cited include lack of established markets for recycled materials, long payback periods due to high initial investment costs.

The reason why most of the responses cited financial constraints as a primary barrier to adopting CE methods may be the perception that implementing CE solutions requires significant investment. Many CE initiatives involve high initial costs that may be challenging for companies to bear, especially in a developing market such as Sri Lanka's construction industry. Responses are secondly presented as lack of investment. They may have mentioned that financial assistance to the public sector is limited. Because it is difficult for the public sector to allocate the necessary funds for these initiatives. Third, limited

access to a trained workforce and modernized equipment was noted, indicating that even when some financial resources are available, lack of trained labor and advanced technology hinders effective adoption of CE practices. Finally, citing low demand, respondents indicated that many businesses see limited market incentives for CE practices and that implementation is not as necessary as other factors in implementing it, as incentives can be increased later.Governments and financial institutions can develop targeted funding programs such as grants, low-interest loans, or tax incentives for CE projects, and have the potential to enhance public-private collaboration. Also, efforts to increase awareness among consumers about the benefits of CE methods can also help increase demand for more sustainable products and services, making CE projects more economically viable in the long run.

4.3.2 Legal and political barriers to implementing CE

Table 8. Legal and political barriers to implementing CE

Main legal and political barrier	Frequency	Percentages
Failure of the public sector to introduce legislative or regulatory drivers to promote private sector participation	30	78.9%
High level of political bureaucracy in project allocation	26	68.4%
Lack of legal framework for investment	20	52.6%

According to the table 8, 78.9% of respondents indicated that the main legal and political barriers to implementing CE principles in the construction industry in Sri Lanka include a significant lack of regulatory support and legislative drivers. Additionally, 68.4% of project allocations indicated high levels of political bureaucracy. Furthermore, 52.6% indicated that there is no clear legal framework for investment in CE.

A lack of clear policies and incentives may be the reason many respondents see lack of regulatory support and legislative drivers as a major barrier to implementing CE principles in the construction sector in Sri Lanka.

As standard legal frameworks and regulations are needed to implement and promote any concept, it can be assumed that the respondents prioritized this. Without these, investing in projects and transitioning to CE practices can be a challenge. Second, they may have highlighted the political bureaucracy of the project allocation and they may have tried to mention the corruption of politicians in Sri Lanka.

This particularly affects proposals and the lack of a clear legal framework for investment to start or expand CE projects has received the least response. The reason for that may be the opinion of many respondents that this is also included in a standard legal framework.

To address these issues, the public sector could develop a stronger regulatory framework that includes clear guidelines and incentives for the adoption of CE principles in the construction sector. This could include introducing specific legislation mandating waste reduction, recycling and resource efficiency measures in all public and private construction projects through the Green Building Council. Simplifying the regulatory process and reducing political bureaucracy and red tape can attract domestic and foreign investment in CE initiatives.

4.3.3 Social barriers to implementing CE

Table 9. Social barriers to implementing CE

Main social barrier	Frequency	Percentages
Lack of awareness, knowledge and understanding on CE concept	35	92.1%
Lack of education on CE strategies	25	65.8%
Lack of customer demand and reluctance to take risks to change	22	57.9%

Increase in initial cost	15	39.5%

According to the table 9, 92.1% of the responses indicate insufficient knowledge and understanding of CE. Additionally, 65.8% indicated a lack of education on specific CE strategies, while 57.9% of responses indicated a lack of consumer demand and reluctance to embrace change, suggesting a cultural resistance to adopting new approaches. Furthermore, 39.5% of respondents cited rising start-up costs as a barrier.

Respondents cited cultural resistance to innovation, social norms and established practices as other social barriers.

Sri Lanka currently uses a LE and many responses have been given to the lack of knowledge and understanding of CE, which may be due to the fact that elements of CE are not fully implemented. Along with this, the lack of programs to provide knowledge about this can be assumed as a hindrance by the respondents. Also, thirdly, it is presented that customers do not want to change. However, those who did not respond to this can be made aware of this through a program like the workshop and think that there is a possibility to get out of this situation. Finally, most of the respondents were of the opinion that the increase in the initial cost is not a major social constraint as the initial cost can attract investment.

To overcome these problems, the public sector can work together with the private sector to educate professionals and attract consumers.

4.3.4 Technological barriers to implementing CE

Table 10. Technological barriers to implementing CE

Main technological barrier	Frequency	Percentages
Lack of advanced technology and equipment to implement CE	30	78.9%
Lack of technical capability and expertise in CE	27	71.1%
Lack of efficient recycling techniques	22	57.9%

As per the table 10, a significant 78.9% of the responses highlighted the lack of advanced technology and equipment required for effective CE implementation. Additionally, 71.1% indicated a lack of technical capabilities and expertise, and 57.9% highlighted the lack of efficient recycling techniques as a barrier.

Respondents cited limited recycling infrastructure, limited recycling infrastructure, inadequate waste tracking and management systems as other technical barriers.

The majority of responses identified a lack of advanced technology and equipment as a barrier to CE implementation, perhaps because many CE practices such as recycling, material repurposing, and energy efficient construction require specialized and sophisticated equipment and technology. It is not easily available in the construction industry in Sri Lanka. Without access to advanced machinery and technologies, adopting sustainable practices and improving efficiency in any sector is challenging. Additionally, the second highest number of responses indicated a lack of technical capabilities and expertise, and even when some equipment is available, there may not be enough skilled professionals to effectively operate or maintain these technologies. A significant response to the third position of responses was inefficient recycling techniques, indicating that current recycling methods are outdated or unable to handle the complexity of modern waste streams, making it difficult to achieve the desired resource recovery, resource wastage and sustainability.

The public sector can invest in training and capacity building programs for the construction workforce to address these challenges. Given the current situation in Sri Lanka, the government could facilitate partnerships between international companies with expertise in CE that would help to share knowledge and best practices, import technical tools needed by the country, and bring in foreign experts at the start-up stage.

Public sector role and responsibilities in CE

4.4.1 Strategies for the public sector to effectively support CE initiatives

Table 11. Actions or strategies that the public sector should prioritize

Strategies	Frequency	Percentages
Introducing stricter rules on waste management and recycling in construction	34	71.1%
Provide financial incentives, subsidies or tax breaks for CE projects	27	71.1%
Fostering public-private partnerships to advance CE solutions	26	68.4%
Invest in research and innovation in sustainable building materials and design	26	68.4%
Promote public awareness campaigns and educational programs	20	52.6%

According to the table 11, majority of the respondents 89.5% stated that the priority is to introduce strict rules regarding waste management and recycling in construction. In addition, 71.1% of respondents prioritized financial incentives, subsidies or tax breaks for CE projects, and 68.4% responded that public-private partnerships foster research and innovation in sustainable building materials and design. Investing. Finally, 52.6% prioritized public awareness campaigns and educational programs.

The majority of responses prioritized the introduction of stricter rules for waste management and recycling in construction, possibly due to the underperformance of existing regulatory frameworks in Sri Lanka that allow for inefficient waste handling and excessive resource use. Stricter regulations can ensure that sustainable practices become the standard across the industry while minimizing waste. In addition, responses secondly indicated that financial incentives, subsidies or tax breaks may be high initial costs for implementing CE practices. However, it has been pointed out that the legal framework is necessary for implementation. As the private sector can invest and build a partnership within a legal framework, 68.6% have mentioned partnership. Some respondents also pointed to the need to sensitize stakeholders to awareness campaigns and that research and innovation can help advance CE.

4.4.2 Improvement of public sector policy frameworks or incentives to encourage CE

Table 12. Public sector policy frameworks or incentives to encourage CE

Policy and frame work	Frequency	Percentages
Developing financial mechanisms to support CE investments	30	78.9%
Implementation of mandatory CE guidelines in PP	28	73.7%
Providing tax incentives for companies that use sustainable construction practices	26	68.4%
Simplifying the approval process for CE projects	20	52.6%
Establish CE performance targets for public infrastructure projects	18	47.4%

According to the graph in Figure 4.15, 78.9% of the responses emphasized the need to develop financial mechanisms to support CE investments, while 73.7% of the responses mentioned the implementation of mandatory CE guidelines in public procurement. In addition, 68.4% emphasized the importance of tax incentives for companies adopting sustainable construction methods, simplification of the approval process for CE projects was mentioned by 52.6% of the respondents and finally, 47.4% of the responses for progress. CE mentions establishing performance targets for public infrastructure projects that can create accountability and benchmarks.

The majority of respondents emphasized the need to develop financial mechanisms to support CE investments, perhaps because of the significant initial costs associated with adopting sustainable practices such as advanced technology and

specialized equipment and training and education programs. Also, the second highest number of responses emphasized the importance of implementing mandatory CE guidelines in public procurement, perhaps because they believe that clear regulations can standardize sustainable practices across the industry. Tax incentives rank third in responses and the view is that reducing the financial burden can encourage the private sector to adopt sustainable practices.

The public sector can contribute in several ways to support the adoption of CE in the construction sector in Sri Lanka. Develop and enforce clear regulations such as circular construction standards and waste management policies to guide industry practices and incentives such as tax breaks and green financing initiatives to guide the adoption of sustainable practices. Collaborative innovation can be facilitated through public-private partnerships, and hubs can be created where stakeholders work together on circular solutions.

Investing in recycling infrastructure and establishing digital marketplaces for reused materials can create a viable market for materials recovery and recycled products.

CONLUSION

This study investigated the role and potential contribution of the public sector as a key stakeholder in implementing Circular Economy (CE) principles within Sri Lanka's built environment. The construction industry is a major consumer of natural resources and a significant contributor to waste, largely due to its dependence on the traditional Linear Economy (LE) model. CE provides a more sustainable alternative by promoting a closed-loop system that encourages reuse of materials and reduction of waste. Achieving CE implementation in Sri Lanka requires coordinated action from both public and private sectors, with the public sector responsible for developing regulations, institutional frameworks, attracting investment, and creating an enabling environment for private sector participation.

The findings indicate that, while CE principles are increasingly being implemented in many countries such as those in Europe, New Zealand, China, and India, their adoption in Sri Lanka is still limited. The research identified key benefits of CE, including the potential to create new industries, improve productivity through resource reuse, reduce waste, and promote ethical practices and environmental responsibility.

However, several barriers hinder implementation, including a lack of awareness and understanding of CE, limited funding, insufficient technology, a shortage of skilled professionals, and resistance to using recycled materials. Bureaucratic inefficiencies and lack of commitment from public institutions were also found to be major challenges.

The study recommends developing comprehensive legal and regulatory frameworks, offering tax incentives, simplifying approval processes for CE projects, and investing in public awareness and education. Strengthening public procurement practices, fostering public-private partnerships, and setting measurable CE performance targets for infrastructure projects would further accelerate adoption.

The research acknowledges certain limitations, particularly its focus on the public sector and the relatively small sample size of ten expert interviews and thirty-eight survey responses, which may not represent all stakeholder perspectives. Future research should examine the effectiveness of public-private partnerships in CE implementation, assess pilot projects, and explore policy frameworks that facilitate or hinder CE adoption. These insights could help policymakers and industry leaders create a clear pathway to embed CE principles in Sri Lanka’s construction sector.

REFERENCES

[1] Adams, K. T., Osmani, M., Thorpe, T., & Thornback, J. (2017). Circular economy in construction: Current awareness, challenges and enablers. Proceedings of the Institution of Civil Engineers – Waste and Resource Management, 170(1), 15–24.

[2] Benachio, G. L. F., Freitas, M. do C. D., & Tavares, S. F. (2020). Circular economy in the construction industry: A systematic literature review. Journal of Cleaner Production, 260, 121046. https://doi.org/10.1016/j.jclepro.2020.121046
[3] Bogovac, J., Dodig, D., & Lugarić, T. R. (2021). Public-private partnership and circular economy—What Croatian students learn at university. Energies, 14(11), 3261. https://doi.org/10.3390/en14113261
[4] Clifton, N., Kyaw, K. S., Liu, Z., & Walpole, G. (2024). An empirical study on public sector versus third sector circular economy-oriented innovations. Sustainability, 16(4), 1650. https://doi.org/10.3390/su16041650
[5] Droege, H., Raggi, A., & Ramos, T. B. (2021). Co-development of a framework for circular economy assessment in organisations: Learnings from the public sector. Corporate Social Responsibility and Environmental Management, 28(6), 1715–1729. https://doi.org/10.1002/csr.2140
[6] Ginga, C. P., Ongpeng, J. M. C., & Daly, M. K. M. (2020). Circular economy on construction and demolition waste: A literature review on material recovery and production. Materials, 13(13), 2970. https://doi.org/10.3390/ma13132970
[7] Giorgi, S., Lavagna, M., Wang, K., Osmani, M., Liu, G., & Campioli, A. (2022). Drivers and barriers towards circular economy in the building sector: Stakeholder interviews and analysis of five European countries' policies and practices. Journal of Cleaner Production, 336, 130395. https://doi.org/10.1016/j.jclepro.2022.130395
[8] Halog, A., & Anieke, S. (2021). A review of circular economy studies in developed countries and its potential adoption in developing countries. Circular Economy and Sustainability, 1(1), 209–230. https://doi.org/10.1007/s43615-021-00017-0
[9] Klein, N., Deutz, P., & Ramos, T. B. (2022). A survey of circular economy initiatives in Portuguese central public sector organisations: National outlook for implementation. Journal of Environmental Management, 314, 114982. https://doi.org/10.1016/j.jenvman.2022.114982
[10] Klein, N., Ramos, T., & Deutz, P. (2020). Circular economy practices and strategies in public sector organizations: An integrative review. Sustainability, 12(10), 4181. https://doi.org/10.3390/su12104181
[11] Klein, N., Ramos, T. B., & Deutz, P. (2022). Advancing the circular economy in public sector organisations: Employees' perspectives on practices. Circular Economy and Sustainability, 2, 759–781. https://doi.org/10.1007/s43615-021-00044-x
[12] Lin, S., Shen, B., Yong, S., Angelo, M. B., Choy, A., & Loong, H. (2019). Prioritization of sustainability indicators for promoting the circular economy: The case of developing countries. https://mpra.ub.uni-muenchen.de/95450/
[13] Manewa, R. M. A. S., Rameezdeen, R., Amaratunga, R. D. G., & Ginige, K. N. (n.d.). Towards sustainable construction through minimising site waste in Sri Lanka.
[14] Owojori, O. M., & Okoro, C. (2022). The private sector role as a key supporting stakeholder towards circular economy in the built environment: A scientometric and content analysis. Buildings, 12(5), 695. https://doi.org/10.3390/buildings12050695
[15] Plebankiewicz, E. (2022). The role of public procurement in implementing the circular economy in construction. Inżynieria Bezpieczeństwa Obiektów Antropogenicznych, 4, 69–78. https://www.inzynieriabezpieczenstwa.com.pl/index.php/iboa/article/view/160
[16] Rezaian, A. (2011). Time-cost-quality-risk of construction and development projects or investment. Middle-East Journal of Scientific Research, 10(2), 218–223.
[17] Selvaraj, S., & Chan, T.-M. (2024). Recommendations for implementing circular economy in construction: Direct reuse of steel structures. Journal of Constructional Steel Research, 214, 108439. https://doi.org/10.1016/j.jcsr.2023.108439
[18] De Valence, G. (2019). Reframing construction within the built environment sector. Engineering, Construction and Architectural Management, 26(5), 740–745. https://doi.org/10.1108/ECAM-02-2018-0088
[19] Victar, H. C., & Waidyasekara, A. S. (2024). Optimising construction waste management in Sri Lanka through circular economy strategies: A focus on construction and renovation and use and operate stages. Engineering, Construction and Architectural Management. https://doi.org/10.1108/ECAM-06-2023-0608
[20] Victar, H. C., & Waidyasekara, K. G. A. S. (2023). Construction and demolition waste management issues in building project life cycle stages: A case of Sri Lanka. 11th World Construction Symposium 2023, 392–406. http://dl.lib.uom.lk/handle/123/21289
[21] Wanaguru, K., Mallawaarachchi, H., & Vijerathne, D. (2022). Circular economy (CE) based material selection: Development of a CE-based '10R' evaluation framework for building construction projects in Sri Lanka. 10th World Construction Symposium 2022, 208–219. http://dl.lib.uom.lk/handle/123/19969
[22] Weerakoon, P., & Thayaparan, M. (2023). Analysis of barriers towards implementing circular economy in Sri Lankan built environment. 11th World Construction Symposium 2023, 162–173. http://dl.lib.uom.lk/handle/123/21265
[23] Weerakoon, T. G., Wimalasena, S., & Fedotova, K. (2023). Economic crisis adaptation in Sri Lankan construction industry: Pathway to prosperity. Baltic Journal of Real Estate Economics and Construction Management, 11(1), 240–256.

[24] Wouterszoon J. B., Van Stijn, A., Gruis, V., & Van Bortel, G. (2020). A circular economy life cycle costing model (CE-LCC) for building components. Resources, Conservation and Recycling, 161, 104857. https://doi.org/10.1016/j.resconrec.2020.104857

A Comparative Study on TiO_2/Graphite–PEG and Graphite/Carbon Fibre-Paraffin Shape Stabilized Phase Change Materials for Thermal Energy Storage Applications

Vimukthi Dananjaya[1], Qinning Wen[2], Chamil Abeykoon[2,*]
[1]School of Engineering, Swinburne University of Technology, Hawthorn, VIC 3122, Australia
[2] Northwest Composites Centre, Henry Royce Institute, National Graphene Institute, and Department of Materials, Faculty of Science and Engineering, The University of Manchester, Oxford Road, M13 9PL, Manchester, UK

ABSTRACT
Shape-stabilized phase change materials (SSPCMs) are promising candidates for latent heat thermal energy storage systems due to their high energy density and ability to prevent leakage during phase transitions. This study presents a comparative analysis of two SSPCM systems: TiO_2/graphite–polyethylene glycol (PEG) and graphite/carbon fibre/graphene–paraffin composites. Both composites were prepared by vacuum-assisted infiltration of molten PCMs into porous expanded graphite networks, with the addition of functional fillers to enhance structural integrity and thermal stability. Scanning electron microscopy (SEM) revealed distinct microstructural features for each system; TiO_2 nanoparticles were uniformly dispersed within the PEG matrix and anchored onto graphite surfaces, while carbon fibres and graphene nanoplatelets formed a hierarchical interconnected network within the paraffin-based composites. Differential scanning calorimetry (DSC) demonstrated that both systems preserved high latent heat storage capacities with slight shifts in phase transition temperatures compared to pure PCMs. Thermogravimetric analysis (TGA) showed improved thermal stability of the SSPCMs relative to neat PCMs, with filler composition significantly affecting degradation onset temperatures. In TiO_2/graphite–PEG composites, DSC analysis showed melting temperatures of 61.4-62.7 °C and solidification temperatures of 53.1-54.0 °C, with latent heats of 185-210 J g^{-1} depending on TiO_2 content. Graphite/carbon fibre/graphene–paraffin composites exhibited melting temperatures of 54.8-55.6 °C and solidification temperatures of 48.9-49.7 °C, with latent heats of 140-160 J g^{-1}. Thermogravimetric analysis revealed improved degradation onset temperatures: TiO_2/graphite-PEG composites showed higher thermal stability compared to pure PEG, while carbon fibre/graphene–paraffin composites exhibited enhanced thermal resistance relative to pure paraffin. The TiO_2/graphite-PEG composites exhibited higher latent heat capacities and enhanced thermal resistance, whereas the graphite/carbon fibre/graphene–paraffin composites provided superior mechanical reinforcement and phase change reliability. These findings offer insight into the design optimization of SSPCMs tailored for specific thermal management applications.

KEYWORDS: *Phase Change Materials, Hybrid Composites, Graphene, Nano titanium dioxide, Paraffin, Polyethylene Glycol, Thermal energy storage*

INTRODUCTION
Latent heat thermal energy storage (LHTES) systems are crucial for improving energy efficiency in applications ranging from renewable energy harvesting to building climate control (Stonehouse & Abeykoon, 2022; S. Wu et al., 2020; G. Yang et al., 2019). Organic phase change materials (PCMs), such as polyethylene glycol (PEG) and paraffin wax, are widely studied due to their high latent heat, chemical stability, and relatively safe handling properties (Kadoono & Ogura, 2014; Wong et al., 2023; Wuliu et al., 2021). However, the practical utilization of these materials is hindered by the inherent problem of PCM leakage during the solid-liquid phase change and limited thermal durability at elevated temperatures, which reduces system lifetime and effectiveness (Ouikhalfan et al., 2022).
Shape-stabilized phase change materials (SSPCMs) address these limitations by physically confining the PCM within a porous and rigid supporting matrix, which prevents leakage and maintains structural integrity throughout repeated thermal cycles (Falahatian et al., 2022). Expanded graphite (EG) is commonly used as such a matrix because of its highly porous structure, excellent thermal conductivity, and chemical compatibility with organic PCMs (Zhou et al., 2022). However, EG-based SSPCMs can benefit from the incorporation of functional fillers

to improve thermal stability, mechanical strength, and phase change reliability.

Metal oxide nanoparticles, such as titanium dioxide (TiO_2), are known to enhance thermal stability and mechanical properties of SSPCMs due to their ceramic nature and strong interactions with polymeric matrices (Venkitaraj & Suresh, 2019; J. Zhang et al., 2024; Zou et al., 2022). On the other hand, carbon-based fillers such as carbon fibres and graphene nanoplatelets contribute to the formation of mechanically robust, conductive networks that can reduce phase separation and improve the overall durability of PCM composites (Chen et al., 2021; Mishra et al., 2022).

Although previous research has separately explored TiO_2-enhanced PEG/graphite composites and carbon fibre/graphene-modified paraffin composites, a direct comparative evaluation of these two SSPCM systems remains limited (Sun et al., 2021; Zou et al., 2022). A comprehensive understanding of how filler type, PCM matrix, and microstructural characteristics affect phase change behaviour and thermal stability is critical for advancing SSPCM design.

This study undertakes a systematic comparison between TiO_2/graphite–PEG and graphite/carbon fibre/graphene–paraffin composites prepared via vacuum infiltration methods. Their microstructures, phase transition characteristics, and thermal stabilities are investigated through SEM, DSC, and TGA analyses. The outcomes elucidate the influence of filler chemistry and morphology on SSPCM performance, providing valuable guidance for tailoring materials to specific energy storage needs.

MATERIALS AND METHODS

Polyethylene glycol (PEG, molecular weight ~8000 g/mol) and commercial-grade paraffin wax were selected as the organic PCMs. Expanded graphite was prepared by thermal exfoliation of natural flake graphite at high temperature, resulting in a lightweight, highly porous network with interconnected pores suitable for PCM impregnation.

Titanium dioxide nanoparticles (~25 nm diameter) were used as ceramic fillers for the PEG-based composites. For the paraffin-based system, carbon fibres (3 mm length) and graphene nanoplatelets were employed to form a multi-scale reinforcing network within the graphite matrix.

The TiO_2/graphite–PEG composites were fabricated by first melting PEG at 80 °C and dispersing TiO_2 nanoparticles at weight fractions ranging from 1 to 5%. The dispersion was achieved by mechanical stirring to ensure uniform distribution. Expanded graphite samples were subjected to a vacuum to remove trapped air and then infiltrated with the molten PEG/TiO_2 mixture. The composite was cooled to room temperature to solidify the PCM within the graphite pores.

Similarly, the graphite/carbon fibre/graphene–paraffin composites were prepared by mixing expanded graphite, carbon fibre, and graphene nanoplatelets in targeted weight ratios. Molten paraffin wax was added under stirring, and vacuum-assisted infiltration was performed to achieve homogeneous filling of the graphite network. The composites were cooled to solidify the paraffin and fix the filler arrangement.

Microstructural analysis was conducted using scanning electron microscopy (SEM) to observe the dispersion and interaction of fillers within the PCM matrix and the graphite network. A Field Emission Gun Scanning Electron Microscope (FEG-SEM, Oxford Instruments) was employed to examine the microstructure of samples with thicknesses between 8 mm and 10 mm; non-conductive samples were sputter-coated with a 6.0 nm thick gold/palladium layer (80:20 ratio) using a 40 mA current for up to 5 minutes to ensure electrical conductivity and prevent static charge buildup during electron irradiation.

Phase change behaviour was assessed via differential scanning calorimetry (DSC), measuring melting and solidification temperatures as well as latent heat of fusion. Differential Scanning Calorimetry (DSC) was performed on samples weighing 2–10 mg sealed in aluminium pots, heated from –30°C to 120°C at a rate of 10°C/min, with thermal properties analyzed and data processed by TA Universal Analysis 2000 software.

Thermal stability was evaluated using thermogravimetric analysis (TGA) under nitrogen atmosphere, identifying degradation onset temperatures and mass loss profiles. Thermogravimetric analysis (TGA) was conducted on approximately 10 mg samples using a Q500 analyzer with a heating rate of 10°C/min from 30°C to 500°C under an inert atmosphere, and data were analyzed using TA Universal Analysis software.

RESULTS AND DISCUSSION

Microstructure

The scanning electron microscopy (SEM) micrographs in Figure 1 reveal distinct morphological features of various carbon-based and composite materials, enabling a clear comparison of their surface structures and potential implications for functional performance. In the upper-left panel, the graphene sample exhibits a characteristic sheet-like morphology with irregular wrinkling and folding. These thin, planar flakes, outlined by the highlighted regions, indicate the two-dimensional nature of the graphene structure, which is consistent with its high aspect ratio and large surface area. Such morphology is expected to enhance electrical conductivity and facilitate electron transport when integrated into composite systems (Fan et al., 2014; Xia et al., 2019).

The upper-right panel reveals the microstructure of carbon fibre, characterized by its long, continuous filamentous morphology. The fibres exhibit significant surface roughness and a highly entangled arrangement, which are critical features for enhancing mechanical interlocking within polymer matrices (Jiang et al., 2024). This roughened, fibrillar surface increases the effective contact area between the fibres and the surrounding matrix, promoting stronger interfacial adhesion (X. Wu et al., 2021). The entanglement and microfibrillar features observed are indicative of a highly irregular surface, which can influence phenomena such as wettability and adhesion at the microscopic scale.

In the lower-left panel, graphite is presented as large, compact aggregates composed of stacked platelets. The morphology displays tightly layered structures with relatively low edge compared to more exfoliated carbon forms. This compact platelet stacking is indicative of a stable layered arrangement, which contributes to the material's inherent structural integrity. The reduced surface area and limited edge sites suggest lower chemical reactivity at the particle surfaces compared to more dispersed carbon allotropes (Tian et al., 2016).

The lower-right panel illustrates the composite of graphite and TiO_2 nanoparticles, where graphite retains its plate-like morphology and is decorated with uniformly dispersed, near-spherical TiO_2 aggregates. The nearly homogeneous distribution of TiO_2 on graphite surfaces suggests effective surface coverage, potentially modifying the surface chemistry and optical properties of the composite (Venkitaraj & Suresh, 2019; Zou et al., 2022). The intimate contact between TiO_2 nanoparticles and graphite platelets may enable combined functionalities related to surface interactions, such as enhanced photocatalytic activity or UV absorption, stemming from the distinct properties of each phase (J. Zhang et al., 2024).

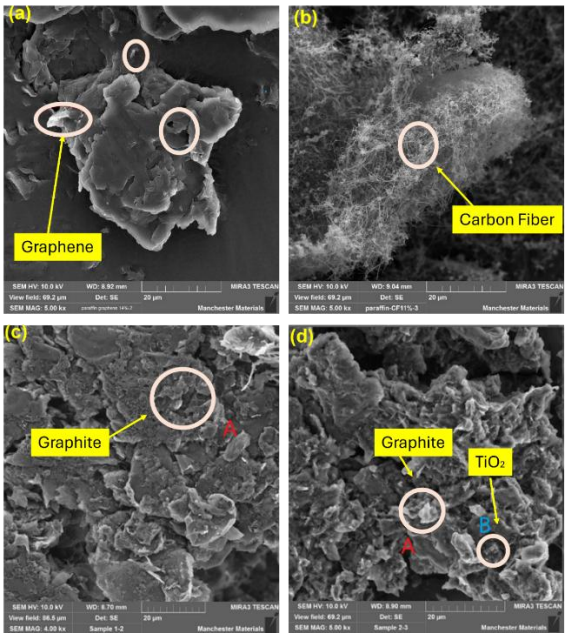


Figure 1. SEM images of (a). Graphene-Paraffin, (b). Carbon Fiber-Paraffin, (c). 35% Graphite 5% TiO_2 PEG and (d). 30% Graphite 10% TiO_2 PEG hybrid composites.

In contrast, the graphite/carbon fibre/graphene–paraffin composites exhibited a hierarchical structure formed by the interplay of carbon fibres bridging larger graphite pores and graphene nanoplatelets

coating the graphite surfaces and paraffin interfaces. This interconnected carbon network appeared to minimize PCM leakage pathways by physically confining paraffin within the graphite pores and DSC measurements showed that the melting temperatures of PEG in the TiO_2/graphite–PEG composites remained close to that of pure PEG, with only minor shifts (approximately 1–2 °C) towards lower temperatures as TiO_2 content increased. The latent heat of fusion for these composites was high, ranging from 185 to 210 J/g depending on filler loading, though a slight reduction was noted at higher TiO_2 concentrations due to the dilution of PCM by the ceramic filler. The narrow melting and solidification peaks indicated stable and reproducible phase transitions, with minimal supercooling effects observed (J. Zhang et al., 2022, 2024)

Thermal Stability

TGA analysis under a nitrogen atmosphere revealed distinct thermal behaviours for the enhancing the composite's structural cohesion (Ohayon-Lavi et al., 2021; W. Yang & Kim, 2023a; D. Zhang et al., 2018). The nanoscopic graphene sheets contributed to sealing microvoids, potentially promoting uniform phase transitions and mitigating paraffin migration (W. Yang & Kim, 2023b).

Phase Transition Behaviour

For the graphite/carbon fibre/graphene–paraffin composites, paraffin melting temperatures were similarly maintained near pure paraffin values, with small decreases of about 1 °C attributed to the presence of fillers. The latent heat ranged between 140 and 160 J/g, with a slight decrease at increasing graphene content above 2 wt%. Notably, the presence of graphene nanoplatelets sharpened the melting and solidification peaks, suggesting enhanced nucleation and improved crystallization kinetics during cooling cycles, which is beneficial for consistent thermal cycling (Chen et al., 2021; Wong et al., 2024). TiO_2/graphite–PEG and graphite/carbon fibre/graphene–paraffin composites as shown in Table 1.

Table 1. TGA data for PCM hybrid composites

Carbon Fibre (wt.%)	Graphene (wt.%)	Graphite (wt.%)	TiO_2 (wt.%)	Onset Temp (°C)	Temp at 5% Loss (°C)	First Degradation Peak	Final Degradation Temp (°C)	Final Mass Residual (%)
4	–	–	–	164.50	206.40	–	359.99	2.173
8	–	–	–	192.06	240.68	–	342.50	7.956
–	4	–	–	172.14	216.55	–	358.22	2.530
–	8	–	–	169.55	216.57	–	376.89	6.147
–	–	35	5	-	-	80.00	270.00	59–65
–	–	30	10	-	-	72.91	292.12	51.83

For the graphite/carbon fibre/graphene–paraffin composites, the onset temperature varied with filler type and loading: 164.50 °C (4 wt.% carbon fibre), 192.06 °C (8 wt.% carbon fibre), 172.14 °C (4 wt.% graphene), and 169.55 °C (8 wt.% graphene). The second peak (temperature at 5% loss) followed the same trend, increasing from 206.40 °C to 240.68 °C for carbon fibre, and from 216.55 °C to 216.57 °C for graphene. The higher loadings of carbon fibre produced the most notable improvements in thermal stability, accompanied by greater final mass residuals (7.956% for 8 wt.% carbon fibre versus 2.173% for 4 wt.%). Graphene-filled samples exhibited higher final degradation temperatures (376.89 °C for 8 wt.% graphene) compared to carbon fibre, suggesting that the continuous carbon network formed by graphene can further delay degradation by restricting oxygen diffusion and dispersing heat more evenly (Jiang et al., 2024; X. Wu et al., 2021). The continuous carbon filler network likely impedes thermal decomposition by restricting oxygen diffusion and dispersing heat more evenly throughout the composite (Zhu et al., 2023).

For the TiO_2/graphite–PEG composites, the first degradation peak (onset) occurred at 80.00 °C. for the graphite 35 wt.%–TiO_2 5 wt.% sample and 72.91 °C for the graphite 30 wt.%–TiO_2 10 wt.% sample. The second degradation peak (corresponding to 5% weight loss) was observed at 270.00 °C and 292.12 °C, respectively. Both composites exhibited earlier degradation onset temperatures compared to neat PEG, which began degrading around 310 °C; however, they demonstrated enhanced thermal stability through increased char residue and sooner major degradation, attributed to the barrier and heat dissipation effects of TiO_2 nanoparticles and porous structure after adding the fillers (Sun et al., 2021; J. Zhang et al., 2024). Final mass residuals were 59–65% and 51.83%, indicating the significant char yield retention of these formulations. This enhancement is attributed to the physical barrier effect and heat dissipation properties of TiO_2 nanoparticles, which slow the release of volatile degradation products.

Comparative Assessment

The TiO_2/graphite–PEG system demonstrates superior latent heat capacity, consistent with PEG's inherently higher heat of fusion compared to paraffin. Furthermore, the ceramic TiO_2 fillers confer improved high-temperature stability, extending the operational range of the PCM. However, increased TiO_2 content reduces PCM loading due to filler volume occupation, which may limit total energy storage.

In contrast, the graphite/carbon fibre/graphene–paraffin composites offer enhanced mechanical reinforcement and phase change reliability due to the multi-scale carbon network. This filler structure effectively suppresses leakage and stabilizes phase transitions, although the overall latent heat is somewhat lower than the PEG-based system.

Choosing between these SSPCMs involves trade-offs: the PEG–TiO_2 composites are preferable where high energy density and thermal endurance are paramount, whereas paraffin–carbon fibre/graphene composites are better suited to applications prioritizing mechanical stability and cycle life.

CONCLUSIONS

This work presents a detailed comparative analysis of two shape-stabilized phase change material systems aimed at thermal energy storage applications. TiO_2/graphite–PEG composites exhibited high latent heat capacities and improved thermal stability, facilitated by ceramic nanoparticle reinforcement within the porous graphite network. Conversely, graphite/carbon fibre/graphene–paraffin composites formed a hierarchical carbonaceous structure that enhanced mechanical integrity and stabilized phase change behaviour, though with slightly reduced latent heat storage.

SEM analysis revealed distinct filler dispersion and matrix interactions in each system, which correlated with the observed phase transition and thermal stability properties. The enhanced thermal resistance in TiO_2-containing composites and the superior structural cohesion in carbon fibre/graphene-filled paraffin composites demonstrate the importance of filler type and morphology in SSPCM design. These findings inform the selection and engineering of SSPCMs for specific thermal management requirements, enabling optimized performance in applications ranging from building temperature regulation to renewable energy systems.

REFERENCES

Chen, X., Cheng, P., Tang, Z., Xu, X., Gao, H., & Wang, G. (2021). Carbon-Based Composite Phase Change Materials for Thermal Energy Storage, Transfer, and Conversion. *Advanced Science*, *8*(9), 2001274. https://doi.org/10.1002/ADVS.202001274

Falahatian, M., Karimzadeh, F., & Raeissi, K. (2022). Preparation of paraffin/silica–graphene shape-stabilized composite phase change materials for thermal energy storage. *Journal of Materials Science: Materials in Electronics*, *33*(16), 12846–12856. https://doi.org/10.1007/s10854-022-08229-2

Fan, L. W., Zhu, Z. Q., Zeng, Y., Lu, Q., & Yu, Z. T. (2014). Heat transfer during melting of graphene-based composite phase change materials heated from below. *International Journal of Heat and Mass Transfer*, *79*, 94–104. https://doi.org/10.1016/J.IJHEATMASSTRANSFER.2014.08.001

Jiang, L., Zhao, L., Zhang, R., Zhang, W., Ma, X., Niu, Z., Chen, G., & Li, M. (2024). Research of the thermal storage properties of thermally conductive carbon fiber-reinforced paraffin/olefin block copolymer composite phase change materials with thermotropic flexibility. *Journal of Energy Storage*, *76*, 109761. https://doi.org/10.1016/J.EST.2023.109761

Kadoono, T., & Ogura, M. (2014). Heat storage properties of organic phase-change materials confined in the nanospace of mesoporous SBA-15 and CMK-3. *Physical Chemistry Chemical Physics*, *16*(12), 5495–5498. https://doi.org/10.1039/c3cp55429e

Mishra, R. K., Verma, K., Mishra, V., & Chaudhary, B. (2022). A review on carbon-based phase change materials for thermal energy storage. *Journal of Energy Storage*, *50*, 104166. https://doi.org/10.1016/J.EST.2022.104166

Ohayon-Lavi, A., Lavi, A., Alatawna, A., Ruse, E., Ziskind, G., & Regev, O. (2021). Graphite-based shape-stabilized composites for phase change material applications. *Renewable Energy*, *167*, 580–590. https://doi.org/10.1016/J.RENENE.2020.11.117

Ouikhalfan, M., Sarı, A., Hekimoğlu, G., Gencel, O., & Tyagi, V. V. (2022). Thermal energy storage properties, thermal conductivity, chemical/and thermal reliability of three different organic phase change materials doped with hexagonal boron nitride. *Surfaces and Interfaces*, *32*, 102176. https://doi.org/10.1016/J.SURFIN.2022.102176

Stonehouse, A., & Abeykoon, C. (2022). Thermal properties of phase change materials reinforced with multi-dimensional carbon nanomaterials. *International Journal of Heat and Mass Transfer*, *183*, 122166. https://doi.org/10.1016/j.ijheatmasstransfer.2021.122166

Sun, X., Yi, M., Feng, B., Liu, R., Sun, L., Zhai, L., Cao, H., & Zou, C. (2021). Shape-stabilized composite phase change material PEG@TiO2 through in situ encapsulation of PEG into 3D nanoporous TiO2 for thermal energy storage. *Renewable Energy*, *170*, 27–37. https://doi.org/10.1016/J.RENENE.2021.01.114

Tian, B., Yang, W., Luo, L., Wang, J., Zhang, K., Fan, J., Wu, J., & Xing, T. (2016). Synergistic enhancement of thermal conductivity for expanded graphite and carbon fiber in paraffin/EVA form-stable phase change materials. *Solar Energy*, *127*, 48–55. https://doi.org/10.1016/J.SOLENER.2016.01.011

Venkitaraj, K. P., & Suresh, S. (2019). Effects of Al2O3, CuO and TiO2 nanoparticles son thermal, phase transition and crystallization properties of solid-solid phase change material,. *Mechanics of Materials*, *128*, 64–88. https://doi.org/10.1016/J.MECHMAT.2018.10.004

Wong, T. Lo, Ma, K., & Abeykoon, C. (2024). Enhancing the thermal performance of polyethylene glycol phase change material with carbon-based fillers. *International Journal of Heat and Mass Transfer*, *220*, 124919. https://doi.org/10.1016/J.IJHEATMASSTRANSFER.2023.124919

Wong, T. Lo, Vallés, C., Nasser, A., & Abeykoon, C. (2023). Effects of boron-nitride-based nanomaterials on the thermal properties of composite organic phase change materials: A state-of-the-art review. *Renewable and Sustainable Energy Reviews*, *187*. https://doi.org/10.1016/j.rser.2023.113730

Wu, S., Yan, T., Kuai, Z., & Pan, W. (2020). Thermal conductivity enhancement on phase change materials for thermal energy storage: A review. *Energy Storage Materials*, *25*, 251–295. https://doi.org/10.1016/J.ENSM.2019.10.010

Wu, X., Shi, S., Wang, Y., Tang, B., Guo, L., Gao, Y., Jiang, T., Yang, K., Sun, K., Zhao, Y., Li, W., & Yu, J. (2021). Polyethylene Glycol-Calcium Chloride Phase Change Materials with High Thermal Conductivity and Excellent Shape Stability by Introducing Three-Dimensional Carbon/Carbon Fiber Felt. *ACS Omega*, *6*(48), 33033–33045. https://doi.org/10.1021/acsomega.1c05186

Wuliu, Y., Liu, J., Zhang, L., Wang, S., Liu, Y., Feng, J., & Liu, X. (2021). Design of bio-based organic phase change materials containing a “safety valve.” *Green Chemistry*, *23*(21), 8643–8656. https://doi.org/10.1039/D1GC02618F

Xia, Y., Zhang, H., Huang, P., Huang, C., Xu, F., Zou, Y., Chu, H., Yan, E., & Sun, L. (2019). Graphene-oxide-induced lamellar structures used to fabricate novel composite solid-solid phase change materials for thermal energy storage. *Chemical Engineering Journal*, *362*, 909–920. https://doi.org/10.1016/J.CEJ.2019.01.097

Yang, G., Yim, Y. J., Lee, J. W., Heo, Y. J., & Park, S. J. (2019). Carbon-Filled Organic Phase-Change Materials for Thermal Energy Storage: A Review. *Molecules*, *24*(11), 2055. https://doi.org/10.3390/MOLECULES24112055

Yang, W., & Kim, J. (2023a). A phase-change material with high through-plane thermal conductivity comprising a palmitic acid–cellulose nanofiber matrix with via acylation an expanded graphite@cellulose nanofiber@carbon nanotube hybrid filler. *Polymer Testing*, *118*, 107880. https://doi.org/10.1016/J.POLYMERTESTING.2022.107880

Yang, W., & Kim, J. (2023b). A phase-change material with high through-plane thermal conductivity comprising a palmitic acid–cellulose nanofiber matrix with via acylation an expanded graphite@cellulose nanofiber@carbon nanotube hybrid

filler. *Polymer Testing*, *118*, 107880. https://doi.org/10.1016/J.POLYMERTESTING.2022.107880

Zhang, D., Chen, M., Wu, S., Liu, Q., & Wan, J. (2018). Preparation of expanded graphite/polyethylene glycol composite phase change material for thermoregulation of asphalt binder. *Construction and Building Materials*, *169*, 513–521. https://doi.org/10.1016/J.CONBUILDMAT.2018.02.167

Zhang, J., Ji, Y., Zhang, Y., Wu, S., Wang, D., Xu, Z., Wei, Q., & Feng, Q. (2024). Preparation and temperature regulation properties of TiO2/CNTs-based form-stable phase change materials. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *685*, 133278. https://doi.org/10.1016/J.COLSURFA.2024.133278

Zhang, J., Sajadi, S. M., Chen, Y., Tlili, I., & Fagiry, M. A. (2022). Effects of Al2O3 and TiO2 nanoparticles in order to reduce the energy demand in the conventional buildings by integrating the solar collectors and phase change materials. *Sustainable Energy Technologies and Assessments*, *52*, 102114. https://doi.org/10.1016/J.SETA.2022.102114

Zhou, J., Fang, M., Yang, K., Lu, K., Fei, H., Mu, P., & He, R. (2022). MIL-101(Cr)-NH2/reduced graphene oxide composite carrier enhanced thermal conductivity and stability of shape-stabilized phase change materials for thermal energy management. *Journal of Energy Storage*, *52*, 104827. https://doi.org/10.1016/J.EST.2022.104827

Zhu, Z., Liu, Y., Zhang, Z., & Sheng, L. (2023). Studying carbon fiber composite phase change materials: Preparation method, thermal storage analysis and application of battery thermal management. *Journal of Energy Storage*, *67*, 107586. https://doi.org/10.1016/J.EST.2023.107586

Zou, L., Li, S., Li, L., Ji, W., Li, Y., & Cheng, X. (2022). Synthesis of TiO2 shell microcapsule-based phase change film with thermal energy storage and buffering capabilities. *Materials Today Sustainability*, *18*, 100119. https://doi.org/10.1016/J.MTSUST.2022.100119

A Numerical Investigation of Valve Timing and Intake Pressure Effects on Performance and Emissions in a Hydrogen Port Fuel Injection Engine

Isuru Wickramaarachchi[1, A *], Indrajith D. Nissanka[1,B] And Sameera Wijeyakulasuriya[2,C]

[1]Department of Mechanical Engineering, University of Moratuwa, Katubedda, Sri Lanka

[2]Convergent Science, Inc., 6400 Enterprise Lane, Madison, WI 53719, USA

[a]wickramaarachchiim.19@uom.lk, [b]nissankai@uom.lk, [c]sameera.wijeyakulasuriya@convergecfd.com

KEYWORDS: *Port Fuel; Valve Timing; Intake Pressure; Performance; Emissions*

ABSTRACT.
Hydrogen internal combustion engines (H_2ICEs) offer a viable low-emission alternative for decarbonizing transport, especially where full electrification is not practical. Among fueling strategies, port fuel injection (PFI) is particularly attractive due to its compatibility with existing engine platforms and simplicity compared to direct injection (DI). Performance and emissions in hydrogen PFI engines are strongly influenced by valve timing and intake boosting strategies. This study presents a computational framework to investigate the coupled effects of valve timing and intake pressure on the performance, thermal efficiency, and NO_x emissions of a hydrogen PFI engine under fuel-lean conditions (ϕ = 0.59). A modified Sandia optical engine geometry was simulated using CONVERGE CFD v4.1, employing detailed chemistry and adaptive mesh refinement. Latin Hypercube Sampling (LHS) was employed to generate 373 design cases that span a wide parametric space. Results show that intake boosting significantly improves performance, achieving a 220% increase in indicated power (up to 43.55 kW) and an 11% improvement in thermal efficiency (up to 48.7%) over the baseline configuration. However, these gains are accompanied by elevated NO_x emissions, particularly at higher valve overlaps. Conversely, the configuration that achieved the lowest NO_x emissions reduced them by 76% compared to the baseline, albeit at the expense of lower power and efficiency. The three configurations representing the most favorable outcomes for power, efficiency, and emissions within the studied parameter space highlight the inherent trade-offs among these objectives. These results provide practical guidance for calibrating hydrogen PFI engines and establish a solid foundation for future studies incorporating formal optimization methods.

INTRODUCTION

The continued rise in global greenhouse gas (GHG) emissions poses a critical challenge for policymakers, industry stakeholders, and researchers worldwide. The Intergovernmental Panel on Climate Change (IPCC) has stressed in its Sixth Assessment Synthesis Report that achieving the 1.5°C target requires rapid and sustained reductions in CO_2 and other GHGs [1]. In 2024 alone, energy-related CO_2 emissions reached 37.4 GT, marking a 410 MT increase from the previous year [2]. The transportation sector remains a significant contributor, accounting for over a quarter of these emissions, primarily from light-duty internal combustion engine (ICE) vehicles [3]. Battery electric vehicles (BEVs) have emerged as the leading strategy for decarbonizing transport. However, their overall environmental impact is closely tied to the carbon intensity of electricity generation. With over 60% of global electricity still derived from fossil fuels in 2024, the actual emissions reduction from BEVs remains limited in many regions [2]. BEVs also face infrastructural, economic, and operational challenges, particularly in high-load or long-distance applications [4].

Hydrogen has gained attention as a complementary energy carrier capable of overcoming some of these limitations. It offers zero tailpipe emissions and high gravimetric energy density, making it suitable for sectors where electrification remains impractical. Hydrogen can be utilized in both fuel cell electric vehicles (FCEVs) and hydrogen-fuelled internal combustion engines (H_2ICEs). While FCEVs offer high efficiency and quick refueling [5], they are hindered by high system costs, durability concerns, and limited fueling infrastructure [6]. In contrast, H_2ICEs offer a cost-effective and technically mature pathway, especially for retrofitting existing vehicle fleets. Hydrogen's favorable combustion properties—including high flame speed, wide flammability range, and low ignition energy—enable lean operation and reliable cold starts. Moreover, H_2ICEs are compatible with existing ICE manufacturing and servicing ecosystems, allowing rapid deployment with minimal

infrastructure changes [7]. Among fueling strategies, port fuel injection (PFI) and direct injection (DI) are commonly employed in hydrogen engines. While DI offers improved mixture control, it involves a higher cost and system complexity. PFI is simpler and more suitable for retrofitting gasoline engines. Recent studies indicate that hydrogen PFI engines can achieve high thermal efficiency and low emissions under ultra-lean conditions [8].

Within this context, two key engine parameters—valve timing and intake pressure—play crucial roles in controlling combustion behavior and emissions. Prior research has shown that both valve timing and intake boosting significantly influence hydrogen engine performance and emissions [9], [10], [11]. However, most existing studies are experimental and limited in parametric scope. High-fidelity CFD studies that investigate the coupled effects of valve timing and intake boosting under realistic operating conditions remain scarce. To address this gap, the present study develops a CFD-based computational framework to systematically explore the combined influence of valve timing and intake charge boosting in a hydrogen-fueled PFI engine. By capturing detailed in-cylinder phenomena, this work aims to identify configurations that improve performance, thermal efficiency, and emissions without relying on complex auxiliary systems such as water injection and EGR.

LITERATURE REVIEW

The Valve timing is a key determinant of combustion dynamics, engine performance, and emissions in hydrogen-fueled spark ignition engines. Park et al. [10] experimentally investigated intake and exhaust valve timing in a hydrogen PFI engine under high-load conditions. Their findings showed that moderately advanced intake valve opening (IVO) enhanced airflow and torque, while excessive advancement caused residual gas interference, diminishing performance. Retarded exhaust valve closing further reduced volumetric efficiency. This study underscored the need for optimal valve timing to enhance power output and thermal efficiency. Lee et al. [9] extended this understanding by demonstrating that retarding IVO to top dead center (TDC) enabled stable ultra-lean operation at $\phi = 0.2$ and achieved near-zero NO_x emissions. Although volumetric efficiency declined, performance was recovered using supercharging, achieving gasoline-equivalent power and thermal efficiency above 39%. This illustrated the synergistic potential of combining valve timing strategies with intake boosting. Huynh et al. [12] examined valve overlap time (VOT) variations from 0^0 to 50^0 CA and found that changes in overlap influenced mixture formation and combustion efficiency. While maximum torque was achieved around VOT = 30^0 CA, shorter overlaps still yielded satisfactory performance when energy input was adjusted. These findings collectively highlight the role of coordinated valve timing in optimizing engine output and reducing emissions under lean-burn conditions.

In parallel, intake boosting methods such as turbocharging and supercharging have proven effective in addressing the volumetric efficiency limitations inherent in hydrogen PFI engines. Jilakara et al. [13] experimentally evaluated a turbocharged, intercooled hydrogen engine designed for ultra-lean operation ($\phi = 0.5$–0.6) in a mini-bus application. Their setup delivered 66.7 kW at 3200 rpm while maintaining low combustion variability. However, turbo lag at low engine speeds remained a challenge. Verhelst et al. [14] explored supercharging options in a single-cylinder PFI engine and compared lean-burn and stoichiometric strategies. Lean supercharging ($\lambda \geq 2$) emerged as the more favorable approach, offering higher thermal efficiency and NO_x levels below 100 ppm across a broad speed range. The study also cautioned against increased NO_x formation at elevated pressures, which could be mitigated through EGR or leaner mixtures. Lee et al. [15] conducted a simulation-based analysis of turbocharged hydrogen engines and observed that moderate boost pressures improved power and airflow but also raised combustion temperatures, potentially increasing NO_x emissions under stoichiometric conditions. Nguyen et al. [16] corroborated these findings through experiments on a small-scale H_2ICE, confirming power enhancement from supercharging, while stressing the importance of spark timing optimization to maintain combustion efficiency.

Despite these contributions, existing studies typically focus on either valve timing or intake pressure in isolation, often within limited experimental setups. Few have examined their combined influence using high-resolution CFD approaches. As hydrogen PFI engines operate under complex, coupled thermochemical conditions, a comprehensive simulation-based exploration of valve timing and intake boosting interactions is essential. This study addresses that gap by providing a detailed CFD framework to inform future H_2ICE development.

NUMERICAL MODEL

The Due to the unavailability of a dedicated experimental research facility and the absence of a publicly available PFI engine geometry with sufficient experimental validation data, the single cylinder optically accessible

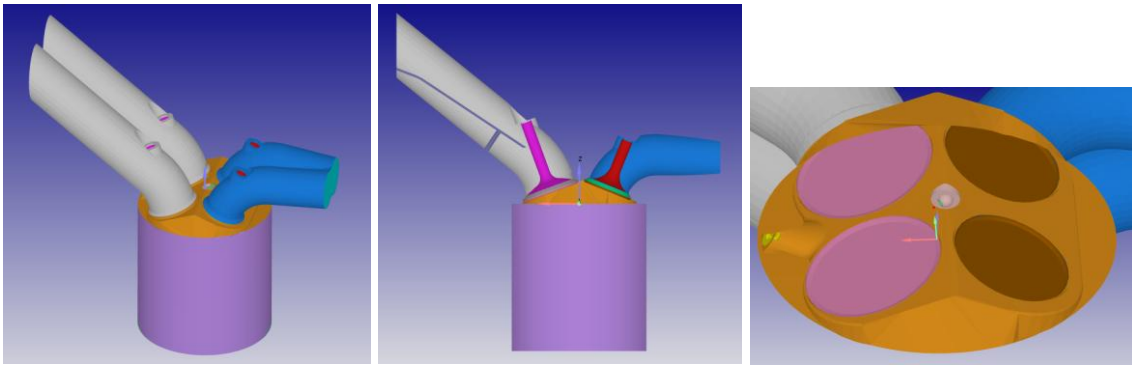
DI hydrogen engine developed by Sandia National Laboratories [12] was selected as the baseline geometry for this study. The engine specifications and operating conditions are summarized in Table 1.

Table 1: Engine specifications and operating conditions

Description	**Value**
Bore	92 [mm]
Stroke	85 [mm]
Displacement	565 [cc]
Compression ratio	10.91
Engine speed	1500 [rpm]
Intake valve timing (IVO/IVC)	346/ -140 [CAD] aTDC
Exhaust valve timing (EVO/EVC)	130/ 346 [CAD] aTDC
Intake temperature	309 K

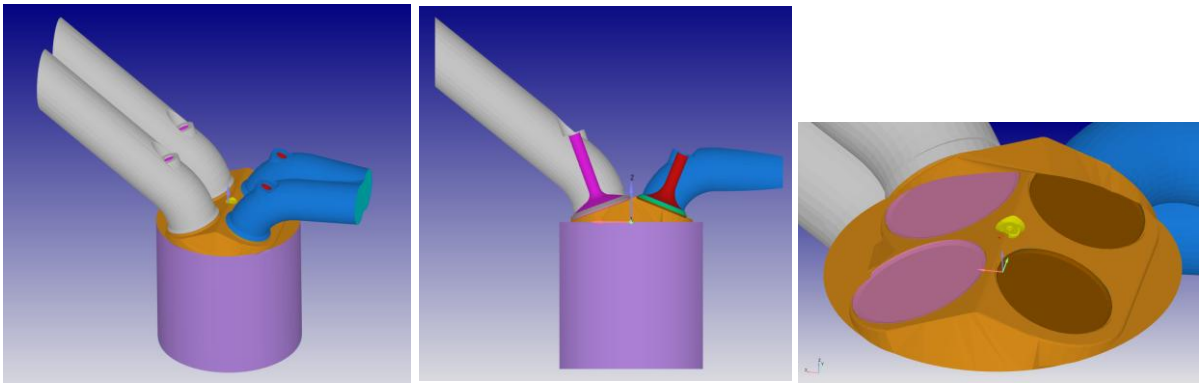
This engine geometry has been employed in numerous prior studies. Scarcelli et al. [13] developed a 3D CFD model to analyze hydrogen jet dynamics, showing that high mesh resolution near the injector is crucial for capturing jet penetration and stratification, with strong agreement to PLIF measurements. Barbato and Cantore [14] conducted full-cycle simulations and validated their adaptive mesh refinement strategy using experimental data, emphasizing the need for accurate time resolution during the injection phase. Wu et al. [15] employed CONVERGE to investigate hydrogen–air mixing under different in-cylinder flow conditions. The original engine geometry is available in CONVERGE [16] and is illustrated in Fig. 1. It consists of distinct intake and exhaust ports (Fig. 1a), a tumble plate integrated into the intake path to enhance in-cylinder motion (Fig. 1b), and a centrally located direct injector with a side-mounted spark plug on the cylinder head (Fig. 1c).

To facilitate the investigation of hydrogen PFI operation, the original DI geometry was modified. The key modifications included removing the direct injector, eliminating the tumble plate, and repositioning the spark plug to the center of the cylinder head. These changes are shown in Fig 2. The intake ports were retained to simulate the entry of a premixed hydrogen–air charge through conventional intake valves. Although these modifications substantially alter the original DI configuration, they are justified within the context of this numerical investigation, which focuses on analyzing relative trends in performance and emissions rather than predicting absolute values.



(a) (b) (c)

Figure 1: (a) Original Sandia DI engine geometry, (b) Cross-sectional view showing tumble plate, (c) Cylinder head showing DI injector and spark plug placement



(a) (b) (c)

Figure 2: (a) Modified engine geometry for PFI operation, (b) Cross-sectional view without tumble plate, (c) Top view showing removed injector and repositioned spark plug

A fully premixed, homogeneous hydrogen–air mixture was assumed to enter the computational domain through the intake ports, consistent with a dry PFI system. To investigate the effects of valve timing on combustion dynamics, valve lift profiles were adapted to allow variation in VOT. Simulations were conducted using CONVERGE CFD v4.1, a state-of-the-art general-purpose CFD software. Each case was simulated over a full engine cycle, beginning at exhaust valve closing (EVC) and continuing through 720^0 crank angle rotations. To accurately capture flow and combustion phenomena, the computational mesh employed adaptive mesh refinement (AMR), which locally increased resolution in regions with steep velocity or temperature gradients, such as flame fronts and flow structures. Fixed embedding was also applied in key regions of interest, including the intake and exhaust valves and the spark plug, where ignition and mixing processes are critical. A grid independence study was performed using a base mesh size of 4 mm, with progressively increased AMR levels based on velocity and temperature criteria. Refinement level 2 was selected as it provided a good balance between accuracy and computational cost. Under these settings, the minimum cell size in the combustion zone was refined to 0.0625 mm during the ignition phase and 1 mm during flame propagation, with the overall cell count reaching a maximum of approximately 0.5 million during peak refinement.

The simulations solved the three-dimensional, transient, compressible Reynolds-Averaged Navier–Stokes (RANS) equations for conservation of mass, momentum, energy, and species transport within a finite volume framework. Turbulence was modeled using the Renormalization Group (RNG) k-ε turbulence model [17], which is well-suited for swirling and recirculating in-cylinder flows. The momentum, energy, and species transport equations were discretized using a second-order central difference scheme within a finite volume framework. Time integration was performed using a first-order implicit Euler scheme with a variable time-stepping algorithm to ensure numerical stability and efficiency across the engine cycle. Pressure–velocity coupling was handled via the PISO algorithm [18], while the gas density was computed based on the energy equation using the Redlich–Kwong real gas equation of state [19]. Wall heat transfer was modeled using the O'Rourke–Amsden model [20], applied to all solid regions. This model estimates instantaneous convective heat transfer coefficients based on in-cylinder conditions and wall geometry.

Turbulent scalar transport was modeled using a turbulent Schmidt number of 0.78 and a turbulent Prandtl number of 0.9. Combustion was simulated using the SAGE detailed chemistry solver [16], coupled with a hydrogen mechanism extracted from the C3MechV3.3 kinetic model [21]. This mechanism includes 12 species and 37 reactions and has been validated for ignition delays (600–2500 K, 0.65–40 atm, φ = 0.06–6) and laminar flame speeds (295–600 K, 1–10 atm, φ = 0.4–2.2). The SAGE solver uses the CVODE integrator from the SUNDIALS package [22] to solve the stiff system of ordinary differential equations governing the chemical kinetics. These equations are solved with internal time steps smaller than the main CFD time step, and the computed reaction source terms are passed to the species transport equations. Molecular diffusivities for each species were calculated using data from a dedicated transport file (transport.dat) [23]. Spark ignition was modeled via multiple staged energy deposition events to replicate spark kernel formation and early flame growth under realistic engine conditions [24]. All simulations were performed under globally lean conditions, with an equivalence ratio of $\varphi = 0.59$ ($\lambda \approx 1.7$). Spark timing was fixed across all cases to isolate the effects of VOT and intake pressure on combustion characteristics and emissions. Thermal NO_x formation was modeled using the extended Zeldovich mechanism [25] under the assumption of local thermodynamic equilibrium.

At the intake ports, total pressure was specified according to the target intake pressure condition for each case

(ranging from atmospheric to 2.61325 bar), along with a total temperature of 309 K. At the exhaust ports, a fixed static pressure corresponding to atmospheric pressure was applied. All combustion chamber walls, piston crown, and cylinder liner were modeled as no-slip boundaries with convective heat transfer calculated using the O'Rourke–Amsden model. The global equivalence ratio was fixed at $\varphi = 0.59$ for all cases, and spark timing was held constant to isolate the effects of VOT and intake pressure.

METHODOLOGY

Design of experiments. Latin hypercube sampling (LHS) [26] was employed to generate a statistically robust and space-filling design of experiments for this study. The sampling strategy focused on two key input parameters: VOT, which varied from 0^0 to 50^0 CA, and intake pressure, which ranged from 1.01325 to 2.61325 bar. Using this method, a total of 373 simulation cases were created, ensuring a comprehensive exploration of the design space. The distribution of the sample cases is illustrated in Fig. 3. Each case was simulated using the CFD setup described earlier. To isolate the effects of VOT and intake pressure, all other parameters including the global equivalence ratio ($\phi = 0.59$, $\lambda \approx 1.7$) and spark timing were held constant across the simulations. The simulations were executed on the CONVERGE Horizon cloud computing platform [27], with each case requiring approximately 5 hours on a 36-core machine. By running 10 machines in parallel, 10 simulations could be completed within 5 hours, enabling efficient execution of the full 373-case design space. This approach ensured the timely completion of the high-throughput CFD campaign while maintaining numerical consistency across all cases.

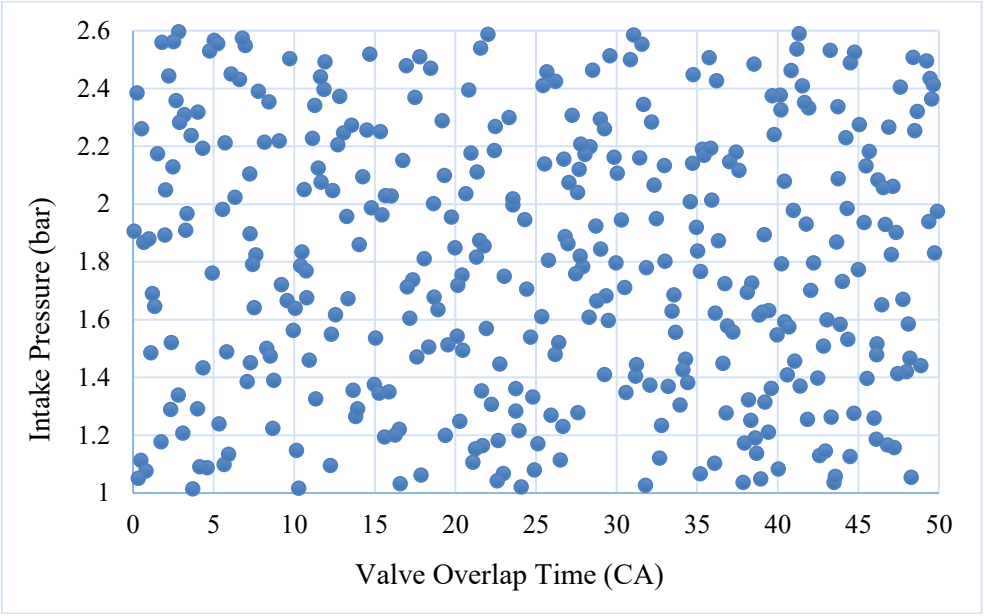


Figure 3: Latin Hypercube Sampling (LHS) distribution of 373 design points in the input parameter space

Performance and emission metrics. Each simulation output was post-processed to extract indicated power (kW), thermal efficiency (%), and NO_x Emissions (mg/cycle). These metrics were chosen to evaluate both engine performance and engine-out emissions.

RESULTS AND DISCUSSION

Effect of valve timing. As shown in Fig. 4, increasing intake pressure results in a clear upward shift in indicated power across the entire VOT range. For each pressure level, power generally increases with VOT, although the rate of increase diminishes at higher overlaps. At an intake pressure of 2.6 bar, the engine reaches peak power levels exceeding 43 kW at VOT > 35^0 CA, while the power remains below 15 kW at atmospheric pressure, highlighting the dominant effect of intake boosting on volumetric efficiency and combustion energy release. Fig. 5 complements these findings by revealing corresponding trends in thermal efficiency. Efficiency increases with both intake pressure and VOT, particularly at mid-range overlaps (15–35^0 CA), where optimized intake-exhaust interaction and mixture quality contribute to better combustion. The highest thermal efficiency

observed is approximately 48.7% at 2.6 bar, significantly outperforming lower intake pressure cases. Notably, while efficiency trends generally mirror power trends, the gains tend to plateau or even slightly decline beyond a certain VOT, suggesting diminishing benefits due to increased overlap-induced pumping losses or reduced volumetric efficiency at extreme settings.

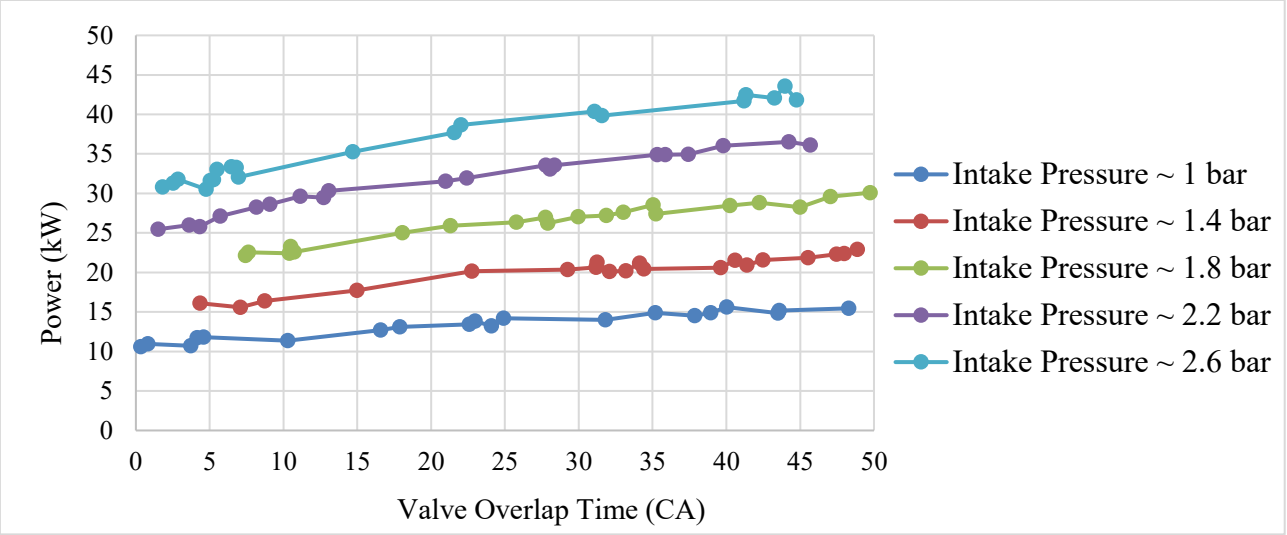


Figure 4: Effect of VOT on indicated power at various intake pressures

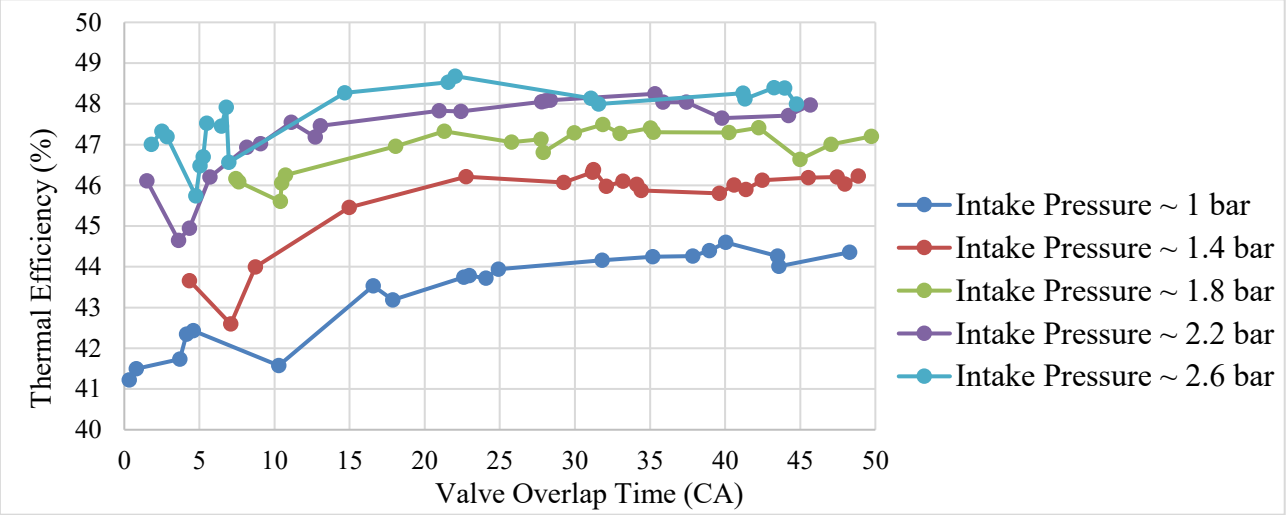


Figure 5: Effect of VOT on thermal efficiency at various intake pressures

Fig. 6 illustrates the variation of NO_x emissions, expressed as total NO_x mass per cycle, with VOT across five intake pressure levels ranging from approximately 1 bar to 2.6 bar. The results demonstrate that NO_x formation increases with both VOT and intake pressure, driven primarily by the elevated in-cylinder temperatures and oxygen availability associated with greater charge density and enhanced combustion intensity under boosted conditions. At low intake pressure (1–1.4 bar), NO_x emissions remain relatively modest and stable across the VOT range, suggesting limited peak temperature rise under lean conditions. However, as intake pressure increases to 2.2 and 2.6 bar, NO_x mass exhibits a pronounced upward trend with VOT, particularly beyond $20^0 – 25^0$ CA. This is attributed to longer valve overlaps promoting higher residual gas temperatures and more complete combustion, which intensify thermal NO_x generation. The highest NO_x values exceed 10 mg per cycle at 2.6 bar and high VOT, marking a significant emissions penalty associated with aggressive intake boosting and extended valve overlap.

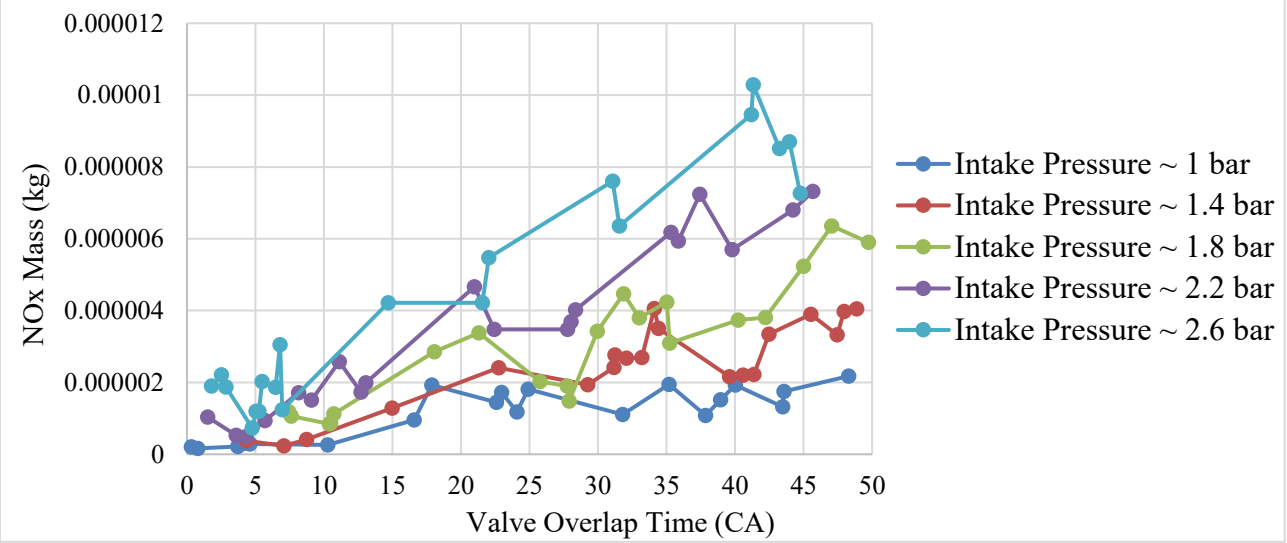


Figure 6: Effect of VOT on NO_x emissions at various intake pressures

Effect of intake pressure. Fig. 7 and 8 show the effect of intake pressure on indicated power and thermal efficiency. As intake pressure increases from 1.0 to 2.6 bar, both power output and thermal efficiency improve significantly due to higher charge density, enhanced volumetric efficiency, and more complete combustion. Intake boosting allows a greater mass of the hydrogen–air mixture to enter the cylinder, resulting in increased combustion energy and improved energy conversion. While power output shows a nearly linear increase with intake pressure, thermal efficiency gains tend to plateau at higher boost levels as secondary losses such as heat transfer and pumping work begin to offset further improvements.

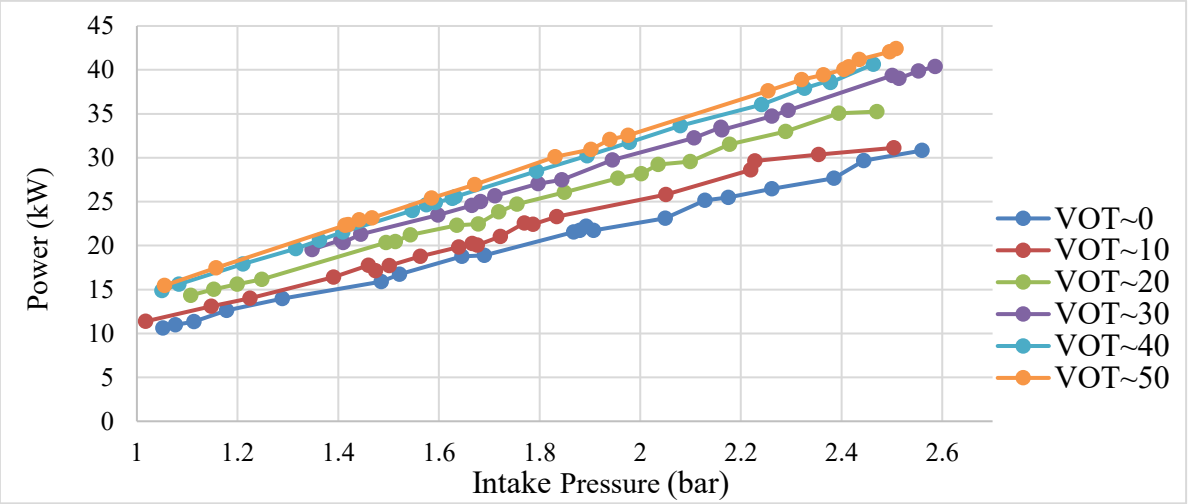


Figure 7: Effect of intake pressure on indicated power at various VOTs

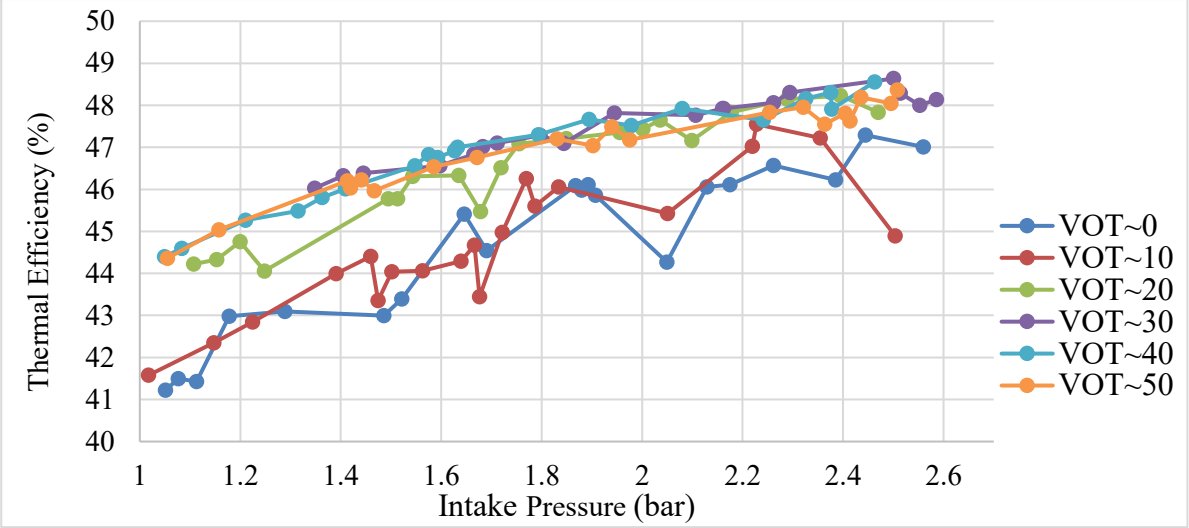


Figure 8: Effect of intake pressure on thermal efficiency at various VOTs

Fig. 9 illustrates the impact of intake pressure on NO_x emissions across various VOTs. As intake pressure increases from 1.0 to 2.6 bar, NO_x formation generally rises, with a more pronounced increase at higher VOT settings. This trend is attributed to elevated in-cylinder temperatures and oxygen availability under boosted conditions, which enhance thermal NO_x production. While low intake pressures result in minimal NO_x emissions regardless of VOT, higher pressures, especially above 2.0 bar, lead to a sharp rise in NO_x, particularly when combined with long valve overlaps.

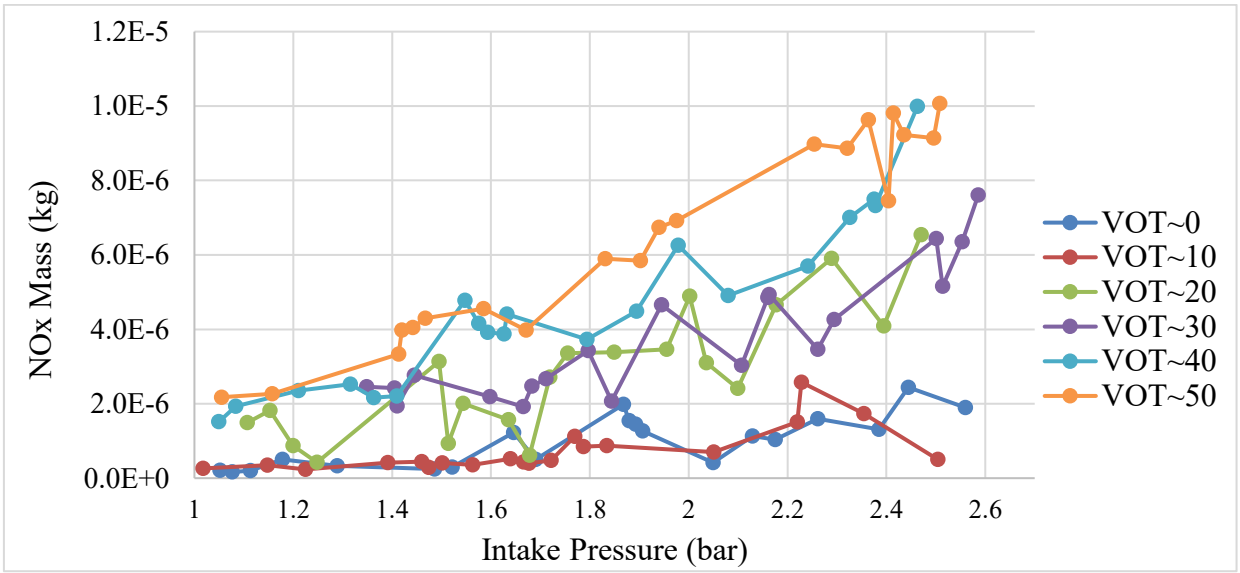


Figure 9: Effect of intake pressure on NO_x emissions at various VOTs

Best outcomes. Within the studied design space, three best operating points were identified, each corresponding to a different objective.

- Case A: Maximum power
- Case B: Maximum efficiency
- Case C: Minimum NO_x emissions

Table 2 presents the corresponding parameter combinations and highlights the trade-offs among power output, efficiency, and emissions. The baseline configuration is defined by a VOT of 30^0 and an intake pressure of 1.01325 bar. This case was selected as it represents a typical naturally aspirated setup and aligns with configurations previously identified as high performing for hydrogen PFI engines across a wide range of operating conditions [28]. Case A delivered the highest indicated power output of 43.55 kW, achieved at a VOT of 43^0 CA and an intake pressure of 2.61 bar. This represents a 221% increase in power compared to the baseline case, along with a thermal efficiency of 48.4%. However, this configuration resulted in elevated NO_x emissions of 8.69 mg per cycle. Case B yielded the maximum thermal efficiency of 48.7% at a VOT of 22^0 CA and an intake pressure of

2.59 bar. Compared to the baseline, this corresponds to an 11% improvement in efficiency and a substantial increase in power output to 38.66 kW, though with moderately high NO_x emissions of 5.47 mg per cycle. Case C prioritized emissions reduction, achieving ultra-low NO_x emissions of just 0.17 mg per cycle, which marks a 76% reduction relative to the baseline. This was accomplished with a near-zero VOT of 0.8^0 CA and a near-atmospheric intake pressure of 1.08 bar. However, this strategy compromised performance, with power output reduced to 10.96 kW and thermal efficiency to 41.5%, illustrating the trade-off necessary to meet stringent emissions targets.

Table 2: Best parameter combinations for maximum power, thermal efficiency, and minimum NO_x emissions

Case	**VOT (CA)**	**Intake Pressure (bar)**	**Power (kW)**	**Efficiency (%)**	**NO_x (mg/cycle)**
Baseline	30.0	1.01325	13.54	43.9	0.72
A	43.0	2.61	43.55	48.4	8.69
B	22.0	2.59	38.66	48.7	5.47
C	0.8	1.08	10.96	41.5	0.17

CONCLUSION

This study developed and applied a detailed CFD-based framework to evaluate the effects of valve timing and intake pressure on combustion performance and NO_x emissions in a hydrogen port fuel injection engine operating under lean conditions. A total of 373 full engine cycle simulations were performed using Latin Hypercube Sampling to span a broad range of intake pressures and valve overlap timings. The results confirm that intake boosting plays a dominant role in enhancing engine performance. The configuration yielding maximum power achieved 43.55 kW, corresponding to a 220% increase over the baseline, while the most efficient case achieved a thermal efficiency of 48.7%, representing an 11% improvement. However, these benefits came with increased NO_x emissions, particularly at high intake pressures and extended valve overlaps. The configuration that achieved the lowest NO_x emissions reduced them by 76% compared to the baseline, although this came at the expense of power and efficiency. These findings highlight the inherent trade-offs between power, efficiency, and emissions in hydrogen port fuel injection engine design. The developed framework provides a solid basis for refining engine calibration and lays the groundwork for future integration of data-driven optimization and control strategies.

FUTURE WORK

Building on the CFD-based framework developed in this study, future work will focus on integrating machine learning (ML)-based surrogate modeling to enable faster and more adaptive engine optimization. Advanced regression techniques and ensemble learning methods will be trained on the existing simulation dataset to predict engine performance and emissions with significantly reduced computational cost. Furthermore, the framework will be extended to incorporate abnormal combustion phenomena, such as backfire and combustion instabilities, which are critical considerations for the practical deployment and calibration of hydrogen PFI engines.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Convergent Science for providing CONVERGE CFD software licenses and technical support. The authors also thank Convergent Science for granting access to the CONVERGE Horizon cloud computing platform, which enabled the high-throughput simulations required for this research.

REFERENCES

[1] K. Calvin et al., "IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.," Intergovernmental Panel on Climate Change (IPCC), July 2023. doi: 10.59327/IPCC/AR6-9789291691647.

[2] IEA, "Global Energy Review 2025 – Analysis - IEA," IEA, Mar. 24, 2025. https://www.iea.org/reports/global-energy-review-2025

[3] IEA, "Global EV Outlook 2025 – Analysis - IEA," IEA, May 14, 2025.

https://www.iea.org/reports/global-ev-outlook-2025
[4] M. K. Anser, S. Ali, M. Umair, R. Javid, and M. Tayab, "Optimizing hydrogen integration in vehicle fuel systems for sustainable development: A step towards economic decarbonization," Int. J. Hydrog. Energy, vol. 98, pp. 321–333, Jan. 2025, doi: 10.1016/j.ijhydene.2024.12.009.
[5] Z. Stępień, "A Comprehensive Overview of Hydrogen-Fueled Internal Combustion Engines: Achievements and Future Challenges," Energies, vol. 14, no. 20, p. 6504, Oct. 2021, doi: 10.3390/en14206504.
[6] Y. He, Y. Zhou, Z. Wang, J. Liu, Z. Liu, and G. Zhang, "Quantification on fuel cell degradation and techno-economic analysis of a hydrogen-based grid-interactive residential energy sharing network with fuel-cell-powered vehicles," Appl. Energy, vol. 303, p. 117444, Dec. 2021, doi: 10.1016/j.apenergy.2021.117444.
[7] K. Naseem et al., "Essential parts of hydrogen economy: Hydrogen production, storage, transportation and application," Renew. Sustain. Energy Rev., vol. 210, p. 115196, Mar. 2025, doi: 10.1016/j.rser.2024.115196.
[8] A. H. Khalid et al., "Hydrogen port fuel injection: Review of fuel injection control strategies to mitigate backfire in internal combustion engine fuelled with hydrogen," Int. J. Hydrog. Energy, vol. 66, pp. 571–581, May 2024, doi: 10.1016/j.ijhydene.2024.04.087.
[9] J. Lee, K. Lee, J. Lee, and B. Anh, "High power performance with zero NOx emission in a hydrogen-fueled spark ignition engine by valve timing and lean boosting," Fuel, vol. 128, pp. 381–389, July 2014, doi: 10.1016/j.fuel.2014.03.010.
[10] C. Park, W. Park, Y. Kim, Y. Choi, and B. Lim, "Effect of Valve Timing and Excess Air Ratio on Torque in Hydrogen-Fueled Internal Combustion Engine for UAV," Energies, vol. 12, no. 5, p. 771, Feb. 2019, doi: 10.3390/en12050771.
[11] S. Verhelst, P. Maesschalck, N. Rombaut, and R. Sierens, "Increasing the power output of hydrogen internal combustion engines by means of supercharging and exhaust gas recirculation," Int. J. Hydrog. Energy, vol. 34, no. 10, pp. 4406–4412, May 2009, doi: 10.1016/j.ijhydene.2009.03.037.
[12] "Hydrogen Engine," Engine Combustion Network. Accessed: June 12, 2025. [Online]. Available: https://ecn.sandia.gov/engines/hydrogen-engine/
[13] R. Scarcelli, T. Wallner, N. Matthias, V. Salazar, and S. Kaiser, "Numerical and Optical Evolution of Gaseous Jets in Direct Injection Hydrogen Engines," presented at the SAE 2011 World Congress & Exhibition, Apr. 2011, pp. 2011-01–0675. doi: 10.4271/2011-01-0675.
[14] A. Barbato and G. Cantore, "3D CFD simulation of a gaseous fuel injection in a hydrogen-fueled internal combustion engine," E3S Web Conf., vol. 312, p. 07001, 2021, doi: 10.1051/e3sconf/202131207001.
[15] B. Wu, R. Torelli, and Y. Pei, "Numerical modeling of hydrogen mixing in a direct-injection engine fueled with gaseous hydrogen," Fuel, vol. 341, p. 127725, June 2023, doi: 10.1016/j.fuel.2023.127725.
[16] Richards, K. J., Senecal, P. K., and Pomraning, E., CONVERGE 4, Convergent Science, Madison, WI (2024).
[17] V. Yakhot and S. A. Orszag, "Renormalization group analysis of turbulence. I. Basic theory," J. Sci. Comput., vol. 1, no. 1, pp. 3–51, 1986, doi: 10.1007/BF01061452.
[18] R. I. Issa, "Solution of the implicitly discretised fluid flow equations by operator-splitting," J. Comput. Phys., vol. 62, no. 1, pp. 40–65, Jan. 1986, doi: 10.1016/0021-9991(86)90099-9.
[19] Otto. Redlich and J. N. S. Kwong, "On the Thermodynamics of Solutions. V. An Equation of State. Fugacities of Gaseous Solutions.," Chem. Rev., vol. 44, no. 1, pp. 233–244, Feb. 1949, doi: 10.1021/cr60137a013.
[20] A. A. Amsden, P. J. O'Rourke, and T. D. Butler, "KIVA-II: A computer program for chemically reactive flows with sprays," LA-11560-MS, 6228444, May 1989. doi: 10.2172/6228444.
[21] S. Dong et al., "A new detailed kinetic model for surrogate fuels: C3MechV3.3," Appl. Energy Combust. Sci., vol. 9, p. 100043, Mar. 2022, doi: 10.1016/j.jaecs.2021.100043.
[22] "SUNDIALS | Computing." Accessed: June 18, 2025. [Online]. Available: https://computing.llnl.gov/projects/sundials
[23] T. Coffee and J. Heimerl, "Transport algorithms for premixed, laminar steady-state flames," Combust. Flame, vol. 43, pp. 273–289, 1981, doi: 10.1016/0010-2180(81)90027-4.
[24] J. B. Heywood, Internal combustion engine fundamentals. in McGraw-Hill series in mechanical engineering. New York: McGraw-Hill, 1988.
[25] G. A. Lavoie, J. B. Heywood, and J. C. Keck, "Experimental and Theoretical Study of Nitric Oxide Formation in Internal Combustion Engines," Combust. Sci. Technol., vol. 1, no. 4, pp. 313–326, Feb. 1970, doi: 10.1080/00102206908952211.
[26] M. D. McKay, R. J. Beckman, and W. J. Conover, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, vol. 21, no. 2, p. 239, May

1979, doi: 10.2307/1268522.
[27] "CONVERGE Horizon," Convergecfd.com, 2025. https://convergecfd.com/products/horizon (accessed Aug. 16, 2025).
[28] T. C. Huynh, J. K. Kang, K. C. Noh, J. T. Lee, and J. A. Caton, "Controlling Backfire Using Changes of the Valve Overlap Period for a Hydrogen-Fueled Engine Using an External Mixture," in ASME 2007 Internal Combustion Engine Division Fall Technical Conference, Charleston, South Carolina, USA: ASMEDC, Jan. 2007, pp. 243–251. doi: 10.1115/ICEF2007-1702.

Diabetic Retinopathy Screening Using Image Processing

W.M.M.P.Wijesekara[a]*, A.E.S.H.Karunarathna[b], H.M.R.M.K.Haluwana[c], S.M.M.S Jayawardhana[d]

[a,b,c,d] *Department of Information and Communication Technology, Faculty of Technology,*
Rajarata University of Sri Lanka, Mihinthale, Sri Lanka
[a]wmmpramodya@gmail.com, [b]sulochanaharshani37@gmail, [d]msjayawa@tec.rjt.ac.lk

ABSTRACT

Diabetic retinopathy, a grave consequence of diabetes mellitus, has emerged as the leading cause of visual impairment worldwide. This ocular condition arises from the deterioration of blood vessels situated behind the retina and progresses insidiously, ultimately leading to blindness. Early detection is paramount in mitigating vision loss among afflicted individuals. In this study, we propose three distinct approaches Support Vector Machines (SVM), Convolutional Neural Networks (CNN), and Residual Networks (ResNets) for the accurate detection of diabetic retinopathy. Our aim is to determine the most effective model for this purpose, thereby improving screening efficiency. Utilizing a pre-processed dataset sourced from Kaggle, we conducted comprehensive experiments to evaluate the performance of each model. This curated dataset was instrumental in optimizing the classification algorithms. Our findings reveal notable disparities in the performance of these models. Through meticulous testing and validation, we sought to identify the model exhibiting the highest accuracy in diabetic retinopathy detection. Leveraging a dataset comprising 2750 retinal images, our experiments yielded accuracy values of 68% for SVM, 74% for CNN, and 63% for ResNet.

KEYWORDS: *Diabetic Retinopathy; Diabetes; Convolutional Neural Network; Support Vector*

INTRODUCTION

Diabetic Retinopathy (DR) is a severe complication of diabetes, leading to vision impairment and blindness globally. With over 400 million individuals affected by diabetes worldwide, a significant portion is at risk of developing DR. The urgency to detect and intervene early to prevent vision loss has sparked research interest in Diabetic Retinopathy screening using image processing.

The motivation for research in this field arises from the critical necessity to devise efficient methods for early DR detection. Often, DR progresses unnoticed until it reaches an advanced stage, limiting treatment options. Hence, regular eye screenings are recommended for diabetic patients to detect DR signs early when interventions are most effective. However, the massive number of screenings required poses a challenge to healthcare systems. Image processing techniques offer a potential solution by automating retinal image analysis, facilitating efficient DR screening and reducing the burden on ophthalmologists. Integrating artificial intelligence (AI) and machine learning algorithms further enhances diagnosis accuracy, enabling timely treatment and preventing vision loss.

Beyond individual patient care, research in Diabetic Retinopathy screening using image processing holds significance for the healthcare industry. Efficient screening methods can mitigate healthcare costs associated with advanced DR treatment and rehabilitation, addressing a critical public health concern and preserving the vision and quality of life of diabetic individuals. Moreover, advancements in this area contribute to broader medical image analysis, AI in healthcare, and the intersection of technology and medicine [1].

1. Stages Of Diabetic Retinopathy

1.1. Mild Nonproliferative Retinopathy

Mild Nonproliferative Retinopathy is the earliest stage of diabetic retinopathy, characterized by tiny swellings in the retinal blood vessels known as microaneurysms. These microaneurysms are small, balloon-like bulges that can leak fluid or blood into the surrounding retinal tissue, leading to subtle changes in vision that often go unnoticed in the beginning. At this stage, damage is usually minimal, and vision may still be normal, but the condition signals the first detectable signs of retinal blood vessel damage caused by diabetes. Early detection through regular eye exams is crucial to prevent progression to more severe stages. These swellings can cause fluid to leak into the retina [2].

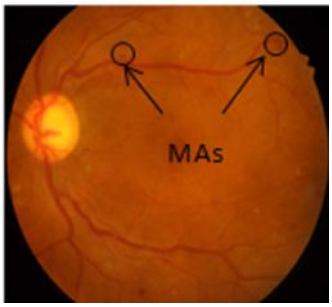


Figure 35: Mild DR (Stage 1)

2.2 Moderate Nonproliferative Retinopathy

Moderate Nonproliferative Retinopathy is marked by noticeable changes in the small blood vessels of the retina. As the condition advances, these vessels swell (microaneurysms) and become twisted or distorted, reducing their ability to transport blood effectively. This impaired blood flow can lead to areas of poor retinal nourishment, causing visible abnormalities in retinal appearance, such as small hemorrhages, leakage of fluid, and early signs of ischemia. At this stage, vision may still be unaffected or only mildly impaired, but the risk of progression to more severe forms of diabetic retinopathy increase [2].

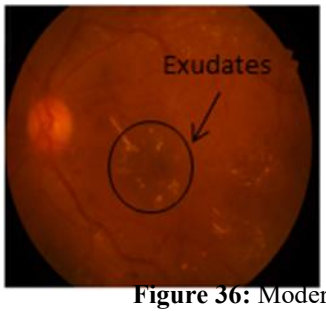


Figure 36: Moderate NPDR (Stage 2)

2.3 Severe Nonproliferative Retinopathy

In this stage, there is extensive blockage of the retinal blood vessels, which cuts off the blood supply to significant areas of the retina. As a result, these oxygen-deprived areas send strong distress signals that trigger the growth of new, fragile blood vessels. This stage marks a critical point in diabetic retinopathy progression, as the risk of

advancing to the more dangerous proliferative stage becomes significantly higher [2]

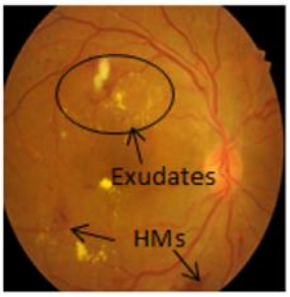


Figure 37: Severe NPDR (Stage 3)

2.4 Proliferative Diabetic Retinopathy (PDR)

Proliferative Diabetic Retinopathy (PDR) is the most severe stage of diabetic retinopathy. In this stage, abnormal new blood vessels begin to grow on the surface of the retina. However, these vessels are fragile and easily rupture, leading to bleeding inside the eye. This bleeding can stimulate the formation of scar tissue, which may pull on the retina and cause retinal detachment. Without timely treatment, these changes can lead to irreversible vision loss. It is the most severe stage of diabetic retinopathy. Here, the growth of new blood vessels is triggered, but these vessels are fragile and prone to leaking and bleeding. This can result in scar tissue forming and causing the retina to detach from the underlying tissue, leading to permanent vision loss.

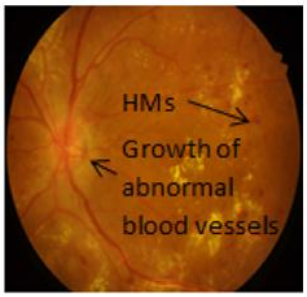


Figure 38: PDR (Stage 4)

LITERATURE REVIEW

The advancements in the late 20th and early 21st centuries have significantly improved the detection and treatment of diabetic retinopathy (DR), thereby reducing the risk of vision loss. Previous studies focused on detecting\various stages of DR through explicit feature extraction and classification using image processing techniques and machine learning algorithms.

In 2010, Singh and Chandra proposed a method for diabetic retinopathy detection based on lesion analysis, achieving an accuracy of up to 92% [3]. Similarly, Agurto et al. introduced a model in the same year that classified images based on lesion types, achieving comparable accuracy [4].

In 2013, Li Yafen et al. proposed a method employing various image processing techniques and an SVM classifier, achieving an accuracy of up to 89% [5].

Saleh & Eswaran introduced a method for automatically detecting no proliferative diabetic retinopathy. They focused on identifying two types of lesions: microaneurysms (MAs) and hemorrhages (HAs). By analyzing important features such as the optic disc, fovea, and blood vessels, they accurately segmented dark spots caused

by these lesions. Their system achieved sensitivity values of 84.31% and 87.53% for detecting MAs and HAs respectively, and specificity values of 93.63% and 95.08% for the same [6].

While these methods achieved high accuracy, explicit feature determination through image processing proved to be complex. With recent advancements in computer vision and the availability of large datasets, deep neural networks have become viable for DR detection and classification, requiring less preprocessing.

In 2016, Joshi proposed a method using deep convolutional neural networks, yielding promising accuracy [7]. Another study by Vishakha Chandore & Shivam Asati trained a deep CNN model on a large dataset, achieving up to 88% accuracy [8].

Diagnosis of Diabetic Retinopathy using Fundus Images and Image Processing Methods. The proposed system utilizes a real-time dataset of approximately 1,000 retinal fundus images from a healthcare institution to diagnose diabetic retinopathy. The system employs image processing techniques such as histogram equalization, top-hat analysis, Canny filtering, and morphological operations to detect and evaluate microaneurysms, exudates, and other optical abnormalities within the images, addressing the binary classification problem [9].

Retinal Image Lesions Assisted Diabetic Retinopathy Screening System Through Machine Learning The chapter discusses the use of deep and transfer literacy models for diabetic retinopathy discovery, fastening on the Kaggle dataset. It highlights challenges in dealing with limited and noisy datasets, emphasizing image preprocessing and data addition. The model integrates deep literacy and underpinning learning for discovery and classification tasks [10].

Specificity (True Negative Rate) Specificity measures the proportion of true negative cases correctly identified by the model among all actual negative cases. It reflects the capacity of the model to recognize cases without diabetic retinopathy accurately [11].

"Comparative Evaluation of Deep Learning Models for Diabetic Retinopathy Classification" This study compared the performance of multiple deep learning models, including CNNs and ResNets, for diabetic retinopathy classification. The study found that ResNet architectures outperformed traditional CNNs in terms of accuracy and sensitivity. The comparative analysis considered factors such as model complexity and training dataset size [12].

"Performance Comparison of Traditional Image Processing Techniques and Deep Learning Models for Diabetic Retinopathy Detection" This study aimed to compare the performance of traditional image processing techniques, such as feature extraction and morphological operations, with deep learning models in diabetic retinopathy detection. The research found that deep learning models, particularly CNNs, consistently outperformed traditional image processing methods in terms of sensitivity and overall accuracy [13].

METHOD AND STRUCTURE

4.1 The Dataset

The Kaggle dataset “Diabetic Retinopathy Dataset (DR Pre-processed Dataset)” **[14]** provided detailed information on retinal images essential for Diabetic Retinopathy detection and diagnosis. The dataset consisted of a total of 2750 images, each categorized into one of five class labels representing different stages of Diabetic Retinopathy (DR).

4.2 Data Pre-Processing

The dataset has already been pre-processed, and it is labeled into five main categories. The dataset underwent resizing to dimensions of 256 x 256 pixels.

Table 29: Diabetic Retinopathy Preprocessed Dataset

Name (Class Labels)	No.of Images	Percentage (%)
Healthy (Not DR)	1000	36.36
Mild DR	370	13.46
Moderate DR	900	32.73
Severe DR	190	6.9
Proliferative DR	290	10.55

4.3 Feature Extraction

Feature extraction plays a pivotal role in Diabetic Retinopathy (DR) screening through image processing, aiming to discern pertinent features from retinal images that distinguish healthy retinas from those afflicted by DR. For microaneurysms detection, techniques like the Hessian matrix or LoG (Laplacian of Gaussian) are commonly employed to identify circular blob-like structures characteristic of microaneurysms. Additionally, for exudates identification, intensity thresholding methods are applied to delineate bright lesions, indicative of exudates, based on intensity variations from the background. Furthermore, shape analysis techniques, including area, perimeter, and circularity measurements, are utilized to characterize exudate lesions. Moreover, for hemorrhages detection, color space transformation is utilized, converting images to different color spaces (e.g., RGB to HSV) to isolate hemorrhagic regions based on their distinct color characteristics. Lastly, for blood vessels segmentation, algorithms such as the Frangi filter or morphological operations are commonly employed to extract blood vessel structures from retinal images.

2. Training The Model

In this research, three different models were developed and trained to classify diabetic retinopathy (DR) stages using a labeled image dataset stored in Google Drive. The models include a Support Vector Machine (SVM), a Convolutional Neural Network (CNN) using EfficientNetB3, and a custom ResNet-style architecture.

The SVM model was trained using raw pixel features extracted directly from resized retinal images, where each image was flattened into a 1D vector. OpenCV was used to load the images, and train_test_split was used to divide the dataset into training and testing sets. The data was then fitted to a linear kernel SVM, and the accuracy_score function was used to determine the accuracy. Passing vectorized images directly into the SVM was computationally simple and allowed the model to operate without the need for handcrafted features. While conventional computer vision pipelines often involve extracting interest points, applying local descriptors (e.g., SIFT or SURF), and then encoding them using methods such as Bag of Visual Words (BoW), Fisher Vectors, or VLAD, this study instead relied on raw pixel intensities to reduce preprocessing complexity. However, more sophisticated handcrafted features such as Hessian-based key points could capture structural changes in retinal vasculature, lesions, or exudates that are highly relevant for DR classification. Hessian features are particularly useful for detecting blob-like structures and vessel crossings in retinal images, which are critical clinical indicators of DR severity.

For deep learning, the CNN model was built on EfficientNetB3 pretrained on ImageNet, with the top layers removed and replaced by custom dense layers, batch normalization, and dropout for regularization. The ResNet-inspired model, constructed from scratch, consisted of multiple convolutional layers followed by pooling operations to extract hierarchical image features. Both CNN and ResNet architectures were trained using the Adamax and Adam optimizers, respectively, with categorical cross-entropy loss, a learning rate of 0.0001, a batch size of 20, and 10 training epochs. Dropout layers with varying rates (0.2–0.4) and L1/L2 regularization were applied to reduce overfitting.

Data augmentation played a crucial role in improving generalization, particularly given the limited size and potential imbalance of the DR dataset. For CNN, the augmentation strategy included ZCA whitening, random rotations up to 30°, and nearest-neighbor filling, while color mode was maintained in RGB to preserve clinical

detail. These augmentations were chosen because DR lesions can occur at varying orientations, and whitening can help decorrelate pixel intensities to emphasize structural patterns such as microaneurysms and hemorrhages. The ResNet implementation used resized images without an equally extensive augmentation pipeline, which may have limited its performance.

Here, preliminary results showed cases where the SVM outperformed the ResNet-style network. This outcome could suggest that the handcrafted or raw feature space, while simple, was sufficient for certain separable patterns in the dataset, or that the ResNet implementation lacked optimization and validation. The absence of extensive augmentation and limited architectural depth in the custom ResNet could have contributed to underfitting. To ensure proper verification of deep models, a rigorous pipeline should include systematic hyperparameter tuning (learning rates, batch sizes, optimizer choices), use of data augmentation across all models, cross-validation to assess robustness, and potentially leveraging pretrained ResNet architectures instead of training from scratch.

By integrating proper augmentation strategies and verifying architectures through ablation studies, a more accurate comparative analysis can be achieved. This would allow for a clearer assessment of when classical machine learning methods like SVM may outperform deep networks in DR classification and when deep learning architectures are likely to yield superior results.

Each model was trained using retinal images resized to a standard size of 224 x 224pixels. The same parameters were applied to all three models to ensure consistency and fair comparison between them. We used Colab platform for model training and used Python as programming language.

Table 30: Training and Testing Parameters

Image Size	224 X 224 pixels
Testing data	30%
Training data	70%

RESULTS

The performance evaluation of the proposed models demonstrated varying levels of accuracy in diabetic retinopathy detection. The Support Vector Machine (SVM) model achieved an accuracy of 67.67%, while the Convolutional Neural Network (CNN) model outperformed it with an accuracy of 74.36%. The ResNet model yielded an accuracy of 63.2%, indicating lower performance compared to the other two approaches.

These results align with the study's aim of evaluating different machine learning and deep learning approaches for diabetic retinopathy classification. The higher performance of CNN compared to SVM suggests that deep feature extraction can capture more relevant patterns in retinal images than traditional feature-based methods. The comparatively lower accuracy of ResNet may be attributed to the limited dataset size, which is insufficient to fully leverage the complexity of deeper architectures.

When compared with findings from previous studies, the reported accuracies are within a similar range, although state-of-the-art approaches have achieved higher results often above 85% using larger, more balanced datasets and advanced preprocessing techniques. This highlights the potential for improvement through enhanced data quality, augmentation, and model optimization. Overall, the presented results provide a solid foundation for further refinement and contribute to the growing body of research on automated diabetic retinopathy detection.

Table 31: Final Results

Model	Training Accuracy (%)
SVM	68
CNN	74
Res Net	63

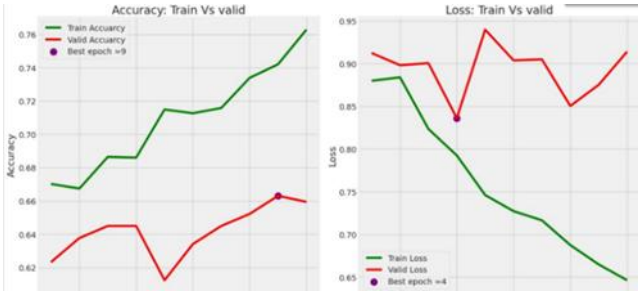


Figure 39: Accuracy and Loss of CNN Model

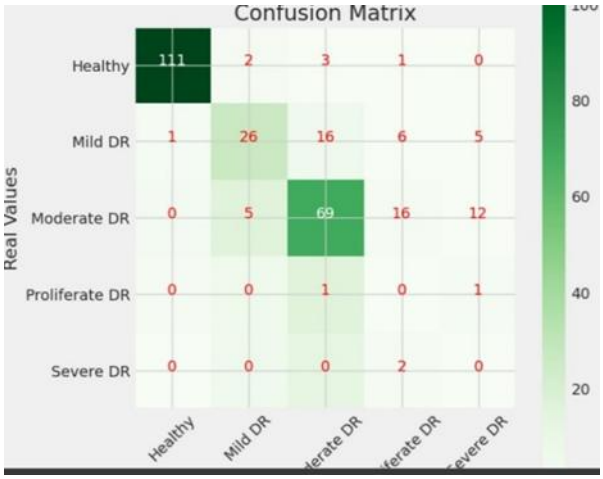


Figure 40: Confusion Matrix of CNN

CONCLUSION

This study explored three approaches for automated diabetic retinopathy detection using fundus images SVM, CNN, and ResNet. With images resized to 224 × 224 pixels and a 70:30 train test split, the SVM and CNN models each achieved an accuracy of 67.67%, while ResNet obtained 63%. Although CNN and SVM outperformed ResNet, all models demonstrated potential for DR classification.

However, several limitations were identified. The dataset size may have been insufficient for deep architectures such as ResNet, and preprocessing was restricted to image resizing without advanced enhancement techniques. Class imbalance was not explicitly addressed, potentially impacting performance on minority classes. Moreover, hyperparameter tuning was limited, and evaluation relied on a single dataset split, raising questions about model generalizability.

Future work will focus on expanding the dataset with diverse sources, incorporating advanced preprocessing (e.g., illumination correction, vessel segmentation), and employing class balancing strategies such as focal loss or weighted sampling. Systematic hyperparameter optimization, the adoption of deeper architectures (e.g., EfficientNet, DenseNet), and the use of k-fold cross-validation are also planned. Additionally, ensemble methods combining SVM, CNN, and ResNet predictions may further improve robustness and accuracy.

ACKNOWLEDGEMENTS

We extend our heartfelt gratitude to our supervisor, Miss S. M. M. S. Jayawardhana, as well as Mr. Wiraj Wickramaarachchi, Head of the Department of Information and Communication Technology, and Ms. Prasadini Uresha Padmasiri, Coordinator of the Research Project, for their unwavering guidance, support, and encouragement throughout our project, which has been crucial to our success.

REFERENCES

[1] "Diabetic Retinopathy," 04 01 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK560805/.

[2] "Fort Lauderdale," [Online]. Available: https://flei.com/retinal-diseases/diabetic-retinopathy/stages-of-diabetic-retinopathy.

[3] a. C. R. Singh N., "Automated early detection of diabetic," *International Journal of Computer Application,* vol. 8, October 2010.

[4] C. e. a. Agurto, "Multiscale am-fm methods for diabetic," *IEEE Transactions on Medical Imaging.,* no. 502-512, 2010.

[5] L. Yafen, "Automated Identification of Diabetic Retinopathy Stages Using Support Vector," 2013.

[6] E. C. Saleh MD, "An automated decision-support system for non-proliferative diabetic," *Comput Methods Prog Biomed.,* 2012.

[7] A. S. D. S. a. P. G. D. Doshi, "Diabetic retinopathy detection using deep convolutional neural networks," *2016 International Conference on Computing, Analytics and Security Trends (CAST),* 2016.

[8] S. A. Vishakha Chandore, "Automatic Detection of Diabetic Retinopathy using Deep Convolutional Neural Network," *International Journal of Advance Research, Ideas and Innovations in Technology.*

[9] A. Deshpande, ""Diagnosis of Diabetic Retinopathy"," vol. 11, p. 5, 2023-05-01.

[10] S. R. E. R. R. M. Y, "Retinal Image Lesions Assisted Diabetic Retinopathy Screening System Through Machine Learning," 2022-12-01.

[11] R. B. H. D. L. Sackett, "The architecture of diagnostic," *British Medical Journal,* vol. 324, p. 7336, 2002.

[12] C. E. o. D. L. M. f. D. R. Classification, "Comparative Evaluation of Deep Learning Models for Diabetic Retinopathy Classification," *Journal of Medical Imaging.*

[13] "Performance Comparison of Traditional Image Processing Techniques and Deep Learning Models for Diabetic Retinopathy Detection," *Journal of Biomedical Engineering.*

Penetratingz the Defenses: An Investigation into the Achilles' Heel of HTTP and SSH

Atharva Prashant Joshi [1], Shalini Bandrevu[2], Navneet Kaur[1], Ishu Sharma[3], M.W.P. Maduranga[4], WAAM Wanniarachchi[5]

Department of Computer Science and Engineering, Lovely Professional University, Phagwara, India. , Department of Computer Science and Engineering, Marri Laxman Reddy Institute of Technology, Hyderabad, India., Chandigarh Group of Colleges Jhanjeri, Mohali, Punjab, India – 140307 Chandigarh Engineering College, CSE-APEX, Dept. of Electrical and Electronic Engineering, Faculty of Engineering, University of Sri Jayewardenepura, Nugegoda, Sri Lanka., Department of Information Technology, Faculty of Computing, General Sir John Kotelawala Defence University.

atharvajoship2003@gmail.com, bandrevushalini0317@gmail.com, navneetphul@gmail.com, ishu.sharma001@gmail.com, pasanm@sjp.ac.lk, ashenw@kdu.ac.lk

ABSTRACT

As the yber threat increases, it becomes necessary for organizations to start securing their digital valuables and infrastructures. This research is mainly about analysing the weaknesses lying in the vulnerabilities of the HTTP and SSH protocols. Investigation here is into how the intruder can escalate his privileges and illegally access computers. Under open-source tools like Netcat and Gobuster, the article examines the vulnerability-assessment methodologies culminating in root access to the target machine. This paper emphasizes the need for proactive security measures and gives recommendations on improving defences against future attacks. The study, as bright as the findings may be, awaits empirical dimensions to affirm the proposed measures.

KEYWORDS: *Cybersecurity, HTTP vulnerabilities, SSH vulnerabilities, privilege escalation, reverse shell, penetration testing.*

INTRODUCTION

Cybersecurity is fast becoming a major issue for companies in all industries, given this digital world. The increasing number of Internet-connected devices, along with the services provided by them, has also initiated a parallel increase in the number and complexity of attacks targeting the fundamental network protocols, namely HTTP and SSH [1]. Such fundamental protocols, which are essential to communication online and secure remote access, are the common targets of cyberattacks that take advantage of misconfiguration, unfixed flaws, or even a lack of systems relying on them [2]. This is why protecting these protocols becomes extremely important in the protection of delicate data and the continuation of operations of contemporary organizations [3]. Although they form a key part of any computing environment, HTTP and SSH still feature several potential exploit vectors. As an example of vulnerability, we can cite the easy targeting of HTTP, which is used so widely in web-based communication and is prone to configuration mishaps or even inadvertent directory listing [4]. In contrast, SSH can be a liability when improperly configured, such as either via the use of weak authentication, or open firewall ports, which may be exploited by malicious users [5]. This code is mostly targeted in order to execute unauthorized access, privilege escalation, and take down whole systems [6]. The risks can hence be minimized given that the method of cyberattacks is always developing with time. The project employs practical tools based on penetration testing to investigate the vulnerability of HTTP and SSH and evaluate the security status of targeted systems with the use of free programs, such as Netcat and Gobuster to detect vulnerabilities that can be exploited in order to gain unauthorized access [7]. The techniques used present practical application on how attackers

can escalate their privileges, and several root access to malicious machines, thus providing useful information on how the exploitation of machines can be done. It is also the goal of the proposed study to improve the general knowledge of efficient protection methods rather than focusing on understanding only the HTTP and SSH vulnerabilities [8]. The other significant contribution made by the research is the illustration of Black-Box vulnerability assessment and exploitation methods in practical life. The study offers good information to a system defender on how to strengthen his system by explaining the methodologies that were followed in the exploitation of HTTP and SSH flaws [9, 10]. Also, this study provides some tactics for strengthening the defence, like the implementation of an effective configuration practice and proactive vulnerability scans [11].

RELATED WORK

RootAsRole David W. Chadwick et al. The paper [1] introduces a Linux Security module that aims to provide administrative privileges on a finer scale than, for example, the su or sudo command. In this module, role-based access is permitted and this means that administrators give certain capabilities to both a user, a group of users as well as applications thus limiting the chances of privilege escalation. It contains management commands to address roles and appraisal of application capabilities, improving control and limiting the possible exploitation of Linux systems. In the meantime, Keshav Kaushik et al. [2] pay attention to the reverse shell attack, under which attacked machines open the connections to the machines of an attacker. Their article describes how to install reverse shells, how to exploit application, network, or configuration vulnerability, and points out the dangers of data leakage and the loss of the entire system. Michael et al. [3] deal with the issue of privilege escalation in Linux-based computing clusters where the issue of security in cluster high-value computing is discussed. In their research, they suggest methods of detecting unauthorized privilege elevation and tools such as NVisionCC that can be used for this process, this action is a strategy commonly utilized by the attackers. Irma Resendez et al. [4] switch to the topic of digital forensics and describe the legal, technical, and procedural issues digital evidence processing rests on. They provide a broad range, including not only traditional computer forensics but also new areas such as network and cloud forensics, and place a particular focus on adhering to legal requirements when preparing and processing evidence and interpreting it during analysis. Kr. Boyanov et al. [5] present experimental contributions that involve practical applications of Netcat in Linux in terms of network security tests as it shows that controlling its operational exploits can be accomplished effectively in port scanning, banner capturing, and reverse shells. Likewise, Rizdqi et al. [6] introduce Sudomy, an efficient tool to enumerate and analyze the subdomains, which are important during reconnaissance stages of penetration testing. Lastly, Kaivan et al. [7] provide a comprehensive overview of the SSH protocol, including its combination of various cryptographic tools including the use of the public key infrastructure, symmetric encryption, and message authentication. They stress the contribution of SSH in facilitating safe communication of current remote and web services.

EXPERIMENTAL PHASE

The exploitation phase comprises four significant stages: reconnaissance, scanning, gaining access, and clearing trails. Information about the target system or network is obtained during the reconnaissance phase by the attackers. Next comes scanning to discover the weaknesses in readiness to be exploited. With such knowledge, the attacker uses the disclosed vulnerabilities and gets access to the district and establishes a foothold in it [14][15]. Finally, the attackers have their trail entered meticulously erased, thus minimizing the risk for detection, to conceal his existence as well as activity. Those four phases make up the foundation of the exploitation process, which enables an attacker to breach and navigate within target systems efficiently.

1.1. Reconnaissance or Information Gathering

Initial Gathering of Critical Information about the Target System Passively and Actively. We used netdiscover to run a search on the network for the Earth computer's IP address. The target's address ip found was 192.168.72.131. We also took nmap to perform a full scan of the open ports and services on this

machine. Two major services were discovered as follows: SSH on port 22 and HTTP on port 8080, which were later attacked. The above Fig 1 shows the steps undertaken to determine the IP address of the Kali Linux machine using the 'ifconfig' command. With this command, there is a detailed view of the IP addressing as well as other network interface related information. In Fig. 2, we used the command "netdiscover -r 192.168.72.0/24" to find the IP addresses within the given networks ranging from 192.168.72.0 to 192.168.72.255. This command actually scans and discovers devices in the network range specified. This is shown in Fig 3, which indicates the results obtained by executing the command in Fig 2, which describes which of the IP addresses were active on the network. After conducting thorough research, it was discovered that 192.168.72.131 is the designated target IP address for our future activities.

1.2. Scanning

After reconnaissance, a more intensive probing phase was launched to assess the vulnerabilities of exposed services. The nmap command was used with -sV, -sC, and -v options to determine service versions and run default scripts [16][17]. This resulted in effective and potentially exploitable information in the HTTP and SSH services identified. Gobuster, a very powerful directory brute-forcing tool, has been used for hidden directory and file enumeration on the HTTP service. Using SecLists' preconfigured wordlist, Gobuster discovered folders such as /admin and /cgi-bin, as well as sensitive files like .htaccess and robots.txt. These findings pointed to web server misconfigurations, which might provide entry points for exploitation. Here is command nmap -sV -sC -v -T4 192.168.72.131 in Fig. 4. This command has multiple options; "-sV" for port version detection, "-sC" for a default script scan, "-v" for activating verbose mode, and "-T4" for setting the timing template to level 4. The result of these scans is shown in Fig. 4 and Fig. 5, which will further give details regarding open ports, vulnerabilities present on the target machine, and the domain names "earth.local terratest.earth.local". These findings serve as a platform for further investigation and exploitation of the system's possible flaws.

1.3. Gaining access

With the target system's vulnerabilities identified, the next step was to acquire illegal access. The gobuster scan showed the existence of an admin panel for the HTTP service. We attempted to access the /admin directory and successfully signed in using the username "terra" and an encrypted password by brute-forcing login credentials obtained from the server. To further abuse the system, we used Netcat to create a reverse shell that allowed us to run commands on the Earth computer remotely. This was done by setting up a listener on the Kali Linux box, then configuring the target machine to connect back to this listener. With the following command nc -e /bin/bash, the reverse shell was launched, giving basic command-line access to the Earth computer. In Fig 6, we utilized the "nano /etc/hosts" command to access and modify the host file. This command enables the addition of IP addresses and corresponding domain names necessary for accessing the Earth Machines website. Fig 7 shows the command output from Fig 6 and the changes applied to the host file. These changes include the addition of the IP address 192.168.72.131 and the domain names "earth.local" and "terratest.earth.local." In Fig 8, we have searched for 192.168.72.131/earth.local.we are actually examining the domain earth.local which we discovered in the fig4 and fig5. In Fig 9, we see a message box, a message key and encrypted messages. In the message box and message key we have typed a random message to ensure that we are getting the output in the form of an encrypted message. These encrypted messages can be used in the further process of exploitation. In Fig 10, the command "gobuster dir -u http://earth.local/ -w /usr/share/wordlist/dirb/common.txt" was used. The '-w' denotes the wordlists used, whereas the '-u' denotes the URL of the target machine. The command result indicated the discovery of the 'admin' and 'cgi-bin' directories linked with the specified URL will be used in the further process of exploitation. In Fig. 11, we are accessing terratest.earth.local, which we have discovered from Fig. 4 and Fig. 5. After accessing that site, we can see that we are getting a warning message that this site is not secure, which means this site is using http protocol. In Fig. 12, the command "gobuster dir -u http://terratest.earth.local/ -k -w /usr/share/wordlist/dirb/common.txt" is used. '-w' indicates

the wordlists used, '-u' denotes the destination computer's URL, and '-k' disables SSL certificate checking. The command result indicated the discovery of the '/.hta', '/.htaccess', '/.htpasswd', '/cgi-bin', '/index.html', '/. robots.txt' directories linked with the specified URL. In Fig. 13, we access the '/.robots.txt' directory from the above Fig. 12. We got a list of subdirectories after examining all the subdirectories. We found that /testingnotes.* is leading us to further process of exploitation. In Fig. 14, we get to know that the encrypted messages in Fig. 9 are using the XOR algorithm, and testdata.txt was used to test encryption, and terra is used as a username in the admin portal, which we have discovered in Fig. 10**.** In Fig 15, we have gained access to the testdata.txt file identified in Fig 14. When we accessed testdata.txt, we discovered the XOR input key, which is critical for the next phases of the exploitation process.

1.4. Privilege Escalation

Once basic access had been achieved, the next step was to escalate privileges to get root access [18][19]. By examining files discovered during the scanning process, including testdata.txt, we discovered an XOR-based encryption mechanism used to safeguard passwords. We used the CyberChef program to decrypt these passwords, revealing the credentials required to elevate access. Further investigation discovered the availability of a vulnerable script (reset_root), which could be used to reset the root password. After getting access to the script, we updated the required files and ran it to successfully reset the root password. This enabled us to switch to the root user using the su root command, giving us complete administrative authority over the Earth computer. In Fig. 16, the CyberChef program is used to decode hexadecimal data using the XOR operation, with UTF-8 as the decryption key. The decryption key is determined from Fig. 15. After decrypting the complete hexadecimal sequence in Figure 8, the password for the username 'terra' is revealed. Fig. 17 shows that access to the earth.local/admin was achieved successfully. The required username, as shown in Fig. 14, has been matched with the correct password, as seen in Fig. 16. In Fig. 18, we successfully secured administrative control of earth.local. Upon logging into the admin portal, we gained command line interface (CLI) access, allowing us to communicate with the earth.local servers.In Fig 19, we ran the simple command 'whoami,' and the username was detected as 'apache.'. In Fig. 20, the command "nc -e /bin/bash 192.168.72.133 4444" is used. This command makes use of a tool known as netcat, which is widely used to establish remote access via a reverse shell or bind shell. Certain command-line restrictions are in place during the exploitation process to improve the security of reverse shell connections. To properly gain reverse shell access, the complete command must be encoded before being executed in the command shell.

In Fig. 21, an echo and base24 scripted message are used. This entails encrypting the command shown in Fig. 20 and then decrypting it using base24. The decryption command is then performed to allow remote access to the machine. In Fig. 22, we configured LVNP with the following options: 'l' for listen mode, 'v' for verbosity, 'n' to disable DNS or service resolution, and 'p' to define the source port. This configuration allowed us to obtain rudimentary terminal access. We also imported the Python modules "bin" and "bash" with the command python -c 'import pty; pty.spawn("/bin/bash")'. While browsing through numerous folders and executable files, we came upon 'usr/bin/reset_root'. However, on attempting to run the 'reset_root' file, it became clear that some triggers were missing. In Fig 23, we went to a separate terminal and used the netcat command on port 3333, which resulted in the successful download of the'reset root' file. In Fig 24, we examined the download status of the'reset_root' file. After confirming the successful download, we used the command "chmod +x reset_root" to grant the file executable permissions. Then, using the command "ltrace./reset_root," we traced the files executed by'reset_root' and discovered the lack of three critical triggers. In Fig. 25, we return to the previous terminal session from Fig. 10. We used the 'touch' command to create the missing files, ensuring that'reset_root' executed properly. In Fig. 26, we run the ' the'reset_root' file, which successfully resets the root password to "Earth". Apparently, in Fig. 27, the 'su root' command is to log in as a root user. With the access obtained, it enabled traversal of all the files on the machine and root capability through the entry of the password shown in Fig. 26.

RESULTS AND DISCUSSIONS

The penetration testing research experiments have portrayed how the commonplace utilization of the HTTP and SSH protocols can be used to earn unauthorized access and further privileges in a system [20][21]. Instead, we started the diversification by collecting information where Net Discover was used to detect the target IP (192.168.72.131), and a Nmap scan was conducted to check the result of finding two open ports, 22 (SSH) and 8080 (HTTP) [24]. This was used as the platform on which later scanning and exploitation were based. Vulnerability discovery employed Nmap to identify the service versions [25] and Gobuster to find hidden folders, including /admin, and/cgi-bin, as well as sensitive files, like. htaccess, all as a result of bad web server set-up [26][27][28]. These results were important to understand the weaknesses of the target system and possible routes to exploit it. Brute-forcing the /admin login with already known credentials was used as the means to gain access, along with a subsequent setup of the reverse shell through Netcat to perform commands on a remote machine [30][31]. Privilege escalation was reached by decrypting the password of the user terra in testdata.txt into CyberChef [32][33], and by abusing a vulnerable script, reset_root, that allowed to reset of the root password and subsequently achieved full control of the system [34][35][36]. Lastly, the logs were wiped after the exploitation to remove all evidence of the intrusion [37][38], which highlights the relevance of a proactive security approach, which includes protecting administrative scripts, properly securing sensitive information, having a solid monitoring and logging system, and so on; otherwise, the attack will go without a trace and it will not be mitigated in time.

CONCLUSION

It is clear from this study that there are important threats that exist within the HTTP and SSH protocols and that basic misconfigurations and risk practices are quite sufficient to cause a complete system breach. We could find some exposed services, exploit them with tools such as Netcat and Gobuster, and finally obtain root access, which was a privilege escalation. These observations indicate why there is a pressing need to constantly monitor and establish a proactive security position. The key learning points are that reconnaissance and vulnerability scanning form the base of any penetration testing process since they give the required information about the weaknesses of systems. This article revealed the mechanism through which publicly available information could be used by people to gain access (by finding exposed directories and improperly configured files). In addition to that, the relative conditions that allow reverse shells and brute-force attacks to be effective highlight the necessity of stricter authentication measures and configuration restrictions. The contribution of the study to the cybersecurity domain is that it provides a real-life case study that reflects the practice of exploitation of general security vulnerabilities in HTTP and SSH. Not only does it provide an example of how such attacks can be carried out, but it also provides workable information in reinforcement of defensive measures. It is advised to introduce strong authentication methods, periodically test the vulnerability of systems, and include stronger encryption along with limiting access to administrative scripts. In the future, it would be interesting to look further into the network defense mechanisms and the scenarios of more advanced attacks to enhance the understanding of threat mitigation. Also, real-time machine learning and automated monitoring systems might be highly beneficial with regard to early detection. With the ever-changing cyber threats, there is a need to balance the level of threat awareness and flexible security mechanisms in the bid to secure systems against complex attacks.

REFERENCES

1. Ahmad Samer Wazan, David W Chadwick, Remi Venant,Eddie Billoir, Romain Laborde, Liza Ahmad, Mustafa Kaiiali. RootAsRole: a security module to manage the administrative privileges for Linux. 2022.
2. Keshav Kaushik, Sakshi Aggarwalb , Shashank Mudgalc , Shubh Saravgid, Vibhor Mathure. A novel approach to generate a reverse shell: Exploitation and Prevention. 2021.
3. Michael Treaster, Gregory A. Koenig, Xin Meng, William Yurcik.Detection of Privilege Escalation for Linux Cluster Security. 2014.
4. Irma Resendez, Pablo Martinez, and John Abraham.An Introduction to Digital Forensics. 2014.
5. Petar Kr. Boyanov.Basic network penetration testing with the network tool Netcat in Linux-based Operating Systems. 2023.
6. Rizdqi Akbar Ramadhan, Redho Maland Aresta, Dedy Hariyadi.Sudomy: Information Gathering Tools for Subdomain Enumeration and Analysis.2020.
7. Kaivan Kaighobadi and Eduardo B. Fernandez.A Pattern for the Secure Shell Protocol.2012
8. Bruce A. Mah.An Empirical Model of HTTP Network Traffic.1997.
9. Toshihiro Yamauchi, Yohei Akao, Ryota Yoshitani , Yuichi Nakamura and Masaki Hashimoto.Additional Kernel Observer to Prevent Privilege Escalation Attacks by Focusing on System Call Privilege Changes.2018.
10. J. N. Goel and B. M. Mehtre, "Vulnerability assessment & penetration testing as a cyber defence technology," *Procedia Computer Science*, vol. 57, pp. 710-715, 2015.
11. Y. Liu, et al., "A novel exploit traffic traceback method based on session relationship," *CS & IT Conference Proceedings*, vol. 13, no. 7, 2023.
12. K. Kujanpää, W. Victor, and A. Ilin, "Automating privilege escalation with deep reinforcement learning," in *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security*, 2021.
13. B. Kallus, et al., "The HTTP Garden: Discovering parsing vulnerabilities in HTTP/1.1 implementations by differential fuzzing of request streams," *arXiv preprint arXiv:2405.17737*, 2024.
14. R. Andrews, D. A. Hahn, and A. G. Bardas, "Measuring the prevalence of the password authentication vulnerability in SSH," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, IEEE, 2020.
15. Y. Liu, et al., "An exploit traffic detection method based on reverse shell," *Applied Sciences*, vol. 13, no. 12, p. 7161, 2023.
16. Z. Hu, R. Beuran, and Y. Tan, "Automated penetration testing using deep reinforcement learning," in *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, IEEE, 2020.
17. A. Johnson and R. J. Haddad, "Evading signature-based antivirus software using custom reverse shell exploit," in *SoutheastCon 2021*, IEEE, 2021.
18. I. Koniaris, G. Papadimitriou, and P. Nicopolitidis, "Analysis and visualization of SSH attacks using honeypots," in *Eurocon 2013*, IEEE, 2013.
19. S. W. Woo, et al., "Modeling vulnerability discovery process in Apache and IIS HTTP servers," *Computers & Security*, vol. 30, no. 1, pp. 50-62, 2011.
20. M. Monshizadeh, P. Naldurg, and V. N. Venkatakrishnan, "Mace: Detecting privilege escalation vulnerabilities in web applications," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014.
21. S. Bugiel, et al., "Towards taming privilege-escalation attacks on Android," in *NDSS*, vol. 17, 2012.
22. S. Bugiel, et al., "Xmandroid: A new android evolution to mitigate privilege escalation attacks," Technische Universität Darmstadt, Technical Report TR-2011-04, 2011.
23. V. Visoottiviseth, et al., "PENTOS: Penetration testing tool for Internet of Thing devices," in *TENCON 2017-2017 IEEE Region 10 Conference*, 2017.
24. N. Gavric, G. Bhandari, and A. Shalaginov, "Introducing the Slowloris E-DoS Attack: a Threat Arising From Vulnerabilities in the FTP and SSH Protocols," 2024.

25. A. Z. Agghey, et al., "Detection of username enumeration attack on ssh protocol: Machine learning approach," *Symmetry*, vol. 13, no. 11, pp. 2192, 2021.
26. W. Alsabbagh, et al., "Hacking the Backbone: Shell Reverse Attacks on IIoT Systems."
27. P. S. Shinde and S. B. Ardhapurkar, "Cyber security analysis using vulnerability assessment and penetration testing," in *2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, 2016.
28. D. Suciu, et al., "Horizontal privilege escalation in trusted applications," in *29th USENIX Security Symposium (USENIX Security 20)*, 2020.
29. G. K. Sadasivam, C. Hota, and B. Anand, "Classification of SSH attacks using machine learning algorithms," in *2016 6th International Conference on IT Convergence and Security (ICITCS)*, 2016.
30. W. Qiang, et al., "PrivGuard: Protecting sensitive kernel data from privilege escalation attacks," *IEEE Access*, vol. 6, pp. 46584–46594, 2018.
31. I. Yaqoob, et al., "Penetration testing and vulnerability assessment," *Journal of Network Communications and Emerging Technologies (JNCET)*, vol. 7, no. 8, pp. 10–18, 2017.
32. S. Calzavara, et al., "Postcards from the post-http world: Amplification of https vulnerabilities in the web ecosystem," in *2019 IEEE Symposium on Security and Privacy (SP)*, 2019.
33. P. M. Cao, et al., "{CAUDIT}: Continuous Auditing of {SSH} Servers To Mitigate {Brute-Force} Attacks," in *16th USENIX Symposium on Networked Systems Design and Implementation (NSDI 19)*, 2019.
34. C. Ntantogian, et al., "NodeXP: NOde. js server-side JavaScript injection vulnerability detection and exploitation," *Journal of Information Security and Applications*, vol. 58, p. 102752, 2021.
35. S. N. H. B. Johari, et al., "Design of IMC-PID controller with fractional-order filter for steam distillation essential oil extraction process," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 2, pp. 801–810, 2021.
36. D. Trizna, et al., "Living-off-The-Land Reverse-Shell Detection by Informed Data Augmentation," *arXiv preprint arXiv:2402.18329*, 2024.
37. A. P. Joshi, N. Kaur, and S. Chauhan, "Encrypting the Unseen: Exploring Steganography Techniques in HTTP Environments," in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2024.
38. A. P. Joshi, N. Kaur, and A. Kaur, "Unveiling the Achilles' Heel: A Comprehensive Exploration of Vulnerabilities in RMI and Bindshell Communication," in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2024.

Construction Dynamics And Digitalization

Panadura Nihal Premachandra
Astrateq Pvt. Ltd, Colombo, Sri Lanka
nihalsl1@yahoo.com

ABSTRACT

The construction industry is at the edge of a decisive transformation, moving away from fragmented, paper-based practices toward an era defined by intelligent digitalization. At the center of this shift is the Digital Twin- a living, data rich model that synchronizes the physical and virtual realms of construction. By Integrating Building Information modeling (BIM), Internet of Things (IoT) sensors, artificial intelligence, and cloud computing, Digital Twins enable Projects to move from reactive monitoring to proactive, predictive control. This paper examines their influence on Construction dynamics, demonstrating how 4D scheduling with Primavera P6 and intuitive dashboards guided by PMBOK-7 principles elevate visibility, collaboration, and decision making. A case study of the Maldives International Airport new terminal illustrates tangible outcomes: real-time clash detection, optimized sequencing, energy efficient design, and measurable carbon emission reductions. Beyond showcasing benefits, the study outlines a pragmatic roadmap for Sri Lanka, stressing the importance of regulatory reform, academia-industry partnerships, and pilot implementations. The findings suggest that Digital Twins are not distant aspirations but present-day necessities for sustainable, data driven construction.

KEYWORDS *: Digital Twin, Construction Dynamics, BIM, Digitalization, PMBOK-7, 4D scheduling, sustainability.*

INTRODUCTION

For decades the construction industry has been constrained by analog practices- characterized by fragmented workflows, reactive decisions, and limited transparency. However, sectors like manufacturing and aerospace embraced digital reinvention decades ago, construction industry remained hesitant, trapped in inertia. That era is ending.

The modern construction site is no longer just concrete and steel; it is hybrid of the physical and digital interlaced through sensors, simulations, and synchronized data streams. At the center of the evolution is the **Digital Twin**: a real time, interactive model that evolves alongside its physical counterpart, enabling visibility, traceability, and proactive control at every phase.

Far from being academic jargon, Digital Twins represent a decisive leap in construction dynamics. Project stakeholders- from site Engineers to facility managers now engage with shared digital truth. Site activities are no longer interpreted through 2D drawings or siloed spreadsheets. Instead, progress is visualized dynamically, allowing deviations be detected and addressed long before they materialized on site.

This paper draws from direct, on the ground implementations of Digital Twins and digitalization strategies across complex infra-structure projects by the Author who is an industry professional with 38 years of design development and construction experience with both local and international contractors. Rather than remaining at the level of theory, this paper provides empirical insights into how digitalization improves construction performance, reduce inefficiencies and acceleration of sustainable delivery in the built environment.

UNDERSTANDING DIGITAL TWIN IN CONSTRUCTION

Definition and Fundamentals

A Digital Twin in construction is not just a digital replica- It is a living, evolving model that integrate real-time data with virtual simulations to monitor the behavior, condition and performance of physical asserts. It combines Building Information Modeling (BIM), IoT sensors, cloud computing and Artificial Intelligence (AI) into a unified platform that spans into design, construction and operation.

Unlike static models, a construction oriented Digital Twin evolves through progressive Levels Of Development (LOD) [1]. At LOD 400, It delivers fabrication-ready precision drawings like fabrication drawings and shop drawings; at LOD500 it becomes an authoritative as-built reference. These models enable direct control and connection with CNC machinery and robotics to execute tasks using geometry extracted straight from the twin, eliminating manual drawing interpretations.[2]

Moreover, tools such as Autodesk Revit function not only as design platforms but also a repositories of embedded intelligence within each component. Doors, ducts, and walls contain embedded data; thermal properties, Environmental information extracted from Environment Product Declarations (EPDs) certificate, about the life cycle of the products used and information required to carry out maintenance of the facility giving manufacturer data, warranty data, and spare parts lists and sources of obtaining such spare parts etc. helping maintenance cycles of the facility. This transforms the model into a functional database that guides both construction and future operations.

Sun path simulations further enhance their value. By modeling solar seasonal trajectories, design teams can optimize glass selection and shading strategies choosing: for instance, where to apply low E solar glazing to cut heat gain and boost daylight harvesting. The result: smarter envelope, lower cooling loads and improved energy efficiency.

Components of a construction digital twin

A fully realized construction digital twin include

- **Physical assert**- Built structure or system
- **Digital Replica**- BIM Modls, material data, MEP configurations, Maintenance data,
- **Data interface layer**- Inputs from IoT sensors, Augmented reality/ Virtual reality (AR/VR), which are immersive technologies that enhances user experiences by blending the digital and physical worlds, drones, and Radio Frequency Identifications (RFID) systems, which are type of wireless technology that uses radio waves to identify and track objects, people or animals.
- **Analytics core**- AI engines that forecast issues, detect anomalies, and optimize decisions.

These components combine to form specific Twin types:

- **Digital Twins**- For performance testing and iteration
- **Construction Twins**- For sequencing, logistics and clash detection
- **Operational Twins**- For energy monitoring and maintenance optimization

Together, they enable lifecycle continuity- from conception to long term facility management, anchoring construction dynamics in digital intelligence.



Fig. 1. Maldives International Airport, new terminal building, initially approved design, west elevation, Image from Digital Twin

CONSTRUCTION DYNAMICS AND THE ROLE OF DIGITALIZATION

Traditional challenges

Construction has long struggled with inefficiencies- poor communication, delays, and fragmented workflows. Decisions are often made late, based on outdated information and tracked through static spreadsheets or paper logs. This reactive model leaves little room for foresight or agility.

4D Optimization through P6 and Digital Twins

When primavera 6 is integrated with Digital Twins, schedule come alive. Each activity ID links to 3D BIM elements generating **4D simulation** where time and space converge. Rather than relying solely on Gantt charts, teams can **visualize the building virtually as it unfolds over time.** Delays in resource deployment, clashes and critical path risks appear in motion prompting immediate action before the site is impacted.

This is not just visualization, it is operational intelligence embedded in the construction flow.

Democratizing control with low code dashboards

Project performance can now be managed without requiring advance data science expertise. With tools like Power BI, schedule and cost metrics detailed in PMI-PMBOK 7[3]- Schedule Performance Index (SPI), Cost Performance Index (CPI), Earned Value (EV), To Complete Performance Index (TCPI), can be transformed into intuitive dashboards. These visuals update in real time, giving everyone, from Engineers to executives, the ability to grasp insights at a glance and respond immediately.

This shift from static data to dynamic interaction empowers faster more informed decisions across all levels of Project hierarchy.

A New Era of Proactive construction

Digitalization is no longer optional- It is central to modern construction dynamics. The fusion of modeling, scheduling and live analytics enable real time adaptation. Teams move from adaptation to leading, Waste drops, coordination improves, and foresight becomes the norm.

In the new paradigm, buildings are constructed with digital precision and managed with continuous clarity.

REAL WORLD APPLICATION- CASE STUDY: MALDIVES INTERNATIONAL AIRPORT, NEW TERMINAL BUILDING

The design and construction of the new international terminal building at Maldives International Airport is a landmark example of how Digital Twins and Digitalization can significantly enhance construction dynamics in a large scale EPC (Engineering, Procurement and Construction) projects.

INTEGRATED DIGITAL TWIN EXECUTION

The Maldives International Airport's New International terminal building exemplifies impact of **Digital Twin first approach** on complex infrastructure project delivery. Separate BIM models for Structural, Architectural and MEP works were federated into a single BIM environment including Airport system supplier's BIM Models like Passenger Boarding Bridges (PBB), Baggage Handling Systems (BHS), Advanced Visual Docking Guiding Systems (AVDGS) where each component precisely linked to the corresponding activity in the baseline schedule.
This integration enabled 4D simulations of construction sequence. Project teams could visualize work progress day-by-day, identifying schedule drift, bottle necks or coordination gaps in real time. The build becomes a live performance, rehearsed virtually and corrected before physical execution.

INTERFACE CONTROL AND CLASH RESOLUTION

Contractual mandates extend digital participation to all trades. Subcontractors and suppliers provided BIM models for their systems and equipment with embedded parameters. Airport systems like Passenger Boarding Bridges (PBB), Baggage Handling System (BHS) conveyors, façade panels were merged into master coordination BIM model.
Using Navisworks, digital clash detection exposed and resolved spatial conflicts pre-construction stage. This proactive digital coordination significantly cuts rework, reduces on site congestion, and prevents waste.



Fig. 2. Maldives International Airport, new terminal building, modified final design, west elevation, Image from Digital Twin

AIRSIDE SIMULATION FOR LAYOUT OPTIMIZATION

Airside planning was digitally simulated with true-scale Aircraft models [3], fuel hydrant points, and Advanced Visual Docking Guiding System (AVDGS). Virtual testing refined Aircraft positioning, nose wheel stops, Apron traffic flows-enhancing safety, reducing congestion, and enabling smooth operations from day 1



Fig. 3. Maldives International Airport, new terminal building, North-East Birds eye view, Image from Digital Twin

Climate responsive design adjustments

Through Digital Twin sun-path analysis, intense heat gain zones were identified on the East and West facades. Data driven energy modeling predicted over 700kg/day Co2 emission reduction from each of East and West faces of the building by mitigating thermal loads with the introduction of new façade. This resulted in 40 million USD worth shading retrofit- supported by simulation evidence rather than speculative assumptions.

Renewable energy integration and energy strategy

The Terminal was pre-engineered for solar panel integration and featured a marine thermal system designed to draw cooling energy load from seawater 25 km away and 5m deep spot below mean sea level at 24 deg. C. New International terminal is designed to serve 7.26 Million passengers per year in 2030 and the thermal energy extracted from sea water is sufficient to meet the energy requirements to operate the facility. This system forms the heart of the facility's **Net Zero Energy Building (NZEB)** road map- reducing long term Co2 emissions through smart utility planning.

Digitalization across the lifecycle

Delivered under full EPC turnkey Contract, the Project embedded digitalization from planning through operation. Digital Twin enabled agile design validation, informed procurement, optimized sequencing, and real time construction monitoring- resulting in timely delivery with lower carbon emissions.

The Maldives Airport project demonstrate how digitally empowered construction can be precise, efficient, and climate aligned- transforming airports from static facilities into **dynamic, data rich ecosystems**.



Fig. 4. Maldives International Airport, new terminal building, North-East Birds eye view, Image from Digital Twin

ROAD MAP TO ADOPTION IN SRILANKA

Building capacity through education

Digital transformation begins in academia. Engineering and Architecture programs must shift from traditional drafting to digital modeling, simulation, and Data Analytics. Universities should partner with Industry to launch joint research, pilot programs and skill-building labs.

The goal is to develop graduates who are not only competent but also digitally fluent

Updating regulatory standards

Sri Lanka's existing BIM guidelines must evolve to support Digital Twin integration. Updated regulations must cover LOD definitions, data interoperability, and lifecycle data continuity.
Incentives such as Green certification points, faster approvals, or carbon credits can accelerate adoption- turning policy into progress.

Launching pilot projects

Early momentum can come from high-impact public projects, Airports, Hospitals, Universities. These can serve as test beds, where the value of digitalization is demonstrated through real data: time saved, Carbon emissions cut, rework avoided.
Success stories will build confidence and catalyze broader national uptake.

CONCLUSION

Digital Twins are no longer futuristic concepts- they are foundational to the future of construction. Unlike static 3D models, Digital twins are dynamic systems that evolve with asserts, connecting design, construction, and operation through live data and intelligent feedback.
This paper has shown how digitalization reshapes construction dynamics- moving the industry from reactive, fragmented process to proactive data-driven control. **Maldives International Airport** showcased model driven coordination, real time sequencing, and energy optimized design.
These aren't isolated experiments, they are blueprints. For Engineers, Architects, and builders, application of these techniques is fast becoming non-negotiable. These tools don't just enhance efficiency, they underpin sustainable delivery, lifecycle intelligence, and digital accountability.
In the construction world of tomorrow, buildings won't simply be constructed; they will be computed, optimized, and continuously redefined by Data. That future is not distant- It is under construction today.

REFERENCES

[1] American institute of Architects (2023), Level Of Development (LOD) specification 2023, BIM Forum, https://bimforum.org/lod/
[2] Grieves, M. (2014), Digital Twin: Manufacturing excellence through vertual factory replication.
[3] Project Management institute (2021), Aguide to the Project mmanagement Body of Knowledge (PMBOK guide)- Seventh edition. Project Management Institute.
[4] International Orgnization for Standardization. (2018). ISO 19650: Organization and Digitization of information about buildings and civil engineering works, inclluding Building Information Modeling (BIM)
[5] McKinsey & Company. (2017). Re inventing construction: A route to higher productivity

Feature Analysis of Blood Spatter Patterns with Image Processing

Thepuli Khemaratne[a]*, Lakmini Malasinghe[a]

[a] *Department of Electrical and Electronic Engineering, Sri Lanka Institute of Information Technology, Malabe 10115, Sri Lanka*

*thepulikhemaratne2003@gmail.com

ABSTRACT

Bloodstain Pattern Analysis (BPA) is a vital component in forensic investigations that aids in reconstructing the sequence of events at a crime scene. It is centralized in and revolves around the categorization of the patterns based on their features, as this is the most significant and critical stage of BPA. Therefore, a preliminary measure of BPA is via the thorough evaluation of images photographed of the crime scene to collect evidence as much as possible to arrive at the correct conclusion and to deduce the relevant details accurately. However, currently existing BPA methods are vulnerable to subjectivity, hence which can lead to pre-assumptions, without thoroughly and completely observing the crime scene, and consequently cause the arrival of incorrect conclusions and discrepancies in BP feature classification. Additionally, other flaws such as unintentional crime scene contamination and evidence tampering exist in these current methods as well. Henceforth, it is imperative that a novel method is constructed to eliminate these issues and arrive at the correct conclusions. This study introduces a robust image-processing-based methodology for extracting and quantifying bloodstain pattern features, thereby enhancing objectivity and reducing human error. The proposed technique encompasses critical stages: image acquisition, preprocessing, segmentation, feature extraction, and analysis. Through the use of image enhancement and segmentation algorithms, essential attributes such as impact angles, tail-to-body ratios, shape irregularities, and distribution densities are computed. The results were validated against original findings and show close agreement in feature values such as convergence area and circularity. The approach demonstrates the potential to integrate with existing BPA tools, facilitating automated, accurate, and reproducible forensic analysis.

KEYWORDS: *Image processing, Forensic science, Bloodstain pattern analysis, Feature extraction, Classification algorithm, Impact angle*

INTRODUCTION

Bloodstain Pattern Analysis is an important forensic technique used to reconstruct and determine the events (including the type of impact, force, weapon, modus operandi, source…etc.) of a crime scene by the analysis of the bloodstain patterns and their features, conducted by forensic scientists, forensic analysists, and investigators (James, Kish, & Sutton, 2005). Effective extraction of these features and their analysis can provide valuable insights into the nature of the incident, the underlying mechanisms, and the aforementioned events/ details of the crime (Siu, Pender, Springer, Tulleners, & Ristenpart, 2017), (Swgstain, 2009). Yet, conventional BPA methods, though widely used, often suffer from subjectivity, potential contamination, and interpretational inconsistencies. Therefore, the need for a robust, impartial, and optimized approach rises, as the reliability of BPA relies heavily on the accuracy and objectivity of pattern interpretation, which remains a challenge due to human bias and potential contamination.

To address the necessity of objectivity in BPA, modern-day developments and advances fixate on incorporating quantitative measures to forensic analysis. Particularly, digital image processing techniques have proven to be powerful tools to enhance the analysis of forensic evidence by the extraction of these aforementioned BP features. These approaches have been successfully applied to various branches of criminology and forensic science, including shoeprint analysis, fingerprint recognition, surveillance footage

enhancement, and DNA matching. Yet, their application in BPA seems to remain underused (Arthur, et al., 2017), (James, Kish, & Sutton, 2005). From this research paper, an image processing methodology (by following IP techniques and an algorithm) is executed whose main focus and aim is to extract crucial BP features in an automated manner, and therefore increasing the accuracy, precision and efficiency of BPA (Arthur, Cockerton, de Bruin, & Taylor, 2015). This methodology encompasses multiple key image processing steps, such as image acquisition, pre-processing, element segmentation, feature extraction, and the evaluation of the results obtained.

Thus, the intended objectives are as follows:

- To devise an image processing methodology to analyze features of a bloodstain pattern with less complexity, less computational power consumption, and minimal resources to produce results with satisfactory accuracy.
- To deduce the area of convergence to determine the vicinity of origin.

In the first step: image acquisition, unless a pre-existing image is obtained from another source, the original image of the bloodstained surface is typically captured to obtain a high-resolution photograph, by using suitable imaging equipment, i.e., digital cameras. After this image acquisition step, in the preprocessing stage where the image is subjected to quality enhancement and reduction of noise via image filtering, contrast enhancement…etc. (Arthur, et al., 2017).

The ultimate goal of this research is to enhance the objectivity of BPA by the provision of reliable numerical data that can back up their conclusions for credibility. By the devising of a foundation of quantizable features, this methodology is able to contribute to the development of automated classification of BP features. Furthermore, this explores the potential of the fusion of this method with currently existing BPA methods for the improvement of accuracy, precision, and consistency. Thus, the rising demand of the entities involved in forensic science is also addressed, as the subjectivity and partiality are decreased by this method. The aim in this research is to develop the field in forensic science by the exhibition of the value of such methods in BPA, which will henceforth contribute to more reliable and accurate crime reconstructions and evidence analysis (Arthur, et al., 2017).

METHODOLOGY

In this research report, as the data input, the image of the bloodstain pattern which was given in the original research paper itself was used, since due to resource limitations, that was the only option available. Also, to conduct the image processing methodology, only the *MATLAB R2021a* software was used to compile the C++ program. The methodology is depicted in the overview diagram illustrated below in Figure 1:

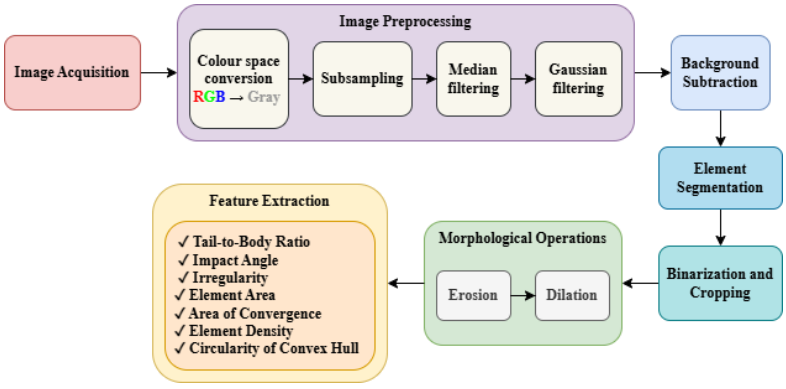


Figure 41. Overview diagram of the Methodology

Image Acquisition

First and foremost, the original image (Figure 2) given in the original/ reference research paper itself was directly extracted due to limitations in resources instead of typically capturing from a new or reenacted/ recreated bloodstained scene using an image capturing device. Secondly, its resolution was measured and compared to that of the original image (Arthur, et al., 2017). Then, it was found that the resolution of this

image was significantly lower than that of the original image, as mentioned in it. Thus, to resize it to its original dimensions and restore it to attempt to increase its resolution, this image was resized to its original dimensions by using the *imresize* function, with the width and height of the original image defined as an array of pixels. (Lazzaro, Murra, & Schwortz, 2013)

Image Preprocessing and Background Subtraction

Secondly, the image pre-processing steps were followed, in order to enhance image quality and suppress noise, ensuring robust feature extraction. First, the extracted image in RGB colour scale was transformed to gray scale, since it eases computational process in image processing, as it is simpler to deal with a single colour plane ranging from 0 – 255 than that of three separate colour planes, and hence increasing the efficiency in the consequent preprocessing steps, reducing the computational load and memory requirements, and easing the visualization (Gonzalez & Woods, 2009).



Figure 42. Extracted image of the bloodstain pattern

Next, it was subjected to subsampling by a factor of 8, where the pixels are removed from the columns and rows eight-folds, and a median filter of a window size of 50 was applied. Then, the subsampled image underwent background estimation via a Gaussian filter with a sigma value of 50. Afterwards, this was subjected to background subtraction from the gray scaled image. The gray scaled image and the background subtracted image are as given under the Results section as Figure 5 and Figure 6 respectively (Arthur, et al., 2017).

Element Segmentation

Thirdly, the brightness/ intensity histogram of the gray scaled image was obtained to perform segmentation of bloodstain elements. Here, the Triangle Thresholding algorithm was used instead of Otsu's Thresholding algorithm, since unlike other thresholding algorithms and typically used Otsu's Thresholding, the Triangle Thresholding also considers weak and subtle foreground intensities, i.e., in this case degraded or faint bloodstains, and performs segmentation on them better as well. This intensity histogram along with the comparison of Triangle and Otsu's Thresholds was acquired as evident in Figure 7 (Arthur, et al., 2017). Additionally, the threshold value for the Triangle Thresholding algorithm was deduced as well.

Subsequently, the gray scaled image was binarized such that the elements are in black colour (0 intensity) and background in white (255 intensity) and cropped along all four borders by a margin of 1 pixel each. Next, the ruler present in the original image was removed by using a white rectangle of appropriate estimated size, resulting in the image given in Figure 8 in the Results section (Arthur, et al., 2017), (Arthur, Cockerton, de Bruin, & Taylor, 2015).

Morphological Operations

Each bloodstain element is typically comprised of a body (the elliptical or circular shaped part) and a tail (thin, elongated end part), as depicted in Figure 3 below:

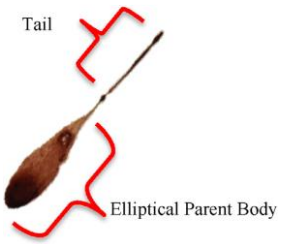


Figure 43. Parts of a typical blood spatter element (a singular bloodstain)

Consequently, the tail parts of each blood spatter element in the pattern was removed by means of morphological operations (Zou & Stern, 2022).
First, the (resultant) cropped binary image was subjected to four iterations of binary erosion using a structuring element of a diamond with the size of 1 pixel defined as $strel$. This eroded image then underwent binary dilation for four iterations as well, using the same aforementioned structing element. This resulted in the removal of the tails of these elements, as evident in Figure 9, which was also restored to its original sizes to preserve its original features as well (Gonzalez & Woods, 2009), (Arthur, et al., 2017).

Feature Extraction

Consequently, the feature extraction steps were followed. Firstly, the tail-to-body ratio of each element was obtained by using equation (1) below. Here, the tail lengths were extracted from the binary cropped image (Figure 8), and the sum of the body length and tail length, i.e., the length of the entire bloodstain element was each obtained from the binary eroded/ dilated image (Figure 9). Consequently, the average tail-to-body ratio of the elements were calculated (Zou & Stern, 2022), (Arthur, et al., 2017).

$$Tail - to - body\ ratio = \frac{Tail\ length}{Body\ length + Tail\ length} \quad (1)$$

Afterwards, the average impact angle (α_{avg}) was calculated from equation (2), by using the width to length ratio. The impact angle is defined as the angle at which a blood droplet has collided/ impacted on a particular surface. It is calculated by incorporating the two axes present/ considered in a bloodstain element: the major axis and the minor axis. The major axis of a bloodstain (element) is defined as the axis across the entire bloodstain, i.e., the axis adjoining the tip of the body and the end of the tail of the bloodstain. Meanwhile, the minor axis can be defined as the width, i.e., maximum distance perpendicular to the major axis adjoining two points of the body of the bloodstain.
The major and minor axes of a typical bloodstain element can be visually depicted according to Figure 3 below:

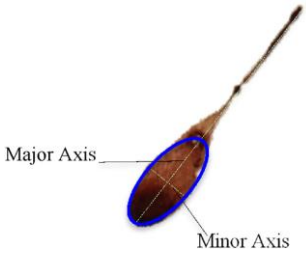


Figure 44. Major and minor axes of a bloodstain element

$$\alpha = \sin^{-1}\left(\frac{Minor\ axis\ length}{Major\ axis\ length}\right) \quad (2)$$

Then, the shapes of the elements were assessed by their irregularity. The irregularity was defined in two ways: by the ratio of the areas of the element and convex hull (Equation (3)), and by the perimeters of the inscribed circle and element body (Equation (4)).

$$Irregularity = \frac{Area\ of\ element}{Area\ of\ convex\ hull} \quad (3)$$

$$Irregularity = \frac{Perimeter\ of\ inscribed\ circle}{Perimeter\ of\ element\ body} \quad (4)$$

The convex hull is defined as the smallest convex region (enclosure comprised of small straight-line segments) that encloses a set of points or a region, which can also be defined as a "rubber band analogy". Alternatively, the convex hull ratio can be described as a degree of how closely the element shape fits/ aligns with its corresponding smooth, convex boundary (Arthur, Cockerton, de Bruin, & Taylor, 2015).
Here, it impartially determines deviations from ideally elliptical/ circular bloodstains which could indicate various types of impacts in several directions or non-linear/ turbulent motion (Arthur, et al., 2017). This ratio ranges from 0 to 1, where ratios closer to 1 indicates smoother/ regular stains that resembles elliptical and circular shapes much more closely. On the contrary, ratios closer to 0 denotes a more irregular/ deviated shape from conventionally elliptical stains.
The inscribed circle can be defined as the circle with the maximum diameter that can be fitted inside the boundary of an element. Similarly, this perimeter ratio also ranges from 0 to 1, where ratios closer to 1 reflects that the element closely resembles circularity/ regularity, whereas ratios closer to 0 signifies that the element shape is irregular, potentially consisting of extrusions/ deformities (Gonzalez & Woods, 2009). Therefore, for each bloodstain element of the pattern, these two ratios were calculated separately. Additionally, the area of each element was calculated by using the foreground pixels of the binary image as well.
To find the area of convergence (Figure 10), thin straight lines of a reasonable length were drawn along the major axes across each bloodstain element to find/ estimate the area of convergence via the congestion of the intersections of these lines, which is defined as the location/ vicinity of the source of blood spatter (James, Kish, & Sutton, 2005), (de bruin, Stoel, & Limborgh, 2011).

Finally, the distribution of the elements within the pattern was assessed. The element distribution is defined in two ways: the circularity of convex hull and the element density. To obtain the circularity value of the convex hull, the binary eroded/ dilated image was used as the input for the function $bwconvhull$, and the area (A) and perimeter (P) of the convex hull were extracted by $regionprops$ function. The circularity is expressed by the Equation (5):

$$Circularity\ of\ the\ convex\ hull = \frac{4\pi A}{P} \quad (5)$$

The element density of the BP was obtained in terms of elements per square cm (Figure 12), and to deduce both the highest and overall element density, a 10 cm × 10 cm box was moved across the BP in 1 cm steps (Zhang & Lu, 2004).

RESULTS

The resolution of the extracted image was found to be **357 × 454 pixels**, and in terms of pixels per mm, it was found to be approximately $\mathbf{3.78\ px./mm}$ **(pixels per mm)**, which was then restored to its original dimensions: **3860 × 4912 pixels (**$\mathbf{10.1\ px/mm}$**)**.
The gray scaled image and the background subtracted image are as follows:

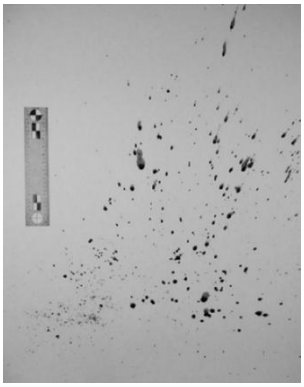


Figure 45. Gray scaled image



Figure 46. Resultant background subtracted image

The obtained intensity histogram of the gray scaled image along with the Otsu's and Triangle Thresholds is displayed beneath:
Additionally, the threshold value for the Triangle Thresholding algorithm was discovered to be **149.**

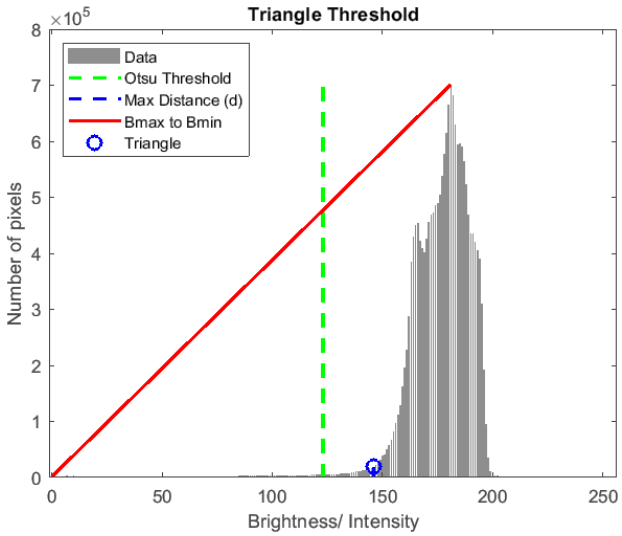


Figure 47. Intensity/ brightness histogram of the gray-scaled image with the Otsu threshold and Triangle threshold indicated

Post binarization and cropping, the image after the elimination of the ruler present is given below:



Figure 48. Binarized and cropped image

The image after the above binarized and cropped image was subjected to the aforementioned morphological operations and restored to its original size is as follows:



Figure 49. Binary eroded/ dilated image

Under the feature extraction steps, the results obtained are as follows:

The tail-to-body ratio of the BP was found to be **0.75**, which indicates that 75% of the body of a bloodstain element is comprised of its tail. The mean impact angle α_{avg} was found to be **46.53°**, which is well within the range of **49.6 ± 17.3°** which was the α_{avg} value obtained in the original source (Arthur, et al., 2017), and can be considered to be approximate to 49.6° as well.

Furthermore, the irregularity in terms of element to convex hull ratio was deduced to be **0.16**, which means that most elements have themselves closely bounded by their convex hulls with lesser blank spaces within those enclosures. Additionally, it was deduced that the mean element area was **0.99 mm^2**.

When straight line segments of significant length were plotted across each bloodstain along their major axes, the following result was obtained. However, it was unable to generate a heat map. Yet, as a rough estimation, the region enclosed below by a box can be considered as this convergence area, and this is quite similar to the area of convergence given per the original source.

Ultimately, the obtained value of the circularity of the convex hull is **0.81**, whereas the original value of the reference was **0.79**, and thus it can be considered that this obtained value is approximate enough to its original value. Also, the highest and overall element densities were **0.96 elements/cm^2** and **0.28 elements/cm^2** respectively, and their respective original values were **1.3 elements/cm^2** and **0.34 elements/cm^2**. Furthermore, it was discovered that the region with the highest element density was at the bottom-left part of the BP.

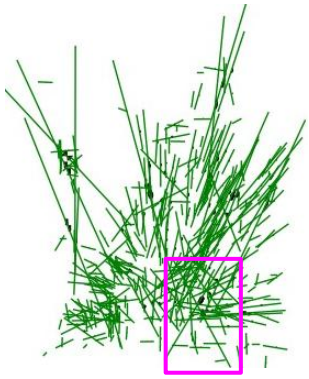


Figure 50. Straight lines drawn across the major axes of the bloodstain elements to deduce the area of convergence

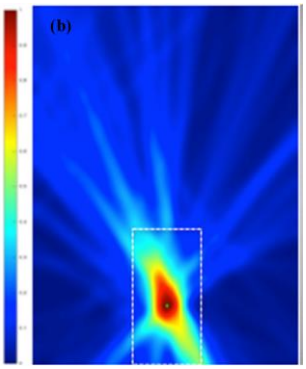


Figure 51. Original heat map with the 'hot spot' for side-by-side comparison and verification of the area of convergence

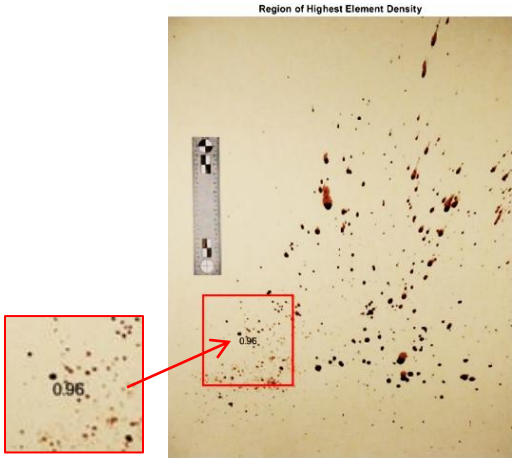


Figure 52. BP with a box enclosing the area with the highest element density

DISCUSSION

The results obtained and depicted above reflect the effectiveness of the adapted image-processing methodology in extracting and analyzing quantitative features of bloodstain patterns. Despite working with an image extracted from a publication (as opposed to a raw high-resolution capture), the feature values achieved closely mirror those in the original study.

In terms of validity and reliability, the close agreement of impact angle, density, and circularity measurements between this study and the original source supports the reliability of the implemented methodology. Minor deviations such as the slightly higher circularity and slightly lower density are likely attributable to image resolution limitations and manual segmentation artifacts. Furthermore, during segmentation, the selection of the Triangle Thresholding algorithm proved effective in isolating faint stains and background elements that would likely have been missed with Otsu's Thresholding algorithm. Additionally, morphological operations preserved the integrity of stain shapes, which was crucial for shape-based feature extraction.

Additionally, in terms of regional context, this methodology was designed in a generalized manner, and hence, it can be implemented anywhere, locally or globally. Also, it has utilized less software tools, where the image resolution was deduced and refinement was conducted in the algorithm itself, and hence, it consumes less computational power and effort, more efficient. Therefore, this can be implemented in rural areas where facilities are scarce as well.

As stated, numerous limitations arose in conducting this methodology, where the preliminary limitation was the inability to generate a heatmap to visualize the convergence area with a highly contrast visual aid.

This was due to the absence of spatial intensity mapping in the implementation. Additionally, because the image was taken from an existing publication (Figure 11) and not acquired through calibrated imaging hardware, there may be scaling inaccuracies despite resizing. Future implementations could overcome this limitation by integrating alternative scaling techniques and adding support for color segmentation and adaptive binarization, improving performance across more diverse datasets.

Other future work for improvement which potentially could affect the process positively include the integration of Machine Learning and Deep Learning (ML/ DL) and Artificial Intelligence (AI) to incorporate multiple inputs for the dataset to train a highly efficient model, which can consequently developed into a non-invasive, automated device/ model that can predict/ determine the modus operandi, crime weapon, or type of impact, as well as a 3D crime scene reconstruction by incorporating stereo imaging for volumetric bloodstain trajectory modeling.

CONCLUSION

In conclusion, throughout this research paper, a unique image-processing procedure is reported on behalf of the assessment of bloodstain patterns and their images acquired. Through this, it was able to obtain quantizable data of the features of both the bloodstain elements and the entire pattern overall as well. The methodology detailed in this is comprehensive and systematic, as the process to extract BP features has been automated. The key stages of this methodology included image acquisition, preprocessing and post-processing, element segmentation, feature extraction, and their analysis and interpretation. Also, the aim of this research, which was to enhance the accuracy, precision, and objectivity and impartiality of the BPA was achieved as well. By incorporating advanced image processing algorithms and techniques, the proposed methodology will greatly contribute to aiding forensic scientists and crime-scene analysts and investigators in crime-scene reconstruction, identifying the MO and weapons, type of impact or trauma, and ultimately to arrive at the correct conclusions.

Furthermore, some of these results were obtained and verified using *Google Colaboratory* for better accuracy as well (Ali, Khan, & Waleed, 2020).

Also, the potential reasons for the minute deviations of the obtained results and that of the original research paper are image quality issues, though the image resolution has been restored to its original resolution/ dimensions. This is possibly due to the image acquisition method, since to conduct this research paper, the image was directly obtained from the original paper instead of creating an arbitrary blood spatter pattern and capturing it with an imaging device such as a digital camera.

REFERENCES

1. Ali, I., Khan, A., & Waleed, M. (2020). A Google Colab Based Online Platform for Rapid Estimation of Real Blur in Single-Image Blind Deblurring. *2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI).* IEEE. https://doi.org/10.1109/ECAI50035.2020.9223244
2. Arthur, R. M., Cockerton, S. L., de Bruin, K. G., & Taylor, M. C. (2015). A novel, element-based approach for the objective classification of bloodstain patterns. (C. Cattaneo, & C. Jackowski, Eds.) *Forensic Science International, 257*, 220 - 228. https://doi.org/10.1016/j.forsciint.2015.08.028
3. Arthur, R. M., Humburg, P. J., Hoogenboom, J., Baiker, M., Taylor, M. C., & Bruin, K. G. (2017). An image-processing methodology for extracting bloodstain pattern features. *Forensic Science International, 277*, 122 - 132. http://dx.doi.org/10.1016/j.forsciint.2017.05.022
4. de bruin, K. G., Stoel, R. D., & Limborgh, J. C. (2011, July 25). Improving the Point of Origin Determination in Bloodstain Pattern Analysis. (M. Peat, Ed.) *Journal of Forensic Sciences, 56*(6), 1476 - 1482. https://doi.org/10.1111/j.1556-4029.2011.01841.x
5. Gonzalez, R. C., & Woods, R. E. (2009). *Digital Image Processing* (Third ed.). Prentice Hall, Pearson Education. https://dl.acm.org/doi/10.5555/1076432
6. James, S. H., Kish, P. E., & Sutton, T. P. (2005). *Principles of Bloodstain Pattern Analysis - Theory and Practice.* Taylor & Francis. https://doi.org/10.1201/9781420039467

7. Lazzaro, P. D., Murra, D., & Schwortz, B. (2013). Pattern recognition after image processing of low-contrast images, the case of the Shroud of Turin. In Z. Duric, & P. Radeva (Ed.), *Pattern Recognition. 46*, pp. 1964 - 1970. Elsevier. https://doi.org/10.1016/j.patcog.2012.12.010
8. Siu, S., Pender, J., Springer, F., Tulleners, F., & Ristenpart, W. (2017). Quantitative Differentiation of Bloodstain Patterns Resulting from Gunshot and Blunt Force Impacts. *Journal of Forensic Sciences, 62*(5), 1166 - 1179. https://doi.org/10.1111/1556-4029.13418
9. Swgstain, S. (2009). Scientific Working Group on bloodstain pattern analysis: recommended terminology. *Forensic Sci Commun, 11*(2), 14 - 17.
10. Zhang, D., & Lu, G. (2004). Review of shape representation and description techniques. *Pattern Recognition. 37'*, pp. 1 - 19. Elsevier. https://doi.org/10.1016/j.patcog.2003.07.008
11. Zou, T., & Stern, H. S. (2022). Towards a likelihood ratio approach for bloodstain pattern analysis. (C. Cattaneo, & C. Jackowski, Eds.) *Forensic Science International, 341*. https://doi.org/10.1016/j.forsciint.2022.111512

Adoption Of A Circular Economy Framework In The Design Phase Of The Refurbishment Project

S Abilash[1], **K A N** Gunarathna[2], C.Kalugala[3]

Liverpool John Moores University, UK[1]
Sri Lanka Institute of information Technology, SL[2]
Skills College of Technolog, SL[3]

ABSTRACT
Sustainable development in the built environment depends on minimizing environmental impact and conserving resources. In refurbishment projects, integrating circular economy (CE) principles during the design stage can significantly reduce waste and enhance resource efficiency. This study presents a tailored CE framework based on the 2D3R model, which emphasizes Design for Disassembly, Design for Adaptability, and the strategies of Reduce, Reuse, and Recycle. The framework was developed through a combination of literature review, expert interviews, and survey analysis to identify drivers and barriers to CE adoption in refurbishment. Key findings highlight the role of digital technologies particularly Building Information Modeling (BIM) in improving material traceability and facilitating lifecycle assessments. These tools support the implementation of circular strategies by enabling better planning and design decisions early in the project. The study concludes that applying the 2D3R approach in initial design phases enhances sustainability, reduces costs, and improves the adaptability and longevity of refurbished buildings.

KEYWORDS: *Circular Economy, 2D3R method, Sustainable construction, Building Information Modeling, Resource efficiency*

INTRODUCTION
The circular economy (CE) is an evolving model of sustainable development that aims to decouple economic growth from resource consumption and environmental degradation. Unlike the traditional linear model of “take, make, use, dispose,” CE emphasizes restorative and regenerative design by prioritizing resource efficiency, renewable energy use, and minimization of hazardous substances (Ekins et al., 2019). The phrase "circular economy" originated in 1990 while creating a new economic model based on the law of thermodynamics by Mr. Pears and Mr. Turner (Ghisellini et al., 2016). The foundation of the contemporary CE evolved through interconnected models such as environmental economy, the green economy, and industrial ecology (Tambovceva & Titko, 2022). Earlier concepts, including Boulding’s “Spaceship Earth” model and Stahel’s “cradle-to-cradle” principle, further established CE's theoretical foundations (Milovic et al., 2024).
As the construction sector remains one of the largest contributors to global environmental degradation, the application of CE in this domain has garnered increasing attention. The industry is responsible for the consumption of approximately 40% of global raw materials and generates nearly 30% to 40% of all solid waste (Wijesinghe et al., 2024). Construction processes also account for 39% of carbon dioxide emissions, with 11% attributed to embodied carbon in materials and products (Dsilva et al., 2023). Moreover, construction-related activities consume significant quantities of energy, fresh water, and other natural resources, exacerbating ecological pressures (Osobajo et al., 2022). These statistics underscore the urgent need for systemic change, prompting the integration of CE principles into the construction and refurbishment sectors.
Refurbishment, defined as the process of modifying, upgrading, or extending existing structures, has emerged as a sustainable alternative to demolition and new construction. Given the escalating costs of

infrastructure and environmental concerns in developing nations such as Sri Lanka, refurbishment offers a cost-effective and environmentally viable solution (De Silva et al., 2023a). It enables the preservation of building assets while addressing issues of deterioration, obsolescence, and changing functional requirements (Ekanayake et al., 2018). When aligned with CE principles such as design for disassembly (DfD), material reuse, and regenerative practices refurbishment can substantially mitigate waste generation, reduce reliance on virgin resources, and support carbon reduction goals (Bimsara et al., 2024).

The design phase of refurbishment projects presents a critical opportunity for CE integration. Design decisions made at this early-stage influence material selection, construction techniques, and lifecycle performance, directly affecting sustainability outcomes (Wijewansha et al., 2021). For instance, DfD facilitates ease of deconstruction, promotes component reuse, and minimizes waste at end-of-life stages (AlJaber, Alasmari, et al., 2023). Digital tools such as Building Information Modelling (BIM) and material passports support these objectives by enhancing traceability, enabling lifecycle assessment, and informing sustainable design strategies (Leindecker et al., 2024). The deployment of such tools during the design stage allows for comprehensive resource planning and optimization.

Numerous CE frameworks and models have been proposed to guide implementation across various scales of the built environment. At the micro level (materials), meso level (buildings), and macro level (cities), frameworks like REVERT (Resource, Envisagement, Validation, Entity, Regulation, and Technology) and ReSOLVE (Regenerate, Share, Optimize, Loop, Virtualize, Exchange) facilitate multi-scale stakeholder engagement and circular transformation (Bostancı et al., 2025; Csete & Esses, 2022). The 3R (Reduce, Reuse, Recycle) and extended 9R models (including Refuse, Rethink, Repair, Refurbish, Remanufacture, and Repurpose) further provide structured approaches to resource efficiency and waste minimization (Rahla et al., 2021). Together, these frameworks support regenerative practices that preserve the value of materials and infrastructure, contributing to long-term environmental and economic sustainability.

Nonetheless, the integration of CE principles in refurbishment design remains limited, particularly in developing countries. Significant barriers persist, including regulatory ambiguities, inadequate policy enforcement, limited financial incentives, and low stakeholder awareness (AlJaber, Martinez-Vazquez, et al., 2023; Fernandes & Ferrão, 2023b). The absence of standardized documentation, weak CDW policies, and insufficient training further hinder CE adoption (Metinal & Gumusburun Ayalp, 2025). In the Sri Lankan context, outdated regulations, a lack of technological infrastructure, and minimal institutional support pose additional challenges (Athapaththu et al., 2016; Weerakoon & Thayaparan, 2023). These systemic issues result in inconsistent practices and limit practical application, even where awareness of CE benefits exists.

Despite these challenges, global and local case studies demonstrate the practical viability and advantages of CE implementation. The Grensen 9b office project in Norway, for example, achieved a 93% reduction in CO_2 emissions by integrating reused materials into its refurbishment process (Baarnes & Liang, 2025). Similarly, in Sri Lanka, University of Ruhuna, Matara, and the Ocean University Galle demonstrates how refurbishment projects have successfully applied DfD and design for adaptability (DfA) principles, achieving material circularity rates exceeding 85% through lifecycle assessment tools (De Silva et al., 2023b). These examples underscore the transformative potential of CE when incorporated at the design phase.

While existing CE frameworks offer a robust theoretical foundation, they are largely oriented towards new construction rather than refurbishment, which presents a distinct set of constraints and opportunities (Fernandes & Ferrão, 2023a). Refurbishment projects often involve existing structural limitations, non-standardized material documentation, and diverse stakeholder requirements. As such, a specialized framework is needed to facilitate systematic CE integration from the design stage in refurbishment contexts (Charef et al., 2021). Doing so can enable better resource planning, material recovery, and carbon management, ultimately contributing to a more sustainable built environment (Wijewansha et al., 2021). This study plans to design a custom approach for using circular economy standards in the design part of refurbishment projects, mainly focusing on making them more sustainable, unique, and with less waste in places like Sri Lanka and similar areas.

RESEARCH METHODOLOGY

This research aims to develop a framework for applying circular economy (CE) principles during the design phase of refurbishment projects. A qualitative literature review was conducted to explore CE's impact on sustainability, resource use, and waste management in construction. A mixed-methods approach was adopted, combining quantitative data from surveys with qualitative insights from stakeholders to assess environmental impacts and identify sustainable strategies. Purposive sampling was used to select 5–7 experienced professionals, including architects, quantity surveyors, and project managers. Structured interviews provided in-depth data. Ethical approval was secured, ensuring confidentiality and anonymity throughout the research process, in alignment with core ethical standards.

DATA COLLECTION AND ANALYSIS

Structured Interview Respondent's Profile

Interviews with construction stakeholders, including architects, quantity surveyors, and project managers, provided valuable insights into the practical challenges, opportunities, and strategies for applying circular economy (CE) principles in refurbishment design. Six structured interviews were carried out for data collection. The respondents' details were gathered through the first section of the semi-structured interview guideline. The summarized composition of respondents is presented in Table 1.

Table 32. The composition of the respondents

Interviewee	Profession	Industry Experience (Year)
Respondent 1 (R1)	Civil Engineer	10 Years
Respondent 2 (R2)	Quantity Surveyor	4 Years
Respondent 3 (R3)	Senior Cost manager	12 Years
Respondent 4 (R4)	Chartered Quantity Surveyor	30 plus Years
Respondent 5 (R5)	Client Cost Consultant and Estimator	11 Years
Respondent 6 (R6)	Charteredd Quantity Surveyor	30 Years

Thematic Analysis

The thematic analysis method is used to analyze the data gathered from structured interviews. Those selected interviewees had experience in refurbishment and were familiar with sustainable approaches, most importantly CE principles. This target group include leading quantity surveyors, cost managers and project managers doing work related to construction projects in Sri Lanka, giving the study a diverse group of perspectives. Selecting these variables in this way made sure that our data was rich, specific to the context and matched the aims of our study.

Theme 1 - Explore CE principles relevant to refurbishment projects

Expert interviews revealed both shared and unique views on the importance of CE principles in refurbishment projects. Respondents R1, R4, and R6 emphasized materials reuse, highlighting examples like rubber dust in road pavement, reused AC ducts and sanitary fittings, and steel, aluminum, and bricks. They stressed that reuse saves costs and benefits the environment, making it a practical approach in Sri Lanka. Other participants highlighted different CE aspects: R2 focused on designing out waste and pollution, reusing furniture, and durable materials; R3 emphasized adapting structures to minimize material use; and R5 pointed to design for disassembly with modular parts for less waste. These diverse perspectives show that circular refurbishment depends on multiple factors reuse, adaptability, waste reduction, and disassembly. This diversity indicates the need for flexible, combined CE training models to effectively address various principles suited to each project. Being aware of the diversity among professionals shows that a flexible and combined model of CE training is needed to address several principles fitting each case. All professionals in this study agreed that CE significantly enhances the sustainability of refurbishment projects. Respondents R2, R3, R5, and R6 emphasized CE's environmental benefits, such as reducing fresh

raw material use, lowering carbon emissions, and minimizing demolition waste. They highlighted the importance of material reuse, repurposing, and life cycle thinking in improving sustainability. For instance, R2 and R3 noted CE's role in decreasing embodied carbon and extending building lifespans, while R5 shared how using similar materials in hotel renovations benefited both the environment and budget. R6 mentioned the use of renewable timber and bamboo supporting green energy.

Respondents R1 and R4 added economic, social, and procedural perspectives. R1 stressed CE's economic value in Sri Lanka, offering sustainable opportunities for young workers. R4 highlighted how organized planning and communication, like using variation orders, influence material reuse and demolition decisions. Overall, experts agreed on CE's positive sustainability impact but emphasized different aspects such as materials, carbon reduction, management, and local relevance.

Theme 2 – Analyze opportunities and issues of implementing a CE in the design stage

The interviews point out important opportunities, challenge specialists with various problems, and look at how regulations play a part in this area. The study brings out the truth about CE, illustrating how sustainable construction is done in real life with its various hurdles and rewards for people working in the industry.

Six industry experts shared common and distinct views on integrating CE principles in refurbishment design. R2, R3, and R4 emphasized material assessment and reuse, highlighting early project decisions to reduce waste and costs. They noted that involving contractors and clients early supports sustainability efforts. Meanwhile, R1 focused on financial benefits for clients and contractors, linking CE adoption to economic viability. R5 stressed designing adaptable retail and interior spaces to minimize waste during layout changes. R6 highlighted environmental priorities, advocating Life Cycle Assessment and local sourcing to lower carbon footprints and supporting regional economies. Overall, CE integration in refurbishment design balances economic, environmental, and functional considerations, with material reuse complemented by adaptability and sustainability strategies tailored to project needs.

Interviews revealed several challenges professionals face when applying CE principles in refurbishment projects. Respondents R1, R2, R3, and R6 highlighted a widespread lack of CE knowledge within companies, with both experienced leaders and younger engineers unfamiliar or hesitant to adopt it. R2 noted contractors and clients often don't understand CE's importance, while R3 pointed to outdated systems and missing guidelines limiting design flexibility. R6 emphasized skill gaps blocking successful implementation. Financial and technical barriers were also significant, as R4 observed clients' reluctance to consider higher upfront costs and Life Cycle Costing. R3 mentioned budget constraints prevent refurbishing many buildings, and R5 noted insufficient documentation complicates reuse. R6 stressed poor material traceability due to regulations and data gaps. Overall, awareness, funding, expertise, and systemic limitations challenge CE adoption in refurbishment projects.

Experts widely agree that regulatory and policy frameworks play a crucial role in supporting or hindering CE adoption in refurbishment design. R1, R3, and R4 emphasized that clear regulations and support from design bodies guide sustainable decisions early in projects. R3 highlighted rules on sustainability and waste reduction encouraging early CE integration, while R4 noted resources from the UDA promote green practices. R1 urged authorities like CIDA and ICTAD to establish consistent, strict CE policies. Conversely, R2, R5, and R6 felt current policies are insufficient, citing limited industry awareness and challenges in implementing standards like BS 8001:2017. They suggested certification programs could help but stressed the need for clearer guidelines, wider dissemination, and stronger enforcement to drive meaningful CE adoption in design processes. Table 2 below shows the expert opinion on policy influences.

Table 33: Summary of Expert responses on policy influence

Respondent	Main Viewpoint	Similar to	Key Comments
R1	Strongly supports regulatory role	R3,R4	Emphasized that professional bodies (CIDA, IESL, IQSSL) must set CE standards, particularly for quality assurance.
R2	Limited awareness, suggests certification	R5,R6	Not very aware of current CE policies; proposed that certifications (like sustainability) could help spread CE adoption.
R3	Favors mandatory CE policies	R1,R4	Stated that clear legislation and sustainability mandates push stakeholders to adopt CE from design stage.
R4	Recognizes UDA's green policy influence	R1,R3	Noted that UDA requires green practices at planning stage, making stakeholders comply early on.
R5	Policies are evolving, but not compulsory	R2,R6	Believes CE regulations in Bahrain are still informal; acknowledged green encouragement but no mandates.
R6	Weak regulatory impact in Sri Lanka	R2,R5	Highlighted limited effect of CE regulations; BS 8001 is rarely used; called for better support and collaboration.

Theme 3 – Summarize current knowledge regarding the implementation of CE

Interviews with experts revealed key strengths and challenges of current Circular Economy (CE) frameworks in refurbishment projects. Most respondents agreed that CE principles promote environmental care and reduce costs by minimizing resource use, waste, and lifecycle expenses. R4 emphasized life cycle cost analysis, while R6 highlighted benefits like lower emissions, local job creation, and use of nearby resources. R3 noted CE encourages sustainable design and innovation, and R1 mentioned longer-lasting, repairable products reduce waste. However, a major weakness is limited guidance and awareness. R1, R2, and R6 pointed to inadequate knowledge, weak training, and unclear risk-taking leading to poor implementation. R5 stressed difficulties adding CE features due to lack of data on older buildings, while R6 noted the absence of a centralized material database hinders reuse. These findings suggest CE's environmental and economic benefits are clear, but improvements in policy, education, and digital systems are necessary for effective adoption in South Asia's refurbishment sector.

Expert interviews reveal common and differing views on CE adoption in refurbishment projects. Respondents R1, R3, and R6 agree that CE use is limited and not fully established, with client willingness and awareness being key factors. R6 noted ongoing reliance on virgin materials and weak policy enforcement in Sri Lanka. R2 and R4 highlighted cost concerns, lack of knowledge, and adherence to traditional methods as barriers, often prioritizing time and budget over sustainability. R4 recommended better training to overcome reluctance and low oversight. R5 pointed out faster CE adoption in regions like Bahrain, driven by lifecycle costing and green supply chains. These findings suggest that targeted efforts, stakeholder collaboration, and institutional support are essential to expand CE implementation in refurbishment design across different contexts. Table 3 in below shows the Professionals' view on CE adoption

Table 3 Professionals' view on CE adoption

Respondent	View on CE Adoption	Key Points	Similar to
R1	**Supportive, conditional**	Willing clients and consultants essential; CE leads to savings and reduces waste	R3, R6
R2	**Skeptical, limited adoption**	CE often ignored due to cost and time pressures; reuse happens rarely; more awareness needed	R4
R3	**Optimistic, slow progress**	Adoption seen in high-profile/government projects; relies on client motivation	R1, R6
R4	**Critical, minimal adoption**	Few CE projects; traditional mindset; workers not trained for reuse	R2
R5	**Positive (International context - Bahrain)**	CE adoption growing; lifecycle costing and sustainable procurement practiced	Unique perspective
R6	**Partially aligned, systemic gaps**	Some progress via case studies; over-reliance on virgin materials; weak enforcement	R1, R3

Interviews revealed key opportunities to strengthen CE frameworks for refurbishment. R1, R2, and R3 emphasized the need to increase awareness among contractors, clients, and designers, noting this as a major barrier to CE adoption. They also suggested creating clear, separate standards to improve implementation. R4 recommended starting projects with surveys to assess recycling potential and apply CE principles early. R5 highlighted improving designers' CE knowledge at the project's outset, while R6 advocated for consistent data, training, stronger regulations, and incentives. They also noted that reusing existing buildings is more challenging than new construction, requiring careful planning. Overall, findings sow that alongside awareness and guidelines, effective CE design demands tools, education, and supportive policies for sustainable refurbishment.

Role of Digital Tools in Supporting CE Integration

Interviewees widely recognized BIM as essential for supporting CE in building renovations. Respondents R2, R3, R5, and R6 highlighted BIM's role in material tracking, lifecycle assessment, sustainable design, and collaboration. R2 noted its use for mapping materials and simulating environmental impact, while R5 applied BIM for reuse identification and facility planning. R3 praised its support for digital twins and audits, and R6 emphasized material control and greener choices. However, R1 and R4 had limited BIM knowledge, seen potential benefits but lacking expertise. This gap may hinder full CE adoption. Overall, BIM is seen as a valuable tool for circular refurbishment, but advanced training is needed for GCC construction professionals to maximize its sustainability potential.

Questionnaire Survey Respondent's Profile

We conducted a survey to see how well those in the construction industry are familiar with, recognize and make use of CE principles in their work. The survey tried to find out what various stakeholders in the construction industry thought about using CE at the design stage by blending the literature review with interviews. Table 4 shows the number of respondents received from the industry professionals through the questionnaire survey.

Table 34 Number of respondents

Professionals	Number of Participant	
	Percentage	No
Project Manager	8.70%	4
Site Engineer	10.90%	5
Civil/Structural Engineer	8.70%	4
Architect	10.90%	5
Quantity Surveyors	50%	23
Others	10.90%	5

Figure 1 shows that 82.6% of survey participants are early-career professionals with less than five years of experience. Only 14% have 6–10 years, while 2.2% each have 11–15 years and over 15 years of experience. This suggests that the survey results on CE awareness, training, and application in refurbishments may lack input from highly experienced professionals.

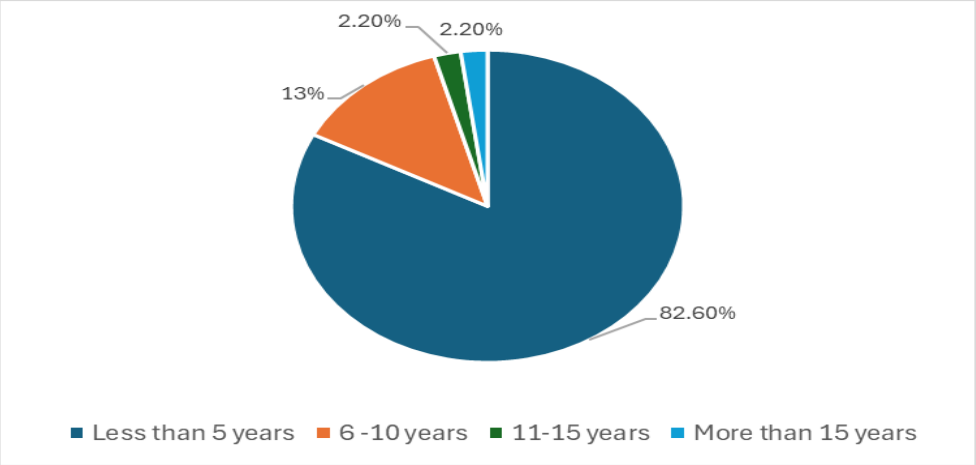


Figure 53 Experience of respondent source data analysis

Awareness and Understanding of CE Principles

Figure 2 shows the level of familiarity with CE concepts in construction.

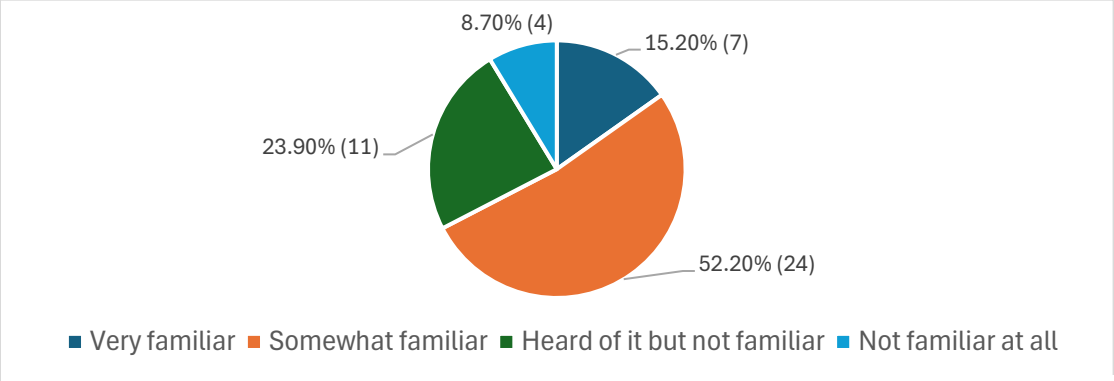
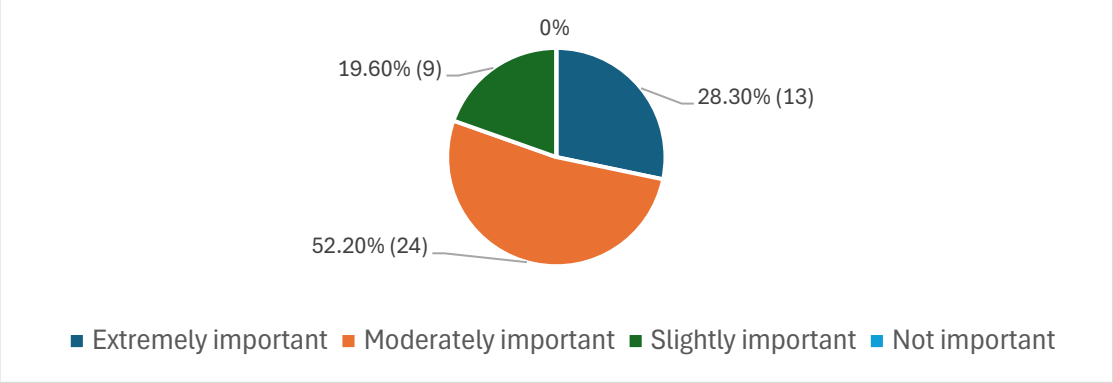


Figure 54 Awareness among participants source data analysis

Figure 3 shows the significance of integrating CE principles at the design stage.



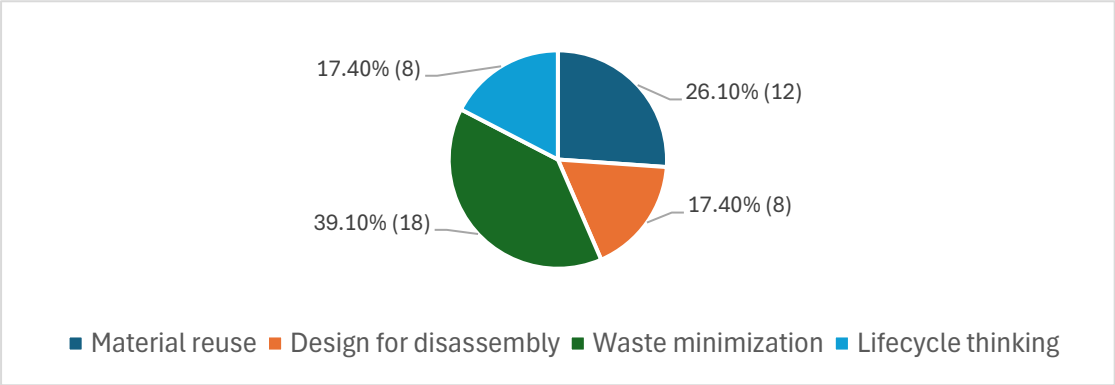


Figure 4 shows the identification of the most relevant CE principles.

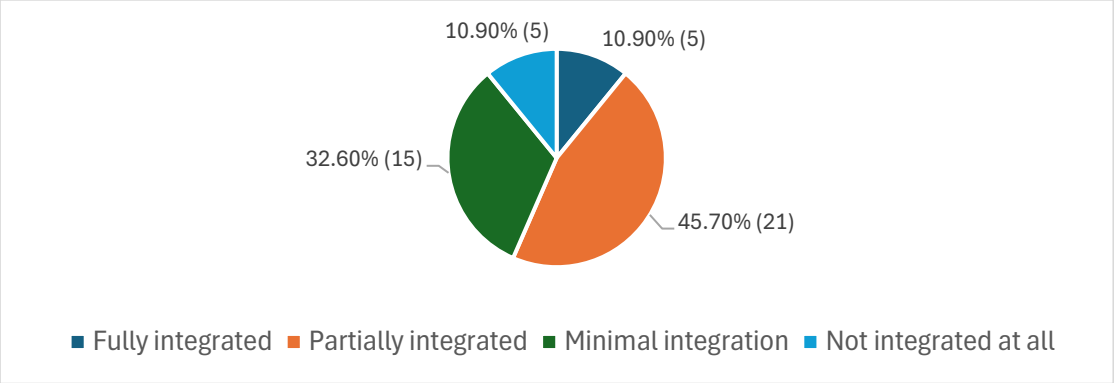


Figure 5 shows the current level of CE integration in the design phase.

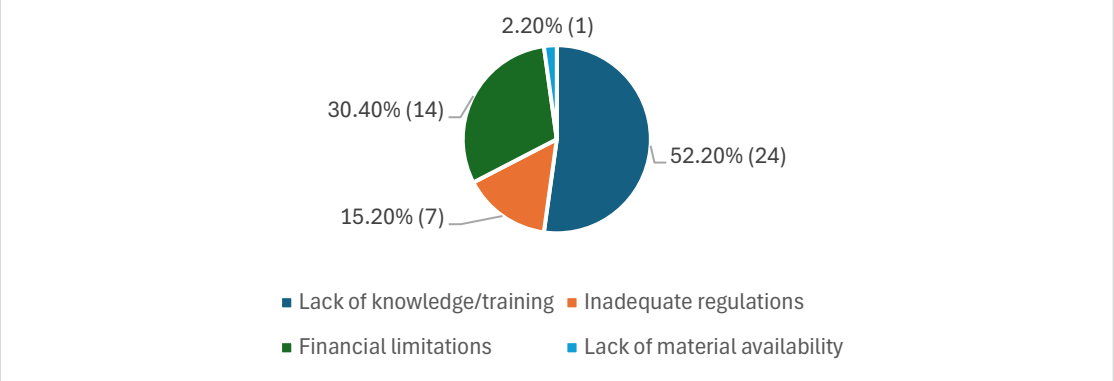


Figure 57 Current level of CE integration source data analysis

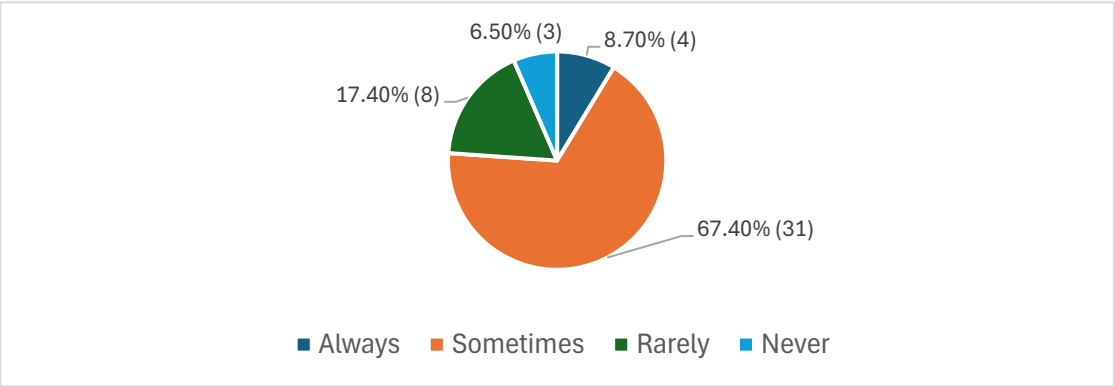


Figure 6 shows the frequency of considering DfD strategies**.**

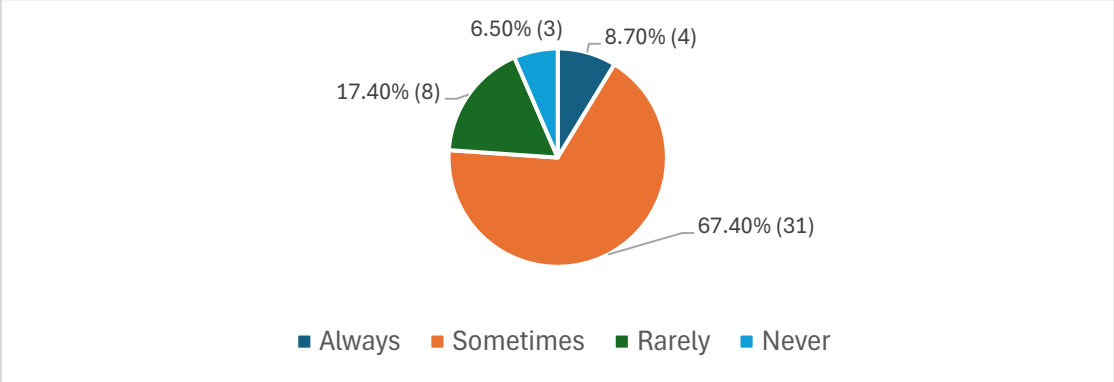


Figure 7 shows the key barriers to CE implem*entation*

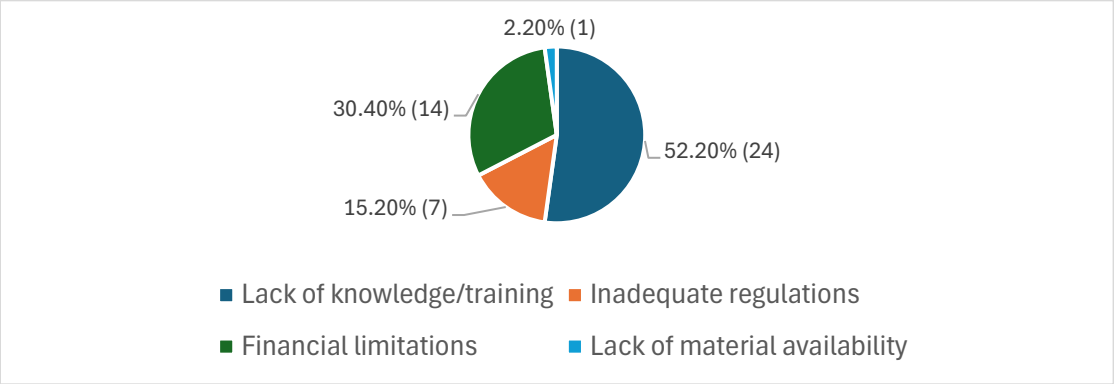


Figure 58 Barrier's source data analysis

Figure 8 shows the greatest opportunity for CE in design.

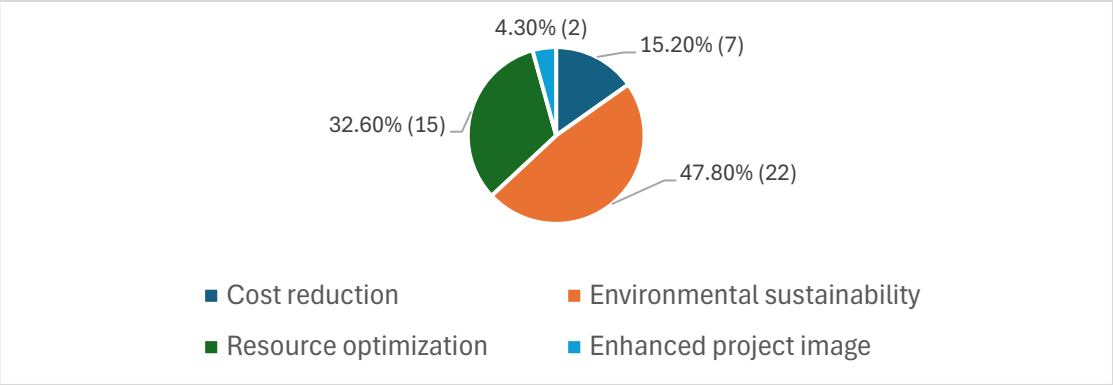


Figure 59 Opportunities source data analysis

Figure 9 shows the impact of the early design stage on the long-term sustainability of refurbishment projects.

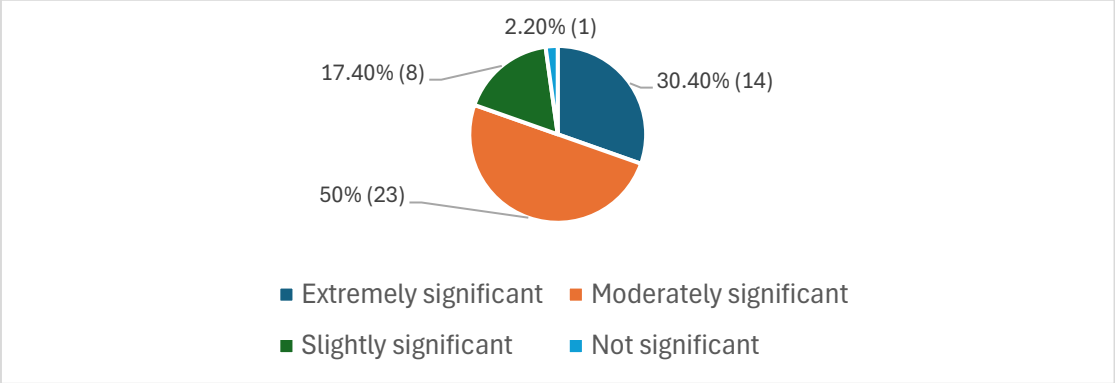


Figure 60 CE impact of early design stages on sustainability

CONCLUSION

This research successfully built a detailed Circular Economy (CE) plan designed especially for the design step in refurbishment projects based on the 2D3R rules: Design for Disassembly, Design for Adaptability, Reduce, Reuse and Recycle. It considers using sustainable methods at the start of design, so that materials can be reused and buildings adapted to live longer with less waste. Data obtained through expert interviews and survey responses revealed several key challenges and opportunities. Participants emphasized the critical importance of early-stage integration of CE strategies, including Design for Disassembly, material reuse, and lifecycle-based design approaches. Identified barriers to implementation included low stakeholder awareness, inadequate policy enforcement, limited training opportunities, and underutilization of digital tools such as Building Information Modelling (BIM). To overcome these constraints, respondents recommended targeted awareness-raising initiatives, the development of clear regulatory frameworks, incorporation of CE practices in the initial stages of design, advancement of digital competencies, and stronger collaboration among industry stakeholders. The findings underscore that effective CE integration in refurbishment projects particularly in developing contexts such as Sri Lanka requires a comprehensive approach involving technological adaptation, stakeholder engagement, and the formulation of context-sensitive regulations. In general, it guides users to include CE in rebuilding designs which supports the move from linear to circular construction. Figure 10 below shows the 2D3R framework.

ACKNOWLEDGEMENTS

I would wish to acknowledge Liverpool John Moores University (LJMU) and Sri Lanka Institute of Information Technology (SLIIT) for the facilitation of this study by providing university resources throughout this research. I owe my sincere thanks to Ms. Nishanthi Gunarathna, my research supervisor, and Ms. Chauri Kalugala, my instructor, for consistently guiding me, giving me valuable feedback and standing behind my work. Their support and advice have played a big role in directing my career. A sincere thanks to all professionals involved in the interview and questionnaire survey. Their time and knowledge were critical in this study, and I really do appreciate them sharing their experiences.
Thank You

REFERENCES

[1]AlJaber, A., Alasmari, E., Martinez-Vazquez, P., & Baniotopoulos, C. (2023). Life Cycle Cost in Circular Economy of Buildings by Applying Building Information Modeling (BIM): A State of the Art. *Buildings*, *13*(7). https://doi.org/10.3390/buildings13071858

[2]AlJaber, A., Martinez-Vazquez, P., & Baniotopoulos, C. (2023). Barriers and Enablers to the Adoption of Circular Economy Concept in the Building Sector: A Systematic Literature Review. *Buildings*, *13*(11), 1–21. https://doi.org/10.3390/buildings13112778

[3]Athapaththu, K., Karunasena, G., & Ekanayake, E. M. A. C. (2016). Sustainable Construction practices of Sri Lankan Contractors. *The 5th World Construction Symposium 2016: Greening Environment, Eco Innovations & Entrepreneurship*, *July*, 489–497.

[4]Baarnes, S., & Liang, E. Z. (2025). A Case Study of Sustainable Resource Management Through Reuse of Building Materials. *Lecture Notes in Civil Engineering*, *237*, 941–954. https://doi.org/10.1007/978-3-031-69626-8_79

[5]Bimsara, K. D. M., Abenayake, D. N., Lingasabesan, V., & Shanika, V. G. (2024). *Adopting circular economy practices in major building materials and elements to minimise waste in Sri Lanka*. *August*, 105–118. https://doi.org/10.31705/wcs.2024.9

[6]Bostancı, H. B., Tanyer, A. M., & Habert, G. (2025). REVERT Framework: Stakeholder Perspective to Enable Circular Transformation of Construction Industry. *The 1st International Conference on Net-Zero Built Environment*, *09 January*, pp 1039–1049.

[7]Charef, R., Morel, J. C., & Rakhshan, K. (2021). Barriers to implementing the circular economy in the construction industry: A critical review. *Sustainability (Switzerland)*, *13*(23), 1–18. https://doi.org/10.3390/su132312989

[8]Csete, M. S., & Esses, D. (2022). Usage of Production Function in Linear Economy. *Periodica Polytechnica Transportation Engineering*, *50*(2), 223–226. https://doi.org/10.3311/PPtr.16616

[9]De Silva, S., Samarakoon, S. M. S. M. K., & Haq, M. A. A. (2023a). Use of circular economy practices during the renovation of old buildings in developing countries. *Sustainable Futures*, *6*(October), 100135. https://doi.org/10.1016/j.sftr.2023.100135

[10]Dsilva, J., Zarmukhambetova, S., & Locke, J. (2023). Assessment of building materials in the construction sector: A case study using life cycle assessment approach to achieve the circular economy. *Heliyon*, *9*(10), e20404. https://doi.org/10.1016/j.heliyon.2023.e20404

[11]Ekanayake, B. J., Sandanayake, Y., & Ramachandra, T. (2018). Application of value management to refurbishment projects: A Sri Lankan case study. *Proceeding of the 34th Annual ARCOM Conference, ARCOM 2018*, *September*, 597–606.

[12]Ekins, P., Domenech, T., Drummond, P., Bleischwitz, R., Hughes, N., & Lotti, L. (2019). The Circular Economy: What, Why, How and Where. In *Background paper for an OECD/EC Workshop on 5 July 2019 within the workshop series "Managing environmental and energy transitions for regions and cities*.

[13]Fernandes, J., & Ferrão, P. (2023a). A New Framework for Circular Refurbishment of Buildings to Operationalize Circular Economy Policies. *Environments - MDPI*, *10*(3).

https://doi.org/10.3390/environments10030051
[14]Fernandes, J., & Ferrão, P. (2023b). Urban Metabolism-Based Approaches for Promoting Circular Economy in Buildings Refurbishment. *Environments*, *10*(1), 13. https://doi.org/10.3390/environments10010013
[15]Ghisellini, P., Cialani, C., & Ulgiati, S. (2016). A review on circular economy: The expected transition to a balanced interplay of environmental and economic systems. *Journal of Cleaner Production*, *114*, 11–32. https://doi.org/10.1016/j.jclepro.2015.09.007
[16]Leindecker, G., Askar, R., Güngör, B., Blázquez, T., Trubina, N., Gómez-Gil, M., Karanafti, A., Bragança, L., & Wolf, C. De. (2024). Material and Building Passports as Supportive Tools for Enhancing Circularity in Buildings. *Circular Economy Design and Management in the Built Environment*, 507–543. https://doi.org/10.1007/978-3-031-73490-8
[17]Metinal, Y. B., & Gumusburun Ayalp, G. (2025). Uncovering Barriers to Circular Construction: A Global Scientometric Review and Future Research Agenda. *Sustainability (Switzerland)*, *17*(4). https://doi.org/10.3390/su17041381

[18]Milovic, T., Laban, M., Stracev-Curcin, A., & Bulatovic, V. (2024). Circular Economy and Its barriers To Implementation In The Construction Sector. *Applied Engineering Science*, *22*(2).
[19]Osobajo, O. A., Oke, A., Omotayo, T., & Obi, L. I. (2022). A systematic review of circular economy research in the construction industry. *Smart and Sustainable Built Environment*, *11*(1), 39–64. https://doi.org/10.1016/j.heliyon.2022.e09169
[20]Rahla, K. M., Mateus, R., & Bragança, L. (2021). Selection Criteria for Building Materials and Components in Line with the Circular Economy Principles in the Built Environment—A Review of Current Trends. *Infrastructures*, *6*(4). https://doi.org/10.3390/infrastructures6040049
[21]Tambovceva, T., & Titko, J. (2022). Introduction To Circular Economy. In *Ekonomikas un kulturas augstskola*. https://doi.org/10.1142/9789811241949_0002
[22]Weerakoon, P., & Thayaparan, M. (2023). Analysis of Barriers Towards Implementing Circular Economy in Sri Lankan Built Environment. *World Construction Symposium*, *1*(July), 162–173. https://doi.org/10.31705/WCS.2023.14

[23]Wijesinghe, E. D. A. T., Jayawickrama, T. S., Tennakoon, G. A., & Wijewickrama, M. K. C. S. (2024). 5S Principles for Sustainable Resource and Waste Management in Sri Lankan Construction Projects. *World Construction Symposium*, *August*, 1–14. https://doi.org/10.31705/WCS.2024.1
[24]Wijewansha, A. S., Tennakoon, G. A., Waidyasekara, K. G. A. S., & Ekanayake, B. J. (2021). Implementation of circular economy principles during pre-construction stage: the case of Sri Lanka. *Built Environment Project and Asset Management*, *11*(4), 750–766. https://doi.org/10.1108/BEPAM-04-2020-0072

Analysis And Optimization Of OLSR And AODV Routing Protocols For Highly Mobile Autonomous Aerial Vehicle Networks: Experimental Performance Evaluation In Various Application Scenarios

Nguyen Duc Tu, Luybov Gorbacheva, Ammar Muthanna

The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, St. Petersburg, 193232, Russian Federation

ABSTRACT

Driven by the rapid evolution of Autonomous Aerial Vehicle (AAV) technology, ad-hoc AAV networks are becoming increasingly significant in diverse domains such as telecommunications, security surveillance, search-and-rescue operations, emergency management, precision agriculture, and cargo transportation. Owing to their flexible deployment and tight coordination capabilities, AAVs enable real-time data exchange, unlocking considerable potential for tasks that demand high accuracy and swift response. Nevertheless, their three-dimensional mobility and continuously changing topology impose substantial challenges on the design of suitable routing solutions, because conventional protocols originally developed for Mobile Ad-hoc Networks (MANETs) are seldom optimized for aerial characteristics. This discrepancy underscores the necessity for a comprehensive analysis and optimal configuration that satisfy the stringent requirements of low latency, high throughput, and reliability in mission-critical AAV applications. Against this backdrop, the present study focuses on the analysis and modelling of two widely adopted routing protocols—Optimized Link State Routing (OLSR) and Ad hoc On-Demand Distance Vector (AODV)—in various deployment contexts of ad-hoc AAV networks. By evaluating the performance of these protocols, we identify their advantages, limitations, and possible enhancement directions, and subsequently propose configuration guidelines that deliver improved link quality.

INTRODUCTION

Against the backdrop of swift advances in telecommunication technologies, AAVs are attracting growing research attention and seeing ever-wider practical adoption. Unlike traditional communication approaches that rely on terrestrial infrastructure, AAVs extend network coverage into the airspace, enabling optimized service areas, reduced deployment costs, and flexible adaptation to contemporary connectivity demands. Establishing telecommunications infrastructure based on AAVs, however, poses several challenges—most notably in routing capability, real-time data exchange, and Quality-of-Service (QoS) assurance. This section analyses the broader context and importance of AAVs with a particular emphasis on telecommunications, while also identifying design obstacles that are specific to AAV networks.

Context and significance of Autonomous Aerial Vehicles

Leveraging breakthroughs in electronics, wireless communication, artificial intelligence (AI), and battery technology, AAVs are recognized as a crucial platform for next-generation telecommunications infrastructure. Moreover, the convergence of Edge-Computing solutions [1], 5G-to-6G technologies, and the Internet-of-Things (IoT) trend empowers AAVs not only to collect information but also to process, relay, and even provide connectivity services to external entities [2].

At present, Autonomous Aerial Vehicles (AAVs) are being integrated into numerous practical telecommunication scenarios, including but not limited to the following:

- Temporary mobile broadcasting nodes—During mass gatherings or in disaster-stricken areas, AAVs equipped with on-board gNodeB (5G) or eNodeB (4G) units can promptly augment network capacity and temporarily extend coverage without the costly construction of cumbersome terrestrial infrastructure [2].

- Last-mile connectivity solutions—In mountainous terrain or regions with under-developed infrastructure, AAVs operate as relay entities between end-user equipment and ground base stations, thereby markedly enhancing link quality and service continuity.

- Monitoring and data acquisition—Owing to their agile maneuverability, AAVs can continuously refresh imagery, video streams, or environmental information and forward these data sets to a processing center, supporting urban management, precision agriculture, and climate-change surveillance [3].

The harmonious convergence of AAV platforms with cutting-edge telecommunication technologies has accelerated the digitalization of diverse socioeconomic activities, yielding improved operational efficiency, cost reduction, and the emergence of novel business models.

In the context of burgeoning applications such as telepresence, virtual reality (VR), smart healthcare, and 5G/6G services, AAVs have further substantiated their role as a flexible, cost-effective, and high-performance solution for monitoring and emergency support [4]. By virtue of sustained radio transmission, real-time data processing, and high mobility, AAVs can rapidly establish communication bridges in regions where the existing infrastructure is impaired or inaccessible. Consequently, rescue teams and command centers can exchange information, high-definition live video, and geospatial data to track evolving situations, localize victims, and implement timely interventions. In telemedicine applications, for instance, AAVs not only facilitate the transport of critical medical data and enable live consultations but also deliver equipment or biological specimens to isolated areas, thereby shortening response times and elevating treatment efficacy. Moreover, the deployment of AAVs obviates the need for extensive terrestrial infrastructure—such as fixed towers or fiber-optic backhaul—conserving financial resources, manpower, and deployment time when compared with conventional solutions. Hence, AAVs constitute a pivotal element within the intelligent-telecommunications ecosystem, underpinning emergency monitoring, enlarging coverage areas, and providing a platform for advanced services—including telepresence, VR, and smart healthcare—in scenarios that demand immediacy and stringent reliability.

Communication and Routing Challenges in AAV-Centric Networking Environments Despite their disruptive potential, AAV systems confront an array of technical hurdles, particularly as requirements for bandwidth, latency, and reliability continue to escalate. These obstacles become increasingly pronounced when AAVs operate in three-dimensional space, exhibit constant mobility, and must uphold stable connectivity.

- Highly dynamic topology in three-dimensional space—Because AAVs move at high velocity while continually changing altitude, radio links are inherently unstable. Conventional MANET-oriented routing protocols struggle to react in a timely manner to sudden link disruptions, necessitating either the fine-tuning of existing algorithms or the development of novel routing schemes that can adapt swiftly to topological perturbations.

- Energy and radio-resource constraints—Beyond sustaining flight, an AAV must uphold uninterrupted data exchange over radio channels characterized by finite capacity. To maximize mission endurance, the system must balance transmission efficiency against power consumption, a trade-off rendered more acute by high-bit-rate requirements such as HD video or real-time services.

- Channel interference and attenuation—Operating at altitude and in open space exposes AAVs to numerous external influences (weather, terrain, obstacles, signal reflections, etc.). The resulting channel instability calls for highly adaptive mechanisms in power control, modulation–coding selection, and routing strategy. Without optimization, fading and interference severely degrade link maintenance capabilities, culminating in elevated packet-loss ratios.

- QoS, low-latency, and security requirements—Many AAV applications, notably in monitoring and emergency support, demand data transmission with exceptional reliability and ultra-low latency. Routing protocols that are not tailored to meet these criteria may incur congestion, violate delay thresholds, or introduce security vulnerabilities.

Against this backdrop, the investigation and assessment of routing protocols for AAV networks are essential for identifying solutions that satisfy responsiveness, connectivity, and resource-efficiency objectives. In this article, we model and analyze two protocols—OLSR and AODV—under varying operational conditions, including AAV velocity, mobility model, fleet size, and mission-area dimension. By measuring throughput, latency, packet-loss ratio, and energy expenditure, we delineate each protocol's merits and limitations and propose configuration-optimization guidelines. The resulting insights facilitate the selection of an appropriate routing paradigm capable of achieving stringent QoS, reliability, and flexibility targets in modern AAV networking environments.

I. Classification of Routing Protocols in Autonomous Aerial Vehicle Networks

In recent years Autonomous Aerial Vehicle (AAV) networks have attracted growing attention, largely because three-dimensional operational space and high vehicle velocities make sustaining a stable communication channel particularly challenging. To address this issue numerous routing strategies have been proposed, all geared toward the agile maintenance and timely update of link information. In general, routing protocols for AAV networks can be grouped into five major families: proactive [5], reactive [6], hybrid [7], geographic [8], and swarm-intelligence/reinforcement-learning-based [9].

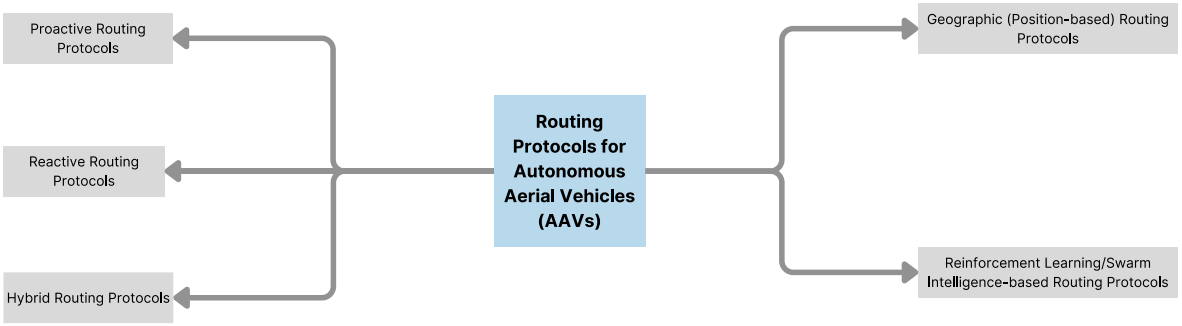


Fig. 1. Classification of routing protocols for Autonomous Aerial Vehicle networks.

1.1 Proactive Routing Protocols

Proactive schemes continuously maintain up-to-date routes to every node by periodically refreshing link-state information or routing tables. Because destination entries already reside in memory, route-establishment latency at data-transmission onset is minimized. Within fast-moving AAV networks these protocols are often favored for their ability to react promptly to incremental topological variations while supplying a global view that simplifies management and monitoring. Nevertheless, when AAVs maneuver incessantly and topology fluctuates rapidly, proactive solutions can consume substantial bandwidth and energy to keep routes current, a drawback that becomes pronounced as network scale increases. Representative examples include OLSR (Optimized Link State Routing), which employs a Multi-Point Relay (MPR) mechanism to curtail broadcast traffic, and DSDV (Destination-Sequenced Distance Vector), which couples classic distance-vector rules with sequence numbers to preserve routing consistency [5], [10].

1.2 Reactive Routing Protocols

In contrast, reactive schemes discover routes only when a data session is initiated, thereby reducing control overhead because topology information is not perpetually maintained. For AAV networks exhibiting bursty traffic or idle periods, this on-demand paradigm affords greater flexibility and conserves both bandwidth and energy. A notable drawback, however, is that initial route discovery can incur additional delay when packets must be dispatched suddenly. Furthermore, at very high AAV speeds the discovery process may be invoked repeatedly, causing network congestion or repeated re-establishment cycles. Widely used reactive protocols include AODV (Ad hoc On-Demand Distance Vector), which leverages sequence numbers to avoid loops, and DSR (Dynamic Source Routing), which relies on a route cache at the source node to reuse previously learned paths [6], [11].

1.3 Hybrid Routing Protocols

To combine the strengths of proactive and reactive approaches, hybrid protocols partition the network into zones or clusters. Within each zone the protocol behaves proactively for immediate responsiveness; between zones it adopts a reactive method to limit unnecessary large-scale broadcasts. For AAV networks this dual strategy mitigates routing-table exchange overhead while retaining prompt local reaction. Nevertheless, maintaining zone boundaries and managing clusters becomes complex when AAVs move continuously or as network size expands. A canonical example is the Zone Routing Protocol (ZRP), originally investigated for MANETs and later proposed for AAV scenarios [12].

1.4 Geographic Routing Protocols

Advances in positioning technologies such as GPS and high-precision inertial sensors have made location-based routing a promising solution for AAV systems. Instead of storing exhaustive link-state data, nodes keep only their current coordinates and those of the destination, forwarding packets via a greedy strategy. This method sharply lowers routing-table maintenance cost, as each node requires knowledge of only its immediate neighborhood rather than an entire multi-hop path. Geographic protocols still face challenges when position information is inaccurate or when voids arise in the three-dimensional space that hinder signal reachability. GPSR (Greedy Perimeter Stateless Routing) is a renowned member of this family, blending greedy forwarding with perimeter traversal to circumvent voids and adapt to diverse AAV scenarios [13], [14].

1.5 Swarm-Intelligence/Reinforcement-Learning-Based Routing

Beyond conventional designs, the application of swarm intelligence (SI) and reinforcement learning (RL) opens new avenues in AAV routing. Rather than relying solely on static routing tables, AAV nodes can learn mobility patterns and channel conditions, autonomously refining forwarding decisions over time. Inspired by the evolutionary behavior of insect or bird swarms, SI/RL schemes allow AAV networks to adapt gracefully to abrupt topological changes and channel dynamics. Yet computational complexity and

the need to synchronize learned information across nodes can impose significant overhead—especially in resource-constrained environments. Ant Colony Optimization (ACO) and bee-inspired protocols use virtual pheromones to guide packet delivery, whereas Q-learning applies reward–penalty mechanisms to optimize objectives such as throughput or latency.

OPERATIONAL PRINCIPLES OF THE OLSR PROTOCOL

The Optimized Link State Routing (OLSR) protocol is grounded in the classical link-state paradigm but achieves greater efficiency through its Multipoint Relay (MPR) mechanism. Every node periodically exchanges link-state and routing information with its neighbors, thereby building a global routing table without waiting for a data-forwarding request. This proactive behavior enables swift packet delivery both in conventional ad-hoc networks and—more critically—in Autonomous Aerial Vehicle (AAV) environments. The key innovation lies in OLSR's selective use of MPRs: each node chooses a minimal subset of neighbors as relays such that every node in the network is reachable through at least one MPR, thus eliminating redundant broadcast transmissions. MPR selection is driven by link-quality and coverage criteria that markedly reduce control-traffic overhead.

OLSR defines three core control messages. HELLO packets discover and refresh neighbor relations; TC (Topology Control) packets advertise selected MPR links; and MID (Multiple Interface Declaration) packets list additional addresses when a node owns multiple interfaces. Parameters such as HELLO_INTERVAL and TC_INTERVAL specify the periodicity of HELLO and TC messages and hence directly influence link-state freshness versus bandwidth consumption. The WILLINGNESS parameter quantifies a node's readiness to forward traffic and thereby shapes MPR selection and the resulting global topology. For highly dynamic AAV applications, careful tuning of these parameters' balances control-update granularity against spectral efficiency. Numerous OLSR enhancements tailored to AAVs have been proposed, concentrating on adaptive MPR selection as node density or velocity grows and on dynamic adjustment of HELLO/TC rates to accommodate three-dimensional mobility while meeting stringent reliability and throughput demands. These efforts collectively aim to minimize delay and energy cost while preserving adaptability under incessant topology changes.

OPERATIONAL PRINCIPLES OF THE AODV

Protocol Ad hoc On-Demand Distance Vector (AODV) routing emphasizes bandwidth and resource conservation in highly mobile scenarios through its reactive (on-demand) philosophy. When a source must transmit but lacks a route, it broadcasts a Route Request (RREQ); any intermediate node possessing a fresh route may respond with a Route Reply (RREP). To preclude loops and guarantee route freshness, AODV incorporates monotonically increasing sequence numbers. Route maintenance is likewise meticulous: upon detecting a link break, the responsible node issues a Route Error (RERR) to affected sources, prompting them to invalidate the stale path. Frequently, AODV invokes local repair to re-establish connectivity near the failure point, thereby curtailing delay relative to network-wide RREQ re-broadcast. Parameters such as ACTIVE_ROUTE_TIMEOUT and ALLOWED_HELLO_LOSS govern route-lifetime thresholds and tolerated HELLO losses, allowing AODV to adapt to frequent AAV maneuvers and tight throughput/latency budgets. Owing to these design choices, AODV remains a viable option for many AAV missions in which routing performance is pivotal to overall system reliability.

II. Performance Evaluation Parameters

To evaluate the performance of OLSR and AODV routing protocols in Autonomous Aerial Vehicle networks, this study uses the following key evaluation parameters:

1. Packet Delivery Ratio (PDR).

PDR quantifies reliability as the fraction of packets that reach their destinations:

$$\mathrm{PDR_{avg}} = \frac{1}{n}\sum_{i=1}^{n}\frac{|P_{recv}^{i}|}{|P_{sent}^{i}|} \times 100\%$$

Where:

- n is the number of traffic sessions,
- $|P_{sent}^{i}|$ is the number of packets issued in session i,
- $|P_{recv}^{i}|$ is the number successfully received.

PDR reflects the level of reliability and efficiency of the routing protocol in ensuring packet transmission from source to destination. The higher the PDR value, the better the protocol performance. In the AAV network environment, PDR is a critical parameter when evaluating routing protocol performance due to high mobility and rapid topology changes that increase the probability of packet loss.

2. End-to-End Delay (EED).

EED measures temporal efficiency:

$$\mathrm{EED_{avg}} = \frac{\sum_{p\in P_{recv}}\left(T_{arr}(p) - T_{dep}(p)\right)}{|P_{recv}|},$$

Where:

- P_{recv} is the set of received packets,
- $T_{dep}(p)$ the departure time,
- $T_{arr}(p)$ the arrival time of packet p.

These metric aggregates routing, queuing, MAC, and transmission delays.

3. Average Throughput.

Throughput reflects data-transfer capacity:

$$\mathrm{Throughput_{avg}} = \frac{\sum_{i=1}^{n}|P_{recv}^{i}|S}{T}\cdot\frac{8}{1000}\ \text{(kbit/s)},$$

Where:

- S the packet size (bytes),
- T the active transmission duration (s).

Jitter (Delay Variation). Jitter captures fluctuation in inter-packet delay and is recursively estimated per RFC 3550:

$$J_i = J_{i-1} + \frac{|D_i - D_{i-1}| - J_{i-1}}{16},$$

Where: D_i and D_{i-1} are successive packet delays and the $1/16$ factor smooths short-term variations.

Collectively, these metrics enable a comprehensive appraisal of routing-protocol effectiveness for AAV networks, balancing reliability, responsiveness, bandwidth utilization, and temporal stability.

III. Performance Evaluation of Routing Protocols for Autonomous Aerial Vehicle Networks

The performance of Autonomous Aerial Vehicle networks is determined by factors such as routing protocol, movement speed, node density, and movement model. In this study, we evaluate two popular routing protocols (OLSR and AODV) in various scenarios to find the optimal solution for the operation of Autonomous Aerial Vehicle networks.

We use the Network Simulator 3 (NS-3.34) to simulate the operation of Autonomous Aerial Vehicle networks in four different scenarios: analysis by number of nodes, analysis by movement speed, analysis by movement model, and analysis by coverage area.

Table 1 General simulation parameters

Parameter	Value
Wi-Fi standard	IEEE 802.11g
PHY mode	ErpOfdm
Data rate	6 Mbit s^{-1}
Tx power	18 dBm
Receiver sensitivity	–85 dBm
CCA threshold	–90 dBm
Packet size	512 bytes
Application rate	128 kbit s^{-1}
Propagation model	LogDistance (exponent = 2.0)
Simulation time	200 s

In these simulations, we use the LogDistance attenuation model with parameter Exponent=2.0 to model signal attenuation in the radio transmission environment of Autonomous Aerial Vehicle networks. The value Exponent=2.0 corresponds to the Free Space Path Loss attenuation model, which is suitable for radio wave propagation conditions of aerial base stations in open space, where there are few obstacles and complex reflection and refraction phenomena. The LogDistance model is a common channel attenuation model based on experimental observation that signal intensity decreases as a logarithmic function of distance, calculated using the formula:

$$PL(d) = PL(d_0) + 10 \times n \times \log_{10}\left(\frac{d}{d_0}\right)$$

Where:
$PL(d)$ - signal attenuation (path loss) at distance d,
$PL(d_0)$ - signal attenuation at reference distance d_0,
n - attenuation exponent, in our simulation taken as 2.0,
d - distance between transmitter and receiver.

2. Performance Analysis Based on Number of Nodes

In this simulation, we analyze the performance of OLSR and AODV protocols when changing the number of Autonomous Aerial Vehicle nodes from 5 to 60. The movement model used is GaussMarkov with a speed of 10 m/s.

Table 2. Simulation Parameters Based on Number of Nodes

Parameter	Value
Number of Nodes	5–60
Speed	10 m/s
Mobility model	GaussMarkov
Flight Area	800 m × 800 m × 100 m

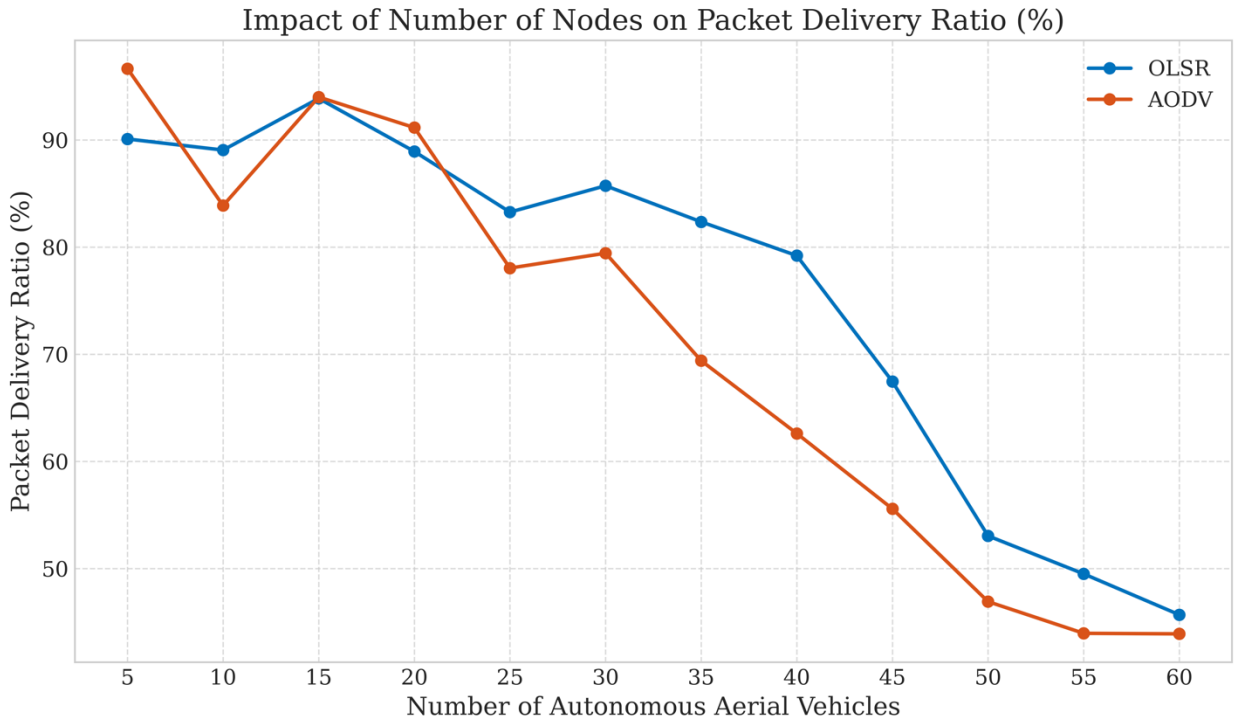


Fig. 2. Impact of Number of Nodes on Packet Delivery Ratio

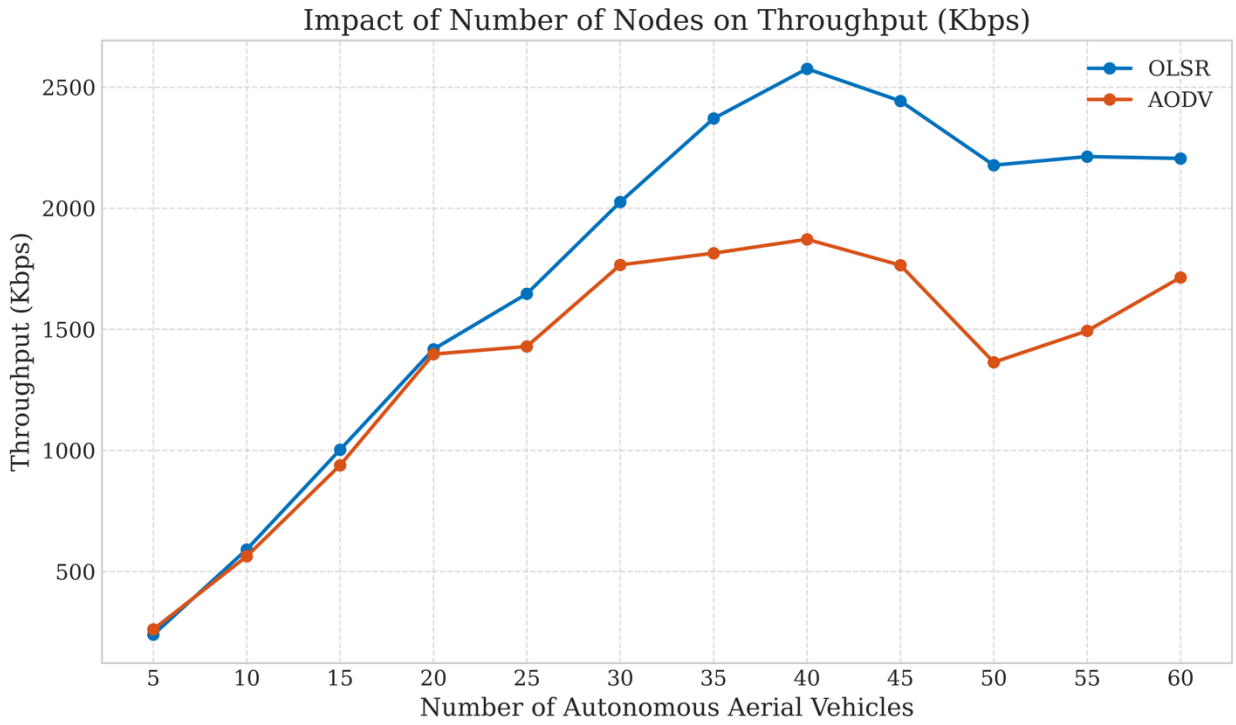


Fig. 3. Impact of Number of Nodes on Throughput

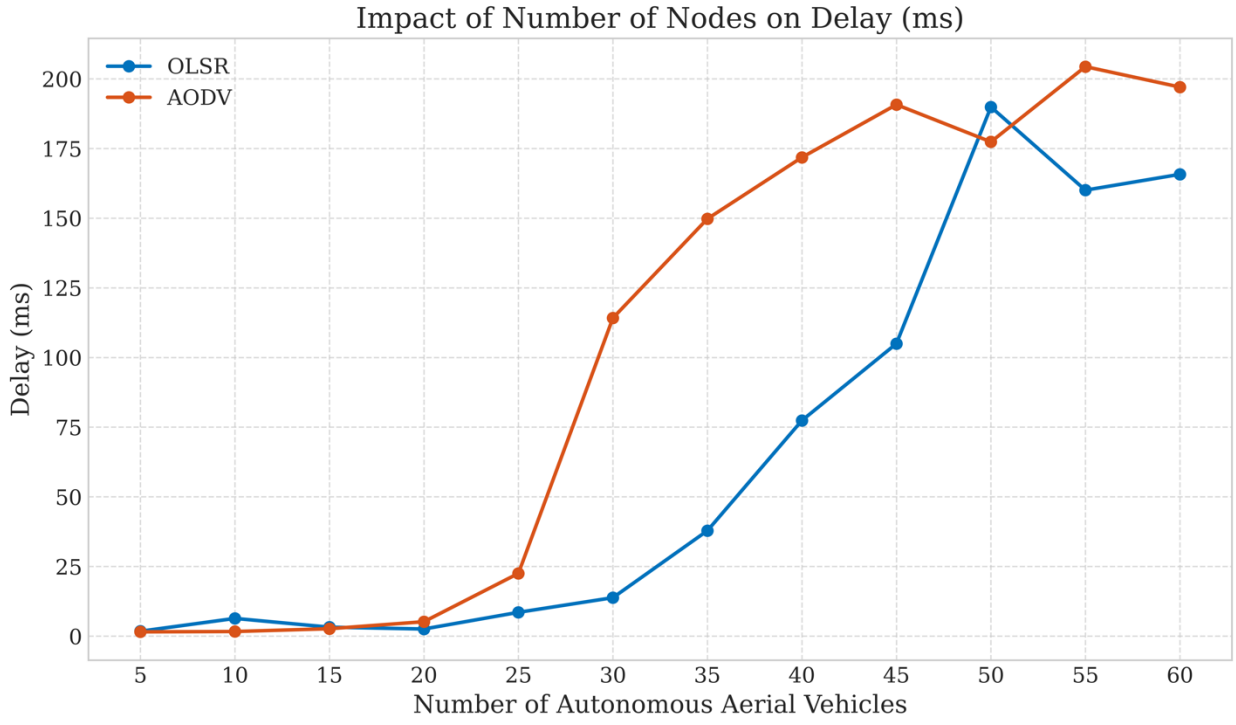


Fig. 4. Impact of Number of Nodes on Delay

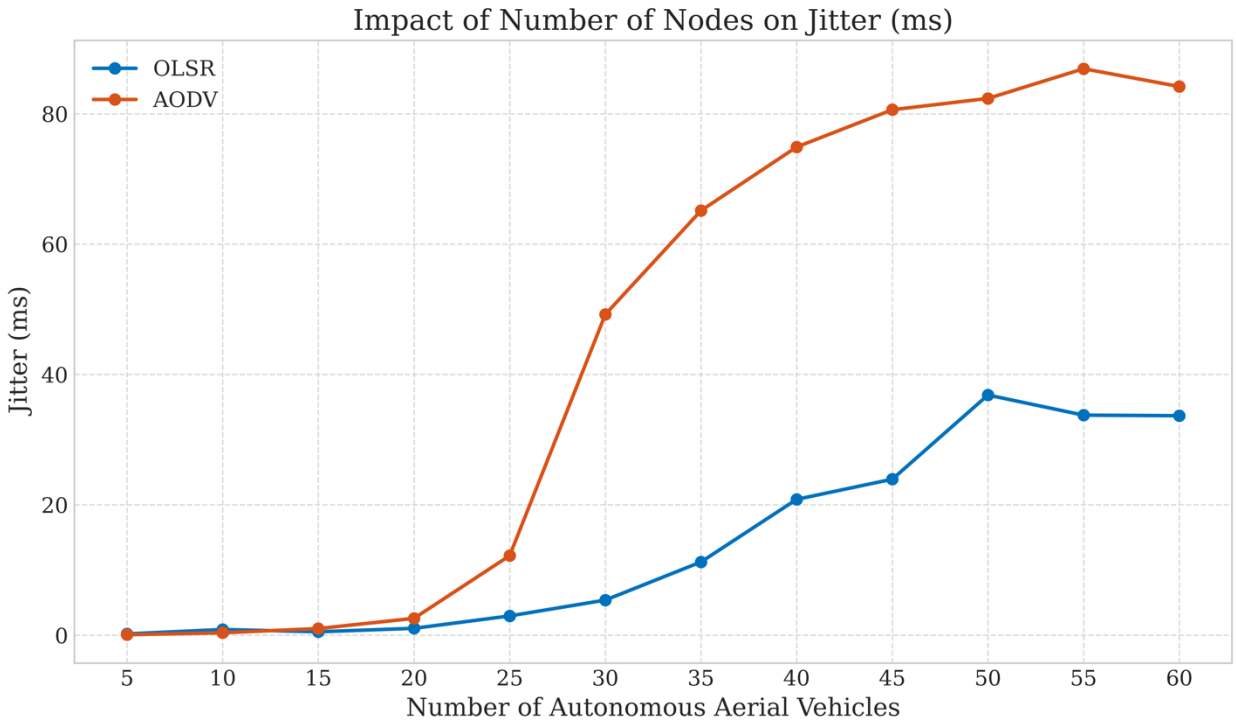


Fig. 5. Impact of Number of Nodes on Jitter

Initially, we compared the performance of protocols for Autonomous Aerial Vehicle networks, AODV and OLSR, in terms of Packet Delivery Ratio (PDR) to determine the appropriate choice for reconnaissance scenarios. In the Autonomous Aerial Vehicle network model, node density is typically sparse, which complicates maintaining connections between nodes. Therefore, we evaluated the performance of AODV and OLSR by gradually increasing the number of nodes during the simulation, thereby determining the impact of node density on network performance.

Fig. 2 illustrates the PDR for each protocol, where the horizontal axis represents the number of nodes, and the vertical axis represents the PDR. The results show that at low node density, AODV achieves a higher packet delivery ratio. In a reconnaissance scenario, the network topology changes rapidly and frequently, allowing AODV (a reactive protocol) to optimize route discovery, when necessary, thereby maintaining a relatively good PDR at small scales. In contrast, OLSR stores a routing table for all nodes, resulting in higher costs to maintain the routing table, reducing PDR in an environment with a small number of nodes.

We also tested the hypothesis that the PDR of Autonomous Aerial Vehicle network protocols would increase with the growth in the number of nodes. In practice, both AODV and OLSR show an increase in PDR up to a certain maximum, followed by a decline. For example, AODV reaches a PDR of about 94% with 15 nodes, after which a decrease is observed; similarly, OLSR reaches 93.87% also with 15 nodes and decreases starting from 20 nodes. This indicates that the optimal node density for both protocols in our simulation environment is around 15 nodes.

Regarding throughput, both AODV and OLSR tend to increase with the number of nodes up to a certain threshold before declining. OLSR reaches a peak throughput of about 2575.86 Kbps with 40 nodes, while AODV reaches approximately 1871.53 Kbps. This result shows that at medium and high node densities, OLSR is capable of processing data more efficiently.

Considering delay, both protocols show a notable increase in delay as the network expands. AODV maintains lower delay at low node densities but grows faster compared to OLSR with a large number of nodes. At 40 nodes, OLSR registers about 77.39 ms, while AODV reaches 171.82 ms. This demonstrates the advantage of OLSR in terms of delay in networks with many nodes.

Finally, the jitter value also increases proportionally to the network size, and AODV typically has higher jitter than OLSR. The reason lies in the "reactive" nature of AODV, which must frequently search for new

routes when the topology is constantly changing.
Findings: In a reconnaissance scenario with low node density, AODV proves to be more suitable. At the same time, OLSR works more efficiently at medium and high node densities. Additionally, the PDR performance of Autonomous Aerial Vehicle network protocols does not increase continuously with the growth in the number of nodes but begins to decline after a certain threshold due to increased costs for exchanging routing information.

PERFORMANCE ANALYSIS BASED ON AUTONOMOUS AERIAL VEHICLE VELOCITY

In this section, we evaluate the impact of node movement velocity on the performance of AODV and OLSR routing protocols. In Autonomous Aerial Vehicle networks, nodes move at medium to high velocities, leading to rapid and frequent changes in network structure. Therefore, we conduct simulations varying node velocities from 5 m/s to 60 m/s to assess the protocols' ability to adapt to changes in network structure.

Table 3. Simulation Parameters Based on Node Velocity

Parameter	Value
Number of Nodes	20
Velocity	5–60 m/s
Mobility model	GaussMarkov
Flight Area	800 m × 800 m × 100 m

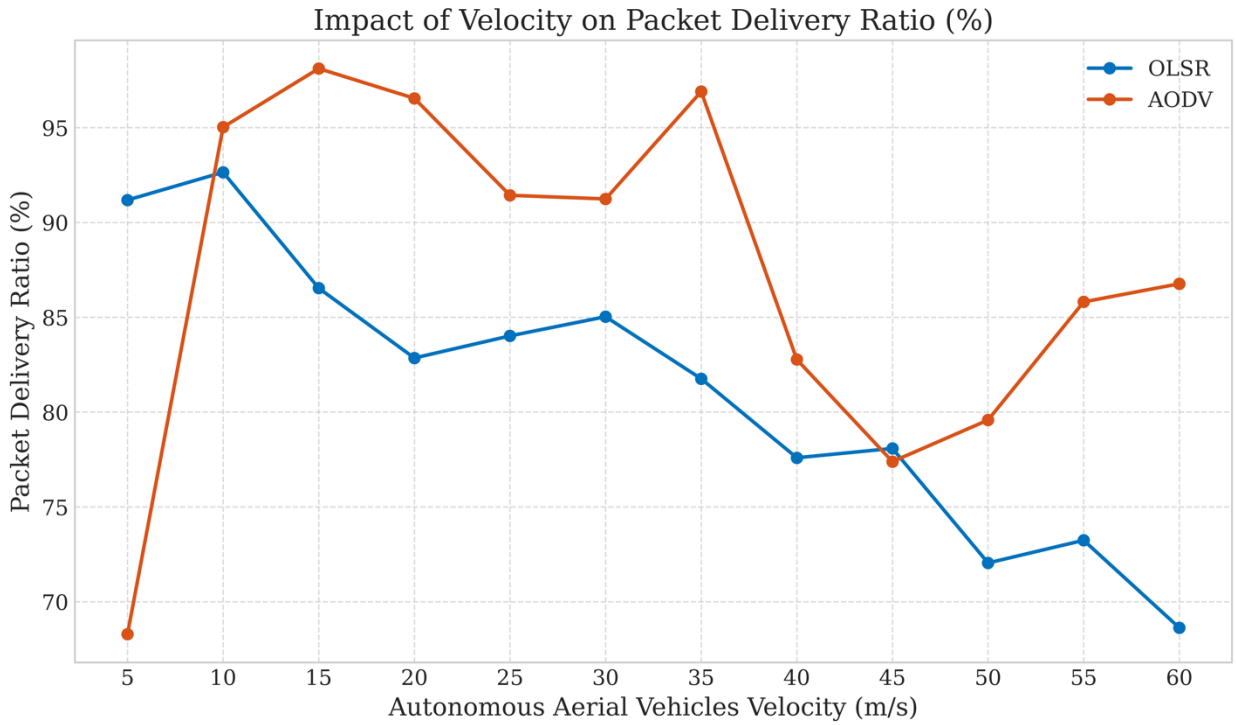


Fig. 6. Impact of Velocity on Packet Delivery Ratio

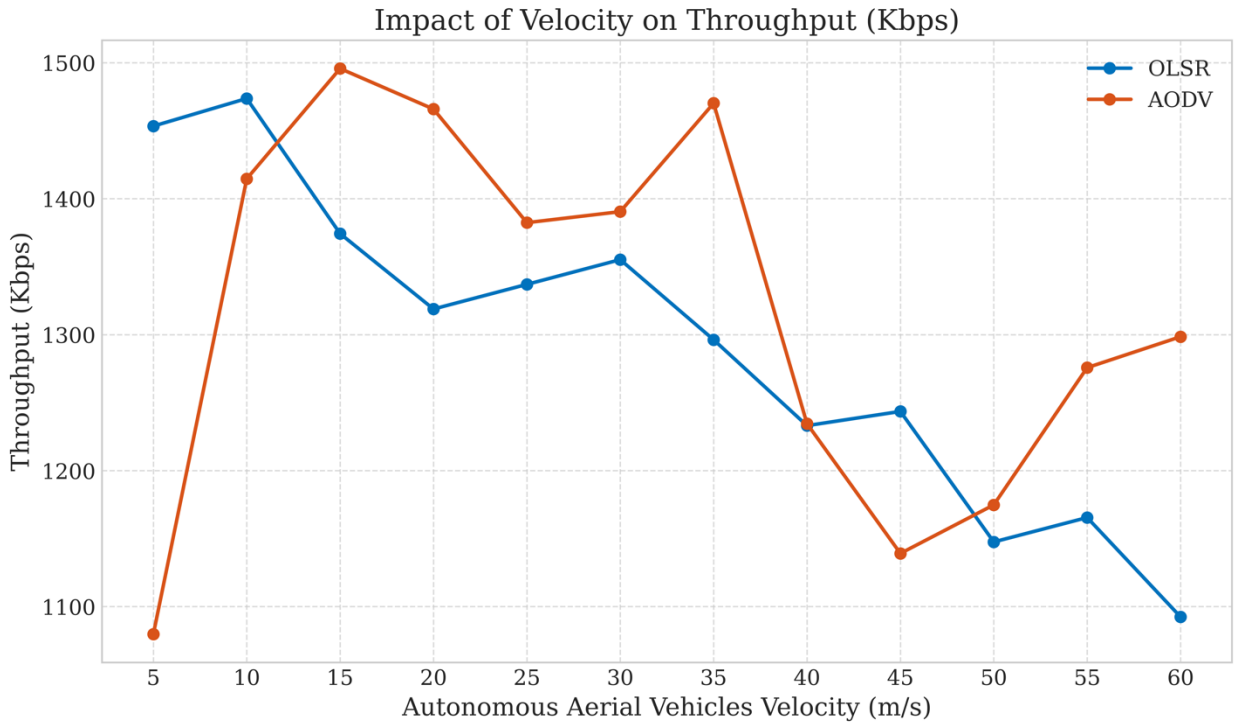


Fig. 7. Impact of Velocity on Throughput

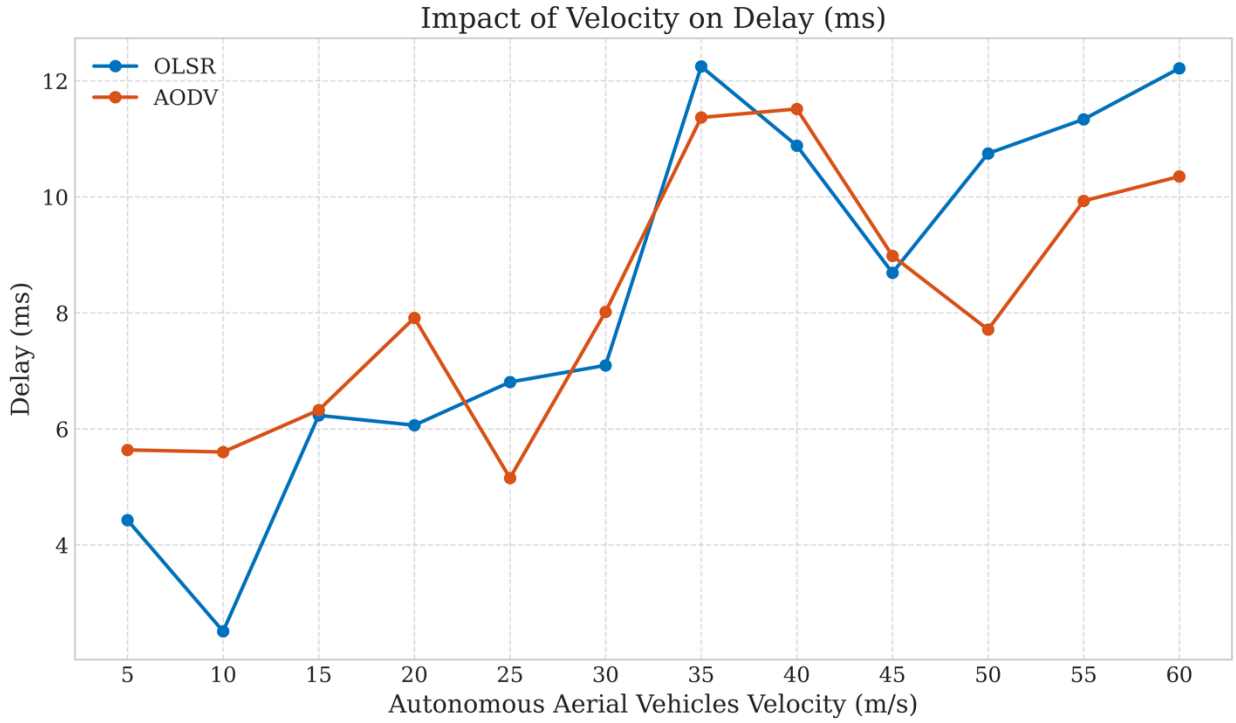


Fig. 8. Impact of Velocity on Delay

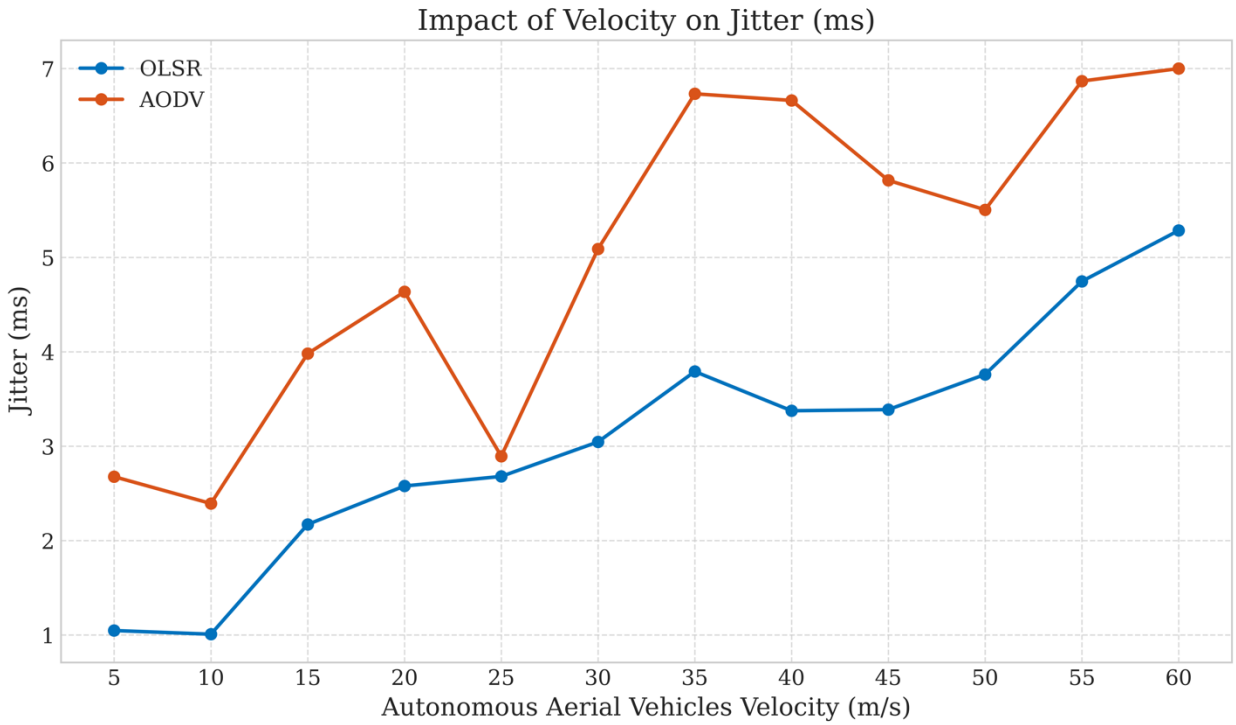


Fig. 9. Impact of Velocity on Jitter

Fig. 6 shows the Packet Delivery Ratio (PDR) as node velocity changes. In the figure, the x-axis represents node velocity (m/s), and the y-axis represents the PDR. From Fig. 6, we observe that the PDR performance of OLSR gradually decreases with increasing node velocity. Specifically, the PDR of the OLSR protocol decreases from 91.18% at a velocity of 5 m/s to 68.63% at a velocity of 60 m/s. This shows that OLSR, a proactive routing protocol, experiences difficulties maintaining an up-to-date routing table at high node velocities.

In contrast, AODV demonstrates better adaptability to high node velocities. The PDR of the AODV protocol remains above 95% at velocities of 15 m/s and 20 m/s, and decreases only to approximately 86.77% at a velocity of 60 m/s. This proves that AODV, a reactive routing protocol, can effectively establish routes on demand, even when the network topology changes rapidly.

Regarding throughput, AODV maintains higher throughput than OLSR at most node velocities. At a velocity of 20 m/s, AODV achieves a throughput of about 1465.91 Kbps, while OLSR reaches 1318.86 Kbps. When the velocity increases to 60 m/s, AODV's throughput decreases to 1298.48 Kbps, while OLSR significantly drops to 1092.45 Kbps.

In terms of delay, we found that AODV has higher delay than OLSR at most velocities. At a velocity of 20 m/s, AODV has a delay of about 7.91 ms, while OLSR has 6.06 ms. When the velocity increases to 60 m/s, AODV's delay increases to 10.35 ms, while OLSR increases to 12.21 ms. This shows that at high velocities, the cost of maintaining OLSR's routing table increases significantly, leading to higher delay.

The jitter of both protocols increases with node velocity, but AODV demonstrates higher jitter than OLSR in most cases. At a velocity of 60 m/s, AODV's jitter is about 7.00 ms, while OLSR's is 5.28 ms. This is due to the reactive nature of AODV, which requires more frequent establishment of new routes in an environment with constantly changing topology.

Findings: The AODV routing protocol demonstrates better adaptability to high node velocities in Autonomous Aerial Vehicle networks. AODV's PDR performance remains at a higher level than OLSR's as node velocity increases, showing that AODV is more suitable for high-mobility reconnaissance scenarios. However, AODV has higher delay and jitter, which may affect delay-sensitive real-time applications.

PERFORMANCE ANALYSIS BASED ON MOBILITY MODEL

In this simulation, we analyze the performance of OLSR and AODV protocols with four different mobility models: GaussMarkov, RandomWaypoint, RandomDirection, and RandomWalk3d. The number of nodes is fixed at 20, and the velocity is 25 m/s.

Table 4. Simulation Parameters Based on Mobility Model

Parameter	Value
Number of Nodes	20
Velocity	25 m/s
Mobility Model	GaussMarkov, RandomWaypoint, RandomDirection, RandomWalk3d
Flight Area	800 m × 800 m × 100 m

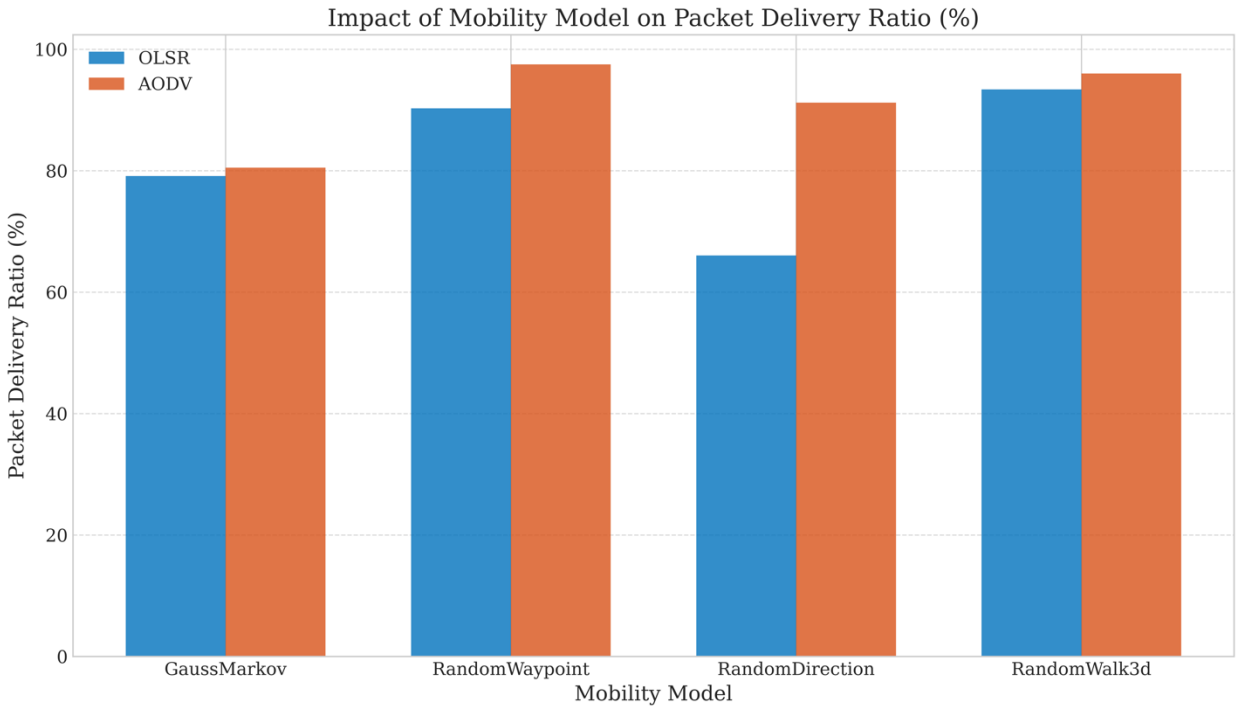


Fig. 10. Impact of Mobility Model on Packet Delivery Ratio

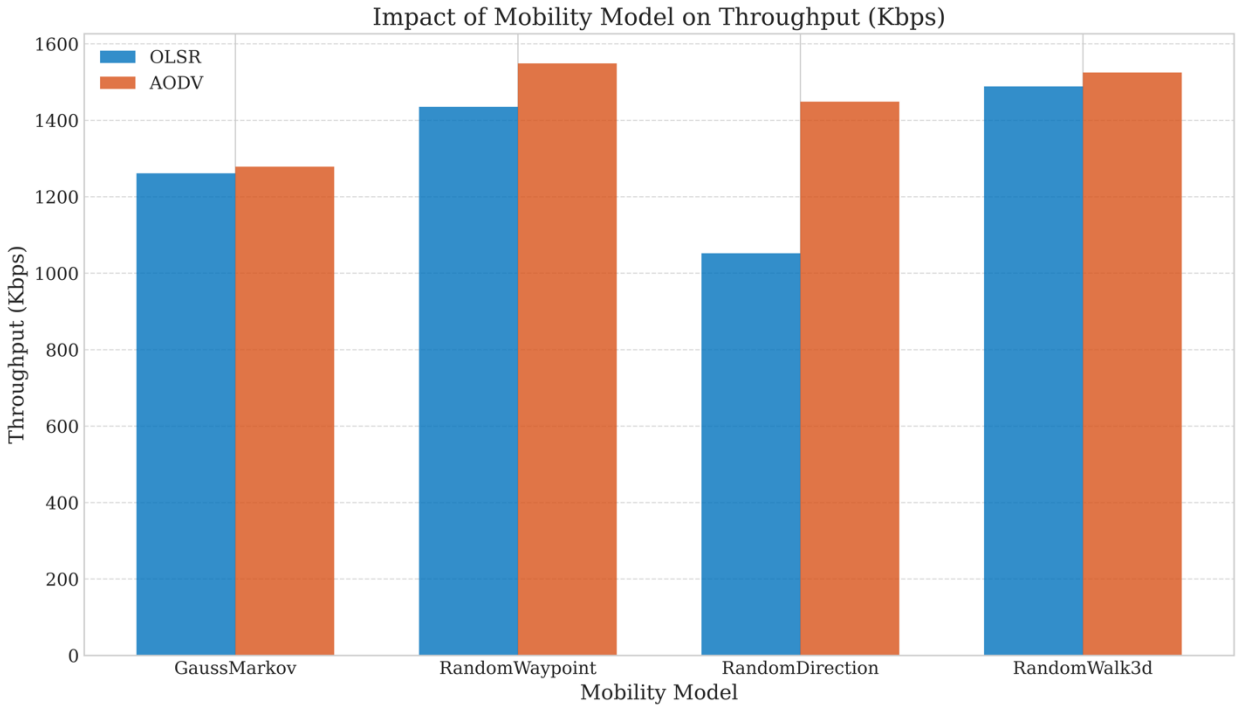


Fig. 11. Impact of Mobility Model on Throughput

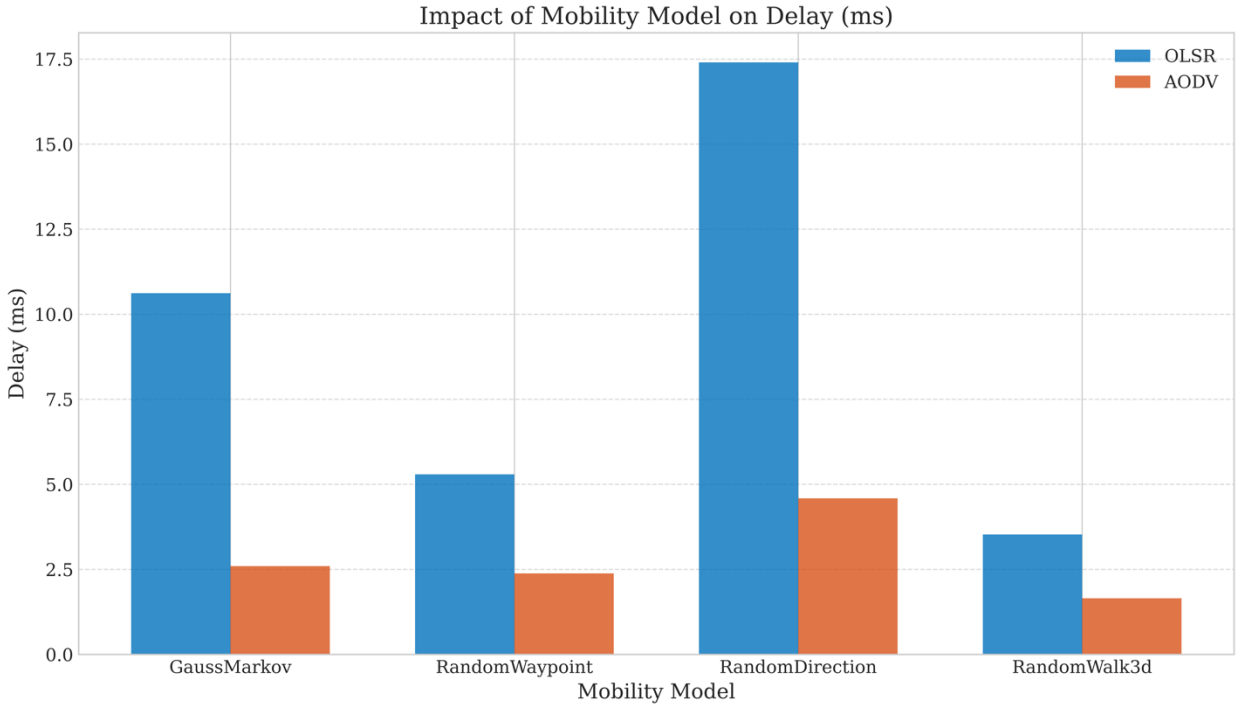


Fig. 12. Impact of Mobility Model on Delay

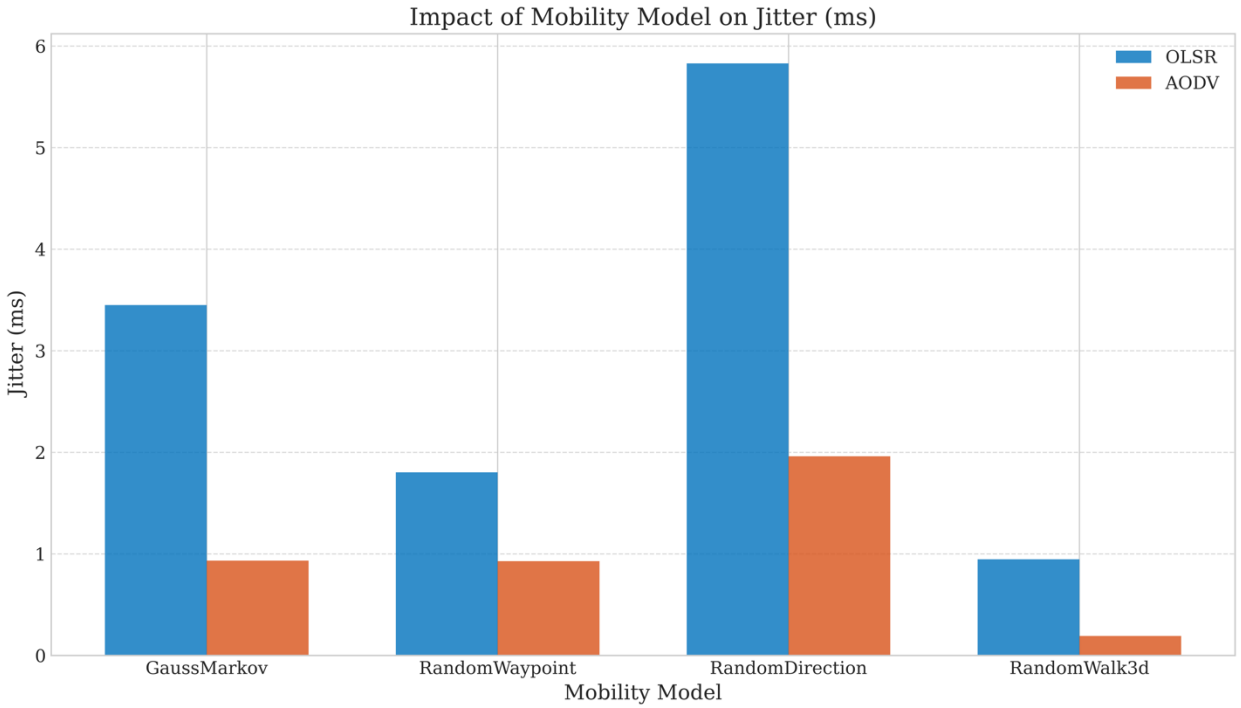


Fig. 13. Impact of Mobility Model on Jitter

From the simulation data, it is evident that the PDR of the OLSR protocol ranges from 66.05% to 93.39% depending on the mobility model. The RandomWalk3d model provides the highest PDR for OLSR (93.39%), followed by RandomWaypoint (90.26%), GaussMarkov (79.12%), and the lowest performance with RandomDirection (66.05%).

For AODV, PDR values vary from 91.21% to 97.52%, with RandomWaypoint providing the highest performance (97.52%), followed by RandomWalk3d (96.00%), RandomDirection (91.21%), and GaussMarkov (80.50%).

Regarding throughput, both OLSR and AODV achieve their highest values with the RandomWaypoint model, 1435.17 Kbps and 1548.92 Kbps respectively. The RandomDirection model gives the lowest throughput for OLSR (1052.12 Kbps), while GaussMarkov gives the lowest throughput for AODV (1278.63 Kbps).

Delay also varies significantly depending on the mobility model. AODV has the lowest delay with RandomWalk3d (1.65 ms) and the highest with RandomDirection (4.59 ms). OLSR has the lowest delay with RandomWalk3d (3.53 ms) and the highest with RandomDirection (17.40 ms).

Findings: These results show that the mobility model has a strong influence on the performance of routing protocols in Autonomous Aerial Vehicle networks. RandomWaypoint and RandomWalk3d models generally provide better performance for both OLSR and AODV, while RandomDirection creates more serious challenges due to frequent and unpredictable changes in network topology.

PERFORMANCE ANALYSIS BASED ON COVERAGE AREA

In this simulation, we analyze the performance of OLSR and AODV protocols when changing network density, maintaining a constant number of nodes (25), but varying the simulation area size from 500 m × 500 m to 2000 m × 2000 m. This is equivalent to changing node density from 100 nodes/km² to 6.25 nodes/km². The purpose of this simulation is to investigate the effect of node density on routing protocol performance. By varying the flight area size from 500×500 m² to

2000×2000 m² with a fixed number of nodes, we create environments with different node densities, from dense (100 nodes/km²) to sparse (6.25 nodes/km²).

Table 5. Simulation Parameters Based on Area Dimensions

Parameter	Value
Number of Nodes	25
Velocity	10 m/s
Mobility Model	GaussMarkov
Flight Area	500×500, 800×800, 1200×1200, 1600×1600, 2000×2000 m²

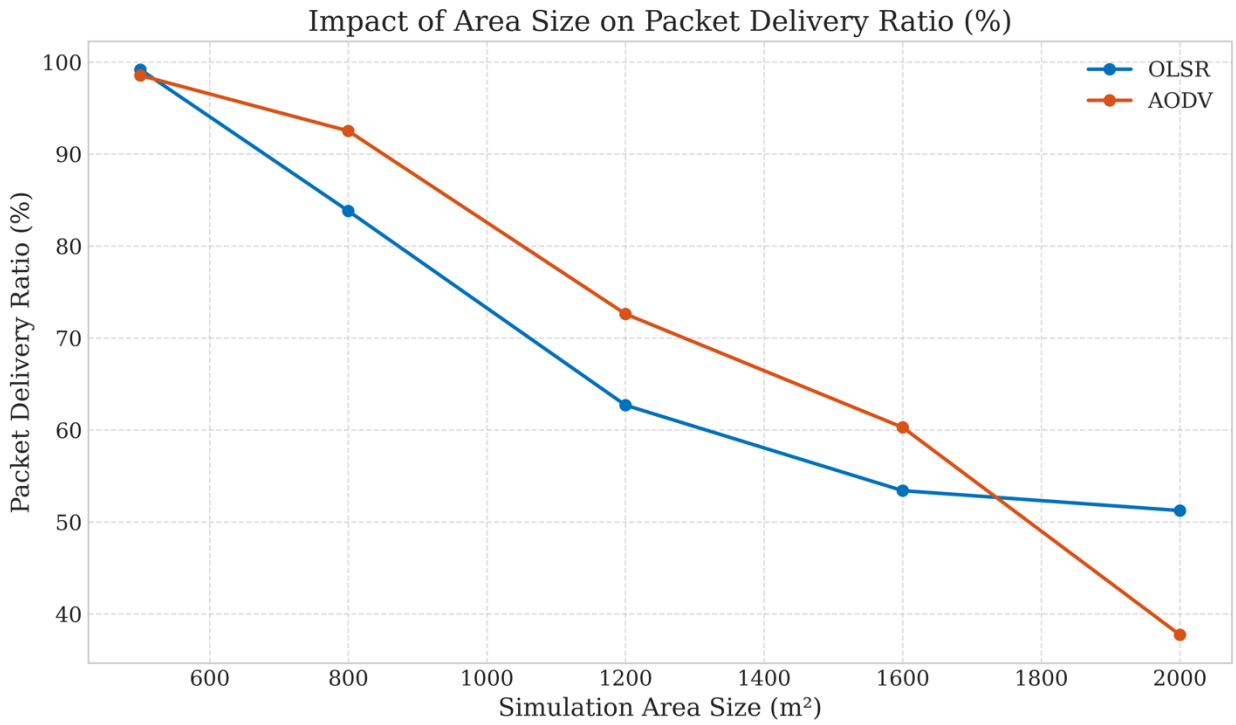


Fig. 14. Impact of Simulation Area Size on Packet Delivery Ratio

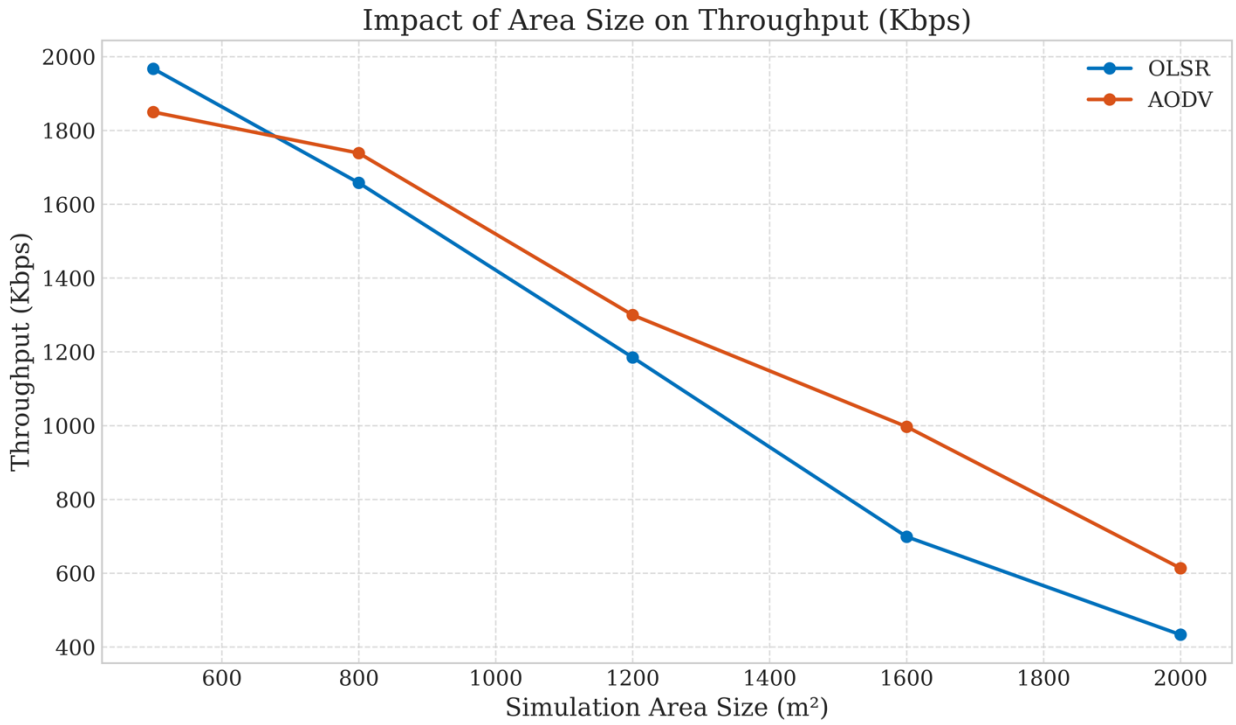


Fig. 15. Impact of Simulation Area Size on Throughput

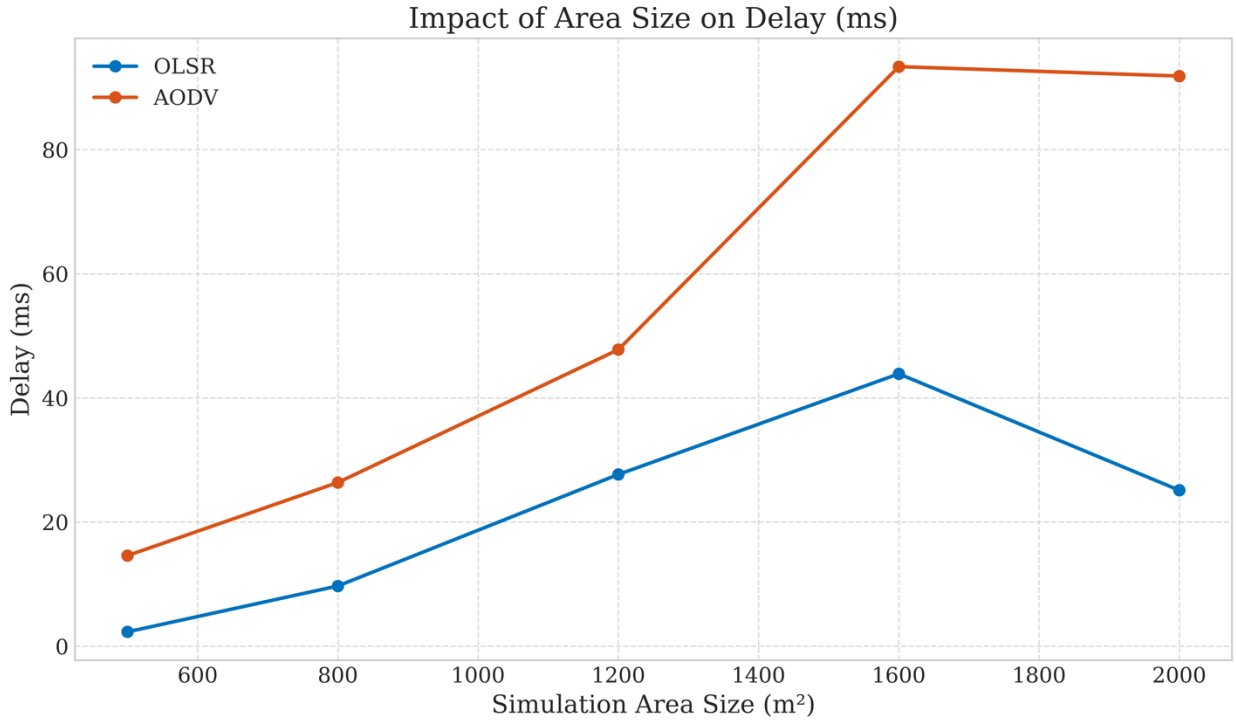


Fig. 16. Impact of Simulation Area Size on Delay

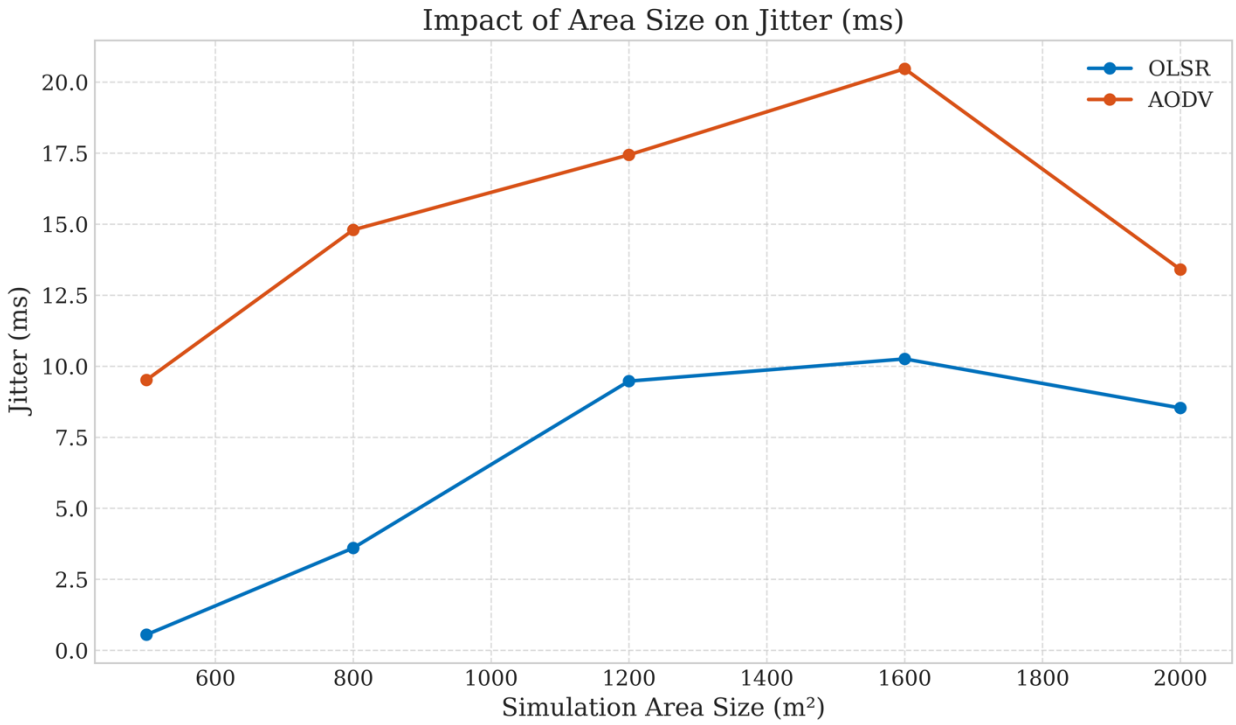


Fig. 17. Impact of Simulation Area Size on Jitter

The purpose of this simulation is to study the effect of node density on routing protocol efficiency. By varying the flight area size from 500×500 m² to 2000×2000 m² with a fixed number of nodes, we create different conditions with varying node densities - from high (100 nodes/km²) to extremely low (6.25 nodes/km²).

The results show that the Packet Delivery Ratio (PDR) in both OLSR and AODV protocols decreases as node density decreases. At the highest density (100 nodes/km²), PDR in OLSR reaches 99.18%, and in AODV - 98.55%. However, at the lowest density level (6.25 nodes/km²), PDR in OLSR decreases to 51.24%, and in AODV - to 37.73%.

Throughput also significantly drops as node density decreases. In OLSR, the maximum value is 1967.18 Kbps at 100 nodes/km² and decreases to 433.18 Kbps at 6.25 nodes/km². Similarly, in AODV, throughput reaches 1849.75 Kbps at maximum density and decreases to 613.46 Kbps at low density.

Regarding delay, an interesting trend is observed: as node density decreases from high to medium, delay increases, but then decreases at extremely low density. In OLSR, delay increases from 2.30 ms (at 100 nodes/km²) to 43.88 ms (at 9.77 nodes/km²) and then decreases to 25.12 ms (at 6.25 nodes/km²). A similar trend is found in AODV: from 14.59 ms (maximum density) to 93.38 ms (9.77 nodes/km²) and a further decrease to 91.87 ms (6.25 nodes/km²).

This phenomenon can be explained by the specifics of routing protocols. At high density, nodes have multiple alternative routes, ensuring low delay. However, as density decreases, the number of available paths decreases, and routes need to be rebuilt more frequently, increasing delay. At extremely low density, many packets do not reach their destination at all due to lack of connectivity, and the average delay only considers successfully delivered packets, usually with a short transmission path.

Comparing OLSR and AODV shows that OLSR generally demonstrates better results at most density levels, especially in terms of PDR and delay at low density. This may be related to the proactive nature of OLSR, allowing it to maintain more stable routing information even in conditions of weak connectivity.

Findings: Node density significantly affects the efficiency of routing protocols in Autonomous Aerial Vehicle networks. As node density decreases, PDR and throughput significantly decrease. OLSR proves to be more adaptive to low node density compared to AODV. In reconnaissance scenarios over a large area

with a limited number of Autonomous Aerial Vehicles, optimization of node placement or implementation of intermediate relays may be required to maintain the required level of network performance.

IV. Practical Applications

Based on comprehensive analysis, we found that neither protocol completely outperforms the other in all situations. AODV and OLSR each have their advantages depending on the deployment scenario. The choice of an appropriate protocol should be based on the specifics of the application, the expected number of deployed Autonomous Aerial Vehicles, mobility velocity, and QoS requirements. Based on simulation results, we propose the following suitable applications for OLSR and AODV in real-world Autonomous Aerial Vehicle scenarios:

Applications Suitable for OLSR:

1. Large-scale Infrastructure Monitoring: In tasks requiring the deployment of multiple Autonomous Aerial Vehicles for monitoring power lines, pipelines, or transportation systems, OLSR demonstrates an advantage in handling a large number of nodes with high performance. The ability to maintain stable throughput above 1500 Kbps at densities of 25-40 nodes makes OLSR suitable for transmitting high-quality images and real-time video.

2. Autonomous Aerial Vehicle Networks with Static or Slow Movement: For long-term monitoring scenarios such as environmental monitoring, precision agriculture, or wildlife research, where Autonomous Aerial Vehicles move at low velocities (5-15 m/s) or fly along fixed trajectories, OLSR provides low delay and reasonable routing information forwarding costs.

3. Deployment in Concentrated Coverage Areas: OLSR is most effective when Autonomous Aerial Vehicles operate in high-density areas such as urban areas, large events, or concentrated combat zones. With PDR above 99% at a density of 100 nodes/km², OLSR ensures reliable data exchange between Autonomous Aerial Vehicles and ground stations.

4. Applications for 3D Mapping and Spatial Modeling: When deploying Autonomous Aerial Vehicles to create 3D models of structures or complex terrain, OLSR with the RandomWalk3d mobility model provides high PDR (93.39%) and low delay (3.53 ms), suitable for continuous transmission of sensor data.

Applications Suitable for AODV:

1. Reconnaissance Missions and Moving Target Tracking: With the ability to maintain high PDR (above 95%) at velocities up to 35 m/s, AODV is suitable for rapid reconnaissance missions, tracking moving targets, or search and rescue operations where Autonomous Aerial Vehicles need to move quickly and flexibly.

2. Deployment in Vast Areas with Low Density: AODV shows greater flexibility when Autonomous Aerial Vehicles operate dispersed in vast areas, especially with the RandomWaypoint model (PDR 97.52%). This is an appropriate choice for missions exploring border territories, forested mountains, or islands.

CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

In this study, we conducted a comprehensive performance evaluation of two representative routing protocols (OLSR and AODV) in Autonomous Aerial Vehicle network environments through four highly practical simulation scenarios. The results show that no protocol has absolute superiority, but each protocol has its advantages depending on deployment conditions. OLSR demonstrates

high performance in high-density and low-velocity environments, while AODV is more suitable for scenarios with high velocity and rapidly changing topology.

Based on the analysis, we proposed several potential improvements for both protocols, as well as a hybrid protocol to overcome existing limitations. In particular, the integration of parameters such as velocity, energy, and link quality into the routing process, as well as the application of reinforcement learning, promise to significantly enhance the performance of Autonomous Aerial Vehicle networks in practical applications.

Directions for future research will focus on implementing and evaluating the proposed improvements through more detailed simulations and practical tests. We plan to:

1. Develop and implement improved versions of OLSR and AODV on the NS-3 platform, evaluating their performance compared to the original protocols.

2. Expand simulations to include more realistic factors such as energy disruptions, environmental interference, and visibility limitations.

3. Study the impact of encryption and security protocols on routing performance in Autonomous Aerial Vehicle networks.

4. Develop analytical models based on machine learning to predict network performance in complex scenarios.

5. Test hybrid routing architecture in real deployment scenarios such as disaster monitoring and support for rescue operations.

Through these efforts, we aim to create optimal routing solutions for Autonomous Aerial Vehicle networks, contributing to the application of Autonomous Aerial Vehicles in many important areas of modern society.

REFERENCES

[1] A. Rovira-Sugranes, A. Razi, F. Afghah, and J. Chakareski, “A Review of AI-enabled Routing Protocols for UAV Networks: Trends, Challenges, and Future Outlook,” Nov. 08, 2021, arXiv: arXiv:2104.01283. doi: 10.48550/arXiv.2104.01283.

[2] Y. Yamazaki, M. Tamaki, C. Premachandra, C. J. Perera, S. Sumathipala, and B. H. Sudantha, “Victim Detection Using UAV with On-board Voice Recognition System,” in 2019 Third IEEE International Conference on Robotic Computing (IRC), Feb. 2019, pp. 555–559. doi: 10.1109/IRC.2019.00114.

[3] N. Ya’acob, M. Zolkapli, J. Johari, A. L. Yusof, S. S. Sarnin, and A. Z. Asmadinar, “UAV environment monitoring system,” in 2017 International Conference on Electrical, Electronics and System Engineering (ICEESE), Nov. 2017, pp. 105–109. doi: 10.1109/ICEESE.2017.8298395.

[4] M. A. Khan et al., “Swarm of UAVs for Network Management in 6G: A Technical Review,” IEEE Trans. Netw. Serv. Manag., vol. 20, no. 1, pp. 741–761, Mar. 2023, doi: 10.1109/TNSM.2022.3213370.

[5] T. H. Clausen and P. Jacquet, “Optimized Link State Routing Protocol (OLSR),” Internet Engineering Task Force, Request for Comments RFC 3626, Oct. 2003. doi: 10.17487/RFC3626.

[6] I. D. Chakeres and E. M. Belding-Royer, “AODV routing protocol implementation design,” in 24th International Conference on Distributed Computing Systems Workshops, 2004. Proceedings., Hachioji, Tokyo, Japan: IEEE, 2004, pp. 698–703. doi: 10.1109/ICDCSW.2004.1284108.

[7] L. Wang and S. Olariu, “A two-zone hybrid routing protocol for mobile ad hoc networks,” IEEE Trans. Parallel Distrib. Syst., vol. 15, no. 12, pp. 1105–1116, Dec. 2004, doi: 10.1109/TPDS.2004.73.

[8] Q. Usman, O. Chughtai, N. Nawaz, Z. Kaleem, K. A. Khaliq, and L. D. Nguyen, "A Reliable Link-adaptive Position-based Routing Protocol for Flying Ad hoc Network," Oct. 12, 2020, arXiv: arXiv:2010.06005. doi: 10.48550/arXiv.2010.06005.

[9] C. Liu, Y. Wang, and Q. Wang, "PARouting: Prediction-supported adaptive routing protocol for FANETs with deep reinforcement learning," Int. J. Intell. Netw., vol. 4, pp. 113–121, 2023, doi: 10.1016/j.ijin.2023.05.002.

[10] C. E. Perkins and P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers," ACM SIGCOMM Comput. Commun. Rev., vol. 24, no. 4, pp. 234–244, Oct. 1994, doi: 10.1145/190809.190336.

[11] D. B. Johnson, D. A. Maltz, and J. Broch, "DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks".

[12] M. R. Pearlman and Z. J. Haas, "Determining the optimal configuration for the zone routing protocol," IEEE J. Sel. Areas Commun., vol. 17, no. 8, pp. 1395–1414, Aug. 1999, doi: 10.1109/49.779922.

[13] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in Proceedings of the 6th annual international conference on Mobile computing and networking, Boston Massachusetts USA: ACM, Aug. 2000, pp. 243–254. doi: 10.1145/345910.345953.

[14] J. Zhou, Y. Chen, B. Leong, and P. S. Sundaramoorthy, "Practical 3D geographic routing for wireless sensor networks," in Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems, Zürich Switzerland: ACM, Nov. 2010, pp. 337–350. doi: 10.1145/1869983.1870016.

Analysis of Thermal Performance of Shell and Tube Heat Exchangers: A Correlation and CFD Based Approach

S.A.P Ushettige[1], W.K. Wimalsiri[2], H.G.S. Hikkaduwa[3]
[1]Faculty of Engineering, School of Civil and Mechanical Engineering, Curtin University,
Bentley, WA 6102, Australia
aperera844@gmail.com ; walallawita.w@sliit.lk
[2]Faculty of Engineering, Mechanical Engineering, Sri Lanka Institute of Information Technology
Malabe, 10115, Sri Lanka
[3]Faculty of Engineering, General Sir John Kothalawala Defence University
Kandawala Road, Dehiwala-Mount Lavinia, 10390, Sri Lanka
hikkaduwahgs@kdu.ac.ok

ABSTRACT
Shell and tube heat exchangers are devices which are widely adopted in thermal systems for the transfer of thermal energy due to both performance and reliability factors. Given their application in systems that are energy intensive, the design and sizing of these devices have become a widely growing field. Traditionally, empirical correlations which were based on experimental results were used for thermal sizing and design. This was replaced by computational fluid dynamics (CFD) modelling given its ability to model and visualize flow expanding the horizon of possibilities for design and performance optimization. As of recent, CFD has been combined with numerical methods such as non-linear least squares regression to develop correlations which can predict the thermal performance based on the input design parameters. However, the application of this integrated method for shell and tube heat exchangers is limited. This study will model a single pass TEMA E-type shell and tube heat exchanger using ANSYS Fluent ®. CFD simulations are used to explore the effect of turbulence on thermal performance by varying both the inlet mass flow rate and the central baffle spacing. As steady state simulations are conducted for four models of six, eight, ten and twelve baffles. The results of CFD modelling are then combined with non-linear least squares regression in the MATLAB Curve Fitter Toolbox ® to develop four sets of correlation in the form of $Nu = C.Re^{a}.Pr^{b}$. Reasonably confident results were obtained in the final fitted data, however, relatively high 95% confidence interval widths were evident for certain fitted coefficients leaving space for improvement in the model. The study highlights that combining CFD with tools such as nonlinear least squares regression aids both engineers and designers in the thermal design process of shell and tube heat exchangers eliminating the need to limit design based on empirical correlations.

KEYWORDS: *Shell and tube heat exchangers, thermal optimization, non-linear least squares regression, central baffle spacing, turbulence, mass flow rate*

INTRODUCTION
Heat exchangers are devices which are commonly used for the transfer of thermal energy among process streams in both commercial and domestic systems (Mangrulkar et al., 2019). A range of different heat exchanger types are used in the industry. Among them, shell and tube heat exchangers are widely adopted due to their versatility to operate in a range of working pressures and temperatures (Sekulić & Shah, 2023). These devices also provide a large area of heat transfer for a given volume with a relatively simple construction (Pal et al., 2016; Aslam Bhutta et al., 2012). In shell and tube heat exchangers, two fluids at relatively different temperatures flow simultaneously. One flows through the tubes (tube side) and other flows outside the tubes or the shell which is a circular enclosed volume around the tubes. The transfer of energy happens due to a combination of conduction and convection from the inside of the tube fluid to the outside shell fluid or vice versa (Bichkar et al., 2018). This is represented in Figure 1 (a) and (b) below.

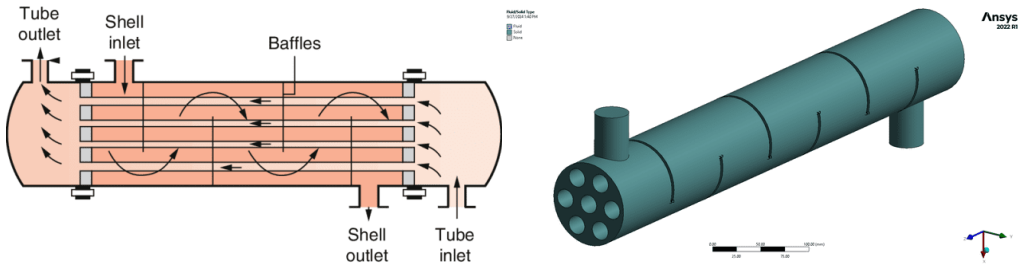


Figure 61: (a) Schematic of a Single Pass TEMA E Type shell and tube heat exchanger (M. Necati Özışık, 1985) (b) Physical model in ANSYS Design Modeler.

The widespread adoption of these devices in industries with significant energy transfers has made the design and sizing of them critical, of which thermal and pressure drop are the key parameters analyzed (Huang et al., 2017). Traditionally, empirical correlations developed by organizations such as the TEMA (Tubular Exchanger Manufacturers Organization) were used. This process was time consuming and restricted design, due to the inability to visualize and identify design weaknesses in the flow (Ozden & Tari, 2010). The adoption of CFD for modelling addressed this challenge given its ability to model and visualize complex flow phenomena.

The thermo-mechanical performance in these devices depends on the level of turbulence present. This is commonly introduced using metal plates known as baffles. A range of different types and configurations can be used to achieve this (Bichkar et al., 2018). Traditionally most models use single segmental baffles. However, the use of these types has two key drawbacks (Ozden & Tari, 2010; Pal et al., 2016; Nedunchezhiyan Mukilarasan et al., 2022; El Maakoul et al., 2016; Abbasi et al., 2020; Zhang et al., 2017). In the region present behind the baffle "Dead zones" or areas of recirculation are created which reduce the amount of heat transfer. Additionally, the contraction and expansion effect for the fluid flow increases the pressure drop, increasing the pumping power required. A common approach of optimizing performance is by adjusting the configuration of baffles, central baffle spacing and the baffle cut (Ozden and Tari, 2010; Abbasi et al., 2020). The central baffle spacing and baffle cut are identified as parameters which significantly affect the flow phenomenon and thermo-mechanical performance. Numerical work states that decreasing baffle spacing is linked to increasing heat transfer and nozzle-to-nozzle pressure drop.

The potential of design and optimization achievable by CFD has been recently combined with tools such as non-linear least squares regression. The integration of the two tools together has allowed for the development of correlations which can relate thermal performance to the input design parameters of the model. However, following a comprehensive survey of the existing literature minimal applications were available for shell and tube heat exchangers. Established applications include those done for a horizontal counter-flow double tube heat exchanger and plate-fin and tube heat exchanger, as done by Alhulaifi (2024) and Marcinkowski et al. (2021) respectively. The study done by Alhulaifi (2024) explored the thermal performance of a nano-fluid flow in the device where a correlation which predicts the *Nu* based on *Pr* and *Re* was developed. While the study by Marcinkowski et al. (2021) involved the development of a set of correlations which can predict the *Nu* of each row based on the respective *Pr* and *Re*. A further development of work was that done by Marcinkowski et al. (2024) which highlighted that varying the modelling scheme had minimal effect on the final outcome.

This study will use the commercial CFD code solver ANSYS Fluent to model both the thermal and mechanical performance of a single pass TEMA E-type shell and tube heat exchanger. Simulations will be conducted in a steady state for the case of six, eight, ten and twelve baffles. In the effort to develop comprehensive correlations which capture the effect of turbulence on thermal performance both the central baffle spacing and the inlet mass flow rate are varied. For each model, simulations are done from the range of 0.50 kg/s to 2.0 kg/s with increments of 0.25 kg/s. The results from the CFD analysis are then fit using non-linear least squares regression in the MATLAB Curve Fitter Toolbox to develop four sets of

correlations in the form of $Nu = C.Re^{a}.Pr^{b}$ where a confidence level of 95 % is used for the purpose of modelling. Overall, the study is an effort to highlight the potential present in combining CFD with tools such as non-linear least squares regressions for the design and optimization of shell and tube heat exchangers.

METHODOLOGY

Physical Modelling of STHX

The shell and tube heat exchangers were modelled using the ANSYS Design Modeler ® tool based on the dimensions shown in Table 1 below (Ozden & Tari, 2010). For simulation purposes, the heat transfer through the thickness of the baffles was neglected.

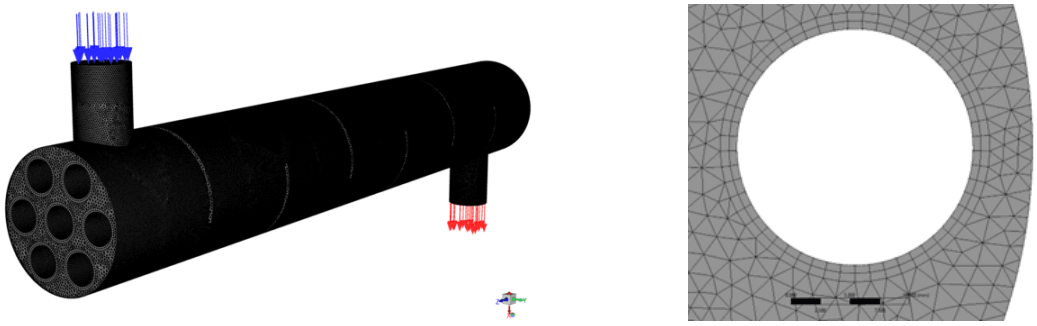
Table 35. Dimensions of the shell and tube heat exchanger (Ozden & Tari, 2010).

Dimensional Parameter	Value / mm
Shell Diameter	90
Outside Tube Diameter	20
Number of tubes used	7 tubes
Length of the heat exchanger	600
Central baffle spacing	86

The working fluid considered is water while, Aluminum is taken as the material for the tube walls. The presence of narrow operating temperature ranges was the reason for not using piecewise linear interpolation for the properties. Additionally, simulations were done in the steady state given similar studies have utilized this assumption (Ozden and Tari, 2010; Nedunchezhiyan Mukilarasan et al., 2022; Zhang et al., 2017).

Computational Domain, Mesh And Boundary Conditions

The shell side volume of the four models is modelled using an unstructured tetrahedral mesh in the ANSYS Meshing tool as shown in Figure 2 below. In order to select the ideal mesh which is a balance of both accuracy and computational memory four grids were generated. The element numbers modelled were ~1.2 million, ~2.2 million, ~3.3 million and ~5.0 million. Given that the discrepancy between the overall heat transfer coefficients for the grid of 3.3 and 5.0 million was less than 2% the grid of element size 3.3 million was selected. This aligns with the level of discrepancy which has been considered reasonable in similar studies, such as that done by Abbasi et al. (2020).



(a) (b)

Figure 62. (a) Unstructured tetrahedral mesh with boundary conditions applied (b) Zoomed in view of the mesh applied at the tube wall.

Analysis of the contours revealed significant amounts of turbulence in the shell side of the heat exchanger. Considering merit factors such as the ability of the model to capture flow detail while maintaining minimal computational time and offering simple closure, the standard k- ε turbulence model was selected (Aslam Bhutta et al., 2012; Ozden and Tari, 2010; Abbasi et al., 2020). Prism layers and a scalable wall function were employed to resolve the near wall flow in combination with the turbulence model (El Maakoul et al., 2016; Shahril et al. 2017) This was done to limit the $y^+ \approx 11.5$ such that the flow maintains its compatibility with the turbulence model. Control of this non-dimensional parameter which represents the distance from the wall to the first node is critical given its ability to define the near wall flow (Shahril et al., 2017). The fluid domain was defined using inlet and outlet nozzle boundary conditions. A standard mass flow and pressure boundary condition were used respectively (Ozden and Tari, 2010, Nedunchezhiyan Mukilarasan et al., 2022; Abbasi et al., 2020; Zhang et al., 2017). Additionally, a fixed thermal condition of 450 K was used for the tube walls while a zero-heat flux condition was assumed for the shell wall.

Non-linear least squares regression for correlation development

The formulations of the parameters in the modelled correlations are similar to those used in established design methods such as the Bell-Delaware, Kern's and flow-stream analysis methods (Jamil et al., 2020). Equations 1-3 below show the basis of calculating the *Re*, a similar analogy is used to evaluate both the *Pr* and *Nu* (Alperen et al., 2019).

$$D_e = \frac{4\left(\frac{P_t}{2} * 0.87 * P_t - 0.5 * \pi * \frac{d_o^2}{4}\right)}{\pi * \frac{d_o}{2}} \quad (1)$$

$$A_s = \frac{(P_t - d_o) * D_s * B}{P_t} \quad (2)$$

$$Re = \frac{\left(\frac{\dot{m}_s}{A_s}\right) * D_e}{\mu} \quad (3)$$

RESULTS AND DISCUSSION

Model verification and thermo-mechanical results

Simulations revealed that increasing mass increased both the overall heat transfer coefficient and the shell side pressure drop. The results align with the trend shown in similar studies, however, differences are observed. This could be attributed to variations in the modelling schemes used. This argument is supported by the agreement visible with the study done by Pal et al. (2016) given they used a similar methodology for modelling.

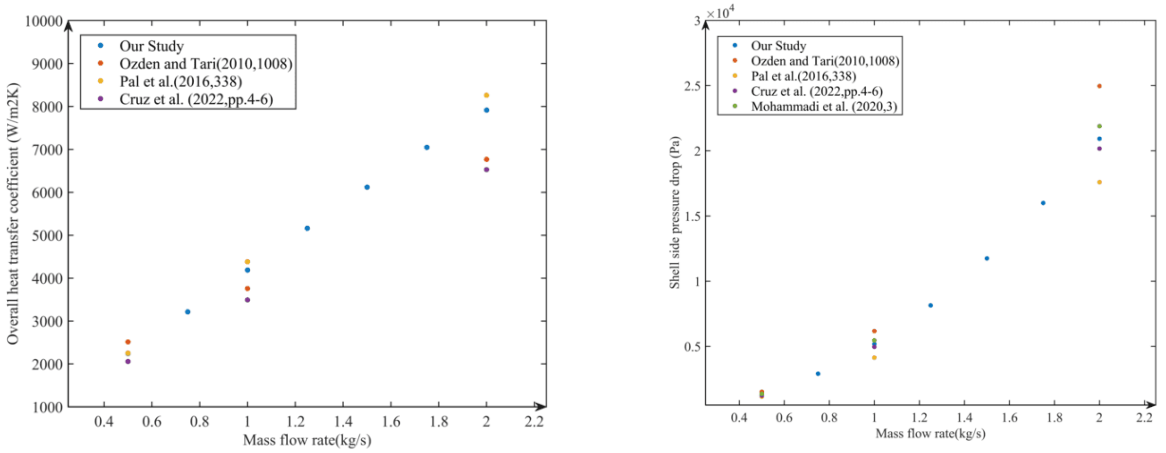


Figure 63: CFD Results for the six-baffle model (a) Overall heat transfer coefficient vs mass flow rate (b) Shell side pressure drop vs mass flow rate.

The overall heat transfer coefficient and shell side pressure were observed to be significantly affected by the increase in number of baffles and the mass flow rate as shown in Figure 4 (a) and (b) below. The overall heat transfer was observed to improve, which could be related to an improvement in the shell side flow phenomena and amount of area available for thermal transfer. Pressure drop was negatively affected, where this would lead to a higher pumping power to maintain the consistent flow of fluid. These phenomena are discussed in the context of the contours in the next section.

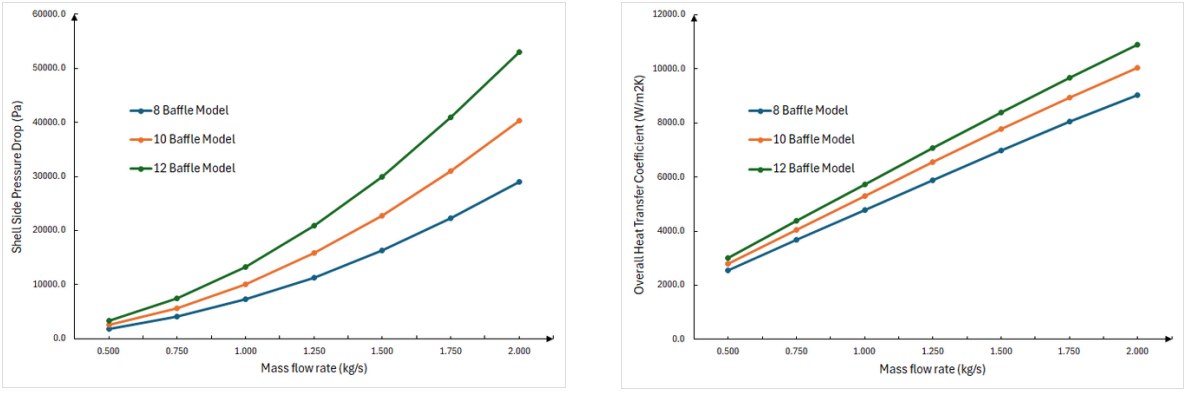


Figure 64. CFD based results with mass flow rate (a) Shell Side Pressure Drop (b) Overall Heat Transfer Coefficient

PATHLINES

The developed pathlines reveal the flow of the fluid in the shell side of the heat exchanger and its dependence on the number of baffles. Fluid follows a zig-zag pattern with regions of dead zones or areas of recirculation observed behind certain baffles. The significance of recirculation areas is high at lower number of baffles as evident in the case of 8 baffles (refer Figure 5 (a)). This is related to a decreased thermal performance given that a higher amount of tube area is bypassed, hence reducing the surface area available for heat transfer (Ozden and Tari, 2010; Alhulaifi, 2024). However, decreasing baffle spacing was observed to improve the flow phenomena with increased flow uniformity and reduced areas of flow recirculation (refer Figure 5 (b) and (c) below). This led to an increased thermal performance as discussed

earlier.

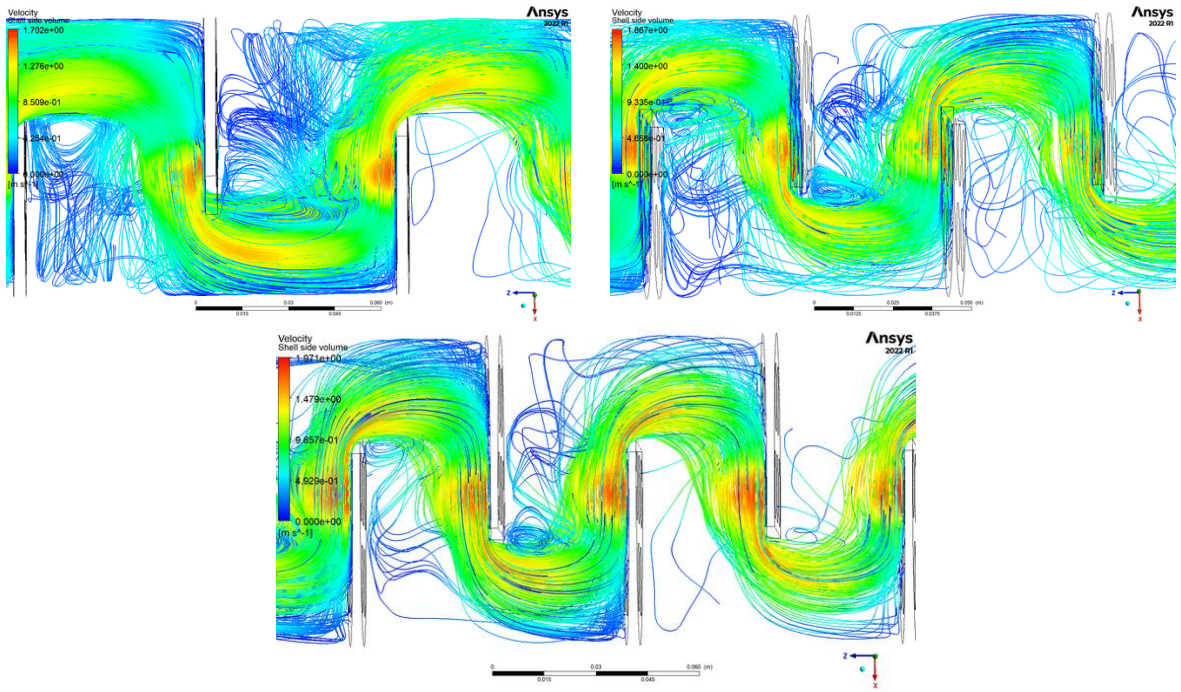
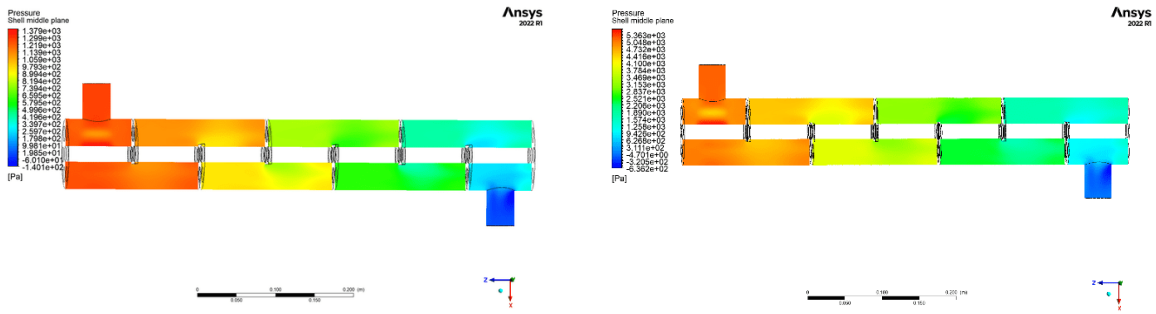


Figure 65. Velocity path lines at 1.0 kg/s (a) 8 baffles (b) 10 baffles and (c) 12 baffles.

Pressure drop along the length

The pressure contours reveal that the pressure drops from the inlet to the outlet along with the length of the heat exchanger. A high pressure is observed at the inlet of the model due to the impingement of the fluid entering at the inlet nozzle. Flow expansion and contraction due to the presence of baffles is a key factor which introduces a significant pressure drop given the barrier to fluid flow is increased as evident in the contours. Additionally, increasing the mass flow rate is related to an increased pressure drop with relative similarity in the contours (refer Figure 6(a) – (c)).



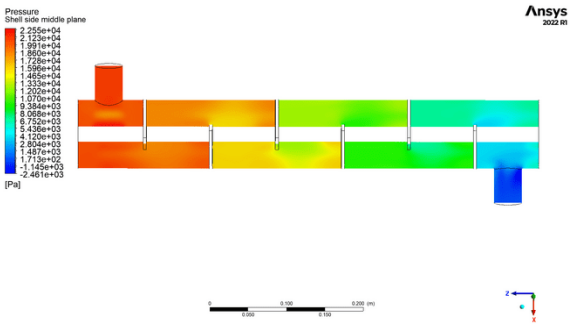


Figure 66. Pressure contours at the middle plane for six baffle model (a) 0.5 kg /s (b) 1.0 kg/s and (c) 2.0 kg/s.

Another key observation is the presence of back pressure at the outlet nozzle; this was a result of the improper flow development which was observed at the outlet plane. A similar observation was present in studies which were done by Pal et al. (2016) and Sauro & Lewis (2012).

CFD-Based Thermal Correlations for the four models

A summary of the correlations which were developed using NLSQR are shown in Figure 7 below.

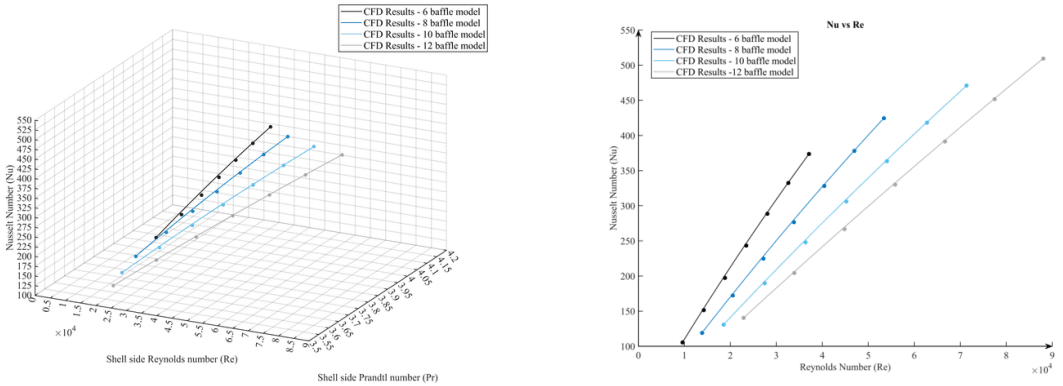


Figure 67 (a) *Nu* variation with *Pr* and *Re* for the four models (b) *Nu* variation with *Re* for the four models.

Nu is observed to increase significantly with the *Re*, which is representative of increasing turbulence. The *Re* is increased as a function of both the mass flow rate and the central baffle spacing. The observed variation is a result of improved flow phenomena as revealed earlier and in similar studies done by Ozden & Tari (2010) and Abbasi et al. (2020). Additionally, the minimal variation of the *Pr* in the observed results could be a result of the slight variation in the thermal properties of the working fluid. A summary of the correlations obtained using NLSQR are shown in Table 2 below.

Table 36. Summary of the correlations developed for the four models.

Baffles	***C***	***a***	***b***	***Re* Range**	***Pr* Range**
6	0.2312	0.9972	-2.1907	9600 < *Re* < 37100	3.68 < *Pr* < 4.12
8	0.139	1.0065	-2.1304	13900 < *Re* < 53400	3.80 < *Pr* < 3.96
10	0.0809	1.0103	-1.9586	18500 < *Re* < 71400	3.66 < *Pr* < 3.81
12	0.0319	0.9964	-1.2735	22900 < *Re* < 88100	3.54 < *Pr* < 3.70

The overall precision and reliability of the predicted correlations were evaluated considering the confidence interval (CI) widths of the predicted coefficients which is a standard methodology utilized in numerical model evaluation according to Sauro and Lewis (2012). A summary of the upper and lower

bounds of the coefficients with a confidence level of 95 % are shown in Table 3 below.

Table 37. Summary of the upper and lower bounds for predicted coefficients with a 95 % confidence level.

Nb	**6**			**8**		
Parameter	**Predicted value**	**Lower bound**	**Upper bound**	**Predicted value**	**Lower bound**	**Upper bound**
C	0.2312	0.014	0.4485	0.1390	0.0158	0.2623
a	0.9972	0.9758	1.0186	1.0065	0.9817	1.0312
b	-2.1907	-3.0117	-1.3696	-2.1304	-2.9624	-1.2914
Nb	**10**			**12**		
Parameter	**Predicted value**	**Lower bound**	**Upper bound**	**Predicted value**	**Lower bound**	**Upper bound**
C	0.0809	0.0263	0.1355	0.0319	0.0207	0.0431
a	1.013	0.9900	1.0307	0.9964	0.9852	1.0076
b	-1.9586	-2.6318	-1.2854	-1.2735	-1.639	-0.908

The observed results show that the CI decreases with the increasing number of baffles and this could again be related to the observed improvement in the flow phenomena. Overall stability of the curve fit was revealed by the significant improvement in the goodness of fit statistics. The sum of squares due to error (SSE) decreased from **0.2064** to **0.1279** for the six and twelve baffle models respectively. Additionally, the root means squared error (RMSE) decreased from **0.1703** to **0.0654**. The results of these statistical parameters indicate that the random component of error present in the model is minimal and that it can be used as a prediction tool with reasonable confidence given the operating ranges are applicable. This can be further verified by performing a practical simulation at the given boundary conditions.

CONCLUSIONS

In this study, a single pass TEMA E-type shell and tube heat exchanger was modelled in a steady state. The effects of turbulence on thermal performance were explored by varying the level of turbulence as a function of both the mass flow and the central baffle spacing for the case of six, eight, ten and twelve baffle models. The key finding was that decreasing baffle spacing improved the thermal performance at a cost of increased pressure drop. Results of CFD were then combined with NLSQR to develop four sets of correlations which relate the Nu to Pr and Re. Given the reasonable goodness of fit parameters and confidence intervals the study reveals that the model can be used as a prediction tool with a reasonable level of confidence highlighting the potential of using such workflows and integration for practical design and optimization purposes. Further research work in the area needs to be conducted to explore the potential by improving the modelling schemes such as the grid quality and convergence tolerances, consideration of the flow in tube sides, flow conditions at inlet and outlets and the application of advanced prediction tools such as ANN (Artificial Neural Networks) to utilize the capability of decision making improving the potential and scope of thermal design and optimization. Another key improvement to judge the accuracy of the developed model would be to conduct practical simulations at the given boundary conditions.

REFERENCES

1. Abbasi, H. R., Sharifi Sedeh, E., Pourrahmani, H., & Mohammadi, M. H. (2020). Shape optimization of segmental porous baffles for enhanced thermo-hydraulic performance of shell-and-tube heat exchanger. *Applied Thermal Engineering*, *180*, 115835. https://doi.org/10.1016/j.applthermaleng.2020.115835

2. Alhulaifi, A. S. (2024). Computational Fluid Dynamics Heat Transfer Analysis of Double Pipe

Heat Exchanger and Flow Characteristics Using Nanofluid TiO2 with Water. *Designs*, *8*(3), 39. https://doi.org/10.3390/designs8030039

3. Alperen, M. A., Kayabaşi, E., & Kurt, H. (2019). Detailed comparison of the methods used in the heat transfer coefficient and pressure loss calculation of shell side of shell and tube heat exchangers with the experimental results. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 1–20. https://doi.org/10.1080/15567036.2019.1672835

4. Aslam Bhutta, M. M., Hayat, N., Bashir, M. H., Khan, A. R., Ahmad, K. N., & Khan, S. (2012). CFD applications in various heat exchangers design: A review. *Applied Thermal Engineering*, *32*, 1–12. https://doi.org/10.1016/j.applthermaleng.2011.09.001

5. Bichkar, P., Dandgaval, O., Dalvi, P., Godase, R., & Dey, T. (2018). Study of Shell and Tube Heat Exchanger with the Effect of Types of Baffles. *Procedia Manufacturing*, *20*, 195–200. https://doi.org/10.1016/j.promfg.2018.02.028

6. El Maakoul, A., Laknizi, A., Saadeddine, S., El Metoui, M., Zaite, A., Meziane, M., & Ben Abdellah, A. (2016). Numerical comparison of shell-side performance for shell and tube heat exchangers with trefoil-hole, helical and segmental baffles. *Applied Thermal Engineering*, *109*, 175–185. https://doi.org/10.1016/j.applthermaleng.2016.08.067

7. Huang, Z., Hwang, Y., & Radermacher, R. (2017). Review of nature-inspired heat exchanger technology. *International Journal of Refrigeration*, *78*, 1–17. https://doi.org/10.1016/j.ijrefrig.2017.03.006

8. Jamil, M. A., Goraya, T. S., Shahzad, M. W., & Zubair, S. M. (2020). Exergoeconomic optimization of a shell-and-tube heat exchanger. *Energy Conversion and Management*, *226*, 113462. https://doi.org/10.1016/j.enconman.2020.113462

9. M. Necati Özışık. (1985). *Heat Transfer*. McGraw-Hill Science, Engineering & Mathematics.

10. Mangrulkar, C. K., Dhoble, A. S., Chamoli, S., Gupta, A., & Gawande, V. B. (2019). Recent advancement in heat transfer and fluid flow characteristics in cross flow heat exchangers. *Renewable and Sustainable Energy Reviews*, *113*, 109220. https://doi.org/10.1016/j.rser.2019.06.027

11. Marcinkowski, M., Dawid Taler, Katarzyna Węglarz, & Taler, J. (2024). Advancements in analyzing air-side heat transfer coefficient on the individual tube rows in finned heat exchangers: Comparative study of three CFD methods. *Energy*, *307*, 132754–132754. https://doi.org/10.1016/j.energy.2024.132754

12. Marcinkowski, M., Dawid Taler, Taler, J., & Katarzyna Węglarz. (2021). Thermal Calculations of Four-Row Plate-Fin and Tube Heat Exchanger Taking into Account Different Air-Side Correlations on Individual Rows of Tubes for Low Reynold Numbers. *Energies*, *14*(21), 6978–6978. https://doi.org/10.3390/en14216978

13. Nedunchezhiyan Mukilarasan, Karthikeyan, R., Ramalingam, S., Damodharan Dillikannan, Ravikumar, J., Sampath, S., & Gopal Kaliyaperumal. (2022). Influence of baffles in heat transfer fluid characteristics using CFD evaluation. *International Journal of Ambient Energy*, *43*(1), 7088–7100. https://doi.org/10.1080/01430750.2022.2063175

14. Ozden, E., & Tari, I. (2010). Shell side CFD analysis of a small shell-and-tube heat exchanger. *Energy Conversion and Management*, *51*(5), 1004–1014. https://doi.org/10.1016/j.enconman.2009.12.003

15. Pal, E., Kumar, I., Joshi, J. B., & Maheshwari, N. K. (2016). CFD simulations of shell-side flow in

a shell-and-tube type heat exchanger with and without baffles. *Chemical Engineering Science*, *143*, 314–340. https://doi.org/10.1016/j.ces.2016.01.011

16. Sauro, J., & Lewis, J. R. (2012). How Precise Are Our Estimates? Confidence Intervals. *Quantifying the User Experience*, 19–39. https://doi.org/10.1016/b978-0-12-384968-7.00003-5

17. Sekulić, D. P., & Shah, R. K. (2023). *Fundamentals of Heat Exchanger Design*. Wiley. https://doi.org/10.1002/9781119883296

18. Shahril, S. M., Quadir, G. A., Amin, N. A. M., & Badruddin, I. A. (2017). Thermo hydraulic performance analysis of a shell-and-double concentric tube heat exchanger using CFD. *International Journal of Heat and Mass Transfer*, *105*, 781–798. https://doi.org/10.1016/j.ijheatmasstransfer.2016.10.021

19. Zhang, X., Han, D., He, W., Yue, C., & Pu, W. (2017). Numerical simulation on a novel shell-and-tube heat exchanger with screw cinquefoil orifice baffles. *Advances in Mechanical Engineering*, *9*(8), 168781401771766-168781401771766. https://doi.org/10.1177/1687814017717665

Circular Economy Practices in Road Rehabilitation and Development in Sri Lanka

L. M. De Alwis[1], **K A N** Gunarathna[2], C.Kalugala[3]

Liverpool John Moores University, UK[1]
Sri Lanka Institute of information Technology, SL[2]
Skills College of Technolog, SL[3]

ABSTRACT

The integration of Circular Economy principles into road rehabilitation and development offers a sustainable alternative to traditional linear construction practices. In Sri Lanka, where road infrastructure is crucial to economic growth and connectivity, the Circular Economy offers opportunities to reduce reliance on virgin materials, minimise environmental impact, and enhance long-term cost efficiency. This study explored applicable Circular Economy strategies, assessed current implementation practices, identified key challenges, and proposed viable solutions to support Circular Economy adoption in the Sri Lankan Road sector. A mixed-methods approach was employed, combining a comprehensive literature review with data from semi-structured expert interviews and a questionnaire survey of construction professionals. The findings indicate that while awareness of the Circular Economy is growing, its practical application remains limited due to barriers such as the absence of standardised technical guidelines, insufficient government incentives, limited stakeholder knowledge, logistical challenges in material sourcing and storage, and reluctance to shift from traditional methods. Data also highlights substantial potential benefits, including cost savings, reduced construction waste, increased material efficiency, and environmental improvements. Participants emphasised the importance of pilot projects, training programs, and policy support in promoting Circular Economy practices. To overcome existing barriers, the study recommends the development of clear Circular Economy specifications, financial incentives, capacity-building initiatives, and the establishment of centralised recycling infrastructure. These strategic actions can facilitate the transition toward a more circular and sustainable approach in Sri Lanka's Road construction and maintenance sectors.

Keywords: *Circular Economy, Road Construction, Road Rehabilitation, Sustainable Infrastructure, Recycled Materials*

INTRODUCTION

Road infrastructure development is a critical component of the development of countries, and it serves as the foundation for economic growth, social connectivity, and enhances the transport logistics system of the country. (Kodithuwakku et al., 2023). The road is a legally permissible way for vehicles and other traffic to traverse. Roads include passageways, side drains, culverts and land needed for the upcoming widening.(Kodithuwakku et al., 2023). Road networks include highways, expressways, urban roads, rural roads, bridges, drainage systems, and pavements, all of which play a vital role in supporting increases in the country's gross domestic product. Each of these road networks is critical to develop and maintain. The highways are designed for long-distance and high-speed travel, facilitating efficient transportation between cities, industrial zones, and ports with enabling the transportation of goods and passengers. Rural roads are essential for connecting remote areas, agricultural zones, and villages to economic centres. Effective drainage and stormwater management systems are crucial for preventing floods and road deterioration. Proper drainage infrastructure includes culverts, stormwater drains, ensuring that excess water is diverted away from the road surface without disturbing road functions. The transport sector contributes 10% to 20% of the Gross Domestic Product (GDP), where road transportation represents between 3% to 5% of the GDP, excluding the inputs of transport equipment of fuel, and infrastructure.(Kodithuwakku et al., 2023).

Traditional road rehabilitation and development are heavily based use of virgin materials. Conventional road construction primarily uses natural soil, crushed rock, and cement concrete in the industry. Not only these, Bituminous materials are heavily used in road construction and the development process.(Kodithuwakku et al., 2023). Bituminous materials are used as surfacing material in all types of road construction. All these materials require significant amounts of energy for extraction, processing, and installation. These processes release substantial amounts of greenhouse gases, particularly carbon dioxide, into the environment and increasing global warming and contributing to climate change.

The **circular economy framework** offers a promising solution to **reduce embodied energy in road construction and rehabilitation. Circular economy** emphasises **resource efficiency and waste reduction** through strategies such as **material reuse, recycling, and optimised design**.

The global rise in construction activity, particularly in developing nations, has led to a critical shortage of high-quality natural aggregates and an increase in material costs, resulting in environmental and economic concerns (Mishra et al., 2023). Simultaneously, growing demands for more durable and resilient infrastructure have intensified pressure on traditional materials, which often fail to meet long-term performance expectations (Plati, 2019). While Circular Economy (CE) strategies offer sustainable alternatives, their adoption in road construction is hindered by high upfront costs, inadequate regulations, and logistical difficulties (Silva & Fernando, 2022). Moreover, public sector agencies face institutional limitations, such as technical gaps and fragmented policies, which restrict their ability to implement sustainable practices effectively (Connolly et al., 2023).

LITERATURE REVIEW

Construction of road infrastructure is one of the main aspects of a country's development and the foundation of economic development, social connectivity, and enhances the nation's transport logistics system.(Kodithuwakku et al., 2023). Conventional road building uses plenty of virgin resources, including aggregates, concrete, and asphalt, and is therefore resource-intensive. These materials require large-scale extraction, processing, and transportation, which translates to a high environmental footprint and high material cost.(Hammond & Jones, 2008). According to the Ellen MacArthur Foundation, CE is restorative by design and aims to keep products, components, and materials at their highest utility and value at all times, distinguishing between technical and biological cycles. (MacArthur, 2015). Benachio et al. and other authors defined CE for the Construction Industry as "the use of practices, in all stages of the life cycle of a building, to keep the materials as long as possible in a closed loop, to reduce the use of new natural resources in a construction project. (Benachio et al., 2020).

Application of CE in Road Construction

Use of Recycled and Alternative Materials

Circular Economy practices promote the substitution of traditional virgin materials with recycled and alternative materials in road construction. Recycled Concrete Aggregate (RCA), derived from demolished concrete, is recognised as a viable substitute for natural aggregates in sub-base and base layers due to its self-cementing properties and high stiffness. (Silva & Fernando, 2022)Crushed brick and glass, originating from construction and demolition waste, also offer utility in sub-base and asphalt applications. (Kodithuwakku et al., 2023)Reclaimed Asphalt Pavement (RAP) is a key strategy to reduce the demand for virgin bitumen, lowering the embodied carbon of road infrastructure. The integration of polymer modifiers like crumb rubber further improves binder performance, reduces rutting, and supports longer pavement lifespans.(Ashtiani et al., 2024). Fly ash and cement kiln dust are sustainable binders used to stabilise recycled aggregates and weak subgrades, offering economic and environmental advantages.(Silva & Fernando, 2022).

Energy-Efficient Road Construction Methods

Energy-efficient construction practices in CE involve reducing emissions across the road's lifecycle—from material extraction to end-of-life disposal. A major strategy includes the use of **blended cement**, where clinker is partially replaced with by-products like fly ash or slag, reducing CO_2 emissions during production. (Worrell et al., 2001). Furthermore, optimising material transport through route planning and double cycling of RAP and HMA (hot mix asphalt) deliveries significantly lowers fossil fuel consumption and construction-related emissions. (Kumari & Kulathunga, 2020 & Ashtiani et al., 2024)Recycling reinforced concrete waste (RCW) at the end of a road's lifecycle further conserves resources and supports circularity, providing environmental and cost benefits(Paranavithana et al., 2024).

Structural Performance of Recycled Materials

A major concern in CE adoption is whether recycled materials can meet structural demands. Studies have shown that pavements constructed using recycled materials can match or surpass the performance of conventional materials. For example, the inclusion of 20% RAP enhances rutting resistance and increases the stiffness of bituminous mixes, while 10% RAP improves resilient modulus from 5000 to over 6700 MPa(Silva & Fernando, 2022). RCA also improves subbase performance, with the potential to exceed 400 MPa in modulus when mixed with 25% crushed bricks. Additionally, incorporating 2% cement helps meet subbase standards under varied moisture conditions.(Silva & Fernando, 2022). Fly ash is another high-performing additive for stabilising expansive soils, achieving a 26% increase in California Bearing Ratio (CBR) at a 15% mix in field tests in Melbourne. HiMA (Highly Modified Asphalt) stands out in terms of performance efficiency it allows for **20% thinner pavement layers** without sacrificing structural integrity. In trials, HiMA outperformed conventional HMA in fatigue cracking and deformation under 20 million ESALs (Kluttz et al., 2017)This contributes to both material savings and lower emissions during production and transportation.

Economic Feasibility of CE in Road Development

The economic viability of CE principles in road development lies in their potential for long-term cost savings, despite initially higher investment. Sustainable pavement design emphasises longevity spanning 30 to 60 years, reducing the need for frequent maintenance, rehabilitation, and associated greenhouse gas emissions. Innovative solutions such as composite pavements and inverted asphalt layers offer high durability at lower lifecycle costs. (Plati, 2019). CE-aligned techniques, such as the use of Cement Stabilised Rammed Earth, provide low-cost construction alternatives, especially suitable for resource-constrained regions. (Silva & Fernando, 2022). Moreover, efficient material management can lower a substantial portion of costs attributed to materials, often exceeding 40% of total project costs, by minimising waste and leveraging recycled inputs. (Kodithuwakku et al., 2023).

Policy and Regulatory Support for CE Adoption

Effective policy and regulatory frameworks are crucial for advancing CE adoption in road construction. Governments can foster circularity by setting mandatory reuse and recycling thresholds for construction projects, thereby boosting demand for secondary materials and reducing reliance on virgin resources.(Ghaffar et al., 2020). Fiscal incentives and formal recognition for sustainable practices also play a pivotal role in attracting investment into CE innovations, especially those aimed at resource recovery from construction and demolition waste. Internationally, several countries have taken proactive steps. For instance, China has set a target to ensure 30% of all new buildings are prefabricated by 2026, as part of its broader CE strategy. It has also enacted policies emphasising the "reduce, reuse, recycle" principle.(Li et al., 2024). Moreover, initiatives like the CERCOM (Circular Economy in Road Construction and Maintenance) project in Europe aim to institutionalise CE within national procurement frameworks. CERCOM promotes a Risk-Based Analysis Framework that evaluates technical, economic, and environmental risks associated with CE transitions, encouraging more robust and structured adoption across road infrastructure projects (Connolly et al., 2023).

CHALLENGES TO CE IMPLEMENTATION

Several authors have identified various types of obstacles and categorised them into six wide groups, namely social, cultural, awareness, technical, economic, and implementation. As helpful as this categorisation is, the present study shall adopt an exploratory approach, taking into consideration all possible obstacles when they emerge without restricting the inquiry to pre-defined categories.(AlJaber et al., 2023).

The lack of standards for recycling and re-use is an impeding factor to the circular construction concept. Other studies also identify this lack of standards as a problem. (Benachio et al., 2020). Furthermore, the lack of standardisation often leads to the need for multiple costly and time-consuming certifications for alternative materials.(World Green Building Council, 2023). Table 1 summarises the challenges to CE implementation.

Table 1: Challenges to CE Implementation - Summary

Challenges	A	B	C	D	E
High Initial/Upfront Costs	✓	✓	✓	✓	
Lack of CE Infrastructure	✓	✓			✓
Limited Knowledge and Expertise	✓	✓		✓	✓
Lack of Government Support/Incentives	✓	✓		✓	✓
Quality of Recycled Materials	✓	✓		✓	
Lack of rules and policies	✓	✓	✓		✓

A.(AlJaber et al., 2023) **B**. (Snigdhya et al., 2024), **C**. (World Green Building Council, 2023), **D**. (Agyemang et al., 2019), **E**.(Weerakoon & Thayaparan, 2023)

RESEARCH METHODOLOGY

With the completion of the comprehensive literature review, semi-structured expert interviews were conducted to identify the challenges and barriers currently affecting the adoption of Circular Economy (CE) principles in Sri Lanka's Road construction sector. Semi-structured interviews offer a balance of flexibility and depth, making them more adaptable than structured interviews and more focused than unstructured ones. The interviews aimed to explore current CE practices, highlight implementation gaps, and capture expert perspectives on viable strategies for future integration. Subsequently, a questionnaire survey was developed based on findings from both the literature review and expert interviews. The survey was distributed among a broader group of stakeholders to validate key challenges, understand stakeholder perceptions, and evaluate the attractiveness and feasibility of CE strategies.

DATA COLLECTION AND ANALYSIS

Qualitative Analysis

Respondent's Profile

Interviews were conducted among Sri Lankan Quantity Surveyors (QSs) and engineers who are in consultancy and contractor organisations. Seven (7) semi-structured interviews were conducted for data gathering. The respondents' details were gathered through the first section of the semi-structured interview guideline. The summarised composition of respondents' profiles is presented in Table 2.

Table 2: The Composition of the Respondents

Interviewee	Designation	Experience (No. of Years)
I-01	Chartered Quantity Surveyor	8
I-02	Chartered Civil Engineer	23
I-03	Planning Engineer (civil engineer)	9
I-04	Project Manager (civil engineer)	20
I-05	Chartered Quantity Surveyor	19
I-06	Chartered Civil Engineer	24
I-07	Road Contractor	6

The semi-structured interviews conducted with professionals from Sri Lanka's Road construction sector revealed several critical insights into the existing status, challenges, and future strategies for adopting Circular Economy (CE) practices. Interviewees were drawn from government, consultancy, and contractor backgrounds, providing a comprehensive overview of the sector's position on CE.

Adoption of Circular Economy (CE) Practices

Expert interviews revealed that although all the participants were familiar with CE practices, their application in Sri Lankan Road construction projects remains limited. Respondents I01, I02, and I03 acknowledged awareness of CE principles but noted that actual implementation is seldom and often informal. For instance, I05 highlighted routine reuse of construction elements such as signage and milestones, while I04 provided examples where materials from demolished roads were reused as base layers in new projects. I01 also mentioned using crushed concrete in granular sub-base layers. These cases reflect a fragmented but emerging pattern of CE adoption.

Key Challenges in CE Implementation

Multiple challenges were identified that hinder the broader adoption of CE in road construction. The lack of technical standards and regulatory guidelines was cited by I01, I03, and I04 as the primary barrier. Engineers are reluctant to use recycled materials without standardised performance benchmarks. Moreover, I02 emphasised that both public and institutional awareness are limited, with misconceptions about the safety and durability of recycled materials. Public scepticism further hinders implementation, particularly in community-sensitive projects. Material availability and quality control were also highlighted. Interviewees I05 and I06 noted that recycled materials vary significantly in quality, and storage of materials such as RCA is logistically difficult without dedicated facilities. In addition, I01 and I05 pointed out the absence of updated testing labs, making verification of recycled material properties challenging. Resistance from stakeholders, including contractors and employers, also emerged as a theme. As I04 and I05 stated, conservative design perspectives and fear of performance issues discourage the use of alternative construction methods.

Current Use of Recycled and Alternative Materials

All experts confirmed using Recycled Concrete Aggregate (RCA), particularly in rural road sub-base applications. Reclaimed Asphalt Pavement (RAP) was also commonly used, especially for binder courses, as noted by I02, I04, and I06. Fly ash and cement kiln dust were used as stabilisers or filler materials in road layers, and I02 reported using quarry dust to reinforce subbase layers. Turfing reuse for drainage embankments was also mentioned. Despite these efforts, the industry faces technical and logistical constraints. Experts cited a lack of design knowledge and material quality assurance, often due to

inconsistent sources of recycled material and the absence of a central recycling supply chain. Approval bottlenecks, noted by I01 and I05, are frequent due to the lack of familiarity of regulatory bodies with non-conventional materials.

Benefits of using CE practice in Road Construction

All respondents affirmed that CE practices reduce material costs and environmental impact. The reuse of materials leads to savings in raw material procurement and reduces construction waste. I02 and I06 noted that CE improves project efficiency by saving time on material sourcing. I01 and I04 highlighted CE's contribution to long-term sustainability and raw material conservation. I05 emphasised an opportunity to develop local markets and circular supply chains, while I06 pointed to the growth of waste management industries as a secondary economic benefit.

Long-Term Cost Savings

All the expert interviewees strongly agreed that CE-based road construction works save considerable long-term costs, despite their comparatively greater initial investment. The consensus was based on some key economic benefits inherent in CE practices. Firstly, the reduction in raw material quarrying was mentioned as a major cost-cutting factor because CE emphasises the recycling and reuse of existing materials such as RAP and RCA, hence avoiding the expense of procuring copious amounts of virgin materials. This not only reduces procurement and transportation costs but also lowers environmental compliance expenses associated with quarrying and extraction. Secondly, it was noted by the respondents that landfill and disposal costs are significantly minimised through material recovery and reuse practices, reducing the economic and environmental cost of waste management. Also, recycling and processing of the recovered materials were described as considerably less compared to the energy and resource-intensive extraction and preparation of new materials.

Regulatory and Policy Needs

Government and institutional policy play a vital role in the effective implementation of CE practices in Sri Lanka's Road and infrastructure sector. I01, I02, I04, I05 and I07 indicate that there is an issue of the absence of lack of standard specification for supporting the use of recycled or alternative materials. Most existing specifications are made for use with virgin materials. Therefore, it is important to update the specifications that are relevant to use with recycling materials, as pointed out by the respondents. Also, I01, I03 and I04 noted that the introduction of Circular Procurement Policies is also crucial. They pointed out that the inclusion of CE-based Policies in Public procurement guidelines, thereby encouraging bidders to incorporate sustainable solutions. In the building construction sector, there is a standardised certification process to assess the building's environmental impact and sustainability, such as LEED and BREEM. I03, I06, and I07 mentioned that the road sector also needs such a standardised certification process to implement CE practice. Moreover, interviewees I01 and I05 emphasised that the Construction Industry Development Authority and Road Development Authority, as the primary regulatory bodies in Sri Lanka's Road Sector, should take the lead in initiating and implementing necessary policy changes to support the adoption of CE practices.

Financial Strategies to Support CE Implementation in Road Construction

The successful adoption of CE practices in road infrastructure projects depends not only on policy changes and technical enhancement but also on robust financial support strategies. In construction industry settings, contractors typically evaluate potential advantages and tangible benefits before adopting new construction methods. The decision to implement innovative approaches such as CE practices is often influenced by considerations of cost-efficiency. All the interviewees noted that it is important to provide tax benefits for

the contractors and clients who implement CE strategies in road projects. The other widely mentioned strategy was the inclusion of CE in bid evaluation and procurement systems. Interviewees I01, I02, I05, and I06 suggested the marking criteria within tender evaluation to favour CE-compliant contractors. I02 proposed applying this, especially in Design & Build project types, which often offer flexibility in material and process selection. I05 emphasised that, as a financial strategy to implement the CE principle in road construction in BOQs, contractors can use additional preliminary items for the storage of recycling materials and recycling plants.

Quantitative Analysis

A Questionnaire Survey was distributed among forty-two (42) professionals in the construction industry, and thirty (30) responses were collected, which accounts for a response rate of 71.4%. These respondents possess hands-on experience in road projects, including consultants, project managers, quantity surveyors, civil engineers, site engineers and academics. Figure 1 illustrates a breakdown of professional categories of respondents.

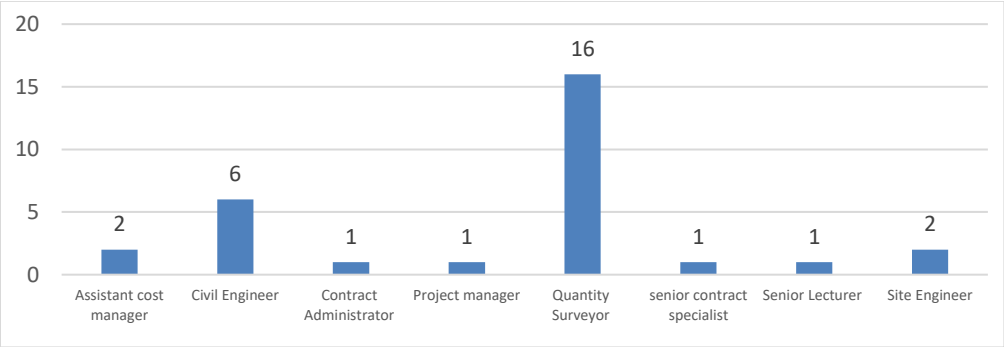


Figure 1: Breakdown of Professional categories of Respondents
(Sources: Quantitative data analysis)

Circular Economy practices implemented in projects

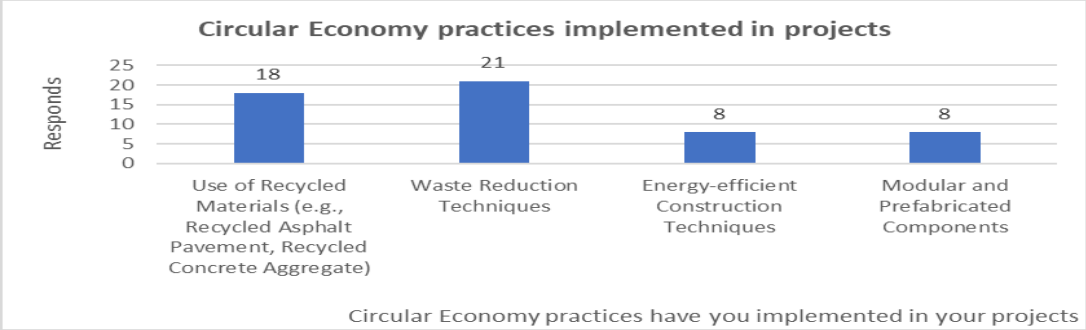


Figure 2: CE practices Implemented in Projects
(Sources: Quantitative data analysis)

The
survey results, as depicted in Figure 2, reveal the most applied CE practice among the participants, being methods in waste reduction, by 21 participants. This is followed closely by 18 participants using recycled materials such as RAP and RCA. Fewer participants (8 participants) reported using energy-efficient building practices and modular or prefabricated components within their activities, as another CE practice implemented.

Barriers faced when adopting CE practices in road development

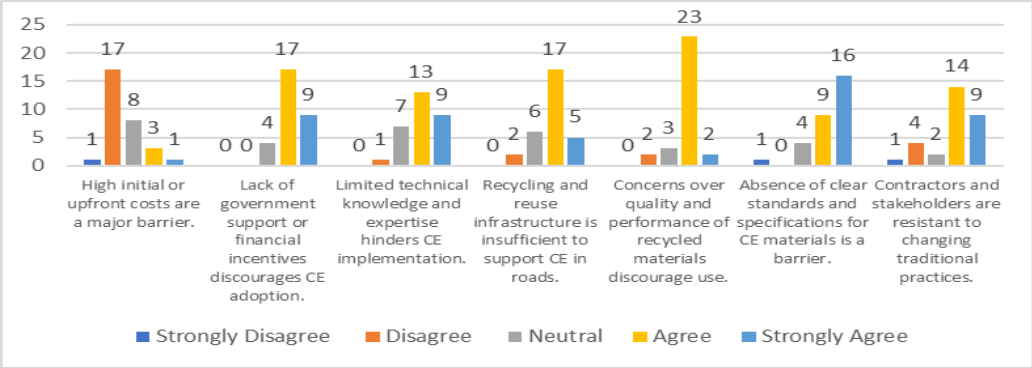


Figure 3: Barriers faced when adopting Circular Economy practices in road development. (Sources: Quantitative findings)

After conducting a thorough Expert interview, these seven (7) major factors, as illustrated in Figure 3, were identified which could be challenges and barriers faced when adopting CE practices in road development. All the participants were asked to rank these factors on a Likert scale from 1 to 5 (1 being strongly disagree and 5 strongly agree) based on the magnitude of the challenges. Sixteen (16) respondents strongly agree, while nine (9) participants agree with the absence of clear standards and specifications for CE materials. Also, twenty-three (23) agree, while two (2) strongly agree, that the concerns over quality and performance of recycled materials are considered the most significant hurdles. This suggests a strong need for robust regulatory frameworks and proven performance data to build confidence in CE materials. Notably, a lack of support and financial support (17 agree) and insufficient recycling and reuse systems (17 agree, 5 strongly agree) make it difficult for CE to happen.
Furthermore, the collected data were analysed and ranked by the weighted mean method using Eq. 01.

Equation 1 Weighted Factor

$$W = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i}$$

Where,

W	=	weighted average
N	=	number of terms to be averaged
$Ꞷ_i$	=	weights applied to x values
X_i	=	data values to be averaged

Table 3: Ranking of the Challenges

Factor	Weighted Mean	Rank
Absence of clear standards and specifications for CE materials is a barrier.	4.300	**1**
Lack of government support or financial incentives discourages CE adoption.	4.167	**2**
Limited technical knowledge and expertise hinder CE implementation.	4.000	**3**
Contractors and stakeholders are resistant to changing traditional practices.	3.867	**4**
Recycling and reuse infrastructure is insufficient to support CE in roads.	3.833	**5**
Concerns over the quality and performance of recycled materials discourage use.	3.833	**5**
High initial or upfront costs are a major barrier.	2.533	**7**

(Sources: Quantitative data analysis)

Table 3 shows the challenges that impact while adoption of CE practices in road construction. These challenges are ranked according to the weighted mean. Accordingly, the absence of clear standards and specifications for CE materials is taken as the major barrier.

Cost effectiveness of using CE-based Materials

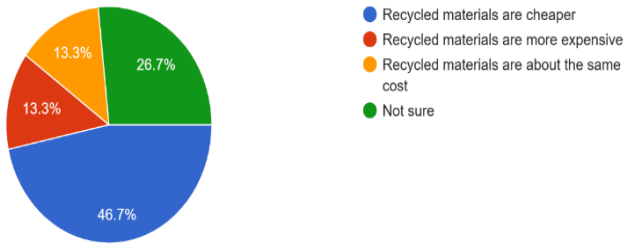


Figure 4: Cost comparison between the use of Recycled and Virgin Materials
(Source: Quantitative data analysis)

Figure 4 shows the cost comparison between the use of recycled and virgin materials in road construction. In accordance with the data, 46.7% of respondents find recycled materials to be cheaper due to raw material extraction, transportation, and landfill disposal savings. On the contrary, 13.3% find recycled materials to be higher in cost, possibly due to extra processing, specialised equipment, or conforming to performance standards. Another 13.3% think the costs are about the same, indicating that in some cases, savings from material reuse are balanced by preparation and performance costs. Notably, 26.7% of respondents are unsure, reflecting a lack of consistent cost data or limited awareness in the industry.

Attractiveness of the financial incentives or policies to encourage CE practices in road construction

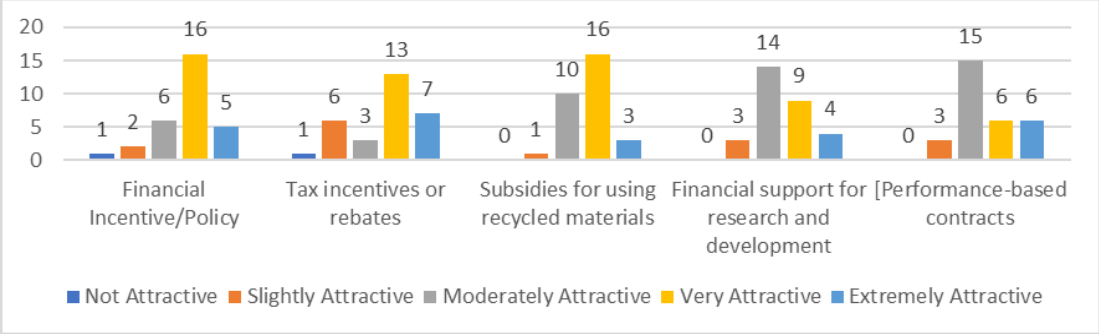


Figure 5: Attractiveness of the Financial Incentives or Policies
(Sources: Quantitative data analysis)

According to the questionnaire survey, financial mechanisms are essential for the adoption of the CE on Sri Lanka's roads. When respondents were asked about financial incentives and policies, the vast majority thought they were extremely attractive. Shortly afterwards, tax incentives and rebates or discounts for using recycled materials were offered, demonstrating that the industry places great significance on directly reducing costs tied to CE. Financial help for research and development, as well as results-based contracts, were appreciated but were placed lower in importance than instant financial incentives. It seems that the industry recognises innovation and quality performance, but currently it is mainly interested in ways to achieve cost savings to help CE take hold. This is clearly illustrated in Figure 5.

Table 4: Ranking of the attractiveness of giving financial incentives

Factor	Weighted mean	Rank
Financial Incentive/Policy	3.733	**1**
Subsidies for using recycled materials	3.700	**2**
Tax incentives or rebates	3.633	**3**
Performance-based contracts	3.500	**4**
Financial support for research and development	3.467	**5**

(Sources: Quantitative data analysis)

Table 4 above shows the ranking of attractiveness of financial incentives or policy changes to encourage the CE principle in road construction. These enables are ranked according to the weighted mean. Accordingly, the most attractive factor is financial incentives, and secondly, subsidies for using recycled materials.

Importance of the policy and regulation strategies

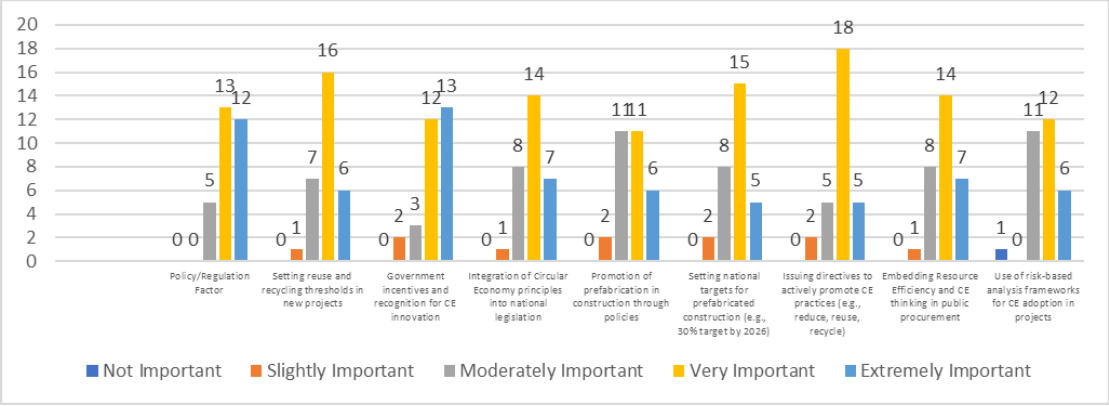


Figure 6: Importance of Policy and Regulation Strategies
(Sources: Quantitative data analysis)

The stakeholder survey revealed a strong consensus on the necessity of policy and regulatory support to promote the adoption of Circular Economy (CE) practices in Sri Lanka's Road construction sector. As illustrated in Figure 6, respondents overwhelmingly rated policy interventions such as clear standards for recycling and reuse, government-backed financial incentives, and public procurement alignments "Very Important" or "Extremely Important," underscoring their critical role in driving CE integration. Notably, the establishment of recycling and reuse standards was identified as a top priority, reflecting the need for consistent technical specifications to facilitate the use of alternative materials. Similarly, financial mechanisms such as grants and tax rebates were seen as key to overcoming initial cost barriers and accelerating industry uptake. Furthermore, integrating CE principles at the planning stage was strongly endorsed, reinforcing the importance of embedding sustainability early in project lifecycles. While national targets for CE and directives aimed at waste reduction received high support, strategies such as prefabrication promotion and risk-based CE evaluation frameworks were seen as moderately important—valuable but not as central as policy and financial levers. This indicates that while technical innovations are helpful, systemic policy reform and robust institutional frameworks are viewed as the most effective catalysts for advancing CE in the road construction industry.

Table 5: Ranking of the importance of the policy and regulation strategies

Factor	Weighted mean	Rank
Policy/Regulation Factor	4.233	**1**
Government incentives and recognition for CE innovation	4.200	**2**
Setting reuse and recycling thresholds in new projects	3.900	**3**
Integration of Circular Economy principles into national legislation	3.900	**3**
Embedding Resource Efficiency and CE thinking in public procurement	3.900	**3**
Issuing directives to actively promote CE practices (e.g., reduce, reuse, recycle)	3.867	**6**
Setting national targets for prefabricated construction (e.g., 30% target by 2026)	3.767	**7**
Use of risk-based analysis frameworks for CE adoption in projects	3.733	**8**
Promotion of prefabrication in construction through policies	3.700	**9**

(Sources: Quantitative data analysis)

The ranking results depicted in Table 5 above highlight that policy and regulatory factors are considered the most crucial drivers for adopting CE practices in road construction, with a weighted mean of 4.233. Government incentives and recognition closely follow (4.200), emphasising the importance of financial and institutional support to encourage innovation.

DISCUSSION

The data gathered from the literature review, semi-structured interviews, and questionnaire survey reveal that while the concept of Circular Economy (CE) is somewhat recognised in Sri Lanka's Road construction sector, its practical implementation remains limited. Both academic literature and expert insights confirm awareness of various recycled and alternative materials such as Recycled Concrete Aggregate (RCA), Reclaimed Asphalt Pavement (RAP), fly ash, cement kiln dust, crushed brick, and glass. Additionally, industry professionals highlighted the use of quarry dust, turfing for embankments, and the reuse of modular prefabricated components like curbs and signboards—demonstrating a foundational but underdeveloped application of CE practices in the sector.

CHALLENGES AND STRATEGIC RECOMMENDATIONS

When integrating findings from both literature and primary data, the study proposes the following strategies to enhance CE adoption, as given in Table 7 below.

Table 7: Strategic Recommendations

Challenges	Proposed Strategies	Expected Outcome
Lack of standard specification	Updating the specification related to the road	Engineers and contractors will have clear guidelines for using recycled materials, increasing confidence and reducing reluctance.
	Incorporate alternative specs and particular specs in the BOQ and tenders.	
Lack of Government Support/Incentives	Providing tax concessions for CE approaches	The industry will be financially motivated to adopt CE, overcoming the barrier of high
	Adopt the CE approach in the bid	

	evaluation criteria for D&B type.	upfront costs and encouraging innovative solutions.
Limited Knowledge and Expertise	Training programs for contractors and suppliers on CE methods	Industry professionals will gain the necessary skills and confidence, while pilot projects will serve as practical examples and learning models.
	Make CE-based pilot projects.	
Reluctant to adopt alternative construction methods	Certification process	Stakeholders will have a formal process to validate sustainable practices, and financial incentives will encourage them to move away from traditional methods.
	Training programs for contractors and suppliers on CE methods	
	Giving tax concessions and rewards	
Availability of recycled material	Make market competition for recycled materials.	A competitive market will ensure a steady and reliable supply of recycled materials, addressing the current logistical challenges.
Storage problem	Use additional preliminary items for storage and recycling process	This will provide dedicated budget and resources for proper storage and handling, improving material quality control and logistical efficiency.

In order to effectively promote a circular economy (CE) in Sri Lanka's road construction sector, this research proposes a number of actionable recommendations for policymakers and practitioners. The study emphasises the urgent need for a regulatory framework, beginning with the development of clear, standardised specifications for the use of recycled and alternative materials. The findings show that without these technical standards, engineers and stakeholders are reluctant to adopt CE practices due to concerns about the quality and performance of recycled materials. Financial incentives are also critical, with a strong preference among industry professionals for **tax benefits, subsidies, and the inclusion of CE principles in bid evaluation criteria** to offset initial costs and reward sustainable approaches. Furthermore, the study suggests that the **Road Development Authority (RDA) and Construction Industry Development Authority (CIDA)** should lead the change by establishing circular procurement policies, supporting pilot projects, and promoting training programs to address the existing knowledge and expertise gaps.

The discussion reveals a substantial gap between theoretical awareness and practical application of Circular Economy practices in road construction in Sri Lanka. While literature from the global and local environment confirms the technical and financial viability of CE, ground-level implementation is hampered by regulation, institutions, and logistics. However, industry professionals demonstrate a willingness to adopt CE, provided that supportive strategies and incentives are put in place.

CONCLUSION

This study concludes that while the awareness of Circular Economy practices in Sri Lanka's Road construction sector is increasing, practical implementation remains limited due to key challenges such as unclear policies, lack of technical standards, limited stakeholder knowledge, and insufficient infrastructure for recycled materials. Despite these barriers, the research highlights significant potential for CE to enhance environmental sustainability, reduce construction costs, and improve long-term infrastructure performance. The findings suggest that with stronger government support, targeted policy reforms, industry training, and

the development of centralised recycling systems, CE practices can be effectively integrated into road development projects. This transition will require collaborative action across public and private sectors to create a more resource-efficient and sustainable construction industry.

ACKNOWLEDGEMENTS

The researcher would like to acknowledge Liverpool John Moores University (LJMU) and Sri Lanka Institute of Information Technology (SLIIT) for the facilitation of this study by providing university resources throughout this research.

REFFERENCE

[1]Agyemang, M., Kusi-Sarpong, S., Khan, S. A., Mani, V., Rehman, S. T., & Kusi-Sarpong, H. (2019). Drivers and barriers to circular economy implementation: An explorative study in Pakistan's automobile industry. *Management Decision*, *57*(4), 971–994. https://doi.org/10.1108/MD-11-2018-1178

[2]AlJaber, A., Martinez-Vazquez, P., & Baniotopoulos, C. (2023). Barriers and Enablers to the Adoption of Circular Economy Concept in the Building Sector: A Systematic Literature Review. *Buildings*, *13*(11). https://doi.org/10.3390/buildings13112778

[3]Ashtiani, M., Lewis, M., Waldman, B., & Simonen, K. (2024). *Strategies to Reduce Embodied Carbon in Roadway Infrastructure*.

[4]Benachio, G., Freitas, M. do C., & Tavares, S. (2020). Circular economy in the construction industry: A systematic literature review. *Journal of Cleaner Production*, *260*, 121046. https://doi.org/10.1016/j.jclepro.2020.121046

[5]Connolly, L., Sheils, E., Lamb, M., O'Connor, A., O'Brien, E., Viner, H., Bailey, H., Ramdas, V., & Varveri, K. (2023). Circular economy in road construction and maintenance. *Transportation Research Procedia*, *72*, 3458–3465. https://doi.org/10.1016/j.trpro.2023.11.770

[6]Ghaffar, S. H., Burman, M., & Braimah, N. (2020). Pathways to circular construction: An integrated management of construction and demolition waste for resource recovery. *Journal of Cleaner Production*, *244*, 118710. https://doi.org/10.1016/j.jclepro.2019.118710

[7]Hammond, G. P., & Jones, C. I. (2008). Embodied energy and carbon in construction materials. *Proceedings of Institution of Civil Engineers: Energy*, *161*(2), 87–98. https://doi.org/10.1680/ener.2008.161.2.87

[8]Kluttz, R. Q., Dongré, R., Powell, R. B., Willis, J. R., & Timm, D. H. (2017). *Long-term performance of a highly modified asphalt pavement and application to perpetual pavement design*. *June*. https://doi.org/10.14311/ee.2016.212

[9]Kodithuwakku, H. D. D. P., Ramachandra, T., & Kajavathani, P. (2023). ALTERNATIVE MATERIALS FOR SUSTAINABLE ROAD CONSTRUCTION IN SRI LANKA. *World Construction Symposium*, *1*, 110–122. https://doi.org/10.31705/WCS.2023.10

[10]Kumari, T., & Kulathunga, U. (2020). *To cite this article : Thakshila Kumari, Udayangani Kulathunga, Thathsarani Hewavitharana & Nandun Madusanka (2020): Embodied carbon reduction strategies for high-rise buildings in Sri Lanka, International Journal of Construction Management, DOI : E*. https://doi.org/10.1080/15623599.2020.1814939

[11]Li, C. Z., Tam, V. W., Lai, X., Zhou, Y., & Guo, S. (2024). Carbon footprint accounting of prefabricated buildings: A circular economy perspective. *Building and Environment*, *258*. https://doi.org/10.1016/j.buildenv.2024.111602

[12]MacArthur, F. E. (2015). Towards a circular economy: Business rationale for an accelerated transition. *Ellen MacArthur Foundation (EMF)*, 20.

[13]Mishra, P., Rajak, F., & Kumar, R. (2023). *Outlook of Circular Economy to High-rise Buildings Carbon Dioxide Outlook of Circular Economy to High-rise Buildings Carbon Dioxide*. *June 2024*.

[14]Paranavithana, G. N., Samarakoon, S. M. K., & Chathurika, R. H. J. (2024). Analysis of road construction waste generation in developing economies: A focus on surfacing and sub-base construction projects in Sri Lanka. *Journal of Cleaner Production*, *457*. https://doi.org/10.1016/j.jclepro.2024.142465

[15]Plati, C. (2019). Sustainability factors in pavement materials, design, and preservation strategies: A literature review. *Construction and Building Materials*, *211*, 539–555. https://doi.org/10.1016/j.conbuildmat.2019.03.242

[`6]Silva, S. De, & Fernando, S. (2022). *Icsbe 2022*.

[17]Snigdhya, N. S., Uddin, S., & Dipu, A. (2024). *Integrating Climate Change Adaptation into Flood Risk*

Management : Global Perspectives. *XII*(August), 1003–1012. https://doi.org/10.51244/IJRSI

[18]Weerakoon, P., & Thayaparan, M. (2023). ANALYSIS OF BARRIERS TOWARDS IMPLEMENTING CIRCULAR ECONOMY IN THE SRI LANKAN BUILT ENVIRONMENT. *World Construction Symposium*, *1*, 162–173. https://doi.org/10.31705/WCS.2023.14

[`19]World Green Building Council. (2023). *The Circular Built Environment Playbook*. 1–97. www.worldgbc.org

Worrell, E., Price, L., Martin, N., Hendriks, C., & Meida, L. O. (2001). Carbon dioxide emissions from the global cement industry. *Annual Review of Energy and the Environment*, *26*(November), 303–329. https://doi.org/10.1146/annurev.energy.26.1.303

Fog Computing Power For Telepresence Suit As User Terminal For Metaverse

Vladislav Mogilatov, Valeria Kaijalainen, Artem Volkov, Ammar Muthanna, Andrey Koucheryavy.

The Bonch-Bruevich St Petersburg State University of Telecommunications

ABSTRACT

Fog computing represents an emerging paradigm in information and communication technologies that is becoming pivotal for meeting the growing performance requirements of modern computing systems, particularly in the context of the Internet of Things (IoT) and real-time services. One of the most promising 2030-network service classes enabled by fog is telepresence. This paper investigates both the theoretical and practical aspects of deploying fog-computing solutions in telepresence networks to minimize latency and enhance overall system performance. The study combines analytical modelling with discrete-event simulations carried out in Any Logic.

INTRODUCTION

The accelerating pace with which new technologies are conceived and deployed in the telecommunications domain has created an imperative for continuous adaptation. While fifth-generation (5G) networks are still maturing—and many layers, from standardisation through production, remain in flux—the rapid advancement of enabling technologies has already brought the notion of 2030 communication networks, commonly referred to as 6G, to the research forefront [1], [2], [3], [4]. This transition is not merely an incremental upgrade; it calls for the creation of genuinely new service categories that will define the coming decade of connectivity.

A. Envisioned Capabilities of 6G Networks

Although detailed specifications and standards for 6G are yet to be ratified, the community broadly converges on several key development vectors:

- Ultra-high data rates: 100 Gbit/s – 1 Tbit/s
- Extreme low latency: ≈ 0.1 ms
- Massive device density: ≈ 100 devices per m³
- Extended spectrum utilization: 1 – 10 THz
- Native integration of artificial intelligence for traffic prediction, resource allocation, and network self-optimization
- Support for novel service verticals such as holographic telepresence, tactile Internet, and immersive extended reality

One paradigm that can satisfy these unprecedented requirements is fog computing. Conceptually a further evolution of cloud computing, fog shifts processing and storage resources closer to the network edge, thereby curbing the latency incurred when data must traverse to centralized clouds and back.

B. Fog Computing for IoT and Real-Time Services

Fog computing enables data processing on edge devices—computers, mobile handsets, sensors, or smart nodes—rather than exclusively in distant data centers. By handling computations at or near the “last mile,” fog mitigates cloud-roundtrip delays and supports instantaneous interactions. It

also allows a substantial share of operations to be executed locally, reducing backhaul congestion. Adoption is expected to accelerate owing to three converging factors:

3. Rising computational power of terminal equipment
4. Mature cloud-storage and processing ecosystems that complement, rather than replace, edge resources

Practical solutions to the scalability and latency bottlenecks inherent in large-scale IoT deployments

C. Reference Architecture Considerations

In 2015, the OpenFog Consortium introduced a reference architecture specification that crystallizes the foundational traits of fog computing. While comprehensive, the OpenFog framework does not explicitly target data-storage locality; it is primarily designed to execute computations near clients to minimize latency, transport costs, and other network-bound constraints while meeting bandwidth requirements. Control elements—including configuration, access management, and network monitoring—are likewise deployed near the edge rather than orchestrated from a central gateway. The architecture further supports local analytics, with results securely synchronized to the cloud for deeper processing or long-term use [5], [6].

OpenFog's architecture emphasizes horizontal scalability and system interoperability, while maintaining high standards of security, elasticity, and flexibility (see Fig. 1.1).

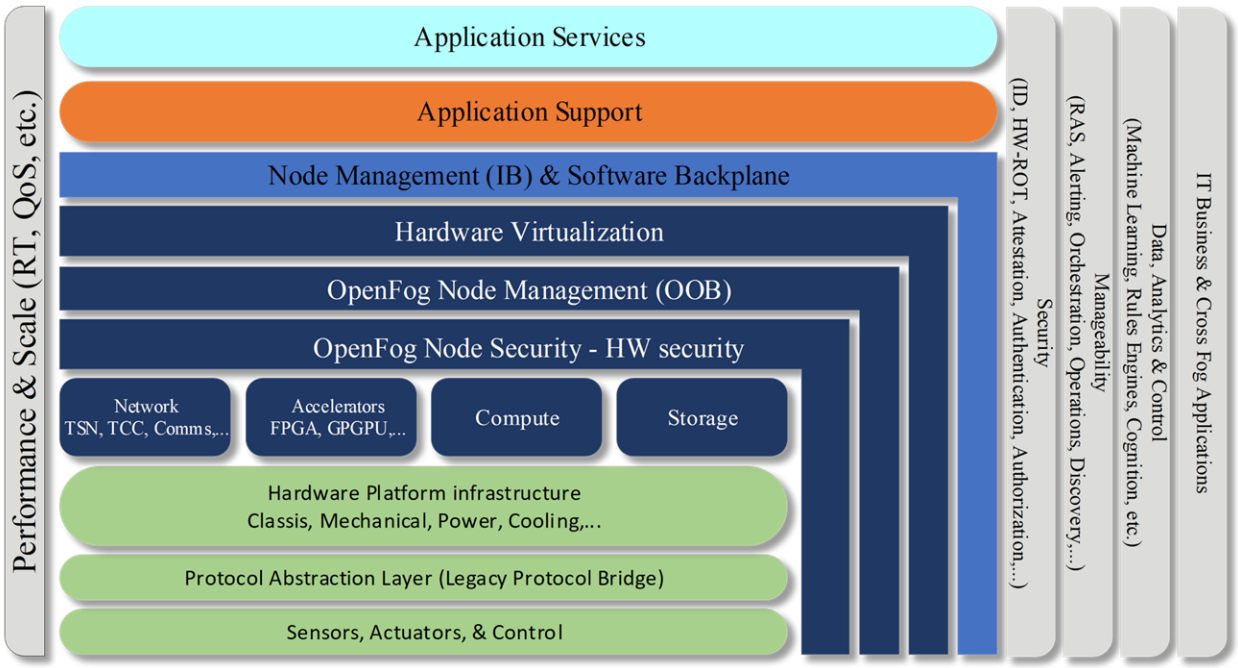


Figure 1.1 — OpenFog Architecture

The reference architecture comprises the following key components [7]:

1. Physical Layer – encompasses all physical devices such as sensors, actuators, and edge-computing nodes.
2. Connectivity Layer – provides communication between physical-layer devices and the rest of the network, including access technologies, protocols, and interfaces.

3. Control Layer – manages resources, tasks, and security within the fog-computing environment; it offers mechanisms for performance monitoring, orchestration, and access control.
4. Processing Layer – dedicated to processing and analyzing data received from the physical layer, including real-time analytics, machine learning, and artificial-intelligence workloads.
5. Application Layer – hosts applications that consume data gathered and processed by the lower layers (e.g., monitoring, management, and optimization services).
6. Data-Management Layer – ensures efficient data handling across all layers, covering acquisition, storage, processing, and analysis.
7. Security Layer – a cross-cutting plane that shields data and devices at every tier from unauthorized access and cyber-attacks.

Fog computing embodies a distributed architecture that accelerates data processing, lowers latency, and locates computational resources close to end-users. By deploying compute nodes at the network edge, data can be processed nearer to its source. The OpenFog framework thus provides the foundation for scalable, flexible, and secure systems capable of real-time operation. In this sense, fog computing is not merely an alternative to cloud platforms but a necessary complement for the next wave of ultra-low-latency technologies. A broad body of literature addresses various facets of fog evolution and implementation.

RELATED WORKS

Table 1.1 — Surveyed Contributions in Fog and Cloud Computing

Author(s)	Year	Contribution
S. Delfin; N. P. Sivasanker; Nishant Raj; Ashish Anand	2019	Fog Computing: A New Era of Cloud Computing [8] – analyses the conceptual foundations of fog computing
Asma N. Elmoghrapi; Ahmed Bleblo; Younis A. Younis	2022	Fog Computing or Cloud Computing: A Survey [9] – compares advantages and drawbacks of the two paradigms
Pooyan Habibi; Mohammad Farhoudi; Sepehr Kazemian; Siavash Khorsandi; Alberto Leon-Garcia	2020	Fog Computing: A Comprehensive Architectural Survey [10] – offers an exhaustive overview from an architectural viewpoint
Muneer Bani Yassein; Ismail Hmeidi; Farah Shatnawi; Saif Rawasheh	2020	Fog Computing: Characteristics, Challenges, and Issues [11] – focuses on fog-computing security concerns
Manoj Muniswamaiah; Tilak Agerwala; Charles C. Tappert	2021	Modern Computing and the Internet of Things (IoT): A Survey [12] – explores IoT integration with cloud and fog computing
Ahmed Salem; Marc St-Hilaire; Muhammad Sajjad Khan	2023	Mathematical Models for SLA-Aware Task Allocation in Fog Computing [13] – proposes analytical models for load balancing and energy efficiency

Author(s)	Year	Contribution
Mani Sharifi; Abdolreza Abhari; Sharareh Taghipour	2021	Real-Time Application-Processor Scheduling in Fog Computing [14] – presents a MILP-based scheduling model
Engr. Urooj Yousuf Khan; Tariq Rahim Soomro	2023	A Review of iFogSim: A Simulation Tool for Future Fog Networks [15]– evaluates the iFogSim simulation software

Delfin et al. [8] trace the origins of fog computing, highlight its principal advantages, catalogue representative use-case families, and discuss pathways for integrating fog with established cloud-centric solutions as the paradigm continues to mature.

Elmoghrapi et al. [9] expand the discussion by juxtaposing cloud and edge-local computing models. Their comprehensive survey assesses how each technology affects day-to-day digital interactions, clarifying that cloud and fog serve complementary—rather than interchangeable—roles in advanced network ecosystems.
Habibi et al. [10] provide an exhaustive architectural panorama, comparing contemporary reference frameworks with application-specific designs. The review dissects algorithmic and technological facets, then maps the distinguishing features of cloud, edge, mobile-edge, and fog paradigms.
Security emerges as a recurrent theme. Bani Yassein et al. [11] analyze fog from a layered perspective, underscoring privacy and protection challenges that must be addressed before large-scale adoption is feasible.
Because IoT workloads generate prodigious sensor data that demand near-instant analytics, Muniswamaiah et al. [12] investigate how tightly coupling fog resources with IoT platforms can unlock performance gains across heterogeneous devices.
Salem et al. [13] contribute formal task-allocation models for service-level-agreement (SLA) compliance in fog clusters. One formulation target load-balancing efficiency, whereas another emphasizes eco-sustainability by minimizing energy consumption.
Sharifi et al. [14] advance this line of inquiry with a mixed-integer linear-programming (MILP) scheduler that delivers optimal dispatching decisions; they benchmark its efficacy against a FIFO baseline under two representative scenarios.
Finally, Yousuf Khan and Soomro [15] evaluate iFogSim, a widely used discrete-event simulator that offers modular templates and class libraries for modelling next-generation fog topologies and application flows.

MODEL NETWORK

The integration of fog computing into the design of telepresence traffic-handling devices provides an effective approach to mitigating several critical challenges, including limited onboard computational capacity, rapid battery depletion, data acquisition constraints and scalability issues, as well as insufficient transmission speeds. By leveraging distributed processing at the network edge, fog-enabled telepresence systems can achieve higher efficiency, more compact form factors, and enhanced user experience due to reduced latency and optimized resource utilization.
Building on a rigorous analysis of traffic offloading mathematical models, we have developed a discrete-event simulation model that emulates data transmission from a telepresence suit to metaverse infrastructures. This simulation enables comprehensive performance evaluation, including delay metrics and load distribution across fog nodes, providing a framework for fine-tuning real-time applications.
The figures included in this study illustrate the prototype architecture and operational workflow of the proposed model, highlighting how fog-assisted offloading can effectively reduce end-to-end delays while improving throughput and energy efficiency in telepresence environments.

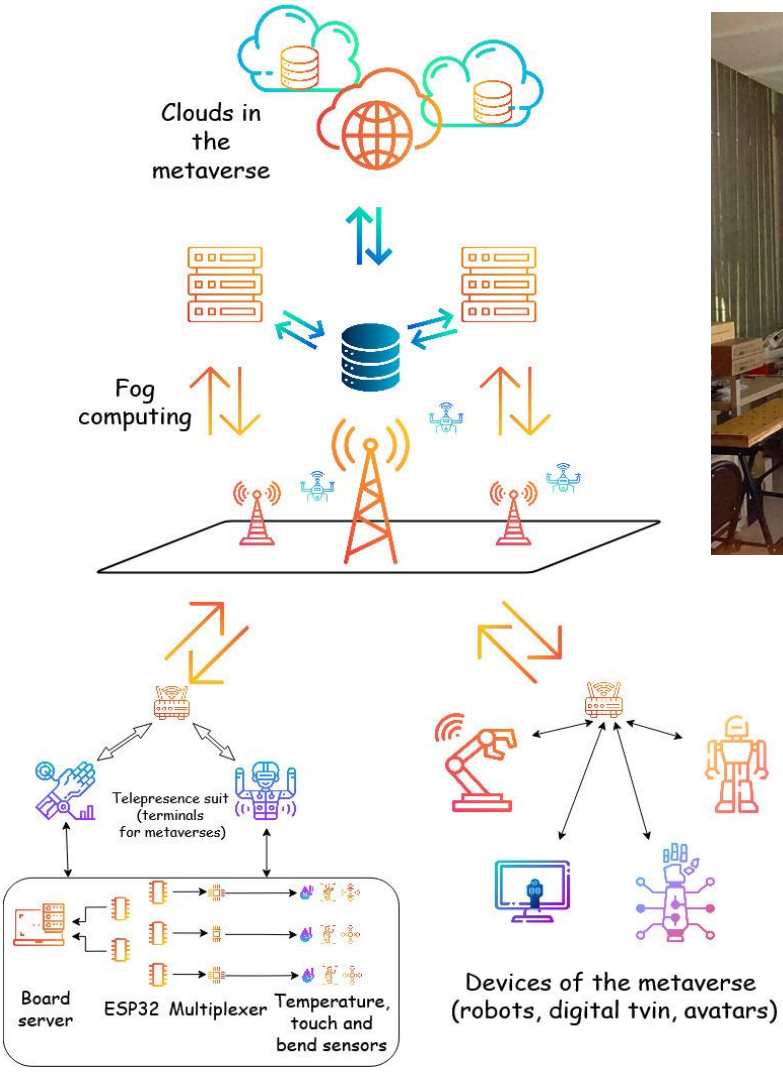




Figure 1.2 — Model Architecture

Mathematical Model

We consider a mobile-computation offloading scheme that comprises a single fog-computing node and a remote cloud-computing facility. The telepresence suit equipped with M processing boards is treated as a multi-server queuing system capable of diverting traffic to both fog and cloud resources. *Figure 1.2* illustrates an exemplary network in which

- M low-power devices represent the on-board processors,
- node $M + 1$ denotes the fog-computing element, and
- node $M + 2$ corresponds to the cloud-computing element.

During operation, the suit's boards consume substantial computing cycles and electrical power. According to the proposed model, a portion of the data-processing workload is offloaded to the fog node. Should the fog segment become saturated, overflow tasks are forwarded to the remote cloud to prevent congestion. Our design goal is to reduce the number of boards, minimize energy consumption, and utilize network resources efficiently while maintaining strict end-to-end latency constraints.

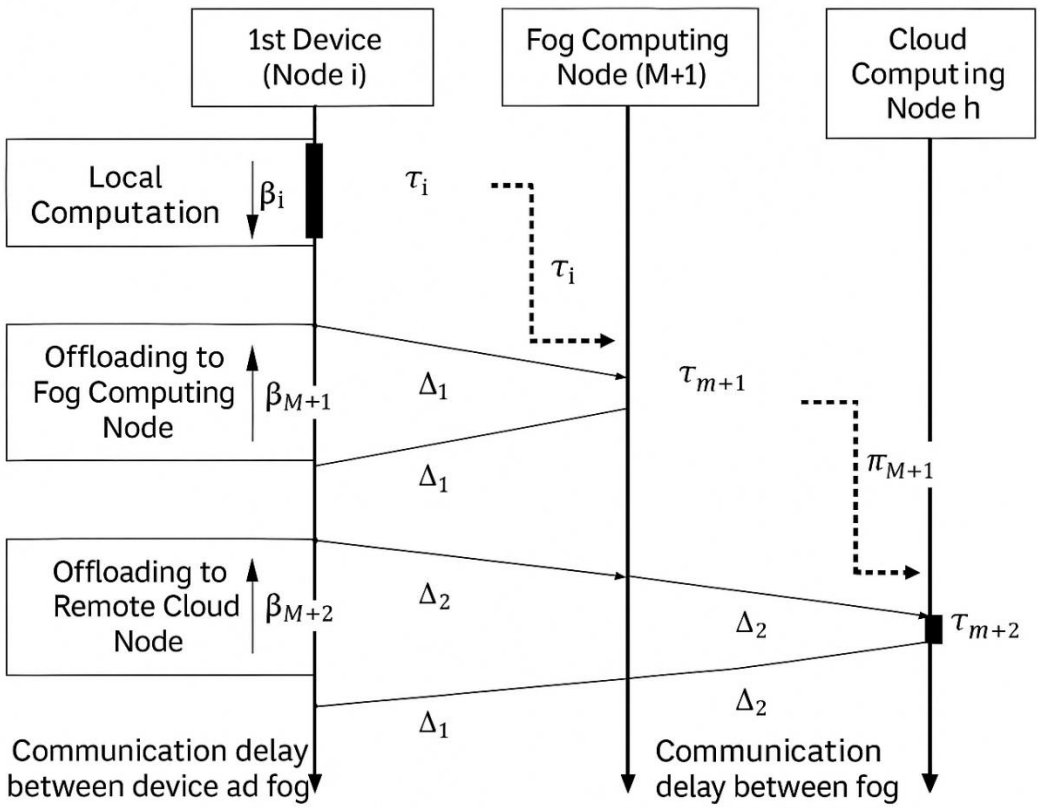


Figure 1.3 — Network-Based Computation-Offloading Model

Let β_i denote the random variable representing the response time of a task originating from board i.

- Upper branch — Local execution. The task is processed entirely on board i; hence the board sustains the greatest computational load, but the response time is minimal and equals the local processing duration:

$$\beta_i = t_i.$$

- Middle branch — Fog-level execution. The task is offloaded via a wireless link to the fog node. Although transmission incurs delay, this option conserves the board's processing cycles and energy. The response time is the sum of the round-trip transmission delay $2\Delta_1$ and the fog-processing time t_{M+1}:

$$\beta_i = 2\Delta_1 + t_{M+1}.$$

- Lower branch — Cloud-level execution. When the fog node is unable to handle the task, it forwards the job to the remote cloud, thereby relieving fog congestion at the expense of additional latency. The response time now includes the round-trip delay to the cloud $2\Delta_2$:

$$\beta_i = 2\Delta_1 + 2\Delta_2 + t_{M+2}.$$

The expressions in (1) capture the latency trade-offs among local, fog, and cloud processing paths, forming the basis for the performance analysis presented in [16].

This computation-offloading scheme can be formalized in terms of queueing networks. Let M denote the number of processing boards—equivalently, the local-computation nodes—within the coverage area of the fog node. A board that acquires and forwards sensor data is indexed as node i,

where $1 \leq i \leq M$. Each such node generates a Poisson stream of requests with rate λ_i [requests s^{-1}]. Every task requires a specific computational workload, measured in operations or floating-point operations. Let $w_i, \quad 1 \leq i \leq M,$ be the non-negative random variable representing the workload of a task originating on board i, with cumulative distribution function $W_i(x) = \Pr\{ w_i \leq x \}$. Denote by μ_i the constant service rate of board i in instructions per second. A task is offloaded to the fog node whenever $w_i > w^*$, where w^* is the workload threshold; otherwise, it is executed locally on the mobile device. The offloading probability for board i is therefore $\pi_i = 1 - W_i(w^*)$. With the foregoing notation, the complete mathematical model of the system is depicted schematically in *Figure 1.3* [14].

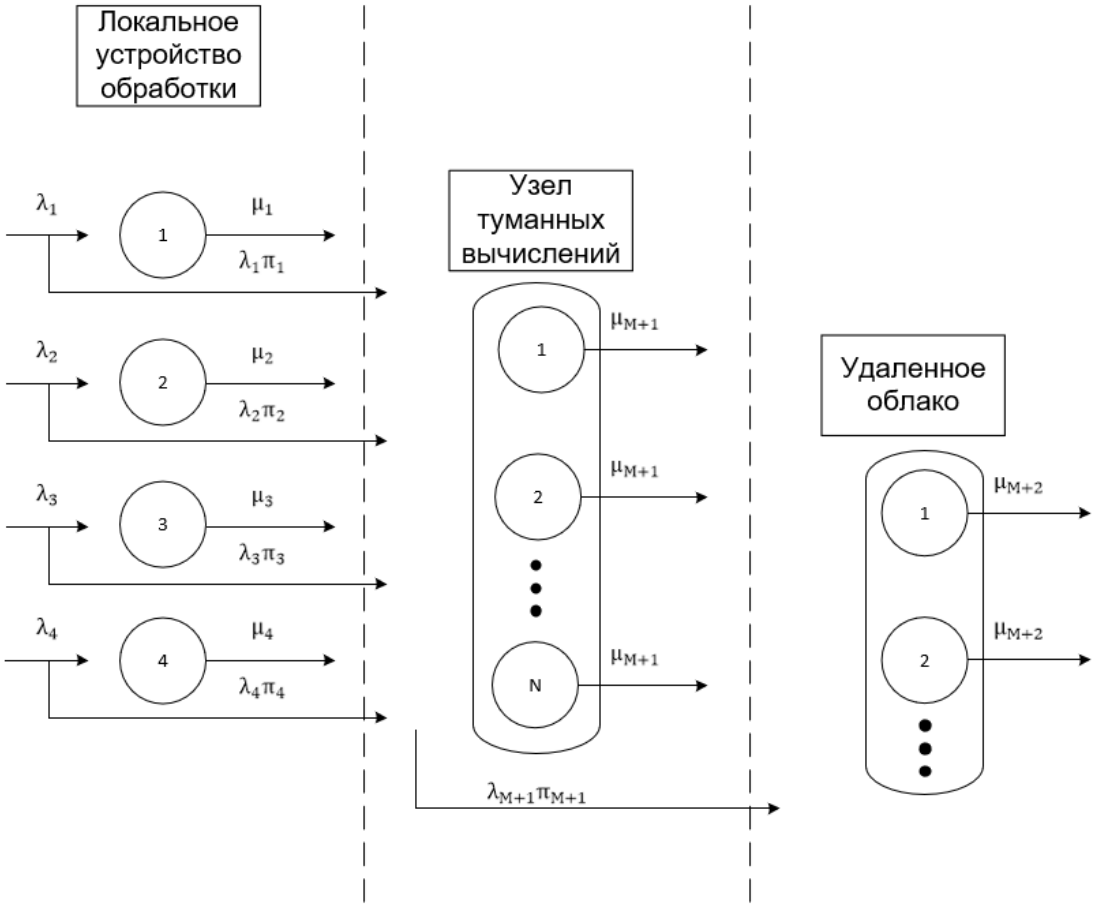


Figure 1.4 — Queueing-System Model for Telepresence-Suit Traffic Offloading

Table 1.2 — Probability Distribution of the Response-Time Random Variable β_i

β_i value	t_i	$2\Delta_1 + t_{M+1}$	$2\Delta_1 + 2\Delta_2 + t_{M+2}$
Probability P	$1 - \pi_i$	$\pi_i(1 - \pi_{M+1})$	$\pi_i \, \pi_{M+1}$

Simulation Model

Building on the mathematical framework outlined above, we developed a discrete-event simulation in AnyLogic and conducted a series of experiments to validate the analytical results.

The simulator implements both fog-assisted and non-fog data-transmission scenarios, allowing direct comparison of end-to-end performance.

The accompanying figure depicts the experimental test-bed layout, while the corresponding table lists the parameter values used in the simulations. This model enables network planners to:

- Forecast system behavior when telepresence traffic is integrated with fog computing,
- Dimension fog clusters appropriately, and
- Estimate the user-side network resources required to sustain real-time service delivery.

These capabilities are instrumental for designing scalable, low-latency telepresence architectures in next-generation networks [17] [18].

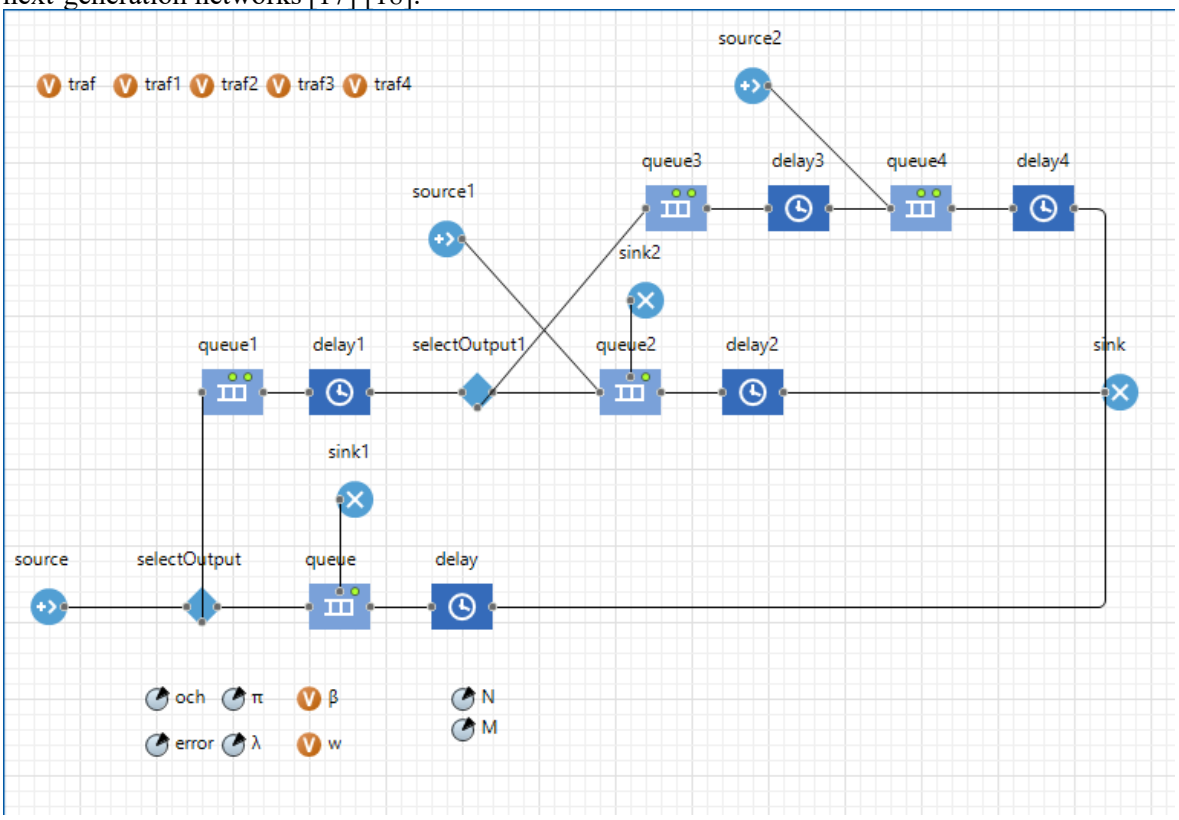


Figure 1.5 — Discrete-Event Simulation Model of the Queueing System

Table 1.3 — Model Parameters Used in the Experiments

Parameter (Name)	Symbol
Flow arrival rate per board	λ_i
Offloading probability	π
Offloaded-workload threshold	w^*
Local processing rate	μ_i
Fog-node processing rate	μ_{M+1}
Cloud processing rate	μ_{M+2}
One-way fog-transmission delay	Δ_1
Additional cloud-transmission delay	Δ_2
Mean response time	β_i

Parameter (Name)	Symbol
Overload events (error)	*error*
Queue length	*och*
Number of processing boards	M
Devices in the fog cluster	N

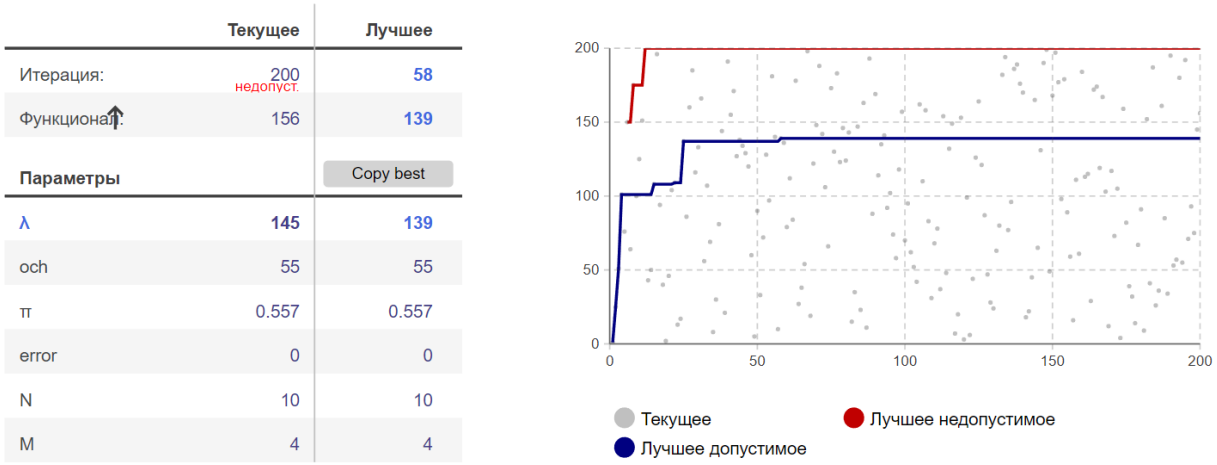


Figure 1.6 — Optimisation-Model Outcome

In the first experiment, we determined the optimal values of the offloading probability and the queue length that minimize the mean system sojourn time. Because timely delivery of every packet is critical for telepresence traffic, the configuration must guarantee zero overload events.

The optimization yielded an offloading probability of $\pi = 0.557$, which is fully consistent with the analytical model, along with the corresponding minimal queue length. This configuration balances the volumes offloaded to fog, processed locally, and forwarded to the cloud, ensuring uniform workload distribution across the boards.

The second and third experiments employed parameter sets derived from the actual telepresence-suit hardware specification and refined through the simulation framework. We established

- the maximum data volume that the boards can collect and process without fog support, and
- the performance gains—in terms of throughput and latency—achieved through fog-computing integration.

Table 1.4 — Experimental Results

Metric	Baseline	Prototype	Fog Deployment
Number of processing boards	20	4	4
Offloading probability	$\pi = 0.559$	$\pi = 0$	$\pi = 0.557$
Mean response time β [s]	0.012	0.021	0.007
Aggregate arrival rate λ [tasks s^{-1}]	80	48	139
Normalized β	1.00	1.75	0.58
Normalized λ	1.00	0.60	1.74

The introduction of fog computing into telepresence service architectures addresses critical challenges—including limited on-board computational power, rapid battery depletion, data-acquisition bottlenecks, scalability constraints, and insufficient transmission throughput. As demonstrated, the fog-enabled configuration triples the data-collection and transfer capacity without adding extra hardware to the suit. Consequently, system efficiency is markedly improved while complexity and cost remain unchanged, enabling a lean yet high-performance telepresence solution suitable for next-generation 6G networks.

CONCLUSION

The application of analytical modelling and discrete-event simulation in AnyLogic has produced several quantitative outcomes: (i) a pronounced reduction in system response time, significantly exceeding the data-processing rates of current solutions; (ii) lower energy consumption owing to task execution at fog nodes; and (iii) a decrease in the amount of required on-board hardware. By integrating fog computing, the traffic intensity can be maximized without a three-fold increase in the number of processing boards, thereby demonstrating substantial economic efficiency.

The simulation model has experimentally validated the theoretical insights obtained from the mathematical analysis. This validation not only confirms the accuracy of the theoretical calculations but also identifies optimal operating parameters for the system under a variety of conditions, providing a reliable basis for further research on fog-cluster sizing.

The experiments clearly indicate the significant advantages of employing fog computing in telepresence devices, including a three-fold improvement in system performance and a corresponding reduction in response time. These findings confirm that the integration of fog technologies is a promising avenue for the advancement of telepresence systems, delivering higher efficiency and greater user convenience for end users.

REFENRENCE:

[1] C.-X. Wang, J. Huang, H. Wang, X. Gao, X. You, and Y. Hao, “6G Wireless Channel Measurements and Models: Trends and Challenges,” *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 22–32, Dec. 2020, doi: 10.1109/MVT.2020.3018436.

[2] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, “6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities,” *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, July 2021, doi: 10.1109/JPROC.2021.3061701.

[3] A. Shi, F. Sun, and M. Li, “An Exploration of Holographic Scenarios in 6G Networks,” in *2022 10th International Conference on Information Systems and Computing Technology (ISCTech)*, Dec. 2022, pp. 450–456. doi: 10.1109/ISCTech58360.2022.00076.

[4] A. Volkov, A. Muthanna, and A. Koucheryavy, “Fifth generation communication networks: on the way to networks 2030,” *Telecom IT*, vol. 8, no. 2, June 2020, doi: 10.31854/2307-1303-2020-8-2-32-43.

[5] abid, “Introducing the Industrial Internet Consortium, Now Incorporating OpenFog!,” Industry IoT Consortium. Accessed: July 21, 2025. [Online]. Available: https://www.iiconsortium.org/2019/02/introducing-the-industrial-internet-consortium-now-incorporating-openfog/

[6] “3. OpenFog Consortium Architecture Working Group, OpenFog architecture overview, OpenFog Consortium, - Google Search.” Accessed: July 21, 2025. [Online]. Available: https://www.iiconsortium.org/pdf/OpenFog_Reference_Architecture_2_09_17.pdf

[7] Thang, D. V., Volkov, A., Muthanna, A., Koucheryavy, A., Ateya, A. A., & Jayakody, D. N. K. (2025). Future of Telepresence Services in the Evolving Fog Computing Environment: A Survey on Research and Use Cases. Sensors, 25(11), 3488.

[8] S. Delfin, N. P. Sivasanker., N. Raj, and A. Anand, "Fog Computing: A New Era of Cloud Computing," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, Mar. 2019, pp. 1106–1111. doi: 10.1109/ICCMC.2019.8819633.

[9] A. N. Elmoghrapi, A. Bleblo, and Y. A. Younis, "Fog Computing or Cloud Computing: a Study," in *2022 International Conference on Engineering & MIS (ICEMIS)*, July 2022, pp. 1–6. doi: 10.1109/ICEMIS56295.2022.9914131.

[10] P. Habibi, M. Farhoudi, S. Kazemian, S. Khorsandi, and A. Leon-Garcia, "Fog Computing: A Comprehensive Architectural Survey," *IEEE Access*, vol. 8, pp. 69105–69133, 2020, doi: 10.1109/ACCESS.2020.2983253.

[11] M. B. Yassein, I. Hmeidi, F. Shatnawi, and S. Rawasheh, "Fog Computing: Characteristics, Challenges and Issues," in *2020 International Conference on Mathematics and Computers in Science and Engineering (MACISE)*, Jan. 2020, pp. 240–245. doi: 10.1109/MACISE49704.2020.00051.

[12] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Fog Computing and the Internet of Things (IoT): A Review," in *2021 8th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, June 2021, pp. 10–12. doi: 10.1109/CSCloud-EdgeCom52276.2021.00012.

[13] A. Salem, M. St-Hilaire, and M. S. Khan, "Mathematical Models for Task Assignment with Service Level Agreement in Fog Computing," in *2022 International Conference on Smart Applications, Communications and Networking (SmartNets)*, Nov. 2022, pp. 1–8. doi: 10.1109/SmartNets55823.2022.9993983.

[14] M. Sharifi, A. Abhari, and S. Taghipour, "Modeling Real-Time Application Processor Scheduling for Fog Computing," in *2021 Annual Modeling and Simulation Conference (ANNSIM)*, July 2021, pp. 1–12. doi: 10.23919/ANNSIM52504.2021.9552113.

[15] Engr. U. Yousuf Khan and T. Rahim Soomro, "Overview of iFogSim: A Tool for Simulating Fog Networks of the Future," in *2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, Nov. 2022, pp. 1–4. doi: 10.1109/MACS56771.2022.10023395.

[16] Zhao, L. H., Zhang, J., Bah, M. J., Li, Z., Ying, J. J. C., Muthanna, A., & Elgendy, I. A. (2025). Energy-Efficient Task Offloading in Multi-server Mobile Edge Computing Networks: A Deep Reinforcement Learning Approach. Arabian Journal for Science and Engineering, 1-22.

[17] Vorobyova D. et al. IoT network model with multimodal node distribution and data-collecting mechanism using mobile clustering nodes //Electronics. – 2023. – Т. 12. – №. 6. – С. 1410..

[18] van Thang, D., Volkov, A., Muthanna, A., Elgendy, I. A., Alkanhel, R., Jayakody, D. N. K., & Koucheryavy, A. (2025). A Framework Integrating Federated Learning and Fog Computing Based on Client Sampling and Dynamic Thresholding Techniques. IEEE Access, 13, 95019-95033.

Highly Efficient 3D Object Transmission System for HTC Services in 6G Networks

Daniil Svechnikov, Artem Volkov, Anastasia Marochkina, Ammar Muthanna, Andrey Kouhceryavy

The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, St. Petersburg, 193232, Russian Federation

danil2001sv@mail.ru

ABSTRACT

In recent years, advancements in technology have brought forth a new frontier in visual communication. Holography is a technique that captures and reproduces threedimensional (3D) images with an unprecedented level of realism and depth, has emerged as a groundbreaking method for conveying visual information. Unlike traditional images and videos, holography recreates scenes with full parallax, enabling viewers to perceive objects from various angles. The transmission of holographic images presents both exciting possibilities and unique challenges. To this end, this article conducts a comparative analysis of a previously developed application system for transmitting dynamic 3D human movements with a ready-made solution for transmitting 2D video streams in order to provide conference calling services. The network characteristics of the systems were collected and compared. The opportunities that programs currently provide and will provide in the future are examined.

KEYWORDS: *Telepresence, Comparative Analysis, MVP, TrueConf, HTC, 6G, Teleconference, 3D images, Conferencing*

INTRODUCTION

With the development of communication networks, users have more and more requests for the provision of various services. This paper will compare two applications from the field of telepresence services, which are still developing. Separately, we can note the direction of Holographic-Type Communication, the development of which is currently happening everywhere in the world [1].

During the pandemic, the need for video conferencing has greatly increased, examples of which have become firmly established in our daily lives: a work meeting, a lecture at the university, friendly communication. Conferencing technology allows for simultaneous communication between multiple users.

At the moment, applications for transmitting 2D images over networks for conferencing are very widespread. One of the popular solutions in this area today is TrueConf - a software platform for video conferencing. It helps employees communicate and work effectively from a distance, as well as easily conduct webinars and remote training.

Another direction of applications for remote work, education, etc. is the remote transmission

of three-dimensional images, which makes it possible to see human movements in more detail from all angles, and also opens up new possibilities for immersive communications [2,3]. Transferring data about three-dimensional objects in its original form, for example in the format of streaming video from an array of cameras, requires significant bandwidth, so it becomes important to use technologies for converting data received about an object into other formats [3, 4].

This paper aims to investigate and compare the network traffic of several telepresence applications to test whether it is possible to provide a more complex conferencing service than is currently available without significant losses in network traffic. The following will be considered: existing commercial solutions and a self-developed system that provides the opportunity to see a person not only in 2D, but also adds a third plane, allowing you to display a 3D model of a person in augmented reality, mixed reality or any holographic method of providing data.

EXPERIMENTS DESCRIPTION

During a superficial analysis of the network characteristics of other programs for providing conferencing using live video stream transmission, similar to TrueConf, it became clear that any of the advanced applications, such as Zoom, Google Meat, Yandex Telemost, have comparable results. This conclusion made it possible to compare the developed MVP system with a single application that is more convenient to use and test - TrueConf [5].

The research involved two applications:

- TrueConf - a unified communications software system from the company with the same name - TrueConf, using streaming video transmission to a flat surface in 2D to provide conferencing.

- MVP - an application system designed to synchronize control points on the Sender's body and an avatar that simulates movements on the Receivers' devices.

TrueConf, like any other software currently on the market, is a means of transmitting streaming video and audio - this is the minimum set of capabilities for providing conference calls. As you know, streaming video puts a lot of load on network equipment due to its weight. At the time of the transition of the whole world to remote communications, network equipment was not ready to provide the resource for transmitting such an amount of traffic. It was necessary to come up with a new way of transmitting data to provide conference calls. For this purpose, the MVP application system was developed.

MVP is an application system consisting of 3 parts (Sender, Server, Receiver). The Sender collects data from the Kinect camera about the control points of the human body (Atlas of body "Key Points"), groups them into a vector database and sends it to the Server. The Server constantly synchronizes network components on all connected devices at regular intervals; if synchronization of a component fails, a timer is started, after which the user will be disconnected as inactive. The network component on the Receiver's side receives updated data from the Server and transmits it to a 3D avatar that simulates the Sender's movements. Receiver client applications are implemented both on the Standalone system (Windows / Linux / MacOS) and for Android. Moreover, playback on any Standalone machine can be projected onto a holographic cube to achieve a three-dimensional effect (the speaker's avatar can be walked around from all sides), and the Android version uses ARCore to create augmented reality, which will allow the user to place the avatar at any convenient point in his room using your phone camera [6].

It is important to note that unlike TrueConf, the MVP application does not transmit video traffic, but uses data transmission instead. This feature allows us to move away from the usual concept of providing conferencing services and get closer to the full transfer of point clouds over communication networks to

support holographic communication [7].

MVP is a platform that can be developed in various directions, from solving problems of reducing network traffic to provide telepresence services using 3D models, to developing a full-fledged holographic application for providing HTC. At the moment, this is a prototype that will allow conducting statistical experiments to test various hypotheses. Comparing the network characteristics of the developed application and a related application (similar application) will give us an understanding in which direction it is worth developing and whether it is worth continuing development at all [8].

Sixteen (16) experiments were carried out, each of which lasted 300 seconds and took place under the same conditions on the same network topology (Fig. 1): Eight (8) experiments for each application (Table 1).

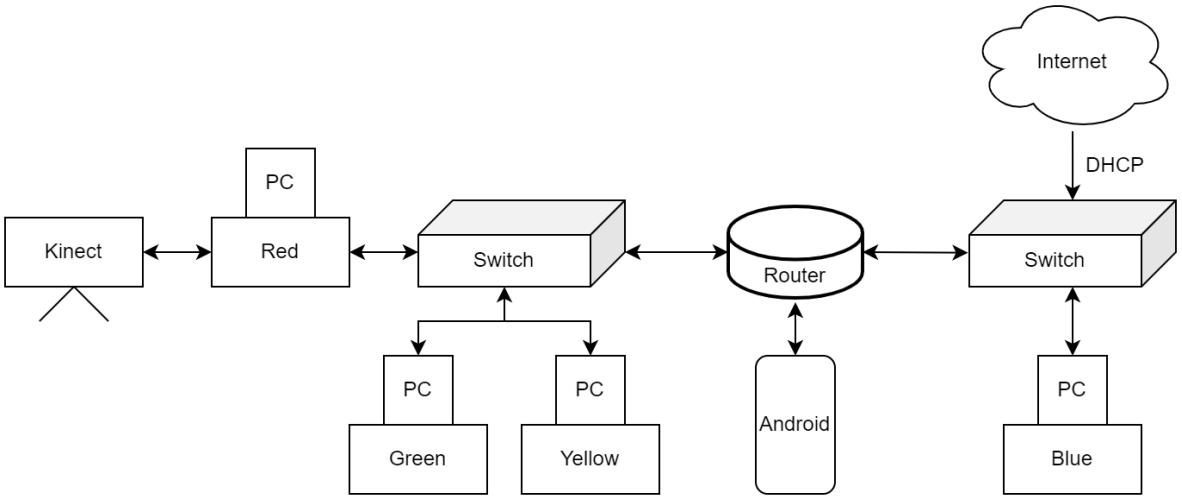


Figure 1. Network Architecture

Red, Green, Blue - different names of computers, corresponding to their color, which is later translated into graphics, Android - designation of a phone with the Android system (indicated in dark green). Idle – a type of experiment in which there is no person in the frame and the picture is stationary, Movement – experiments in which a person actively moves in the frame.

Table 1. Experiments

	Idle	Movement
1 Client	Yellow PC	Yellow PC
2 Clients	Yellow PC + Green PC	Yellow PC + Green PC
3 Clients	Yellow PC + Green PC + Blue PC	Yellow PC + Green PC + Blue PC
3 + 1 Clients	Yellow PC + Green PC + Blue PC + Android	Yellow PC + Green PC + Blue PC + Android

EXPERIMENTS DIFFERENCES

Before moving on to a direct comparison of network characteristics, it is important to note several points that affect the generated traffic, which we will need later.

The big difference between the "Idle" experiments comparing the two applications is caused by the way the applications generate network traffic. The TrueConf application converts the 2D image received from the camera into packets and sends it over the network every unit of time, as in the "Movement" experiment, MVP requires a small bandwidth just to maintain client connections to the server (0-8 Kbps), if this connection is lost, the server will automatically disconnect the user. This feature will be clearly visible in the presented graphs, and on some you may get the feeling that there is no traffic at all, but this is not the case, it is simply extremely small compared to the rest [9].

The experiment presents various devices with different hardware: "Yellow PC" and "Blue PC" interface - "Realtek Gaming GbE Family Controller", "Green PC" interface - "Realtek PCIe 2.5GbE Family Controller", the used phone is "Xiaomi Mi 8" with "Android 9". The network components of the MVP application for smartphones and for PCs are completely identical.

Also, the devices are located in different places in the network topology (Fig. 1); in all experiments, the server and sending client were deployed on the "Red PC", after which the traffic went to the "Aruba 2930F JL259A" switch device, to the "Green PC" and " Yellow PC", then comes the "Xiaomi Mi AIoT AX3600" router, it transmits data to Android; from the router, network traffic goes to the next "Catalyst 3750G" switch, from switch to the "Blue PC" end device.

COMPARISON OF NETWORK CHARACTERISTICS OF DIFFERENT SYSTEMS

The catalyst for developing a functionally new video conferencing application was the urgent need to reduce network video traffic, given the transition of a large number of people to remote work and education.

In the realities of the need to write an application "here and now" many auxiliary systems and libraries were used that lower final performance for the sake of faster application writing.

The developed system does not utilize all the capabilities of the physical layer of the OSI model, specifically the Ethernet frame length. In most cases (62%), only half of the Ethernet packet is used, which is approximately 700-750 bytes (Fig. 2). This underutilization of the packet means that, at any given time, we have the potential to transmit more useful information than we currently do (a path for transmitting more points), without compromising Quality of Service (QoS) and network characteristics. In contrast, a more optimized TrueConf application uses the full packet length of 1550 bytes in 54% of cases, and cannot transmit a greater amount of data in a single packet, i.e., it cannot increase the ratio of useful information to the total packet length with headers (Fig. 3).

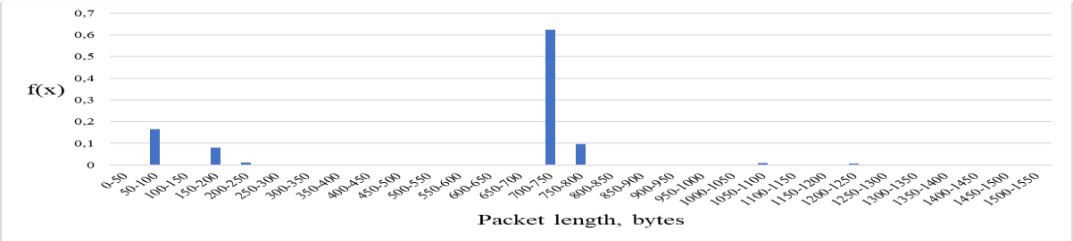


Figure 2. MVP, 3 + 1 Clients, Packet length

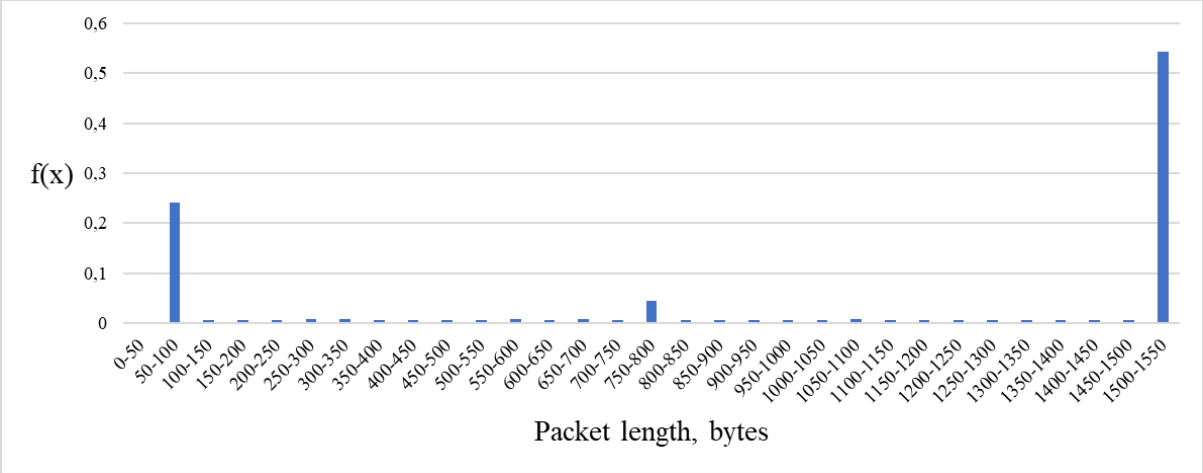


Figure 3. TrueConf, 3 + 1 Clients, Packet length

The described features can be observed in the comparative graph of the Average Packet Size in Bytes (Fig. 4). The average packet size values, as well as their distributions, have a small margin of error and are independent of the number of connected devices.

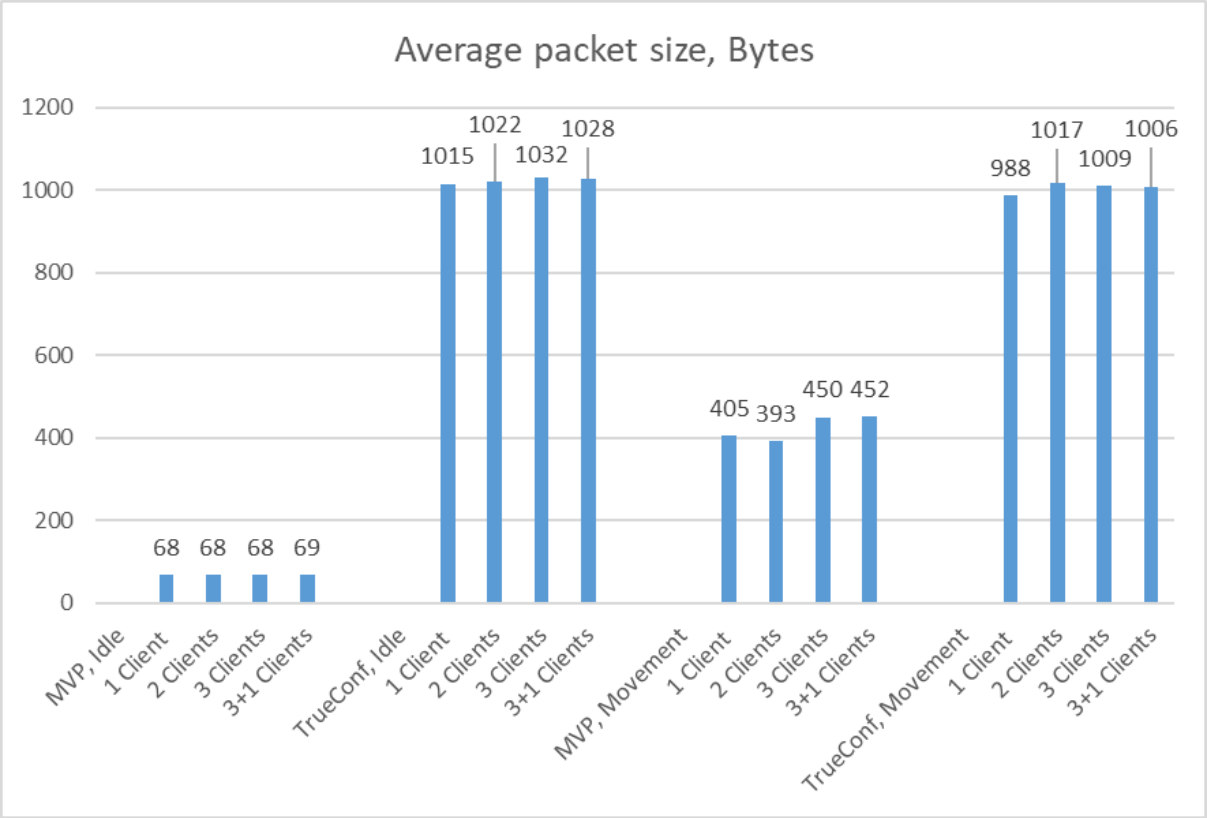


Figure 4. Average packet size, Bytes

The intensity refers to the number of packets transmitted per unit of time, in this case, per second (Fig. 5). The graph shows that the developed system imposes a lighter load on the network compared to TrueConf, and this difference is particularly noticeable when there is no person in the frame, as mentioned earlier.

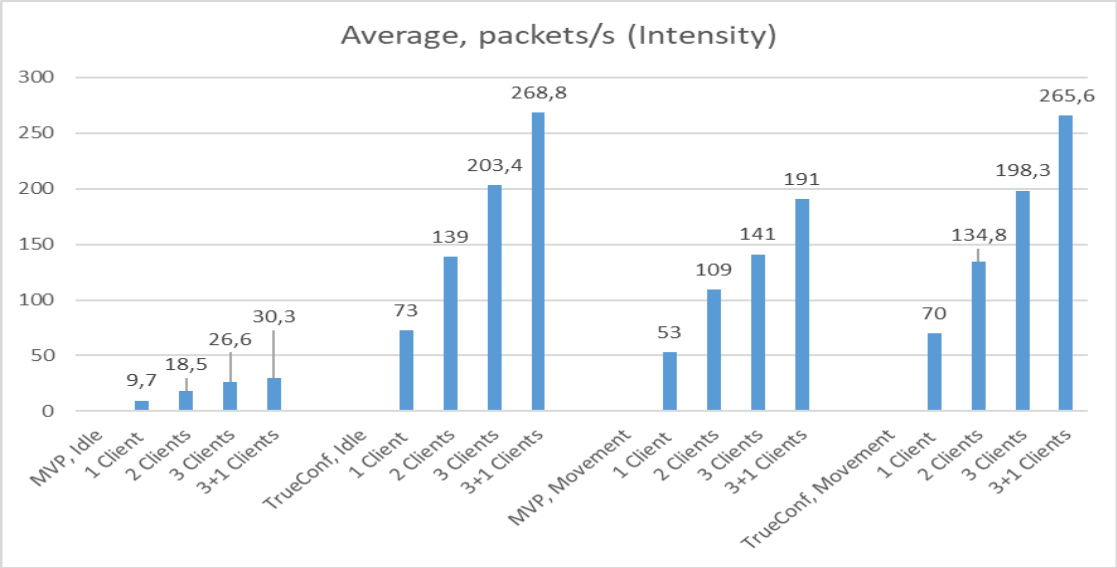


Figure 5. Average, packets/s (Intensity)

Data transfer rate, which measures the quantity of data transmitted within a specific timeframe, in this case, the number of bits sent per second, already exhibits a substantial advantage in terms of the volume of information transmitted (Fig. 6). This advantage is contingent on the fact that the developed MVP system conveys data about the 3D model, which can be processed and displayed to the user in various ways, in contrast to TrueConf, which only transmits 2D video.

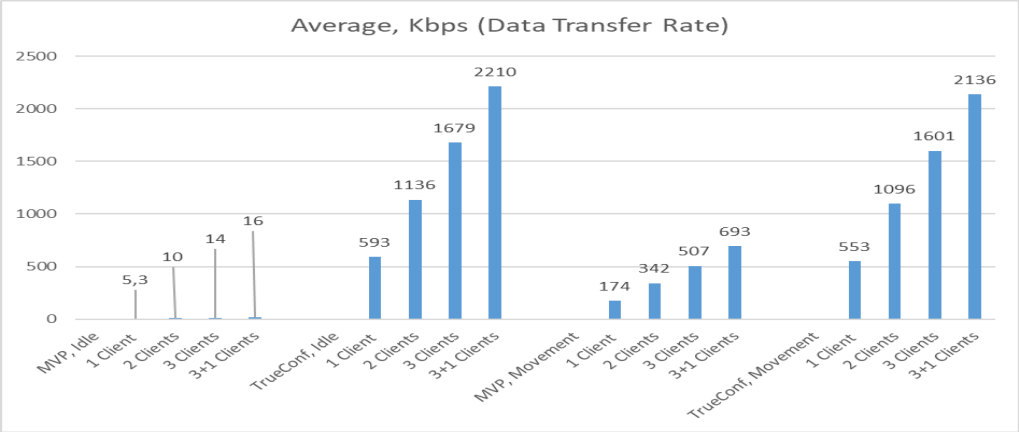


Figure 6. Average, bps (Data Transfer Rate)

The chart representing the transmitted data (Fig. 7) displays the volume of bytes transmitted throughout the entire five-minute application lifecycle, encompassing startup, connection, data transmission, and disconnection. This graph effectively illustrates the reduction in data transmitted over the network. When considering only the "Movement" experiments, the data reduction factor consistently exceeds 3. This implies that regardless of the number of connected devices, when a person is in the frame, the developed system transmits three times less information over the network than TrueConf. As for the "Idle" experiments, in these cases, the factor hovers around a value of 130, which is noteworthy but less significant, given the absence of the transmission of useful information.

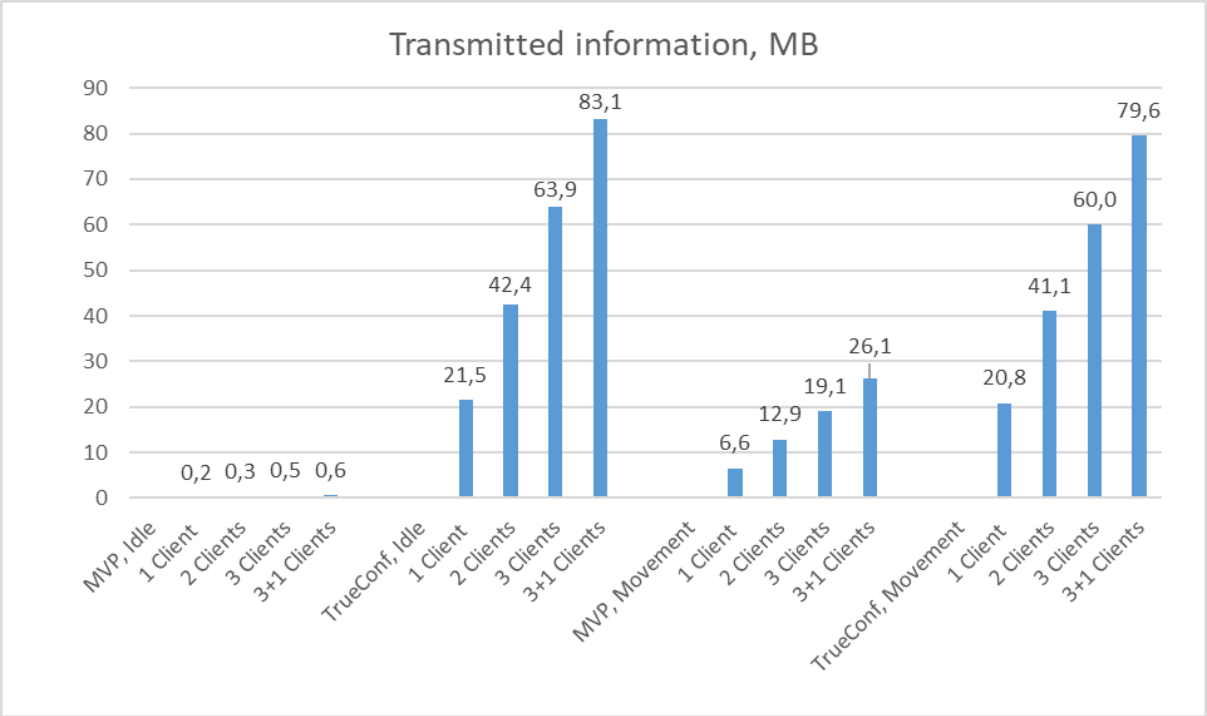


Figure 7. Transmitted information, Bytes

The chart (Fig. 7) illustrates that the volume of transmitted data increases linearly with the number of connected devices, both for the developed application and TrueConf. This information will be valuable for future network load forecasting when necessary.

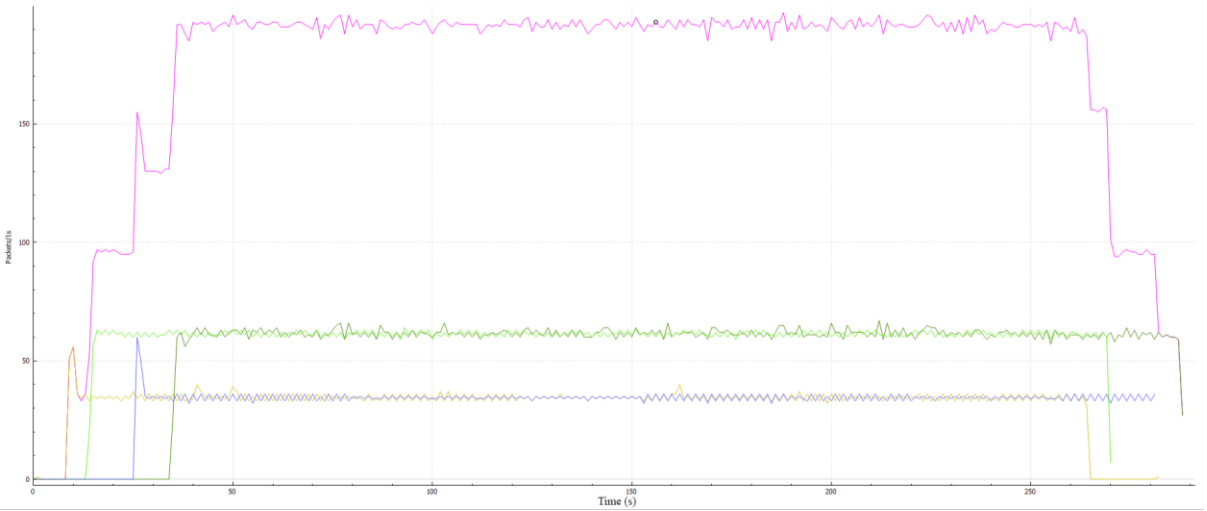


Figure 8. MVP, 3 + 1 Clients, Movement, Packets per second

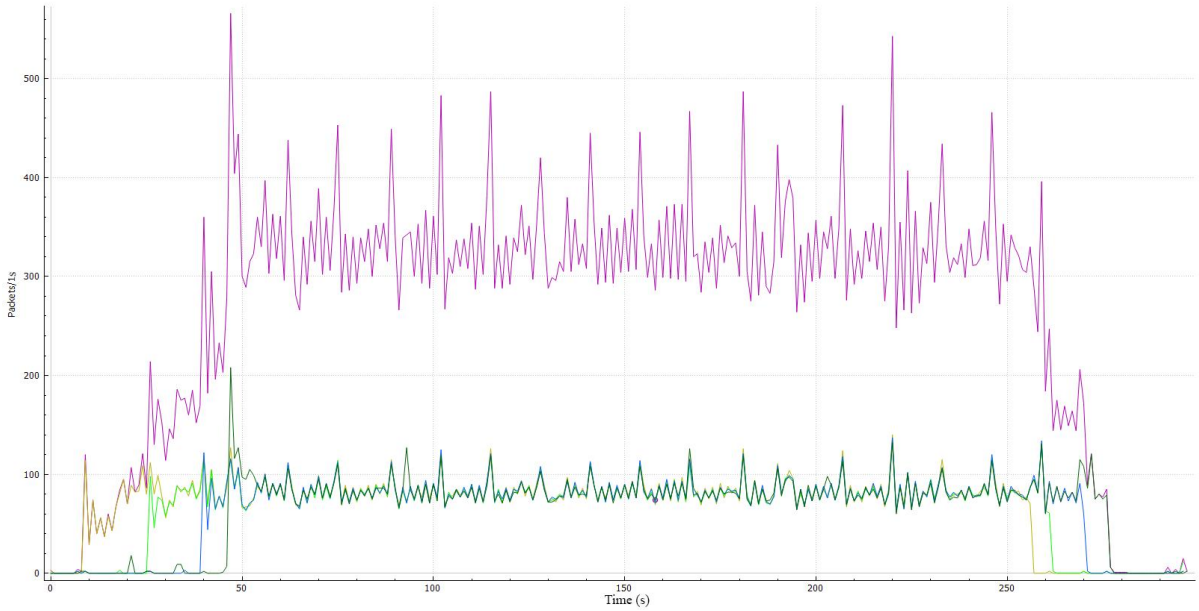


Figure 9. TrueConf, 3 + 1 Clients, Movement, Packets per second

The figures 8 "MVP, 3 + 1 Clients, Movement, Packets per second" and 9 "TrueConf, 3 + 1 Clients, Movement, Packets per second" display the graphs depicting the relationship between the number of received packets per second over time for the two most demanding experiments, one for each of the compared applications. It's immediately evident that the developed application uses the UDP protocol for data transmission, resulting in a "smooth" traffic pattern, while TrueConf uses TCP, as shown in the second graph, resulting in a "choppier" traffic pattern. Additionally, there's a notable difference in the scale of values between the two graphs, indicating that TrueConf utilizes the available bandwidth more aggressively.

The primary achievement of MVP at this stage is its low network resource consumption compared to applications providing 2D video communication services. The developed system not only outperforms in terms of network characteristics but also introduces revolutionary new capabilities for holographic communication services by transmitting control points of a person. There is potential for further development to enable the transmission of complete holograms or a cloud of points representing an entire individual [7].

CONCLUSION

In this paper, we have described the results of a comparative analysis of two conferencing applications: the self-developed MVP application and the commercial product TrueConf. The research has clearly shown that MVP surpasses TrueConf not only in offering more extensive services, including the transmission and display of a 3D model of a person for augmented reality and holography but also in terms of network characteristics.

The key findings of this comparative analysis are as follows: MVP's data transmission efficiency reduces the required amount of transmitted information by a factor of 3 compared to TrueConf. This would indicate a more optimized data transmission if both applications were transmitting video traffic. However, the MVP system was developed with the goal of reducing network load and replacing such traffic with a data stream, which can be particularly valuable in conditions of limited network bandwidth. The number of packets transmitted by MVP is also lower in this comparison, reducing the number of transmitted packets

by 1.3 times. This parameter can be further improved by more efficient utilization of network resources and packet filling. This study touched upon the issue of QoS, but it was not fully explored. Further research should address the question of minimizing the risk of potential jitter during conferencing.

These results underscore the importance of developing and optimizing the next generation of conferencing applications, especially in cases where data transmission and functionality requirements are critical. MVP demonstrates the potential to provide a wider range of services and better network characteristics, which can be beneficial in various fields, including education, healthcare, and business. This development is ready to lead the way for the future of networks in 2030 and the world of HTC.

REFERENCES

1. Akyildiz, Ian F. and Hongzhi Guo. “Holographic-type communication: A new challenge for the next decade.” ITU Journal on Future and Evolving Technologies (2022)

2. Clemm A. et al. Toward truly immersive holographic-type communication: Challenges and solutions //IEEE Communications Magazine. – 2020. – Т. 58. – №. 1. – С. 93-99.

3. Ateya, A. A., Muthanna, A., Gudkova, I., Vybornova, A., & Koucheryavy, A. (2017, July). Intelligent core network for Tactile Internet system. In Proceedings of the international conference on future networks and distributed systems (pp. 1-6).4. Cao C., Preda M., Zaharia T. 3D point cloud compression: A survey //The 24th International Conference on 3D Web Technology. – 2019. – С. 1-9.

5. Volkov, Artem, et al. "Optimized Data Transmission and Signal Processing for Telepresence Suits in Multiverse Interactions." Journal of Sensor and Actuator Networks 13.6 (2024): 82.

6. Kontogianni, E.; Anthopoulos, L. Towards a Standardized Metaverse Definition: Empirical Evidence from the ITU Metaverse Focus Group. IEEE Eng. Manag. Rev. 2024.

7. Cao C., Preda M., Zaharia T. 3D point cloud compression: A survey // The 24th International Conference on 3D Web Technology. 2019. pp. 1-9.

8. Svechnikov, D., Dunaytsev, R., Muthanna, A., & Aziz, A. (2023, December). Holographic Images Delivery Model Toward 6G Telepresence Services. In International Conference on Next Generation Wired/Wireless Networking (pp. 338-345). Cham: Springer Nature Switzerland.

9. Ateya, A. A., Abd El-Latif, A. A., Muthanna, A., Volkov, A., & Koucheryavy, A. (2025). Enabling Metaverse and Telepresence Services in 6G Networks. CRC Press.

Improving Post-Harvest Rice Drying Efficiency through a Low-Cost Halogen Dryer Design for Rural Communities

Supun Sadeepa and Roshan Thilakarathna*
Department Mechanical and Manufacturing Technology, Faculty of Technology, Wayamba University of Sri Lanka.
fot209185@kul.wyb.ac.lk, roshant@wyb.ac.lk

ABSTRACT

Small-scale rice farmers in Sri Lanka often depend on traditional sun-drying methods, which are inefficient, weather-dependent, and contribute to significant post-harvest losses. This research focuses on the conceptual design and evaluation of a low-cost wet rice dryer using halogen lamps as the heat source, aimed at improving drying efficiency prior to milling. Field surveys were conducted to identify the common challenges faced by rural farmers, such as uneven drying, weather interruptions, and grain rejection by millers due to high moisture content. Based on the survey results, key user requirements were identified, including low operating cost, simple structure, and potential for multi-crop drying. A conceptual design was developed accordingly, with a drying chamber and tray system optimized for 1 cm thick rice layers. The full assembly was modeled in 3D using CAD software, allowing for virtual evaluation of airflow, heat source positioning, and accessibility. Finite Element Analysis (FEA) was applied to simulate the mechanical response of the tray under typical loads, confirming its structural soundness. Preliminary thermal experiments were conducted using a controlled test box setup to evaluate the heating performance of a 1000W halogen lamp. The system successfully achieved drying temperatures up to 82°C, suitable for surface moisture reduction. Temperature trends were recorded over time, and manual quality checks showed promising results for further development. These findings indicate the technical feasibility of the design and its potential to improve post-harvest efficiency in rural settings. The study provides a foundation for future stages of prototype fabrication, sensor integration, and field validation.

Keywords: *halogen lamp dryer, low-cost design, post-harvest technology, prototype development, rice moisture reduction, small-scale rice processing, wet rice drying*

INTRODUCTION

Rice *(Oryza sativa)* plays a vital role in the diets of millions of people around the world and is especially important in Sri Lanka as the staple food and a source of livelihood for many farming communities. From cultivation to consumption, rice goes through several stages of processing, and drying is one of the most critical among them. Proper drying ensures the preservation of rice quality, prevents spoilage, and prepares the grain for milling. According to this (Thakur *et al*., 2024). However, for many small-scale farmers, especially in rural areas, access to reliable and efficient drying methods remains a challenge.

In Sri Lanka, it is still common to see farmers relying on open sun drying, where rice is spread on the ground in open spaces. This method is heavily dependent on weather conditions and often results in uneven drying, contamination, and loss of grain due to insects, birds, or rainfall. When rice is not dried adequately, it cannot be milled effectively, as high moisture content can lead to mechanical issues in milling machines, such as clogging, blade damage, or system failure. (Chua *et al*., 2003) According to the mentioned this many millers refuse to process wet rice, forcing farmers to bear significant post-harvest losses.

Recognizing this problem, the aim of this research is to design and develop a low-cost wet rice dryer that can be fabricated using locally available materials and operated by small-scale farmers. The proposed dryer is intended to reduce moisture content in rice before milling, using a controlled heating system based on

halogen lamps. The project also explores ways to improve drying performance through airflow enhancements and proper chamber insulation according Wang *et al*., 2025).

By offering a practical and affordable solution, this research seeks to support rural agricultural communities by minimizing grain loss and improving the overall efficiency of the rice milling process. The work combines design, fabrication, and experimental testing to create a prototype dryer that is both accessible and effective.

METHODOLOGY

The methodology of this research followed a structured process, beginning with a field survey and data collection from local farmers to identify challenges related to drying wet rice under small-scale conditions, as shown in this (Xangsayasane *et al*., 2019). Based on the insights gathered, design considerations were defined, focusing on energy efficiency, low cost, and user-friendly operation according to the mention (Nath *et al*., 2024). A conceptual layout of the dryer was then developed to guide the structural planning and heat source arrangement. This was followed by detailed 3D modeling using CAD software, where each component, such as the frame, heating unit, and tray system, was designed and assembled digitally for compatibility and precision according to this (Mihret *et al*., 2023). Each design and development stage was aligned with practical constraints and existing literature on small-scale post-harvest technologies.

Survey and Data Collection

To initiate the research process, a field survey was conducted targeting small-scale rice farmers and local mill operators. A total of eight subjects were engaged in the survey to gather relevant insights. According to Table 2 can see the sample of the field survey. The primary aim was to understand the practical difficulties faced during the post-harvest phase, especially regarding the drying of wet rice before milling. Data was collected through informal interviews and direct observations in rural farming areas, where it was found that many farmers rely on open sun drying methods, which are weather-dependent and often ineffective (Dissanayaka *et al*., 2020). Most respondents reported that wet rice often leads to machine damage during milling or complete rejection by mill owners, resulting in financial loss and wastage, as explained in this (Introduction of Rubber Roller Milling Machine for Improving the Domestic Milled Rice Quality, n.d). The survey also revealed that the average quantity of rice processed daily ranges between 40 to 50 kg, but individual milling instances are typically around 5 kg per batch. This information was critical in determining the design capacity and layout of the proposed drying system. The feedback collected directly shaped the direction of the conceptual and structural design to ensure the final solution would be practical, affordable, and well-suited to the local context.

Table 1: survey data collection sample

Question point	Description
Subject 01	
1. Rice flour amount per day	20kg - 30kg
2. Price for 1kg of rice flour	80/= - 100/=
3. Method of drying	Dry on a cloth
4. Time for drying the wet rice	30min
5. Challenges faced when drying the wet rice	Weather issues, uneven drying
6. Problems faced while milling wet rice?	Grain breaking, machine clogging

7. How does wet rice threshing affect machine breakdowns during rice threshing?	When milling the wet rice flour
8. If designing a dryer, what features should it include?	Multi-food drying, that is, drying other foods using this dryer.
9. Are you willing to spend money on a dryer?	15000/= - 20000/=
10. How will it affect your business if rice is dried by a machine rather than the rice drying method you are using?	The drying time of wet rice is usually 30 minutes, so its milling efficiency can be increased. If so, this method is a very suitable solution for that.
11. Would you consider using this machine if it could dry other crops or foods other than rice?	That is a good idea for this mechanism.
12. How do you measure the moisture content of the rice you take for drying?	Wet rice is heavier than normal. Also, if you touch wet rice with your hand, the rice will stick to your hand. They use that method to measure moisture.
13. What is the average moisture content of the rice before you dry it?	No
14. Do you monitor the moisture content of the rice after drying it? If yes, how?	Yes, He touched the rice. If you keep the rice in your hand while taking it out of the rice. Then he decides that the rice is dry. Otherwise, the rice is not dry.
15. Who are you primarily producing rice flour for?	For business

Design consideration and Conceptual design

The wet rice dryer was designed with a focus on efficiency, low cost, and ease of use. Controlled heat sources like halogen lamps were selected to ensure even drying, while a moisture monitoring system is planned to avoid over-drying, as explained in this (Chua *et al*., 2003). Food-grade and durable materials were chosen to improve safety and lifespan (Kim, 2013). The overall structure was kept simple to allow easy maintenance and local fabrication, according to reviewed this (Bunyawanichakul, 2006), Figure 1. (a) and (b) illustrate the conceptual design and some main components of the machine, providing a clear understanding of its structure and overall drying process.

3D MODEL DESIGNING

A 3D model of the wet rice dryer was developed to visualize and validate the design before fabrication. Each component, such as the frame, trays, and heating unit, was modeled separately and assembled to ensure proper fit and function. This process helped identify design improvements early, particularly in airflow and heat distribution, according to reviewed this (Bennamoun *et al*., 2003). The final model supports both construction and future optimization.

\

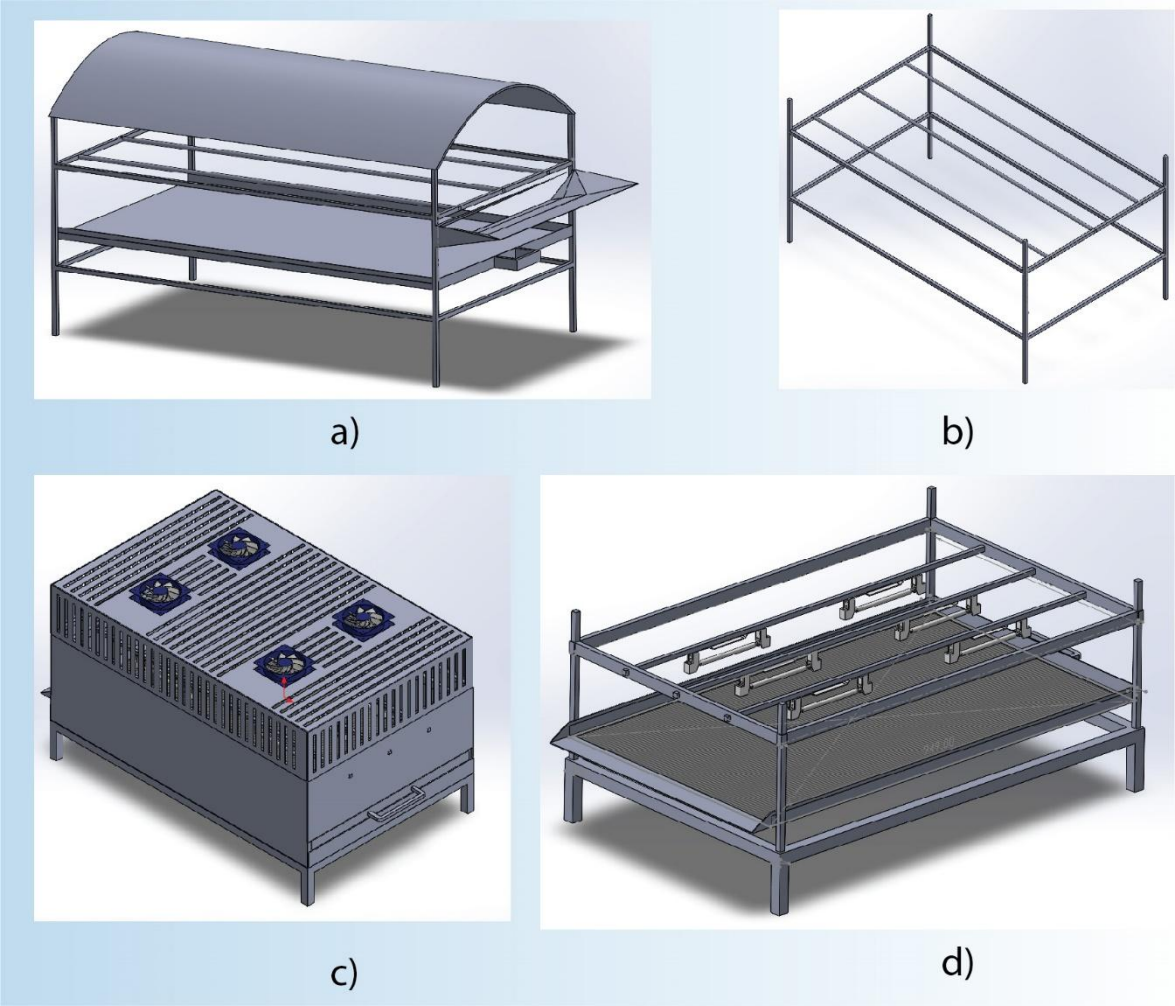


Figure 68.Conceptual sketch and 3D design of the proposed machine. (a) Conceptual design of the dryer; (b) conceptual design of the halogen light bar; (c) 3D design of the final proposed dryer; (d) base part of the dryer with tray and light bars.

According to Figure 1(a) and (b), all major components of the dryer were designed and assembled. The 3D model of the rice dryer was fully completed, with each part, such as the frame, trays, and lamp holders, designed individually to ensure accuracy and compatibility. Once all components were finalized, they were assembled into a complete model to visualize the overall structure and functionality. The design was refined based on insights gained from earlier experiments, allowing for practical improvements in airflow and lamp positioning. Halogen lamps were selected as the heating source due to their high heat output, fast response, and cost-effectiveness, making them suitable for small-scale drying. A power rating of 1000 W per lamp was chosen to achieve rapid heating to the target temperature range while maintaining efficient energy use.

RESULTS & DISCUSSION

Calculation Of Design

For this study,(Pathiraje et al., 2010) BG 352 rice was selected as the preferred variety for rice flour production due to its suitability for industrial milling applications. When evaluating material properties, two key density values were considered: bulk density and true density. The bulk density of rice typically ranges from 0.77 to 0.87 g/cm³, while the true density, which refers to the mass per unit volume of the rice grain itself (excluding air spaces), ranges from 1.17 to 1.41 g/cm³. In contrast, bulk density accounts for the air gaps between grains when stored or handled in bulk. For design calculations in this research, the bulk density value was selected and converted to 800 kg/m³ to suit volume and mass estimations within the drying chamber.

New Tray Area Calculation

1. Volume Needed

$$V = \frac{5kg}{800kg/\mathrm{m3}} \quad (1)$$
$$= 0.00625\, m3$$

2. Tray Area Calculation (for 1.2 cm layer thickness)

$$A = \frac{V}{Layer\ thickness} \quad (2)$$
$$A = \frac{0.00625\, m3}{0.012}$$
$$= 0.520\, m2$$

Possible Tray Dimensions m^3

- 70 cm × 88 cm (0.7 m × 0.9 m)
- 103.6 cm ×0.593 cm (1.036× 0.593 m)

We selected 104 cm × 595 cm (1.04 mm 0.595 mm)

According to Eqs. (1) and (2) did this calculation. Then we can get the total area = 0.614 m^2. This area is enough for the rice in the tray.

Survey result

To ensure the proposed dryer meets the real needs of end users, a field survey was carried out among small-scale rice farmers and mill operators. The objective was to gather practical insights on current drying methods, common challenges, and typical processing capacities. The data collected from this survey served as a foundation for shaping the design requirements and functional goals of the dryer.

Based on the findings from the field survey, several important factors related to small-scale rice drying and milling were identified. The typical drying thickness used by farmers was around 1 cm or less, and the average drying time using traditional sun drying methods was approximately 30 minutes. Respondents reported that the average quantity of rice milled per day was around 50 kg, with an individual consumption of about 5 kg per person per session. Additionally, the average temperature of rice before milling was recorded at approximately 30°C. When discussing affordability, the survey indicated that the average acceptable cost for a locally built drying machine was around LKR 27,000, making low-cost solutions like the proposed dryer both feasible and relevant for the target users.

According to my surveys, I could get a different temperature from dry rice samples. Those I represented

using this bar graph

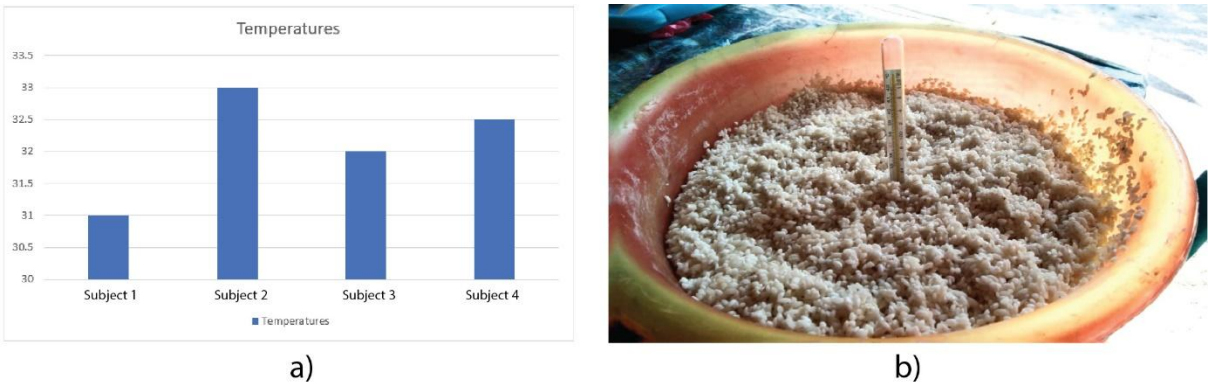


Figure 69. Survey results of temperatures for four selected subjects. (a) Different rice sample temperature levels; (b) method of wet rice temperature measurement.

During the survey, temperature data of wet rice samples were collected from four different locations. It was observed that the temperature varied slightly depending on the environmental conditions of each area, with values ranging between 30°C and 33°C. As shown in Figure 2. (b), the recorded temperatures were used to create a graphical representation, which highlights a maximum temperature of **33°C** and a minimum of 31°C. The measurements were taken using a standard thermometer, as illustrated in Figure 2. (b)

Experiments and testing

I have started conducting experiments for my research to evaluate the performance of halogen lamps in achieving the required drying temperature.

Experiment 01

An experimental setup was prepared to evaluate the heating efficiency of the halogen lamps used in the drying system according to Figure 3. (a). In the first experiment, a 1000W halogen lamp was tested, and the setup successfully achieved a maximum temperature of 82°C. These initial results demonstrated the lamp's capability to generate sufficient heat within a short period. Based on this outcome, further experiments are planned to optimize the heating conditions and determine the most suitable parameters for efficient and uniform rice drying.

Experiment 02

An experimental setup was developed to evaluate the heating efficiency of the halogen lamps used in the dryer design according to Figure 3. (a). In the initial trial, a 1000W halogen lamp was tested, and the system successfully reached a maximum temperature of 80.7°C, indicating strong heat generation potential according to Figure 3. (b). For this experiment, the ambient room temperature was recorded at 32.5°C according to Figure 3. (d). To simulate actual drying conditions, as Figure 3. (c), a 100g rice sample was prepared by soaking it in water for 40 minutes, after which the excess water was drained before placing the sample into the insulated test chamber. These preliminary results will guide future experiments focused on optimizing lamp performance and determining the most effective conditions for moisture removal in the drying process.

During the heating process, temperature measurements were taken at regular intervals to monitor the drying performance. The experiment began by placing the prepared rice sample evenly inside an insulated box, maintaining a uniform thickness of 1 cm across the surface according to figure 3. (e). A 1000W halogen lamp was then activated as the heat source. The initial temperature of the box was recorded at 32.5°C, and over the course of 5 minutes, as shown in figure 4. (b) the temperature gradually increased. By the end of the heating period, the final temperature reached 68.3°C, indicating a significant and consistent rise in thermal energy within the chamber suitable for the rice drying

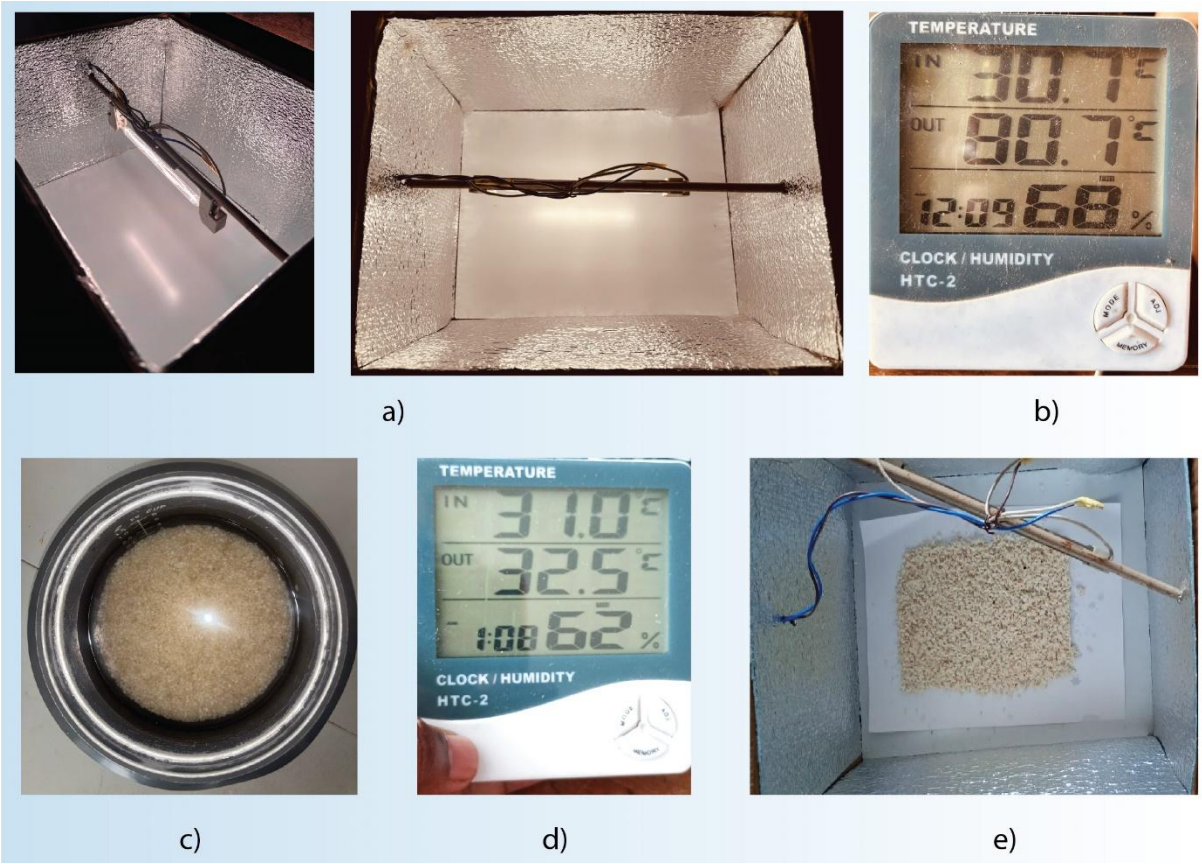


Figure 70. Results of Experiments 1 and 2. (a) Testing the experimental dryer box with different light intensities; (b) experimental box reaching maximum temperature; (c) rice soaking; (d) rice soaking temperature; (e) 1 cm thickness of rice sample during the drying process.

Quality Check

After turning off the halogen lamp, the dried rice sample was manually inspected, as shown in Figure 4. (a) The rice was found to be neither sticky nor wet to the touch, indicating that effective surface drying had been achieved. Based on this manual evaluation, the drying quality was considered satisfactory for the intended application.

Observations & Results

During the experiment, a steady increase in temperature was observed within the first five minutes of operation. The halogen lamp effectively provided sufficient heat for removing surface moisture from the rice sample. No signs of overheating or burning were detected, indicating that the drying process was

controlled and safe under the tested conditions.

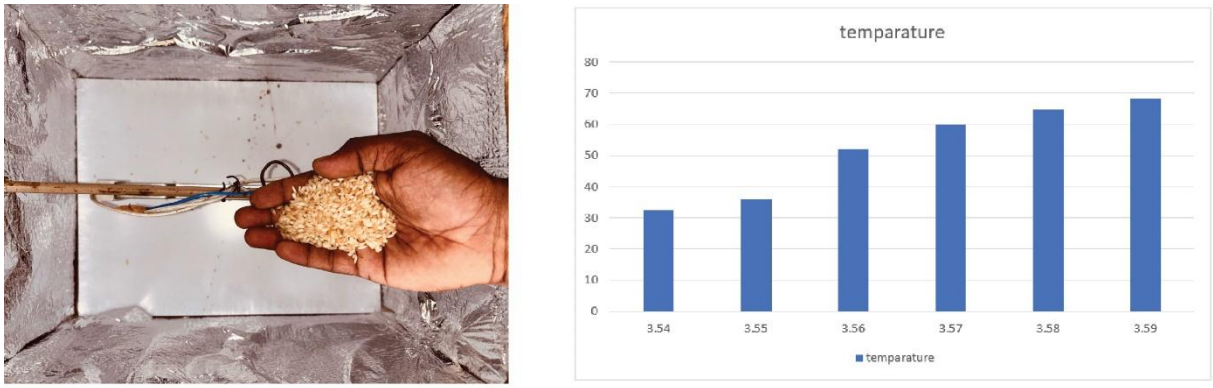


Figure 71.Temperature monitoring results. (a) Inspection of the rice sample after the drying process; (b) temperature changes over time.

ound.. The halogen lamp was then switched on, and the drying process was initiated while monitoring the temperature using a measuring instrument, as shown in *Figure 69*(b). Temperature and time readings were recorded and presented in Figure 4(b). After 5 minutes of drying, the condition of the rice sample was examined, as depicted in *Figure 4*(a).

Phototype and Testing

Phototype Testing 01

This was the first performance test conducted after completing the development of the prototype. The aim was to evaluate the dryer's ability to reduce the moisture content of soaked rice efficiently while monitoring temperature and moisture trends. A 100 g rice sample was soaked in water for 1 hour, resulting in an initial moisture content of approximately 40%. The sample, with an initial temperature of 29 °C, was evenly spread on the drying tray *Figure 5*(e). Halogen lamps served as the heat source, and temperature and moisture content were continuously measured using a thermocouple and moisture sensor connected to an Arduino Nano. Data were logged in real time to Excel for analysis.

The drying process lasted 18 minutes, during which the temperature increased steadily to 60 °C and the moisture content decreased to an expected 14%. Manual quality checks confirmed that the rice was dry and free from surface moisture *Figure 5*(b), and the final moisture content was recorded *Figure 5*(c). The complete prototype used in this test is shown in *Figure 5*(d). The graphical data demonstrated a consistent temperature rise accompanied by a corresponding decrease in moisture, confirming the dryer's effective performance.

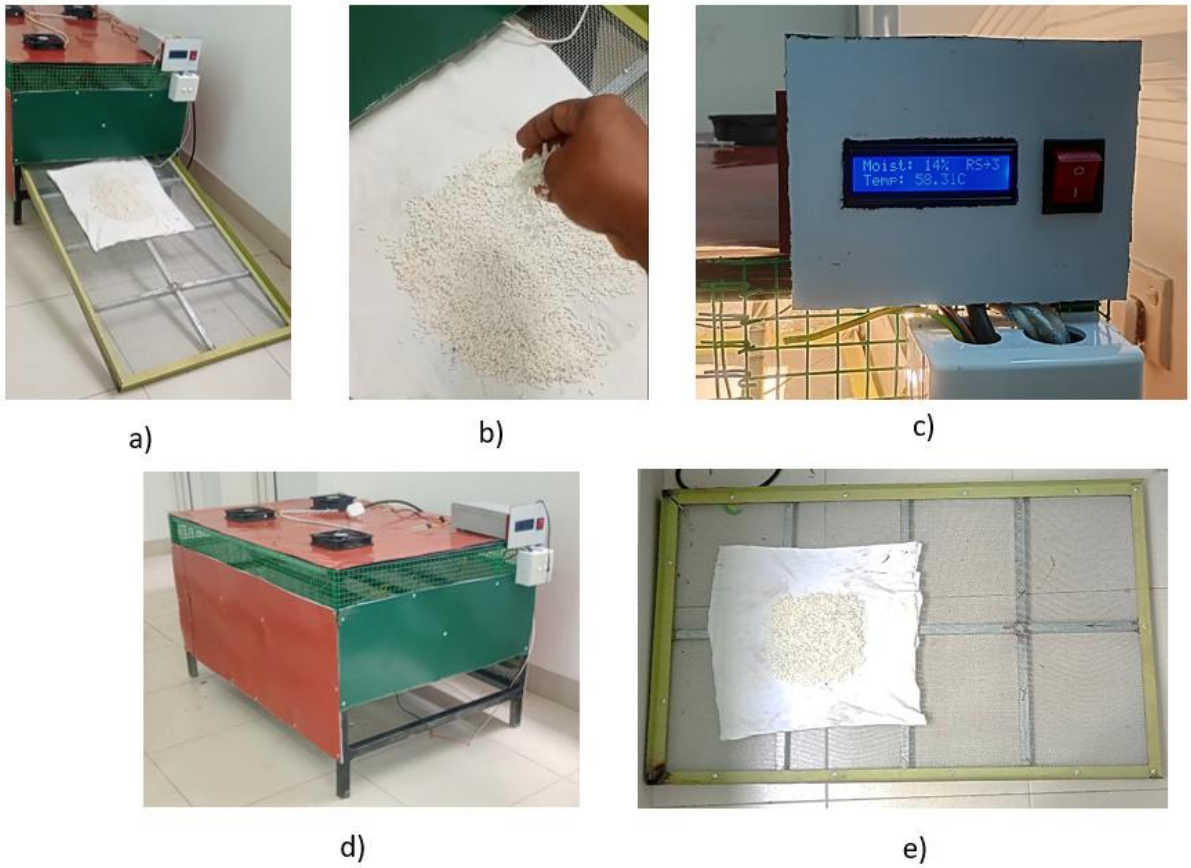


Figure 72.Testing the rice sample on the phototype dryer. (a) Rice sample on the drying chamber; (b) Quality checking the rice sample;(c) Final moisture content level after drying; (d) Final phototype of dryer; (e) spreading rice sample on dryer

r's operation under full-load conditions. In this test, a 5 kg rice sample was soaked in water for 2 hours *Figure 6*(d). After soaking, the water was drained, and the rice was evenly distributed on the drying tray *Figure 6*(c). The initial temperature and moisture content were recorded as 29 °C and 42%, respectively.

The drying process was initiated using four halogen lamps as the heat source. Temperature and moisture content were continuously monitored using a thermocouple and a moisture sensor connected to the Arduino Nano, with real-time data logging to Excel. The process continued until the moisture content reached the target range of 14%.

The total drying time recorded was 42 minutes, during which the temperature rose steadily while the moisture content decreased, as shown in *Figure 6*(b). A manual quality check was conducted at the end of the process, *Figure 6*(a), confirming that the rice was dry, non-sticky, and suitable for milling. This test demonstrated that the dryer maintained effective performance and drying uniformity even under maximum load conditions.

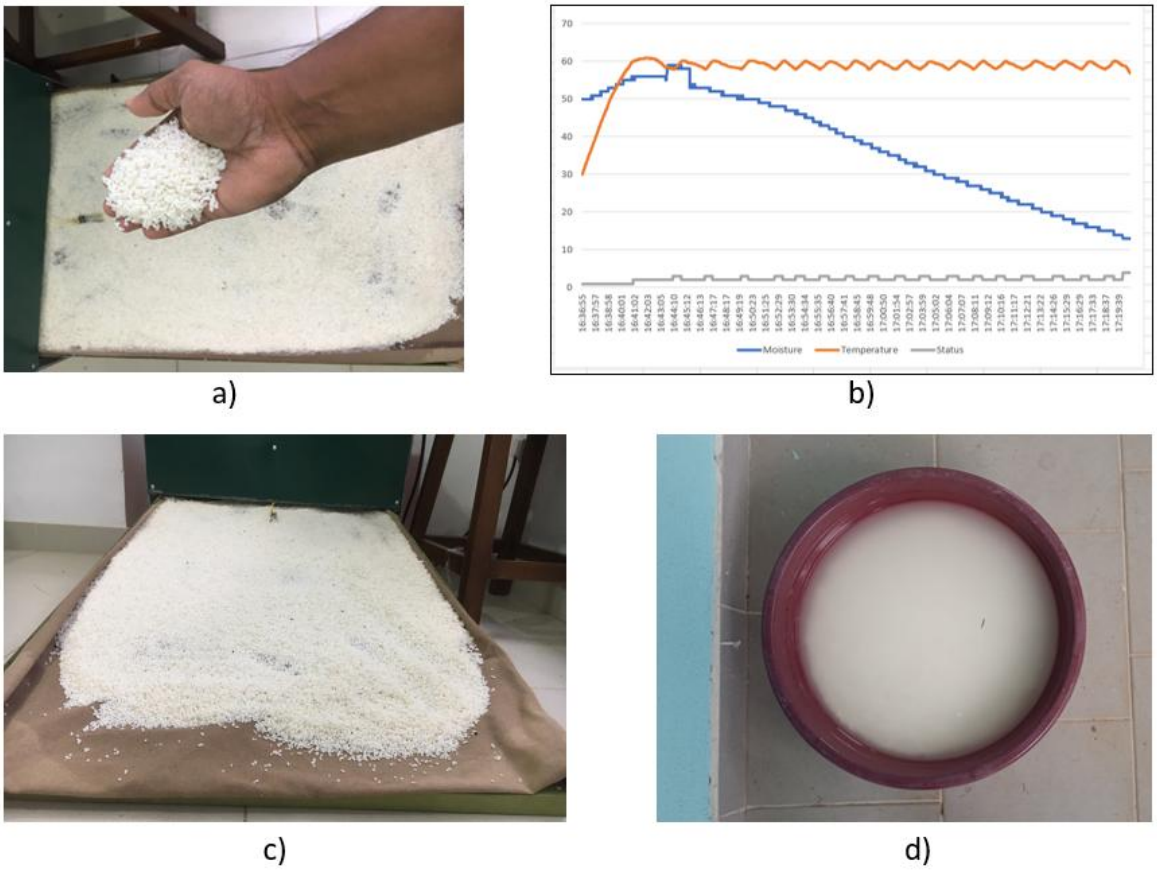


Figure 73.Testing the rice sample of maximum capacity (a) checking the drying quality of rice sample; (b) Temperature vs moisture content graph; (c) Place the rice on drying tray; (d) soaking rice sample

Simulation of Drying Tray

To ensure the structural integrity of the drying tray, a simulation was carried out using **SolidWorks simulation tools**. The tray was modeled using **steel material**, with a **50 N force** applied to simulate the load during drying, and fixed supports were applied at the rail positions to reflect actual mounting conditions. As illustrated in **Figure 7(a)**, the setup accurately represented the tray's behavior under stress. The simulation produced a **maximum stress of 98,530 N/m²**, which is well below the steel's yield strength, confirming the structural safety of the design. The **maximum displacement** observed was **0.275 mm**, and the **strain** was minimal. The detailed results are shown in the **displacement diagram Figure 7(b)**, **strain distribution diagram Figure 7(c)**, and **stress diagram Figure 7(d)**. These results confirm that the tray design is both stable and mechanically suitable for use in the drying system.

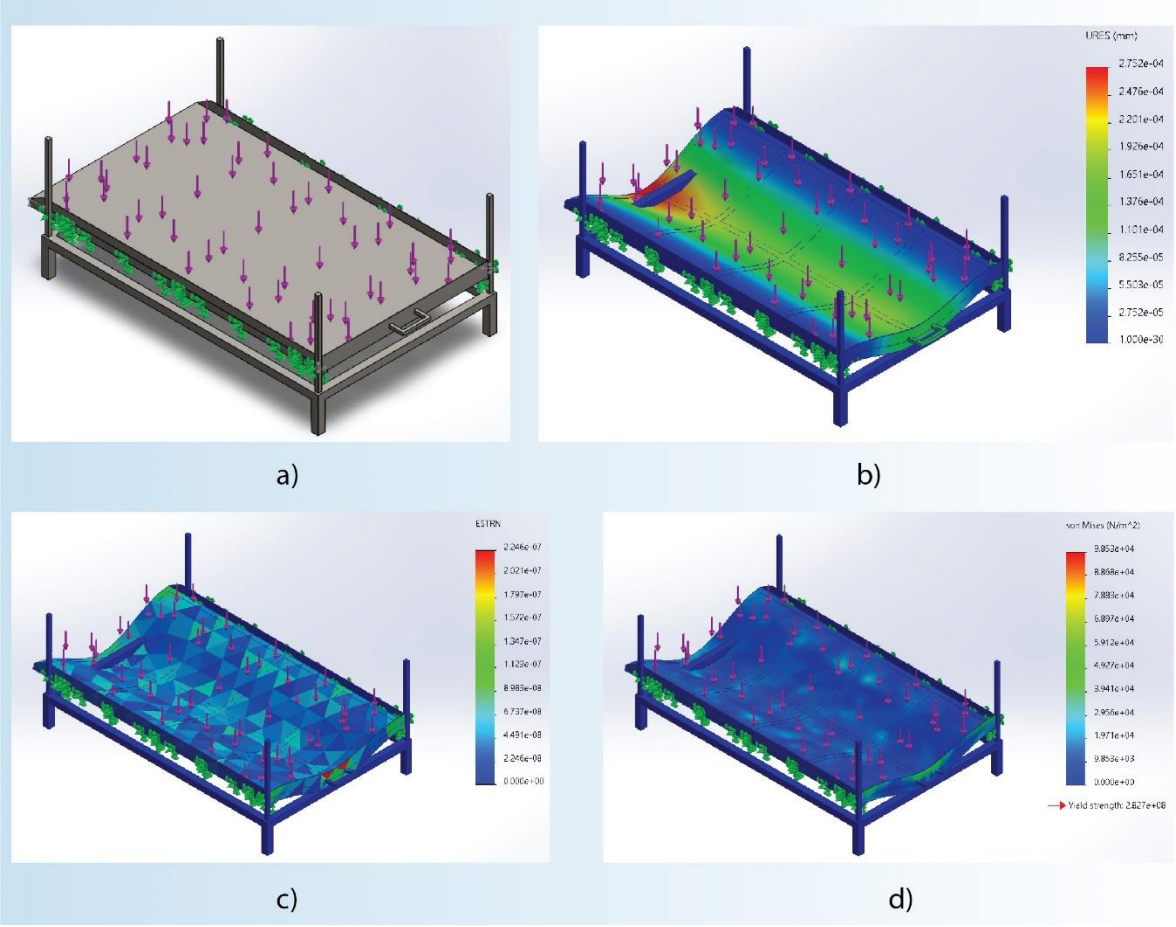


Figure 74. Finite Element Analysis (FEA) of the dryer tray. (a) Force distribution on the tray; (b) displacement results; (c) strain results; (d) stress results.

CONCLUSION & FUTURE WORK

This study presented the design and development of a cost-effective wet rice dryer to address the practical issues faced by small-scale rice processors. The system was designed with simplicity, affordability, and functionality in mind. It uses halogen lamps as the heat source and includes a manually accessible tray system to ensure ease of use. Field surveys confirmed the need for such a device, especially in rural areas where weather-dependent drying methods often lead to grain loss and reduced milling efficiency.

Initial experiments showed that the prototype was capable of reaching the required drying temperatures in a short time, and the drying quality met acceptable standards through manual evaluation. Based on these outcomes, further enhancements are planned, including the optimization of drying efficiency and an increase in the drying capacity to meet larger processing demands. Additionally, the current quality assessment method after drying relies on manual hand evaluation; future work will focus on developing an objective, measurable quality detection system. Other improvements will include adding controlled airflow using exhaust fans and integrating moisture sensors to automate the drying process more precisely. Additional testing with larger samples will also be carried out to evaluate performance and long-term

reliability. Overall, this system shows promise as a practical and affordable solution for improving post-harvest rice handling at the small-scale level.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of all lecturers and laboratory technicians of the Department of Mechanical and Manufacturing Technology, Faculty of Technology, Wayamba University of Sri Lanka, in completing this work successfully. Furthermore, the authors express their sincere gratitude to the editors and anonymous reviewers for their insightful suggestions and constructive comments, which have greatly improved the quality of this research paper.

REFERENCES

Bennamoun, L., & Belhamri, A. (2003). Design and simulation of a solar dryer for agriculture products. *Journal of Food Engineering*, *59*(2–3), 259–266. doi: 10.1016/S0260-8774(02)00466-1

Bunyawanichakul, P. (2006). *DEVELOPMENT OF A CYCLONE RICE DRYER*.

Chua, K. J., & Chou, S. K. (2003). Low-cost drying methods for developing countries. In Trends in Food Science and Technology (Vol. 14, Issue 12, pp. 519–528). Elsevier Ltd. doi: 10.1016/j.tifs.2003.07.003

Dissanayaka, D. M. S. K., & Wijayaratne, L. K. W. (2020). Common Agricultural Practices and Post-Harvest Losses in Trincomalee, Sri Lanka. In Journal of Tropical Forestry and Environment (Vol. 10, Issue 02).

Introduction of Rubber Roller Milling Machine for Improving the Domestic Milled Rice Quality. (n.d.). Retrieved from https://www.researchgate.net/publication/380929315

Kim, M. H. (2013). Review on Rice Flour Manufacturing and Utilization. *Journal of Biosystems Engineering*, *38*(2), 103–112. doi: 10.5307/jbe.2013.38.2.103

Mihret, Y. C., Delele, M. A., & Hailemesikel, S. T. (2023). Design, development, and testing of rice-husk fueled mixed-flow rice dryer for small-scale rice producer farmers. *Heliyon*, *9*(7). doi: 10.1016/j.heliyon.2023.e18077

Nath, B., Chen, G., O'Sullivan, C. M., & Zare, D. (2024). Research and Technologies to Reduce Grain Postharvest Losses: A Review. In Foods (Vol. 13, Issue 12). Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/foods13121875

Pathiraje, P. M. H. D., Madhujith, W. M. T., Chandrasekara, A., & Nissanka, S. P. (2010). The Effect of Rice Variety and Parboiling on in vivo Glycemic Response. In Tropical Agricultural Research (Vol. 22, Issue 1).

Thakur, A., Singh, A., & Bharat, N. K. (2024). Conservation and development of rice germplasm for natural farming. In Climate-Smart Rice Breeding (pp. 45–61). Springer Nature. doi: 10.1007/978-981-97-7098-4_3

Wang, Y. J., Li, Y. F., Pan, S., Qin, M. Y., Yuan, Y. D., Yuan, W. J., & Li, C. (2025). Rice drying quality using gas-catalytic infrared equipment: an experimental study. *IOP Conference Series: Earth and Environmental Science*, *1500*(1), 012047. doi: 10.1088/1755-1315/1500/1/012047

Xangsayasane, P., Vongxayya, K., Phongchanmisai, S., Mitchell, J., & Fukai, S. (2019). Rice milling quality as affected by drying method and harvesting time during ripening in wet and dry seasons. *Plant Production Science*, *22*(1), 98–106. doi: 10.1080/1343943X.2018.1544463

Macaranga peltata Leaf Extract Mediated Green Synthesis of Iron Nanoparticles and Their Application in Organic Dye Removal

Dissanayake D.M.K.N.[a], Perera M.A.D.[a], Karunaratne M.S.A.[a], Pahalagedara M.N.[b]

[a] Department of Materials Engineering, Faculty of Engineering, Sri Lanka Institute of Information Technology, Malabe 10115, Sri Lanka.
[b] Department of Basic Sciences, Faculty of Allied Health Sciences, University of Sri Jayewardenepura, Nugegoda 10250, Sri Lanka
kasuni0620@gmail.com, ann000461@gmail.com, mudith.k@sliit.lk, madhavi.pahalagedara@gmail.com

ABSTRACT
This study presents an eco-friendly method for synthesizing iron nanoparticles (FeNPs) using *Macaranga peltata* leaf extract, evaluating their potential in degrading the organic dye methyl orange (MO). The synthesis exploits phytochemicals in the leaf extract as natural reducing and stabilizing agents. The synthesized FeNPs were characterized using UV-Vis spectroscopy, FTIR, XRD, and SEM, confirming amorphous structure and particle sizes ranging from 34–94 nm. Catalytic activity was evaluated via MO degradation experiments, achieving 85.16% efficiency within 200 minutes. The study demonstrates a sustainable solution for wastewater treatment through green nanotechnology.

KEYWORDS: *Green synthesis, Iron nanoparticles, Macaranga peltata, Methyl Orange, Dye degradation, Wastewater treatment.*

INTRODUCTION
Nanotechnology has made revolutionary changes in many fields, such as medicine, electronics, and environmental science by introducing advanced material properties at the nanoscale (ElHelaly, 2013; Yusuf, 2019). Among the broad applications of nanotechnology, the treatment of environmental pollutants is one of the most promising fields where nanoparticles have already shown their outstanding capability in degrading contaminants (Ashraf et al., 2018).

The synthesis methods of nanoparticles can be divided into three main approaches: physical, chemical, and biological. However, most of the conventional synthesis methods such as chemical and physical methods involve the use of toxic chemicals and energy-intensive processes, raising concerns about their overall environmental impact (Jiang et al., 2022). In recent years, green synthesis has emerged as a sustainable, environmentally friendly alternative to produce nanoparticles. This method employs biological entities, like plant extracts, bacteria, or fungi, for nanoparticle synthesis, thereby negating the use of toxic chemicals. Among these, plant extracts have gained a lot of interest due to the abundance of bioactive compounds present in them such as ketones, aldehydes, polyphenols, caffeine, isoverbascosides and carbohydrates, serving as natural reducing and stabilizing agents to convert the metal ions into nanoparticles in the synthesis process, simplicity in preparation, and cost-effectiveness (Ebrahiminezhad et al., 2018).

Plant-Mediated Green Synthesis of Nanoparticles
There are significant studies that impress various plants involved in the process, the sizes of the nanoparticles synthesized, and their applications. A prominent example of plant-mediated synthesis involves *Cupressus sempervirens* to produce zero valent iron nanoparticles. Ebrahiminezhad et al. (2018) reported the formation nanoclusters with diameters from 9-31 nm, synthesized using Mediterranean cypress *(Cupressus sempervirens)* leaf branches. The reaction between the iron solution and the leaf extract was observed by noticeable color change indicating the high reduction potential of phytochemicals in this plant extract. This study highlighted iron nanoparticles' potential for dye removal from aqueous environments. The bark extract of *Pinus eldarica* has also been used in the green synthesis of silver nanoparticles, which

has been reported by Iravani and Zolfaghari (2013). This study underscored *Pinus eldarica* bark extract contains polyphenolic compounds characterized by hydroxyl and carboxyl groups making it an excellent candidate for nanoparticle synthesis. It has been revealed that the exposure of *Pinus eldarica* bark extract to Ag+ ions resulted in a color change in the reaction mixture from yellowish brown to dark brown, indicating the successful formation of silver nanoparticles. In this study, the effect of pH on the size and morphology of nanoparticles was highlighted, and it was shown that at lower pH values, larger nanoparticles were synthesized. In another study, Ahmed and Mustafa (2019) published a wide-ranging review of many studies involving the use of lemongrass *(Cymbopogon flexuosus)* and green tea *(Camellia sinensis)* in the synthesis of gold and silver nanoparticles. The study identified various phyto-constituents in these plants including tannins, terpenoids, flavonoids, ketones, aldehydes, amides, and carboxylic acids which are responsible for reduction of silver ions. Among these, tannins were found to play a crucial role in both the reduction and capping of silver nanoparticles.

Overview of Macaranga peltata

Macaranga peltata belongs to the family Euphorbiaceae, is one of the common pioneer arboreal species that is mainly distributed in Southeast Asia, particularly in regions such as India and Sri Lanka. This species is more popularly known as "kenda" in Sri Lankan and "chandada" in India. It may reach up to 10 meters in height exhibiting a resinous tree structure with its young parts covered with velvety hairs. Its leaves are alternately arranged, measuring between 20 to 50 centimeters in length and 12 to 21 centimeters in width, characterized by a broad ovate shape and a distinct palmately 9-nerved pattern. The flowering period occurs in the months of January to February, with yellow-green flowers clustered in long panicles within the axils of the leaves, and the fruit is a spherical capsule that carries one single black seed (Megha M et al., 2024).



Figure 75. *Macaranga peltata* leaf

A review of the phytochemical profile of *Macaranga peltata* shows that it is highly composed of secondary metabolites such as flavonoids, stilbenes, tannins, and terpenes, isolated from leaves. Preliminary phytochemical screenings have revealed the presence of some bioactive compounds like carbohydrates, alkaloids, phenolic compounds, tannins, saponins, and steroids in ethanolic extracts of the plant (Megha M et al., 2024). These bioactive compounds have been attributed to several pharmacological activities like antimicrobial, antioxidant, anti-inflammatory, anticancer, antifungal action, and other properties such as wound healing and hepatoprotection (Megha M et al., 2024). The potential therapeutic applications of these bioactive compounds point out the importance of *Macaranga peltata* in traditional medicine and its prospective application in modern therapeutic practice (Megha M et al., 2024). In addition, recent studies have undertaken the potential use of the extracts from *Macaranga peltat*a for the green synthesis of nanoparticles by capitalizing on its bioactive compounds.

Overview of Iron Nanoparticles (FeNPs)

Over the last few years, great attention has been paid to iron nanoparticles (FeNPs) because of their unique properties and huge application potential in biomedicine, environmental remediation, and catalysis. There are different synthesis methods of FeNPs that include chemical reduction, various physical methods, and green synthesis using plant extracts. Among the outstanding methods for synthesizing iron nanoparticles

includes chemical reduction, wherein iron ions (Fe^{3+}) are reduced to zero-valent iron (Fe^{0}) using reducing agents like sodium borohydride ($NaBH_4$). This method has demonstrated that the size of the synthesized nanoparticles is around 100 nm, and it was indicated by the black color formation during synthesis-evidence of iron nanoparticles (Abdeen et al., 2013).

However, green synthesis methods have emerged as an environmentally friendly alternative to produce iron nanoparticles. This method represents a way of using natural extracts derived from plants, which are rich in various phytochemicals that tend to reduce iron ions to form nanoparticles and stabilize them. For example, green synthesis involves *Cupressus sempervirens* to produce zero valent iron nanoparticles. Ebrahiminezhad et al. (2018) reported the formation nanoclusters with diameters from 9-31 nm, synthesized using Mediterranean cypress *(Cupressus sempervirens)* leaf branches. The reaction between the iron solution and the leaf extract was observed by noticeable color change indicating the high reduction potential of phytochemicals in this plant extract. This study highlighted iron nanoparticles' potential for dye removal from aqueous environments. In another study, it was found that green tea leaf extract (GT-Fe NPs) which contains polyphenols acts as reducing and capping agent mediated the synthesis of iron nanoparticles within the size range of 40-60 nm. Fe NPs were identified by the formation of strong black precipitate during the synthesis process. The synthesized particles showed very good removable capabilities and therefore, potential application in the field of environmental remediation (Shahwan et al., 2011).

Properties of Iron Nanoparticles Relevant to Dye Degradation

Some physiochemical properties of iron nanoparticles (Fe NPs) were identified by Wanakai et al. (2019). These nanoparticles' properties had enabled them to act efficiently as catalysts in the degradation of dyes and especially methylene blue. That study discussed green synthesis, high surface area, oxidation states, fenton and fenton-like mechanisms and catalytic efficiency. The iron nanoparticles were synthesized using aqueous leaf extracts from the plants such as *Galinsoga parviflora, Conyza bonariensis* and *Bidens pilosa*, according to the study was done by Wanakai et al. (2019). This method is ecofriendly and leads to the synthesis of nanoparticles that are non-toxic and stable also which is quite essential when used for the degradations of dyes. Also, the study emphasized the high surface area. Due to the high surface area to volume ratio and iron nanoparticles are good catalysts which reflect the ability to interact with dye molecules more effectively which enhances their degradation. The oxidation states which were noted in the study also fall under the properties of iron nanoparticles relevant to dye degradation. It was emphasized that Iron is one of the versatile in mechanical engineering, which exists in more valencies/oxidation states. Those are ferrous and ferric. This characteristic is rather important since it determines the degree of effectiveness of the nanoparticles in catalyzing the reactions with reference to the use of hydrogen peroxide in the degradation process. Furthermore, Wanakai et al. (2019) has reported that the fenton and fenton-like mechanisms as one of the properties. The study mentioned that methylene blue degradation takes place due to the Fenton and Fenton-like reactions. Fenton reaction is the reaction between Fe ions and H_2O_2 and yields hydroxyl radical (-OH). These radicals are highly reactive and therefore they have an important function in the cleaving of the dye molecules (Wanakai et al., 2019).

Application of Fe NPs in Organic Dye Removal

As a response to the challenge of dye contamination, various remediation strategies have been contrived. Notably, among all the remediation methods, the incorporation of nanomaterials has come up as a promising approach. Nanomaterials have a large surface area-to-volume ratio and unique chemical properties that make them very effective in the adsorption and degradation of dye molecules in wastewater (Ruan et al., 2018a). For example, it has been demonstrated that nano iron particles promote the decolorization of mono azo dyes via advanced oxidation processes that produce reactive species that can degrade dye structures (Raman & Kanmani, 2018). The green synthesis of iron nanoparticles (Fe NPs) has been given significant attention, owing to their environmentally friendly nature and potential for application in wastewater treatment, especially in dye removal. Several plant extracts have been used in the biosynthesis of Fe NPs, which then effectively show the degradation of various dyes such as Methylene Blue (MB), Methyl Orange (MO), and Malachite Green (MG). This synthesis route provides not only a green alternative to chemical

methods but also enables fine-tuning of the catalytic properties of the nanoparticles so that they could be appropriate for various environmental remediation applications.
Giun Tan et al. (2018) have researched the preparation of iron oxide nanoparticles using plantain peel extracts and found the particles to be very effective in the removal of both Methylene Blue and Methyl Orange dyes. The study focused on the catalytic and magnetic properties of green synthesized iron oxide nanoparticles (G-FeONPs) that showed efficacy in wastewater treatment. The percentages of degradation efficiency reported for Methylene Blue and Methyl Orange are quite high, indicating that G-FeONPs have the potential for use in practical applications. In a similar vein, Ebrahiminezhad et al. (2018) have studied the green synthesis of iron nanoparticles using *Cupressus sempervirens* that have yielded iron nanoclusters with a mean diameter of 19 nm. Their work was focused on optimization conditions for the removal of Methyl Orange and their results showed that the green synthesized iron nanoclusters showed a 95% decolorization efficiency in 6 h time. The results showed the nanoparticles effectively degraded Methyl Orange, and therefore, the potential in practical applications in dye removal during wastewater treatment applications. Shahwan et al. (2011) extended their knowledge on Fe NPs produced from green tea leaf extract, demonstrating their role in Fenton-like catalysts in the degradation of dyes. This study has shown a rapid degradation of Methylene Blue and Methyl Orange, in which kinetic analysis suggested the degradation process of Methylene Blue fitted a second-order reaction, while the degradation of Methyl Orange fitted a first-order model. In another study, Khoirotin et al. (2023) investigated the green synthesis of Fe_3O_4 nanoparticles employing green betel *(Piper betle L.)* leaf extract, showing ferromagnetic properties and efficient adsorption ability. This study identified the most effective adsorption ability of Fe_3O_4 against methylene blue at a contact time 120 minutes with 98.75% MB degradation.
Most previous studies have focused on different metal nanoparticles synthesized using different plant extracts and their application in environmental remediation, especially in the removal of dyes. Although the application of green-synthesized nanoparticles in organic dye removal is widely published, the specific role of iron nanoparticles synthesized through *Macaranga peltata* leaf extract has remained underexplored. The given gap in current literature provides an opportunity to investigate the effectiveness and possible mechanisms through which iron nanoparticles, derived from this plant extract, may enhance the removal of organic dyes, hence address environmental pollution and satisfying the demand for sustainable materials in the domain of nanotechnology.

METHODOLOGY

Preparation of *Macaranga peltata* leaf extraction

As an initial step, the leaves of *Macaranga peltata* were collected and washed multiple times to remove any dirt and unwanted materials. Then, 10 g of *Macaranga peltata* leaves were measured using an analytical balance. After that, the leaves were cut into small pieces and chopped them using a motor and pestle and mixed them with 50ml of distilled water. Next, the sample was stirred at 80 ºC temperature for 30 minutes to extract bioactive compounds efficiently. After 30 minutes, the flask was taken out from the hot plate and kept for a few minutes until it came to room temperature. Finally, the resultant extract was filtered using Whatman filter papers to remove leaf debris.

Green Synthesis of Iron Nanoparticles

Green synthesis was done using *Macaranga peltata* leaves extract by adding 10ml *Macaranga peltata* leaf extract to 5ml of 0.1 M ferric chloride solution in 2:1 (v/v) ratio in a conical flask. Then, the solution with ferric chloride was stirred using a mechanical stirrer for 3 hours while adding the extract of leaves drop wisely. In this time, there was an immediate color change after adding the leaf extract to the iron chloride solution. A color change from pale brown to blackish brown was observed indicating the formation of iron nanoparticles (Fe NPs). Then, the solution undisturbed at room temperature for 24 hours to allow the nanoparticles to settle at the bottom of the beaker. After 24 hours, synthesized nanoparticles were separated

by using the centrifuge at 5000 rpm for 15 minutes. Then the separated nanoparticle sample was washed with water two times and accordingly washed with ethanol by one time. Finally, the synthesized nanoparticle sample was dried at room temperature and stored in a sealed container for subsequent analysis.

Kinetic Experiment of the Degradation of Methyl Orange Dye

First, the calibration curve for methyl orange dye was prepared. Then, a 50 mg/L solution of methyl orange was prepared as a model organic dye. 2 mL of dye solution was withdrawn from the 50 mg/L dye solution and diluted it up to 10 mL by adding water into a 10 mL volumetric flask. This solution was transferred to a sample vial to record initial UV-Vis data at t=0. 50 mg of FeNPs were added to 45 mL of 50 mg/L dye solution, along with 5 mL of 10% H_2O_2. The mixture was kept on a magnetic stirrer at room temperature and stirred to ensure proper mixing. 2mL samples were withdrawn from the mixture at time intervals of 20, 40, 60, 80, 100, 120, 140, 160, 180, 200 minutes and recorded the absorbance at 464 nm (characteristic wavelength for methyl orange) using UV-Vis spectrophotometer. The same procedure was followed without adding Fe NPs, serving as a control experiment to investigate the degradation efficiency of Fe NPs.

The dye degradation efficiency was calculated using the following equation.

$$Degradation\ Efficiency\ (\%) = \frac{C_i - C_f}{C_i} \times 100\%$$

In this equation, C_i denotes initial concentration of dye solution, and C_f denotes concentration of dye solution after time t.

Characterization of Synthesized Nanoparticles

1. Ultraviolet-Visible (UV-Vis) Spectroscopy

Periodic optical absorbance scans were conducted to monitor the formation of nanoparticles and confirm their stability.

Absorbances were taken during the experiment in the UV range of 200-800 nm using SHIMADZU UV-1900 UV-Visible spectrophotometer.

2. Fourier Transform Infrared (FT-IR) Spectroscopy

The infrared spectra of samples were analyzed by Bruker Vertex 80 Fourier Transform Infrared Spectrometer (Bruker alpha (10031705) with specifications of scan range 4000-650 cm-1; 32 scans and 8 cm-1 resolution).

FTIR analysis was used to identify functional groups present in the *Macaranga peltata* leaf extract responsible for reducing ferric irons to iron nanoparticles (FeNPs), elucidating the interaction between the leaf extract phytochemicals and synthesized nanoparticles.

3. Powder X-ray Diffraction (XRD)

The crystalline nature and phase composition of the synthesized FeNPs were examined using Bruker D8 Focus X-ray Diffraction Spectrometer.

4. Scanning Electron Microscopy (SEM)

SEM analysis was carried out using Hitachi SU6600 Scanning Electron Microscope to analyze the surface morphology, size and distribution of the synthesized nanoparticles.

RESULTS & DISCUSSION

UV-Visible Spectroscopy Results

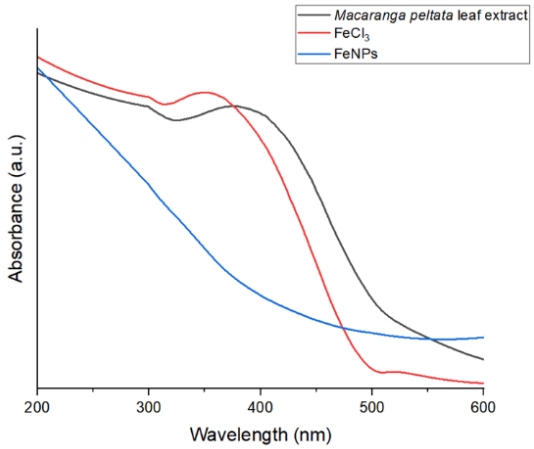


Figure 76. UV-Visible absorption spectra of a) *Macaranga peltata* leaf extract, b) $FeCl_3$ solution c) Synthesized FeNPs

Figure 2 gives the UV-vis spectra of a) $FeCl_3$, b) *Macaranga peltata* leaf extract and c) Fe nanoparticles in the 200-600 nm range. The formation of NPs was indicated by the color change of pale brown to blackish brown which occurred after the addition of the leaf extract to $FeCl_3$ solution. Leaf extract shows peaks at 280 and 450 nm while $FeCl_3$ shows a peak around 350 nm. The UV-vis spectra of the Fe NPs showed a continuous absorption from 200-600 nm. This well agrees with literature reported UV-vis spectra of amorphous Fe NPs.

FTIR Results

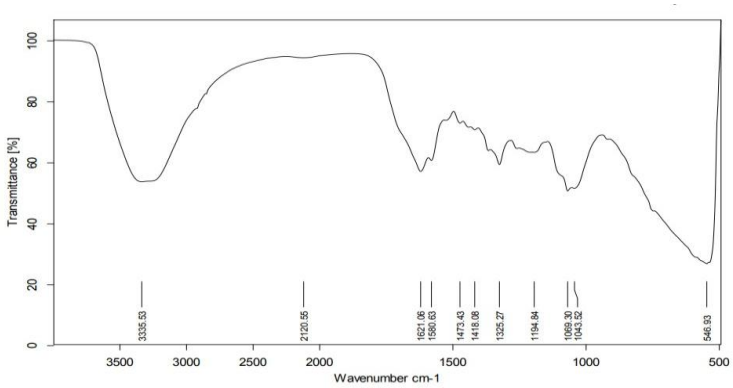


Figure 77. FTIR spectrum of Fe NPs synthesized using *Macaranga peltata* leaf extract

Figure 3 represents the FTIR spectrum of the prepared Fe NPs, indicating presence of organic compounds from *Macaranga peltata* leaf extract. This can be used to identify the functional groups of the phytochemicals responsible for the reduction of the metal precursor. It shows a prominent band at 3335 cm^{-1} due to intramolecular hydrogen bonded OH groups. In addition, the peaks at 1621 cm^{-1}, 1418 cm^{-1}, and 1100 cm^{-1} - 1030 cm^{-1} can be attributed to the C=O, in plane bending vibrations of O-H, and C-O-C stretching respectively.

XRD Results

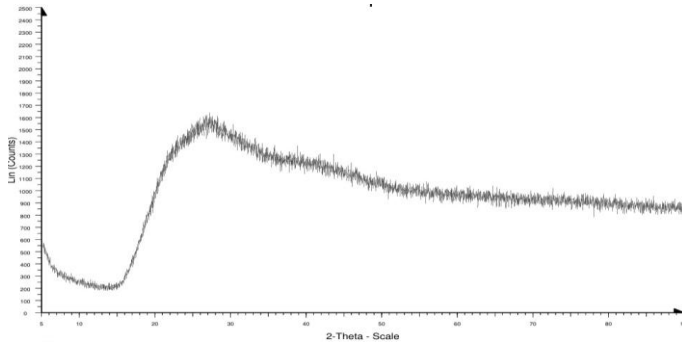


Figure 78. XRD pattern of Fe NPs synthesized using *Macaranga peltata* leaf extract

Figure 4 shows that the X-ray Diffraction of the synthesized iron nanoparticles using *Macaranga peltata* leaf extract. The pattern is deficient in distinct diffraction peaks, suggesting that the iron nanoparticles are amorphous in nature. The broad hump present at about $2\theta = 25^0$ is attributed to organic materials from *Macaranga peltata* leaf extract which can act as reducing and capping agents absorbed on the Fe NP surface.

SEM Results

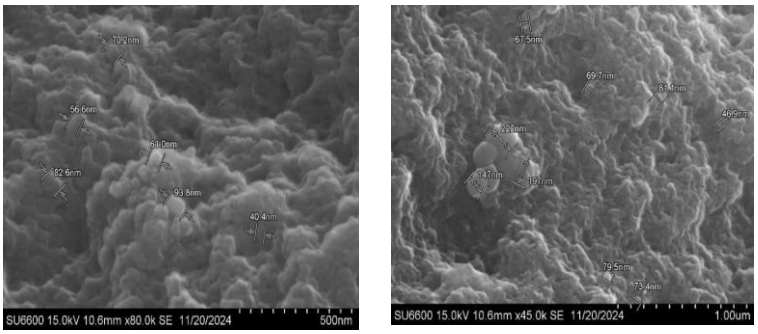


Figure 79. FE SEM Images of Fe NPs synthesized using *Macaranga peltata* leaf extract

The FE-SEM images of Fe NPs are shown in Figure 5. These images suggest that the particles agglomerate to form irregular clusters. This is due to the capping of Fe NPs with phytochemicals present in *Macaranga peltata* leaf extract. The diameter of these clusters' ranges from 34-94 nm.

Preparation of the Calibration Curve

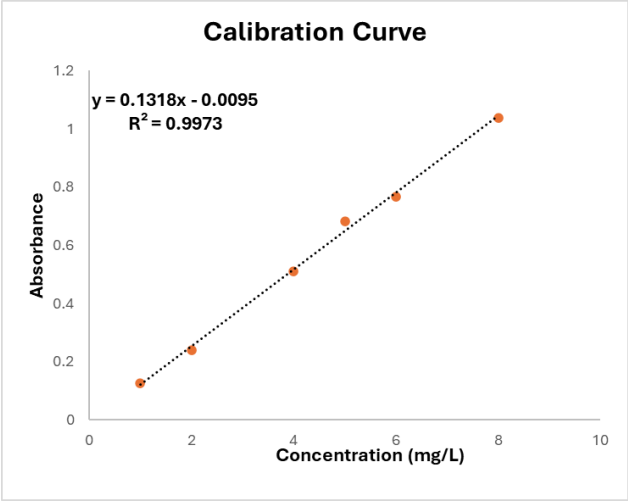


Figure 80. Calibration curve of methyl orange (concentration range: 1mg/L -8mg/L)

Figure 6 shows the calibration curve of methyl orange (MO), and the curve was prepared by taking known concentrations of 8,6,5,4,2,1 mg/L of MO as x-axis and absorbance at 464 nm as y-axis. The curve obeys Beers law in the studied concentration range.

The linear regression of the calibration curve is represented by the equation y = 0.1318x - 0.0095; where y is absorbance and x are the concentration of MO (mg/L). The coefficient of correlation (R^2) of the regression analysis is 0.9973. The calibration curve was found to be linear at the maximum wavelength and hence, suitable for the estimation of concentration of MO.

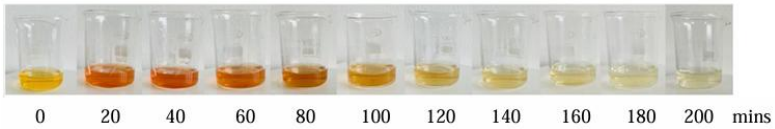


Figure 81. Degradation of MO in the presence of FeNPs

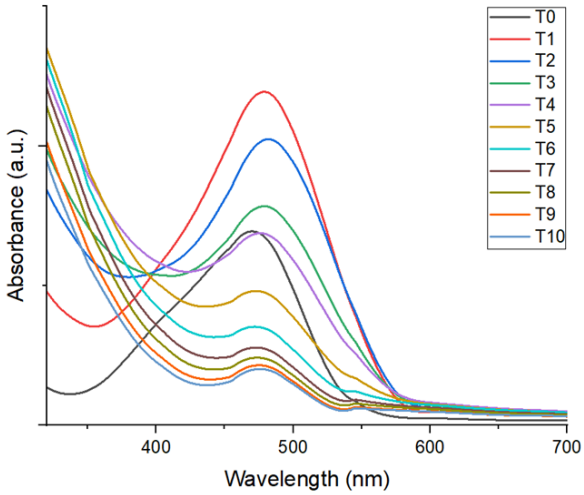


Figure 82. Kinetics of degradation of MO in the presence of the green catalyst-Fe NP (200-800 nm)

The degradation of MO using FeNPs and H_2O_2 was monitored using 2 mL samples that were

withdrawn from the mixture prepared in a previous step at time intervals of 20, 40, 60, 80, 100, 120, 140, 160, 180, 200 minutes by measuring absorbance over time. Figure 8 shows a gradual decrease in absorbance at 464 nm, indicating effective degradation. A significant reduction can be observed over 200 minutes, with final absorbance 0.185 and it corresponds to a decrease in MO concentration, demonstrating the catalytic activity of FeNPs. The degradation process followed pseudo-first-order kinetics, with MO concentration diminishing consistently over time.
Figure 7 shows the visible fading of orange color, and the solution becomes significantly lighter by the end of 200 minutes, providing a qualitative representation of dye degradation process. The progressive color change visually supports the catalytic activity of FeNPs in enhancing dye degradation.
In the initial stages of the reaction in Figure 7, a slight increase in absorbance was observed before the gradual decrease associated with dye degradation. This transient rise can be attributed to the formation of coloured intermediate species during the partial oxidation of methyl orange that absorb at the same or other wavelengths and/or increased light scattering from suspended catalyst particles and oxygen bubbles generated by the H2O2 decomposition. As the reaction progressed, these intermediates were further oxidized and turbidity reduced, leading to a steady decline in absorbance corresponding to the removal of the dye

MO decolorization was observed under the same conditions, with just H_2O_2 and without Fe NPs, to assess the capability of H_2O_2 alone in dye degradation. No significant decolorization was observed for hydrogen peroxide alone (Figure 9).

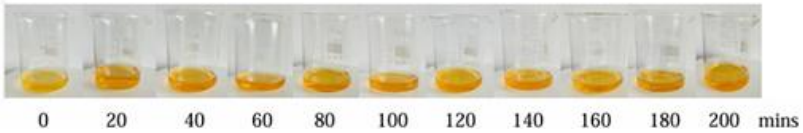


Figure 83. Degradation of MO in the absence of FeNPs

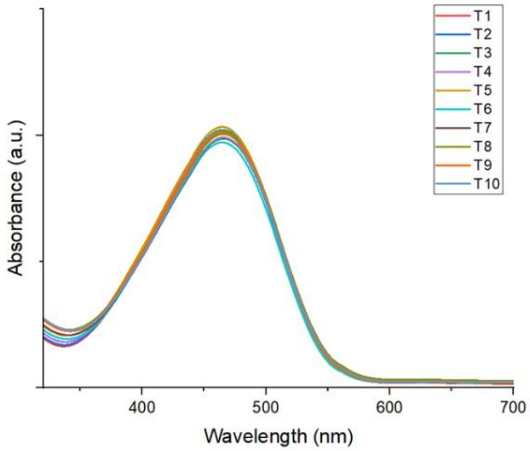


Figure 84. Kinetics of degradation of MO with just H_2O_2 and without Fe NPs (200-800 nm)

Figure 10 illustrates the variation of absorbance over different wavelengths during the control experiment using H_2O_2 without FeNPs. The absorbance remains relatively constant, fluctuating around 1.023-1.049 with minimal reduction over 200 minutes. Unlike the experiment with FeNPs, no notable reducton in the absorbance at the characteristic wavelength of 464 nm is evident. This indicates that without FeNPs, H_2O_2 alone does not significantly contribute to the degradation of MO. Any observed changes in absorbances may be due to experiments noise or minor interactions between MO and H_2O_2. This graph does not exhibit exponential decay characteristic of pseudo-first-order kinetics.

Figure 8 shows no visible change in orange color throughout the 0-200 minutes duration. The orange hue remains prominent, indicating that H_2O_2 alone is ineffective in degrading methyl orange.

Degradation of MO over time with Fe NP and MO with H_2O_2 only (without Fe NP)

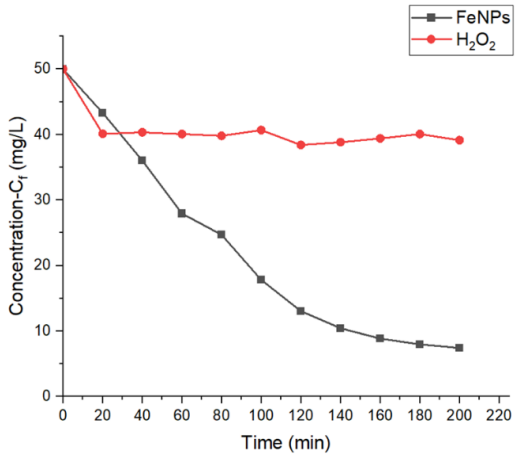


Figure 85. a) Degradation of MO over time with Fe NPs + H_2O_2 b) Control- MO with H_2O_2 only (without Fe NPs)

In Figure 11, the "With FeNPs" curve shows a steep decline, visually representing the catalytic effect of FeNPs in breaking down MO.

Dye degradation Efficiency Evaluation in the presence of FeNPs

The initial and final concentration of MO can be determined using the calibration curve and the degradation efficiency (%) was found to be 85.16 % in the presence of the Fe NPs.
This indicates high efficiency of the Fe NPs in catalyzing MO degradation. Degradation efficiency of MO in the absence of Fe Nps remains relatively flat or slightly fluctuate, indicating the limited impact of H_2O_2 alone.

Dye degradation Efficiency Evaluation in the absence of FeNPs

The percentage degradation in the absence of FeNps was calculated to be 22.00 %. This low efficiency further confirms the ineffectiveness of H_2O_2 alone in degrading MO without the catalytic activity of Fe NPs.
The combined graph (Figure 11) clearly illustrates the enhanced degradation kinetics when Fe NPs are present, validating their role as an effective green catalyst.

CONCLUSION

This work successfully demonstrated the green synthesis of FeNPs using *Macaranga peltata* leaf extract and assessed their catalytic efficiency in the degradation of methyl orange dye. The results proved that *Macaranga peltata* leaf bioactive compounds, mainly flavonoids and tannins, are effective both reducing and stabilizing agents for FeNPs. The green synthesized nanoparticles were amorphous in nature with sizes ranging from 34–94 nm, optimal for catalytic applications.
FeNPs synthesized in this work degraded 85.16% efficiency under optimized conditions within 200 minutes (3 hours and 20 minutes). Therefore, synthesized FeNPs may act as a cost-effective and environmentally friendly material for use in wastewater treatment. The study filled a critical research gap by exploring the

underutilized *Macaranga peltata* for nanoparticle synthesis, linking sustainable nanotechnology with practical environmental applications. However, the following limitations like, single dye focus, concerns of reusability of nanoparticles and toxicity of by-products need to be addressed.

ACKNOWLEDGEMENTS

We would like to acknowledge the Sri Lanka Institute of Information Technology (SLIIT) for providing necessary resources for this research. Additionally, special thanks go to Ms. Thavishi Chandratilake, for her provision of lab resources and guidance in lab work, and to the Laboratory Manager, Mr.Sehan Jayasinghe from the Faculty of Humanities and Sciences at SLIIT, for providing access to lab resources unavailable in our faculty. We also wish to express our great appreciation to the Sri Lanka Institute of Nanotechnology (SLINTEC) in offering analytical services for our samples.

REFERENCES

1. Abdeen, S., R.S, R. I., Geo, S., S, ornalekshmi, Rose, A., & P.K, P. (2013). Evaluation of Antimicrobial Activity of Biosynthesized Iron and Silver Nanoparticles Using the Fungi Fusarium Oxysporum and Actinomycetes SP. on Human Pathogens. *Nano Biomedicine and Engineering*, *5*(1). https://doi.org/10.5101/nbe.v5i1.p39-45
2. Ahmed, R. H., & Mustafa, D. E. (2019). Green synthesis of silver nanoparticles mediated by traditionally used medicinal plants in Sudan. *International Nano Letters*, *10*(1), 1–14. https://doi.org/10.1007/s40089-019-00291-9
3. Ashraf, M. A., Peng, W., Zare, Y., & Rhee, K. Y. (2018). Effects of Size and Aggregation/Agglomeration of Nanoparticles on the Interfacial/Interphase Properties and Tensile Strength of Polymer Nanocomposites. *Nanoscale Research Letters*, *13*(1). https://doi.org/10.1186/s11671-018-2624-0
4. Ebrahiminezhad, A., Taghizadeh, S., Ghasemi, Y., & Berenjian, A. (2018). Green synthesized nanoclusters of ultra-small zero valent iron nanoparticles as a novel dye removing material. *Science of the Total Environment*, *621*, 1527–1532. https://doi.org/10.1016/j.scitotenv.2017.10.076
5. ElHelaly, M. (2013). Nanotechnology, Occupational Health and Safety Concerns. *Journal of Postgenomics Drug & Biomarker Development*, *01*(03). https://doi.org/10.4172/2329-6879.1000116
6. Giun Tan, W., Ming Ng, W., Kang Lim, J., & Xin Che, H. (2018). Plantain Peel Mediated Green Synthesis Iron Oxide Nanoparticles, Surface Functionalization, and Them Performance towards Methylene Blue and Methyl Orange Dye Removal. *International Journal of Engineering & Technology*, *7*(3.36), 101. https://doi.org/10.14419/ijet.v7i3.36.29087
7. Iravani, S., & Zolfaghari, B. (2013). Green Synthesis of Silver Nanoparticles UsingPinus eldaricaBark Extract. *BioMed Research International*, *2013*, 1–5. https://doi.org/10.1155/2013/639725
8. Jiang, Z., Li, L., Huang, H., He, W., & Ming, W. (2022). Progress in Laser Ablation and Biological Synthesis Processes: "Top-Down" and "Bottom-Up" Approaches for the Green Synthesis of Au/Ag Nanoparticles. *International Journal of Molecular Sciences*, *23*(23), 14658. https://doi.org/10.3390/ijms232314658
9. Khoirotin, Nuhaa Faaizatunnisa, & Munasir Munasir. (2023). Green Synthesis of Fe3O4 Nanoparticles Using Green Betel Leaf Extract for Methylene Blue Adsorption. *Natural and Life Sciences Communications*, *22*. https://doi.org/10.12982/nlsc.2023.042
10. Megha M, Kavana D K, Rithin K, & Ramdas Bhat. (2024, October 20). *A Review On Phytochemistry And Pharmacology Of Macaranga Peltata» Article*. STM Journals. https://journals.stmjournals.com/article/article=2024/view=178897/
11. Raman, C. D., & Kanmani, S. (2018). Decolorization of Mono Azo Dye and Textile Wastewater using Nano Iron Particles. *Environmental Progress & Sustainable Energy*, *38*(s1), S366–S376. https://doi.org/10.1002/ep.13063

12. Ruan, W., Hu, J., Qi, J., Hou, Y., Zhou, C., & Wei, X. (2018a). Removal Of Dyes From Wastewater By Nanomaterials : A Review. *Advanced Materials Letters*, *10*(1), 9–20. https://doi.org/10.5185/amlett.2019.2148
13. Shahwan, T., Abu Sirriah, S., Nairat, M., Boyacı, E., Eroğlu, A. E., Scott, T. B., & Hallam, K. R. (2011). Green synthesis of iron nanoparticles and their application as a Fenton-like catalyst for the degradation of aqueous cationic and anionic dyes. *Chemical Engineering Journal*, *172*(1), 258–266. https://doi.org/10.1016/j.cej.2011.05.103
14. Wanakai, S. I., Kareru, P. G., Makhanu, D. S., Madivoli, E. S., Maina, E. G., & Nyabola, A. O. (2019). Catalytic degradation of methylene blue by iron nanoparticles synthesized using Galinsoga parviflora, Conyza bonariensis and Bidens pilosa leaf extracts. *SN Applied Sciences*, *1*(10). https://doi.org/10.1007/s42452-019-1203-z
15. Yusuf, M. (2019). Silver Nanoparticles: Synthesis and Applications. *Handbook of Ecomaterials*, 2343–2356. https://doi.org/10.1007/978-3-319-68255-6_16

Predicting Cognitive Test Performance from New Onset Behavioral and Personality Changes in Adults over 50 using Post-Selection Boosted Random Forest Classifier

Mervyn M.[a*], Welhenge A.[b], Creese B.[c]

[a]Dept of Computer Systems Engineering, University of Kelaniya, Sri Lanka
[b]EECMS, Curtin University, Perth, Australia
[c]Dept of Life Sciences, Brunel University of London, United Kingdom.

meesham@kln.ac.lk*, anuradhi.welhenge@curtin.edu.au, byron.creese@brunel.ac.uk.

ABSTRACT

Mild Behavioral Impairment (MBI) refers to neuropsychiatric symptoms of various severity levels that might not be discovered by conventional psychiatric nosology. These symptoms should persist for more than six (06) months. MBI is typically observed in adults of age 50 and above. This study investigates the prediction of cognitive test performance of cognitive and behavioral changes in adults over 50 years of age using a post-selection boosted Random Forest (RF) Classifier. The baseline cognitive aging data of the Simple Reaction Time (SRT) metric and Mild Behavioral Impairment Checklist (MBI-C) from the ongoing PROTECT study in the United Kingdom was used to classify the participants' cognitive ability into five classes. Using the post-selected boosted RF classifier, the study obtained an accuracy of 96.26% which was an improvement compared to the 95.52% accuracy obtained by the RF classifier. These findings suggest that machine learning-based prediction models can provide valuable insights into analyzing the cognitive decline of adults of a late age.

KEYWORDS: *Mbi, Cognitive Performance, Lasso Method, Deep Learning*

INTRODUCTION

Mild Behavioral Impairment (MBI) is a phrase generally used in a medical context to describe subtle but noticeable changes in a person's behavior and personality that indicate early signs of cognitive decline. This neuropsychiatric syndrome is more prominently observed in late ages, in people with higher cognitive impairment, with a percentage of frequency between 10% and 15%. The key symptoms observed in a person suspected are mood fluctuation, a noticeable lack of interest and motivation, signs of social withdrawal, and a sudden shift in personality. This study aims to find whether machine learning techniques can be used to predict simple reaction time (SRT) from MBI-C.

BACKGROUND

Mild Behavioral Impairment and its Association with Cognition

Mild Behavioral Impairment (MBI) describes a spectrum of late-life onset, sustained neuropsychiatric symptoms (NPS) that includes apathy, affect, impulse dyscontrol, social inappropriateness and abnormal thoughts and perception. MBI may be the first manifestation of an underlying neurodegenerative disease, or occur alongside, or after subjective cognitive impairment and Mild Cognitive Impairment (MCI) (Ismail et al., 2016). That MBI represents a risk state for dementia and is a symptom of underlying neurodegenerative disease is supported by observational cohort studies linking MBI to incident dementia, and AD biomarkers (Creese & Ismail, 2022).
Because of the epidemiological evidence linking MBI to incident dementia and cognitive decline, there is an increased research need to establish the methods that offer the most reliable prediction, which could in turn lead to early identification of dementia in those with MBI.
(Creese et al., 2019) conducted a study to find whether MBI was associated with neuropsychological performance in subjects without significant cognitive impairment. They measured attention, reasoning, executive function and

working memory of the older adults participated in the study and MBI was recorded using the MBI- C. The study found that MBI could be an early marker of neurodegenerative disease compared to MCI.

The International Society to Advance Alzheimer's Research and Treatment - Alzheimer's Association (ISTAART-AA) has a research diagnostic criterion to measure MBI, Mild Behavioral Impairment Checklist (MBI-C) is specifically designed to assess MBI construct from the research criteria. This has 34 items, and the development process included the five domains of MBI which are 1. Decreased motivation; 2. Emotional Dysregulation; 3. Impulse dyscontrol; 4. Social inappropriateness; and 5. Abnormal perception or thought (Ismail et al., 2017). Studies were conducted to validate the Mild Behavioural Impairment Checklist (MBI-C) that aligned with the diagnostic criteria established by the ISTAART-AA. Findings indicate that nearly 25% of individuals attending Primary Care Centres with Subjective Cognitive Complaints (SCCs), who do not have a history of neuropsychiatric disorders, receive a diagnosis of MBI (Mallo et al., 2019).

Machine Learning and Deep Learning based Improvements

In the past few years, studies have shown that Mild Behavioral Impairment diagnosis can be enhanced using Machine Learning (ML) and Deep Learning (DL), focusing on detecting early signs of Alzheimer's disease. Studies conducted in recent years focused on developing ML and DL algorithms to determine the probability of a patient in later ages developing MBI which has a possibility of morphing into dementia or Alzheimer's disease. Combining ML techniques with expert diagnosis from medical professionals can significantly improve the accuracy level of the diagnosis. Anderson (Anderson, 2019) suggested that there is an urgent need for standardized cognitive tests in order to benefit from the recent discoveries in cognitive neuroscience, which can lead to the development of heightened sensitive measures of MCI.

ML techniques have shown promising results in predicting early signs of dementia. By using NPS to train the algorithms, it is possible to improve the accuracy of ML models (Gill et al., 2020). MBI being identified as a marker of pre-dementia is due to Neuropsychiatric symptoms (NPS) being commonly identified across multiple cognitive disorders. These symptoms are often overlooked due to their similarity with other commonly diagnosed cognitive disorders, especially observed in older adults. Gill S. concluded that baseline NPS categorized by domain and duration for MBI could predict diagnosis change using ML. The risk factor of MBI developing into dementia is significantly higher in individuals over the age of 50. They often exhibit several cross-over symptoms such as anxiety, depression, and apathy. An observational study used a multi-class emotional classifier with Random Forest classification to detect cross-over symptoms in adults of 70-75 years with diagnosed MCI (Zhou et al., 2023).

The ADNI dataset is used in several ML and DL centered studies conducted on cognitive decline. One study compared between two-class and three-class ML classification of cognitive features optimized for detecting early Alzheimer's stages. The two-class model performed better with a minimal assessment time compared to the three-class model (Kleiman et al., 2021).

A number of studies utilize deep learning techniques such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) to analyze structural MRI (Magnetic Resonance Imaging) data. Conventional Machine Learning approaches need a considerable portion of feature engineering in order to optimize the performance of the model. DL models show improved accuracy in differentiating between several psychiatric disorders by identifying patterns that may not be apparent using conventional ML models. A study which showed this fact was conducted in 2019, where Artificial Neural Networks was used to predict the efficiency of certain medications before they were prescribed to patients with depressive disorders in order to develop personalized treatment plans (Chang et al., 2019). Another such research developed an interpretative DL framework using CNN to classify the Alzheimer's Disease probability of patients. The study suggested the possibility of adopting the framework as a clinical strategy for generalized pathophysiological disease discovery (Qiu et al., 2020). Grouping some cognitive assessment features into classes which are optimized for detecting cognitive disorders can significantly improve routine screening. F. Yi et al. used the patient data from the ADNI database to develop an interpretable framework combining SHAP and XGBoost for multi-classification of normal cognition, MCI, and Alzheimer's disease (Yi et al., 2023).

Machine Learning and Deep learning applications in MCI have shown promising results in both prediction of cognitive decline as well as improving the accuracy of diagnosis. However, similar methods could be used to predict MBI using different parameters as inputs and this area has received limited attention in the literature. In this paper we are predicting the MBI-C by using SRT as an input parameter.

METHODOLOGY

Dataset

This study employs a cross-sectional analysis of data from the ongoing online PROTECT Study (www.protectstudy.org.uk), a longitudinal cohort study focusing on 14 201 non-demented individuals aged 50 and over in the United Kingdom. Launched in November 2015, the aim of this study is to investigate cognitive aging and identify risk factors for dementia within a large, UK-based population of >50 years. The study utilizes a digital platform to collect self-reported data on lifestyle, health, and cognitive performance, enabling large-scale participation and frequent data collection. Inclusion criteria required participants to be aged 40 years or older, reside in the UK, possess a good working understanding of English, and have the ability to use a computer with internet access. Individuals with an established diagnosis of dementia were excluded, as the study focuses on understanding cognitive changes in healthy aging and early risk factors for cognitive decline.

This research focuses on classifying the cognitive ability of the participants aged 50 and above using the Simple Reaction Time (SRT) metric into five (05) classes using the MBI-C.

Random Forest (RF) Classifier

In this research an RF Classifier was employed to predict SRT values against the baseline cognitive performance of the participants. The MBI-C values were used as input features, and the SRT values were predicted into five groups classified from low to high cognitive performance. The prediction was conducted class-wise rather than using exact values to improve prediction accuracy and reduce the training complexity of the model. The RF classifier was implemented on the dataset after finding the principal components from the Principal Component Analysis (PCA).

In ensemble learning samples and variables are sampled which is the basis of the popular RF algorithm (by Breiman) which groups multiple decision trees. During classification, the results were calculated based on the majority vote method. The RF implementation is as follows.

M – Training set, p - variables

1. Generating a sample set M* of size Q using bootstrap sampling from the training set M.
2. Generating a tree based on the sample set M* in RF, the following steps are repeated for each node. The process continues until a node reaches the predetermined size.
   - n variables were randomly selected from p variables.
   - Pick the best splitting variable and splitting point from r variables.
   - Split the current node into two child nodes according to the selected splitting variable and splitting point.
3. For a new observation z, the predicted value is calculated as

$$T(z) = majorityvote T_v(z) 1^V \qquad (1)$$

Post-Selection Boosted Random Forest (PBRF) Classifier

As the second part of this study post-selection boosting random forest algorithm (PBRF) was implemented. The Post-Selection Boosting Random Forest (PBRF) algorithm is an advanced ensemble learning technique that combines the strengths of Random Forest (RF) with the Lasso method to improve predictive performance while reducing model complexity. In this study, PBRF is implemented after applying Principal Component Analysis (PCA) to the dataset, further optimizing the model's performance (Wang & Wang, 2021). Principal components representing 90% of the variance have been used in this study.

The implementation of the PBRF algorithm was as follows.

1. Use Bootstrap Sampling to generate a dataset D* with sample size N from data set D
2. Construct a tree T* by randomly selecting r variables from all p variables for each node.
3. Do Lasso Regression to T*, which compress the weight coefficients of some trees to 0.
4. The final prediction is computed as a weighted average of the predictions from the selected trees.

$$y = \frac{1}{m}\sum_{n=1}^{m} \quad T_n(z) \qquad (2)$$

RESULTS

Random Forest (RF) Classifier

The RF classifier was used to analyze the impact of various feature splits on the outcome of the classification, with each split affecting a specific percentage of the selected sample. Figure 1, Figure 2 and Figure 3 depict the features, the percentage of the sample covered, and the accuracy predicted for each specific class.

The overall accuracy achieved by the RF classifier was 95.52%

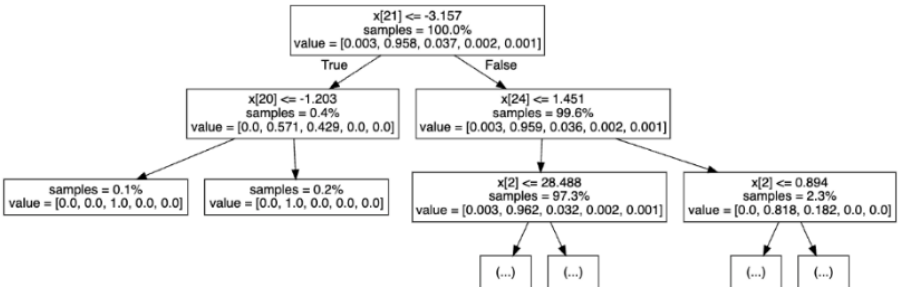


Figure 1: Split of the feature x[21]

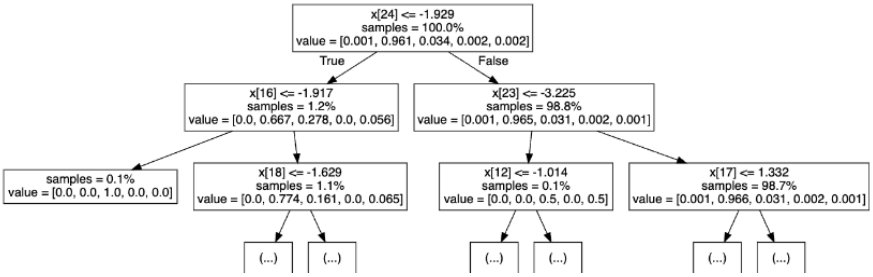


Figure 2: Split of the feature x[24]

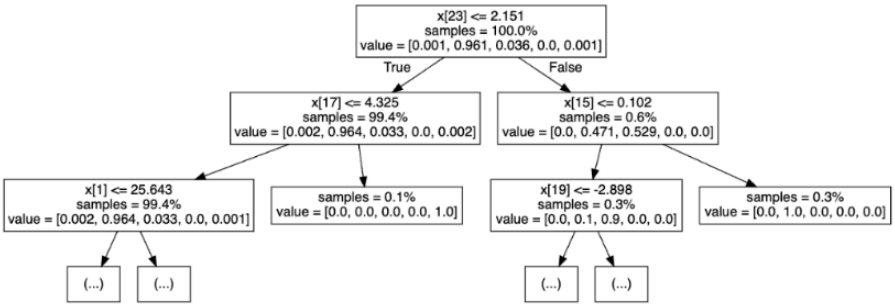


Figure 3: Split of the feature x[23]

Post-Selection Boosted Random Forest (PBRF) Classifier

The analysis conducted in this research was further expanded by using the PBRF classifier. Similar to the analysis conducted using the RF algorithm, the sample was split into nodes based on selected features to refine the classification, and the accuracy was predicted for the specific classes. Figure 4, Figure 5, and Figure 6 depict the selected features and their percentage of the sample size covered, along with the class accuracy predicted.

The overall accuracy obtained by the PBRF classifier was 96.26%.

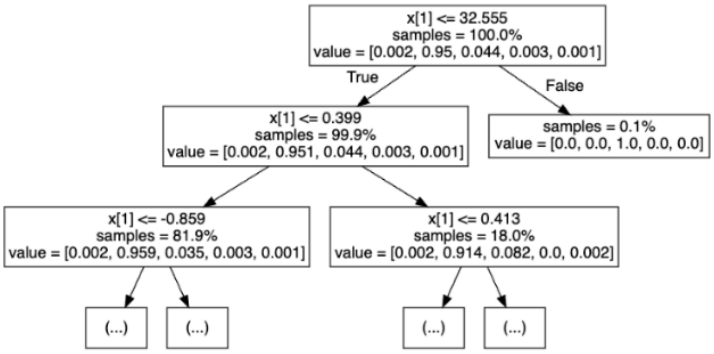


Figure 4: Split of the feature x[1]

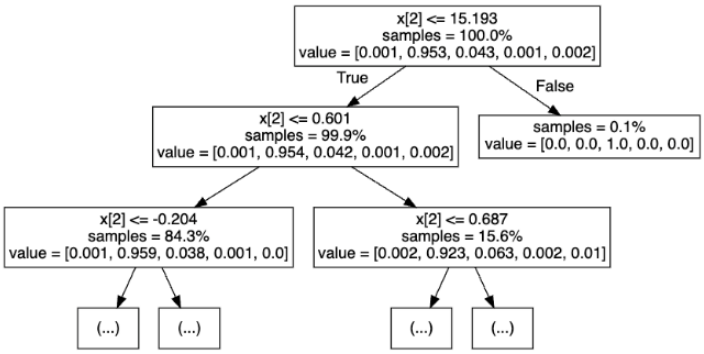


Figure 5: Split of the feature x[2]

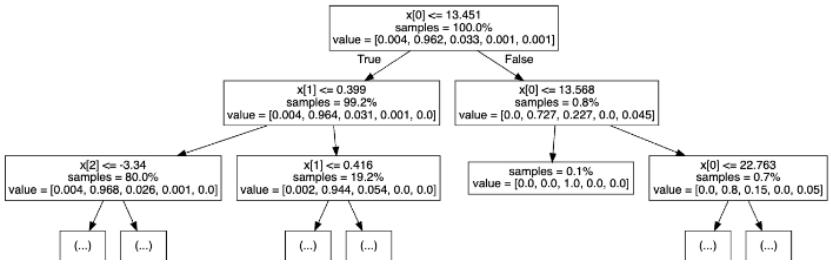


Figure 6: Split of the feature x[0]

Boosting helps the model separate classes better, making more accurate predictions. Overall, the model uses key features to classify data well, though some classes may be less balanced in certain areas. This result shows that PBRF performs better than the RF.

CONCLUSION AND DISCUSSION

Mild Behavioral Impairment (MBI) is a symptom of cognitive decline observed in adults over the age of 50. Using a cross-sectional analysis of the PROTECT dataset from the UK the baseline cognitive ageing was evaluated by using the Simple Reaction Time (SRT). The study used an improved Random Forest Classifier to predict the range of SRT values against the baseline of cognitive aging of the participants. The performance of the improved RF classifier was compared with the accuracy level obtained using a traditional RF algorithm, and it was observed that the improved RF classifier performed better exhibiting a higher accuracy of prediction. This research provides further evidence to support the claim that ML and DL methods will aid in the early detection of cognitive decline during late adulthood by predicting behavioral and personality changes in individuals.
Studies underscore the significance of utilizing standardized assessment tools like the MBI-C in clinical settings to

identify individuals at risk for cognitive decline and dementia. MBI research has gained traction in recent years, and it is recognized as a neurobehavioral syndrome associated with dementia and Alzheimer's. However, the underlying mechanisms linking MBI to cognitive decline are not fully understood. Understanding these mechanisms could significantly improve diagnosis of neurodegenerative disease.

The precise relationship between MBI and cognitive function is an area of ongoing research. Some studies have suggested that MBI is associated with poorer cognitive performance, particularly in the domains of executive function and memory (Petersen & Negash, 2008). However, the nature of this relationship is complex, and there is a need for more longitudinal studies to clarify the trajectories of cognitive and behavioural changes in individuals with MBI. To improve the knowledge of MBI, it will be essential to fill in these gaps through standardised tests, longitudinal studies, studies on a variety of populations, and integration with neuroimaging.

ACKNOWLEDGEMENTS

This PROTECT study was supported by the National Institute for Health and Care Research (NIHR) Exeter Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. This research was also supported by the NIHR Collaboration for Leadership in Applied Health Research and Care Southwest Peninsula.

REFERENCES

1. Anderson, N. D. (2019). State of the science on mild cognitive impairment (MCI). *CNS Spectrums*, *24*(1), 78–87. https://doi.org/10.1017/S1092852918001347
2. Chang, B., Choi, Y., Jeon, M., Lee, J., Han, K.-M., Kim, A., Ham, B.-J., & Kang, J. (2019). ARPNet: Antidepressant Response Prediction Network for Major Depressive Disorder. *Genes*, *10*(11), 907. https://doi.org/10.3390/genes10110907
3. Creese, B., Brooker, H., Ismail, Z., Wesnes, K. A., Hampshire, A., Khan, Z., Megalogeni, M., Corbett, A., Aarsland, D., & Ballard, C. (2019). Mild Behavioral Impairment as a Marker of Cognitive Decline in Cognitively Normal Older Adults. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *27*(8), 823–834. https://doi.org/10.1016/j.jagp.2019.01.215
4. Creese, B., & Ismail, Z. (2022). Mild behavioral impairment: Measurement and clinical correlates of a novel marker of preclinical Alzheimer's disease. *Alzheimer's Research & Therapy*, *14*(1), 2. https://doi.org/10.1186/s13195-021-00949-7
5. Gill, S., Mouches, P., Hu, S., Rajashekar, D., for the Alzheimer's Disease Neuroimaging Initiative, MacMaster, F. P., Smith, E. E., Forkert, N. D., & Ismail, Z. (2020). Using Machine Learning to Predict Dementia from Neuropsychiatric Symptom and Neuroimaging Data. *Journal of Alzheimer's Disease*, *75*(1), 277–288. https://doi.org/10.3233/JAD-191169
6. Ismail, Z., Agüera-Ortiz, L., Brodaty, H., for the NPS Professional Interest Area of the International Society of to Advance Alzheimer's Research and Treatment (NPS-PIA of ISTAART), Cieslak, A., Cummings, J., Fischer, C. E., Gauthier, S., Geda, Y. E., Herrmann, N., Kanji, J., Lanctôt, K. L., Miller, D. S., Mortby, M. E., Onyike, C. U., Rosenberg, P. B., Smith, E. E., Smith, G. S., Sultzer, D. L., & Lyketsos, C. (2017). The Mild Behavioral Impairment Checklist (MBI-C): A Rating Scale for Neuropsychiatric Symptoms in Pre-Dementia Populations. *Journal of Alzheimer's Disease*, *56*(3), 929–938. https://doi.org/10.3233/JAD-160979
7. Ismail, Z., Smith, E. E., Geda, Y., Sultzer, D., Brodaty, H., Smith, G., Agüera-Ortiz, L., Sweet, R., Miller, D., Lyketsos, C. G., & ISTAART Neuropsychiatric Symptoms Professional Interest Area. (2016). Neuropsychiatric symptoms as early manifestations of emergent dementia: Provisional diagnostic criteria for mild behavioral impairment. *Alzheimer's & Dementia*, *12*(2), 195–202. https://doi.org/10.1016/j.jalz.2015.05.017
8. Kleiman, M. J., Barenholtz, E., Galvin, J. E., & for the Alzheimer's Disease Neuroimaging Initiative. (2021). Screening for Early-Stage Alzheimer's Disease Using Optimized Feature Sets and Machine Learning. *Journal of Alzheimer's Disease*, *81*(1), 355–366. https://doi.org/10.3233/JAD-201377
9. Mallo, S. C., Ismail, Z., Pereiro, A. X., Facal, D., Lojo-Seoane, C., Campos-Magdaleno, M., & Juncos-Rabadán, O. (2019). Assessing mild behavioral impairment with the mild behavioral impairment checklist in people with subjective cognitive decline. *International Psychogeriatrics*, *31*(2), 231–239. https://doi.org/10.1017/S1041610218000698

10. Petersen, R. C., & Negash, S. (2008). Mild Cognitive Impairment: *An Overview*. *CNS Spectrums*, *13*(1), 45–53. https://doi.org/10.1017/S1092852900016151
11. Qiu, S., Joshi, P. S., Miller, M. I., Xue, C., Zhou, X., Karjadi, C., Chang, G. H., Joshi, A. S., Dwyer, B., Zhu, S., Kaku, M., Zhou, Y., Alderazi, Y. J., Swaminathan, A., Kedar, S., Saint-Hilaire, M.-H., Auerbach, S. H., Yuan, J., Sartor, E. A., … Kolachalama, V. B. (2020). Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain*, *143*(6), 1920–1933. https://doi.org/10.1093/brain/awaa137
12. Wang, H., & Wang, G. (2021). Improving random forest algorithm by Lasso method. *Journal of Statistical Computation and Simulation*, *91*(2), 353–367. https://doi.org/10.1080/00949655.2020.1814776
13. Yi, F., Yang, H., Chen, D., Qin, Y., Han, H., Cui, J., Bai, W., Ma, Y., Zhang, R., & Yu, H. (2023). XGBoost-SHAP-based interpretable diagnostic framework for alzheimer's disease. *BMC Medical Informatics and Decision Making*, *23*(1), 137. https://doi.org/10.1186/s12911-023-02238-9
14. Zhou, Y., Han, W., Yao, X., Xue, J., Li, Z., & Li, Y. (2023). Developing a machine learning model for detecting depression, anxiety, and apathy in older adults with mild cognitive impairment using speech and facial expressions: A cross-sectional observational study. *International Journal of Nursing Studies*, *146*, 104562. https://doi.org/10.1016/j.ijnurstu.2023.104562

Solar Hotspot Detection Using VHDL-Simulated Fixed-Point SVM: A Methodology Toward FPGA Realization Solar Hotspot Detection via FPGA-SVM

[a]Nayomi Fernando, [b]Lasantha Seneviratne, [c]Nisal Weerasinghe, [d]Namal Rathnayake*, [e]Yukinobu Hoshino

[a,b,c,]Department of Electrical and Electronic Engineering, Faculty of Engineering, Sri Lanka Institute of Information Technology, Malabe 10115, Colombo, Sri Lanka;
[d]Marine-Earth System Analytics Unit, Yokohama Institute for Earth Sciences, JAMSTEC, 3173-25 Showa-machi, Kanazawa-ku, Yokohama 236-0001, Kanagawa, Japan;
[e]School of Systems Engineering, Kochi University of Technology, 185 Miyanokuchi, Tosayamada, Kami 782-8502, Kochi, Japan;
[a]nayomi.f@sliit.lk , [b]lasantha.s@sliit.lk , [c]nisal.w@sliit.lk, [d]namalr@jamstec.go.jp , yukinobu@kochi-tech.ac.jp

The early and accurate detection of thermal hotspots in photovoltaic modules is critical to ensure the efficiency, safety, and longevity of solar power systems. This study presents a complete end-to-end methodology for implementing a fixed-point Medium Gaussian Support Vector Machine classifier using Very High-Speed Integrated Circuit - Hardware Description Language, optimized for Field Programmable Logic Array. The approach begins with feature extraction from thermal images, focusing on MPEG-7 descriptors and blue chrominance. The SVM model is trained in MATLAB and converted into a fixed-point Q1.15 format for hardware compatibility. Key parameters, including support vectors, Lagrange multipliers, bias, and kernel scale, are extracted and verified in a custom Python environment to ensure numerical alignment with MATLAB results. The validated model is then implemented in synthesizable VHDL and verified using GHDL and GNU Tool Kit waveform viewer, confirming bit-accurate hardware behavior. Results show classification accuracy exceeding 99.3% with negligible performance loss due to quantization. The design achieves deterministic latency based on FSM structure and parallel feature processing, completing classification within 2702 clock cycles for a 300-support-vector, 222-feature system. Unlike floating-point models, this approach enables low-power, real-time inference on edge platforms such as drones.

CCS CONCEPTS • **Hardware ~ Electronic design automation ~ Logic synthesis • Computing methodologies ~ Machine learning ~ Learning paradigms ~ Supervised learning • Hardware ~ Integrated circuits ~ Reconfigurable logic and FPGAs • Computer systems organization ~ Real-time systems**

ADDITIONAL KEYWORDS AND PHRASES: *FPGA, ML, VHDL, Solar PV, MPEG-7*

INTRODUCTION

The increasing global demand for renewable energy has led to the widespread deployment of photovoltaic (PV) systems, particularly in utility-scale solar farms and rooftop installations. As these installations scale in size and complexity, maintaining their optimal performance becomes a significant operational challenge.

A key factor that can compromise efficiency and long-term reliability is the formation of thermal hotspots. That is the localized region of elevated temperature, which is typically caused by partial shading, soiling, mismatched cells, or physical defects such as micro-cracks. Hotspots not only reduce the power output of the system but also accelerate module degradation, potentially leading to irreversible damage.

Advances in thermal imaging, computer vision, and Machine Learning (ML) have enabled automated detection and classification of such anomalies. Among various techniques, feature extraction based on blue chrominance and MPEG-7 texture descriptors have proven effective in distinguishing hotspot-affected areas from healthy PV cells. A recent study [1] demonstrated that integrating these features with conventional classifiers, notably the Medium Gaussian Support Vector Machine (SVM), achieves high classification accuracy on drone-captured thermal datasets. Medium Gaussian SVM offers a compelling trade-off between accuracy, model complexity, and computational efficiency. In the referenced study, this approach achieved near-perfect accuracy of 99.3% with an inference time of 18 seconds. However, models developed in MATLAB or Python typically rely on floating-point arithmetic and high-level structures, which are unsuitable for real-time, low-power edge applications due to resource and power constraints. This creates a practical demand for hardware-efficient deployment, especially in remote or mobile inspection platforms such as drones or IoT-enabled monitoring systems. Field-Programmable Gate Arrays (FPGAs) offer a viable solution due to their inherent parallelism, deterministic timing, and energy efficiency attributes, well-aligned with the real-time requirements of edge-based PV monitoring.

However, translating high-level ML models into FPGA-friendly architecture poses several challenges. Most notably, floating-point operations must be converted into fixed-point arithmetic to ensure synthesis compatibility and minimize resource usage. To address this, the current work presents a complete fixed-point SVM implementation using the Q1.15 format, which offers a practical balance between precision and FPGA logic utilization. The workflow starts with extracting hyperparameters such that support vectors, Lagrange Multipliers, labels, kernel scale, and bias from the MATLAB-trained Medium Gaussian SVM model in Classification Learner. These parameters are then implemented in a Python-based fixed-point simulation to verify numerical consistency. The validated model is implemented in synthesizable VHDL using Q1.15 arithmetic. The VHDL design comprises three core modules: the SVM decision logic, a constants package, and a test bench. Simulation and validation were performed using GHDL for logic-level execution and GTKWave for waveform inspection. This allowed for cycle-accurate debugging and ensured the correctness of the decision function. The verified VHDL modules are designed with synthesis readiness in mind, enabling straightforward integration into toolchains such as Intel Quartus or Xilinx Vivado for full FPGA implementation.

This work is based on our previous study, which compared ML and DL models for UAV-based PV hotspot detection, applied Explainable AI methods for interpretability, benchmarked classification accuracy and computational cost, and recommended optimal models for UAV deployment. The present extension focuses on hardware-oriented implementation and contributes to the literature in four specific areas:

- Development of a complete pipeline that spans MPEG-7 feature extraction in ACE (Automatic Content Extractor), SVM-based classification, and FPGA-ready VHDL code generation.
- Conversion of trained SVM parameters into the Q1.15 fixed-point format in order to minimize FPGA resource usage without compromising classification accuracy.
- Verification of the VHDL implementation for bit-accurate agreement with MATLAB predictions and deliver a synthesis-ready FPGA design for seamless integration into FPGA synthesis tools.

This paper is organized as follows: Literature Review investigates recent literature on solar hotspot detection, emphasizing feature extraction and FPGA-based classification approaches. Methodology describes the proposed procedure, including SVM training in MATLAB, fixed-point quantization, and VHDL implementation. Results present simulation outputs, and Discussion section discusses classification

accuracy, hardware efficiency, and suitability for real-time UAV deployment, along with statistical analysis. Finally, the paper concludes with the study and outlines directions for future work, including hardware synthesis and live system integration.

LITERATURE REVIEW

The detection of thermal hotspots in photovoltaic panels has gained significant momentum with the convergence of infrared (IR) thermography, machine learning algorithms, and embedded hardware platforms. This section reviews the current state of research across three main domains: image-based hotspot detection techniques, the application of Support Vector Machines for PV fault classification, and FPGA-based implementation of ML algorithms. The aim is to highlight the evolution of methodologies and identify the key limitations that shape the motivation for the present work.

Image-Based Hotspot Detection in PV Panels

Traditional PV system monitoring relied on electrical parameters such as voltage and current irregularities; however, these approaches often fail to detect early-stage anomalies [2], [3]. The emergence of infrared thermography, especially from drone-mounted platforms, has enabled the identification of localized thermal variations, which are the primary indicators of potential hotspots [4], [5].

The study titled "Efficient Hotspot Detection in Solar Panels via Computer Vision and Machine Learning" proposed a framework that combines MPEG-7 texture descriptors, such as edge histograms and color-based features like blue chrominance and YCbCr contrast, to classify fault regions in PV panels. Their experiments across multiple real-world datasets demonstrated that the Medium Gaussian SVM outperformed other classifiers like RUSBoosted Tree and Binary GLM (Generalized Linear Model) Logistic Regression in terms of classification accuracy and robustness to feature variation. This paper serves as a foundational reference for identifying discriminative features in thermal images, and its findings strongly motivate the use of Medium Gaussian SVM as a core classifier in hardware deployment scenarios [1].

Support Vector Machines in Fault Classification

Support Vector Machines have consistently been favored in classification tasks where feature space dimensionality is high and non-linearity is prevalent [6]. Their ability to define optimal hyperplanes with maximal margin allows them to generalize well, even on limited datasets. The Radial Basis Function (RBF) or Gaussian kernel used in SVMs adds further flexibility by enabling the mapping of inputs into higher-dimensional spaces where linearly inseparable data become separable [7].

Numerous works have leveraged SVMs in PV diagnostics. For example, study [6] employed an SVM for string-level fault detection in PV arrays using electrical signatures, achieving notable precision. More recent approaches utilize thermal image features such as SVM inputs, showing better real-time applicability and generalization under outdoor conditions [8] [9]. While software implementations are widely studied, hardware realization of SVM-based models remains relatively unexplored due to challenges in representing floating-point operations and non-linear kernel functions in hardware [10], [11].

FPGA-Based Machine Learning and Fixed-Point Design

FPGAs offer a reconfigurable and power-efficient platform for deploying ML inference models. Their parallel processing capabilities make them ideal for latency-sensitive applications such as real-time image classification [12]. However, due to hardware constraints, floating-point arithmetic is often avoided in favor of fixed-point implementations, which offer simpler logic synthesis, lower resource usage, and faster execution times [13].

Existing literature shows that deploying ML on an FPGA typically requires either:

- High-level synthesis tools like Xilinx Vitis or Intel HLS, which abstract HDL generation from C/C++ code [14].
- Manual HDL coding with fixed-point models, offering better control and efficiency [10].

The study [14] proposed a fixed-point neural network implementation on an FPGA for object detection, showing that quantization has minimal impact on accuracy while drastically improving performance metrics. Similarly, research [13] implemented a hardware-efficient kernel SVM classifier for gesture recognition using VHDL, successfully validating the feasibility of kernel-based models on hardware.

Despite these advances, very few studies focus on deploying SVMs with Gaussian kernels, especially for solar hotspot detection using VHDL implementations. The complexity of exponential function evaluation and vector distance computations poses challenges for hardware realization. This gap is addressed in the current work, wherein a MATLAB-trained Medium Gaussian SVM is translated into a fixed-point, synthesizable VHDL model. The design is developed using VS Code, simulated using GHDL, and its functional correctness is verified through waveform analysis in GTKWave. This simulation step ensures that the algorithm behaves identically to its floating-point MATLAB counterpart before it is synthesized for real-time FPGA deployment [10], [19].

Research Gap and Contribution Justification

The reviewed literature confirms that most PV monitoring solutions rely on software-based ML models with limited focus on edge deployment. While studies highlight the motivation and complexity of implementing Gaussian kernel SVMs in hardware [15], [16], existing works on DNN and SVM hardware flows [17], [18] lack a complete fixed-point MATLAB-to-VHDL pipeline. Furthermore, a clear gap persists in the realization of fixed-point, non-linear SVMs on FPGA platforms [19]. By addressing these gaps, the present work contributes a novel end-to-end approach that combines ML precision with hardware readiness, setting the stage for energy-efficient, real-time PV health monitoring systems.

In summary, the literature demonstrates substantial progress in image-driven PV fault detection using SVM classifiers, particularly those incorporating Gaussian kernels. While many software-based approaches have shown promising results, limited attention has been paid to fixed-point SVM implementations on an FPGA for real-time applications. The reviewed studies emphasize both the potential and the challenges of hardware realization, especially for non-linear models. This justifies the current research focus on translating a MATLAB-trained Medium Gaussian SVM model into an efficient VHDL-based implementation, bridging the gap between ML accuracy and embedded hardware feasibility for PV monitoring systems.

METHODOLOGY

This section presents the comprehensive methodology adopted to develop and validate a synthesis-ready Medium Gaussian SVM classifier for real-time solar hotspot detection on FPGA hardware through simulation and pre-synthesis verification. Aligned with the feature choices motivated in the literature, the workflow starts from thermal-image feature extraction using MPEG-7 descriptors and blue chrominance (YCbCr), with model training in MATLAB, and spans fixed-point conversion, VHDL development, and simulation using GHDL. Key design steps, including floating-to-fixed-point adaptation in Q1.15, FSM (Finite State Machine)-based control logic, and a hardware-friendly Gaussian kernel approximation, are discussed to demonstrate how high-level machine-learning models can be translated into efficient, synthesizable hardware architectures suitable for UAV-based PV panel monitoring. The pipeline explicitly extracts SVM parameters such that support vectors, Lagrange multipliers, labels, bias, and kernel scale.

Further, this validates a fixed-point MATLAB reference and implements dot-products and kernel logic in VHDL, with cycle-accurate verification in GHDL and GTKWave to ensure bit-accurate parity prior to FPGA synthesis.

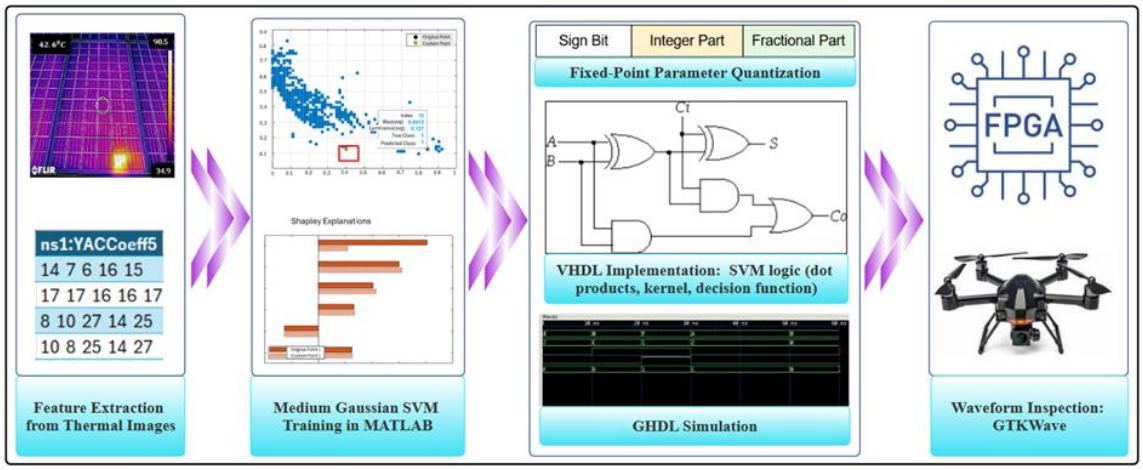


Figure 1. End-to-end workflow for implementing the Medium Gaussian SVM classifier on FPGA: starting from feature extraction and training in MATLAB, through fixed-point parameter quantization, VHDL coding, accurate simulation with GHDL, and culminating in FPGA synthesis and deployment-ready integration on UAV platforms for real-time hotspot detection

Figure 1 illustrates the complete workflow for implementing a Medium Gaussian SVM classifier in an FPGA-compatible form for real-time PV hotspot detection. It begins with feature extraction from thermal images using YCbCr descriptors, followed by training in MATLAB. From the trained model, we explicitly extract support vector coefficients, labels, Lagrange multipliers, bias, and kernel scale. A custom MATLAB implementation of the SVM decision function is then built to replicate the behavior of prediction, ensuring an exact software reference. These parameters and computations are converted to fixed-point Q1.15, with explicit sign, integer, and fractional parts for hardware compatibility. The SVM logic is implemented in VHDL with dot products, Gaussian kernel approximation, and decision, followed by cycle-accurate simulation in GHDL. This produces waveforms for GTKWave inspection to confirm bit-accurate agreement with MATLAB. The verified design is delivered as a synthesis-ready module, compatible with standard FPGA synthesis toolchains like Intel Quartus or Xilinx Vivado for future deployment on UAV platforms.

The subsequent discussion expands on these stages, detailing how MATLAB SVM models are bridged to synthesizable VHDL, including the conversion of floating-point parameters into fixed-point, hardware-friendly forms, the use of fixed-point arithmetic for FPGA compatibility, logic gate-level simulation for solar hotspot classification, and the implementation of Gaussian kernel and decision logic, all intended for FPGA deployment.

Bridging MATLAB SVM Models to Synthesizable VHDL: From Floating-Point SVM to Fixed-Point Hardware-Friendly Implementation

Built-in SVM functions like fitcsvm and predict in MATLAB use floating-point arithmetic and object-oriented code, making them incompatible with HDL Coder and FPGA workflows. To overcome this, a custom SVM decision function is manually implemented in MATLAB using extracted parameters such that alpha, gamma, bias, and support vectors. This allowed conversion to fixed-point using fi() function in

MATLAB, enabling hardware-friendly computation. The manual approach offers precise control over word length, precision, and overflow behavior, aiding in debugging, validation, and matching with the software model. It also simplifies the creation of HDL test benches and serves as a reliable bridge toward synthesizable VHDL modules for solar PV hotspot detection.\

Fixed-Point Arithmetic for FPGA Compatibility

FPGAs lack native support for floating-point arithmetic, and using floating-point IP cores can introduce significant resources and latency overhead. To overcome this, fixed-point representation, expressing numbers as scaled integers, is commonly used. It reduces computational complexity and aligns better with FPGA architecture, making it the preferred choice for efficient, real-time machine learning implementations on FPGAs.

To ensure compatibility with FPGA hardware and maintain numerical precision, a fixed-point format of Q1.15 was adopted, where the total word length is 16 bits comprising 1 sign bit, 1 integer bit, and 15 fractional bits. This configuration allows representation of values in the approximate range of −1 to 0.99997 with a high precision of 0.00003. Such a format is particularly well-suited for ML inference tasks, including SVM decision functions, where values typically lie within a normalized or small range. The Q1.15 format strikes a favorable balance by providing fine-grained resolution while remaining efficient for FPGA logic implementation. Among common 16-bit fixed-point formats (such as Q4.11 or Q8.8), Q1.15 is often preferred when high precision is critical and the dynamic range of values is limited, as is frequently the case in embedded AI applications like solar hotspot classification.

Logic Gate-Level Simulation for Solar Hotspot Classification

To evaluate the feasibility of real-time deployment in an Unmanned Aerial Vehicle (UAV) platform, the trained SVM model was translated into hardware using VHDL. The logic implements binary classification to detect the presence (Class 1) or absence (Class 0) of solar panel hotspots from MPEG-7 feature vectors. The VHDL module is structured around a three-state FSM: IDLE, COMPUTE, and FINISH, and tested using a test-bench.

In the IDLE state, control signals, index counters (i, j), and accumulator registers are initialized, with the system waiting for the start signal to be asserted. Once activated, the FSM transitions to the COMPUTE state, where it iteratively calculates the squared Euclidean distance between the normalized input feature vector and each stored support vector using fixed-point signed arithmetic. This distance value is then used in a hardware-friendly Gaussian RBF kernel approximation, implemented via a rational function or lookup table to avoid floating-point operations. The resulting kernel values are multiplied by the corresponding Lagrange multipliers (α_i) and class labels $(label_i)$, and the products are accumulated into a decision variable (fx_{accum}) according to $fx_{accum}+\alpha_i * label_i * kernel_i$. Upon completing all iterations, the FSM enters the FINISH state, where the bias term b is added to the accumulated decision value, the sign of the result is evaluated, and a binary prediction is generated on the predicted signal. All state transitions are synchronized to the rising edge of the clock signal, ensuring deterministic and cycle-accurate execution.

Fixed-Point Arithmetic and Kernel Logic

The use of Q1.15 signed fixed-point representation (signed (47 downto 0) for accumulators and kernel computation) avoids the overhead of floating-point units, making this implementation resource-efficient for FPGA targets.

- SVM Decision Function Gaussian (RBF) Kernel: The classifier computes the decision function as:

$$f(x) = \sum_{i=1}^{N} \alpha_i . y_i . K(x_i . x) + b \qquad (1)$$

where x_i: support vector, y_i: label of support vector, α_i: Lagrange multiplier, b : bias

$$K(x_i.x) = \exp\left(-\gamma.\left||x - x_i\right|\right|^2\right) \qquad (2)$$

where γ: kernel scaling factor, $K(x_i.x)$: Gaussian kernel value between the input vector x and i^{th} support vector (x_i) approximated for hardware implementation.

It is required to normalize the input feature vector x such that $x' = \frac{x-\mu}{\sigma}$ by Z-score normalization, where x': Standardized value, x: Original data value, μ: Mean of the data, σ: Standard deviation of the data. Support vectors, labels, and Lagrange multipliers are essential in computing the kernel sum. γ defines the kernel function shape. μ and σ must be used to preprocess input features before classification.

- Distance squared:

$$dist_{sq} = \sum_{j=0}^{N}(x_j - sv_{i,j})^2 \qquad (3)$$

is calculated in pipelined j iterations. $dist_{sq}$ is Squared Euclidean distance between the input feature vector x and i^{th} support vector. x_j is the j^{th} element of the normalized input feature vector. $sv_{i,j}$ is j^{th} feature of the $i-$ th support vector in the trained SVM model. N is the total number of features in the input feature dimension vector. i is the index of the current support vector being processed and j is the index of the current feature element in the input vector.

- Gaussian kernel approximation:

$$K(x, x_i) \approx \frac{10}{1+\gamma.dist_{sq}} \qquad (4)$$

where $K(x_i.x)$: Gaussian kernel value is implemented in the exponential approximation function using division-based rational approximation, suitable for hardware. $dist_{sq}$ is Squared Euclidean distance and γ is kernel scaling factor. Decision value is accumulated as in equation (1) across all support vectors and compared to a fixed threshold for a binary decision.

The final stage of the workflow involves hardware simulation and validation to ensure deployment readiness. The fixed-point VHDL implementation of the SVM classifier is simulated using GHDL, with signal waveforms analyzed in GTKWave to verify correct state transitions, timing behavior, and numerical equivalence with the MATLAB floating-point model. Successful validation confirms that the FPGA design can reliably perform real-time solar hotspot classification, enabling seamless integration into UAV-based PV inspection systems.

In summary, the methodology bridges the gap between MATLAB-trained SVM classifiers and their FPGA implementation through a structured process involving parameter extraction, fixed-point conversion in Q1.15 format, and custom VHDL coding. The classifier is realized as a state-machine-based logic module that performs kernel evaluations and accumulates decision values using resource-efficient arithmetic. Simulation results validate the equivalence of fixed-point hardware behavior with the original floating-point model, ensuring deployment readiness for embedded solar hotspot detection applications.

RESULTS

This section presents the experimental evaluation of the Medium Gaussian SVM model trained on MPEG-7 feature descriptors for detecting thermal anomalies in PV modules. The classifier was assessed across multiple datasets of varying sizes to analyze its performance in terms of accuracy, F1-score, precision, recall, and execution time. Furthermore, logic-level hardware simulation of the fixed-point SVM classifier was conducted to verify its effectiveness in an FPGA implementation suitable for real-time UAV-based PV inspection. The results are discussed under several subsections. Multi-Metric Evaluation Across Datasets

validates classification robustness across different dataset scales. Execution Time Analysis assesses the feasibility of near real-time deployment. Precision-Recall Tradeoff and Model Consistency examines the balance between false positives and false negatives in hotspot detection. Logic Gate-Level Analysis and Simulation of SVM Classifier ensures accurate functional mapping of the software model into hardware. Simulation Insights interprets timing waveforms and state transitions, confirming correct classifier operation in FPGA environments.

Multi-Metric Evaluation Across Datasets

Table 1 summarizes the training and testing performance of the Medium Gaussian SVM classifier across five datasets. The classifier maintained high accuracy and generalization throughout.

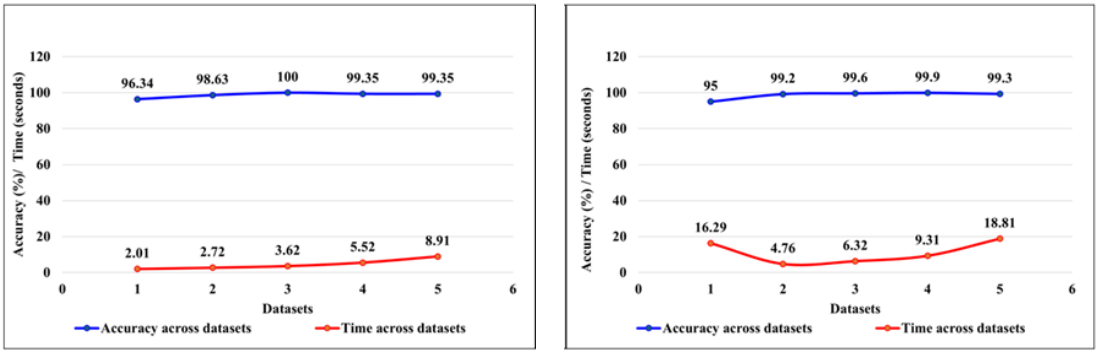
Table 1: Training and Testing Performance across Multiple Datasets

Phase	Dataset Size	Time (seconds)	Accuracy (%)	F1-Score	Precision	Recall
Training						
Dataset 1	724	16.29	95.00	0.9490	0.9570	0.9419
Dataset 2	1302	4.76	99.20	0.9920	0.9908	0.9938
Dataset 3	1836	6.32	99.60	0.9950	1.0000	0.9920
Dataset 4	2746	9.31	99.90	0.9984	1.0000	0.9970
Dataset 5	4114	18.81	99.30	0.9920	0.9920	0.9930
Testing						
Dataset 1	82	2.01	96.34	0.9640	0.9520	0.9760
Dataset 2	146	2.72	98.63	0.9860	0.9860	0.9860
Dataset 3	204	3.62	100.00	1.0000	1.0000	1.0000
Dataset 4	306	5.52	99.35	0.9930	1.0000	0.9870
Dataset 5	458	8.91	99.35	0.9930	0.9910	0.9960

It was observed that the Medium Gaussian SVM achieved up to 100% testing accuracy, with F1-scores greater than 0.99 for larger datasets, demonstrating strong model robustness and generalization.

Execution Time Analysis

To assess deployment viability, we examined the total runtime during inference. Even as dataset sizes increased, inference times remained within practical bounds, making the model suitable for onboard FPGA-based inference in real-time UAV inspections.



(a) (b)

Figure 2. Dataset-wise trade-off between accuracy and time (a) training, (b)testing

As shown in Figure 2-a, the training time exhibited a gradual increase with growing dataset size yet remaining under 20 seconds even for the largest dataset considered. This indicates efficient scalability in the training process. Similarly, as illustrated in Figure 2-b, the testing time grew almost linearly with the dataset size, demonstrating a predictable and manageable computational load. Importantly, there was no significant trade-off observed between computational speed and classification accuracy, that the model maintained high performance across all datasets without sacrificing efficiency. The scalability and low computational overhead make the fixed-point SVM model viable for real-time solar hotspot detection.

Precision-Recall Tradeoff and Model Consistency

Despite dataset variations, the model maintained high precision and recall balance, as per the observations. The precision consistently greater than 95.00% with minimal false positives. Recall up to 99.70%, conveying rare cases of missed detections. F1-score peaked at 1.0000 on dataset 3 with perfect harmonic mean between precision and recall.

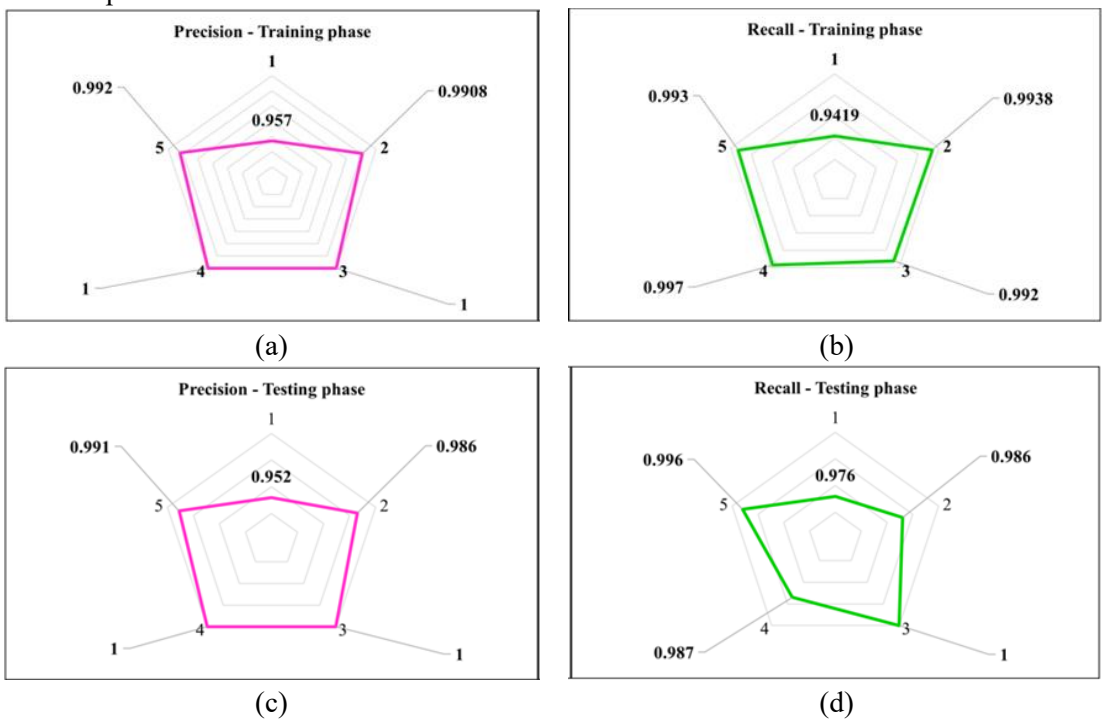


Figure 3. Radar plot of precision and recall across five datasets (a) and (b) training, (c) and (d) testing

Figure 3 (a) and (b), illustrate the radar plots for precision and recall metrics across five datasets during the training phase. The training precision shows a consistent upward trend, with dataset 1 yielding a moderate precision of 0.9570, and datasets 3 and 4 achieving perfect precision of 1.0000, indicating zero false positives. Recall values also improved across datasets, starting at 0.9419 for dataset 1 and reaching 0.9970 for dataset 4. This steady improvement is mirrored in the F1-scores and accuracy, which peak at 99.90% in dataset 4. Notably, training times are proportionate to dataset size, with larger datasets requiring more computation, culminating in 18.81 seconds for dataset 5. Figure 3 (c) and (d), present precision and recall radar plots during the testing phase. Despite the smaller dataset sizes, the model demonstrates robust generalization, with precision values ranging from 0.9520 to 1.0000, and recall from 0.9760 to 1.0000. The peak testing performance is recorded on dataset 3, where the model achieves 100.00% accuracy, perfect F1-score, and zero misclassifications. Datasets 4 and 5 also maintain strong consistency with high precision

0.9910 and recall 0.9960. These results validate the strong performance of fixed-point SVM model, even in unseen data, reinforcing its suitability for edge inference scenarios.

Therefore, it is justifiable to determine that the Q1.15 fixed-point Medium Gaussian SVM delivers high detection accuracy with low execution time across varying dataset sizes. Feature-level fusion of MPEG-7 descriptors amplifies hotspot differentiation capability. Therefore, the results support FPGA deployment for real-time onboard PV inspection.

Logic Gate-Level Analysis and Simulation of SVM Classifier

To transition the trained Medium Gaussian SVM model into a deployable embedded format suitable for UAV edge devices, a hardware-level implementation was developed using VHDL. The module svm_core.vhd simulates the classification behavior using fully deterministic finite-state control logic and signed fixed-point arithmetic.

The VHDL implementation of the SVM predictor in this design follows a structured pipeline that uses basic arithmetic logic and state machine control to realize binary classification for solar hotspot detection. The core architecture leverages fundamental logic gate operations including addition, subtraction, squaring, and multiplication, all orchestrated using a FSM with states: IDLE, COMPUTE, and FINISH. The inner product calculation, which is the kernel distance, is done using repeated subtraction and squaring, followed by accumulation of the weighted kernel contributions. The decision logic uses a final comparison gate to determine whether the output exceeds the bias threshold, which is used to assign the prediction class (1 or 0). The design computes the RBF kernel directly via arithmetic means, which is FPGA-synthesizable and offers hardware interpretability. This method ensures efficient SVM evaluation without memory overheads from precomputed exponential values.

Simulation Insights

The testbench mimics real-world activation by applying start and reset signals. Simulation output provides internal signal visibility via report statements. Example outputs include:

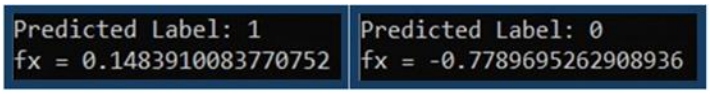


Figure 4. Prediction 1 and 0

Figure 4 shows that the model correctly identifies a hotspot prediction = '1' and '0', when fx_{total} exceeds the threshold. This aligns with expectations from software SVM classification using the same support vectors. Figure 5 and Figure 6 represent the respective waveforms from the GTKWave interface.

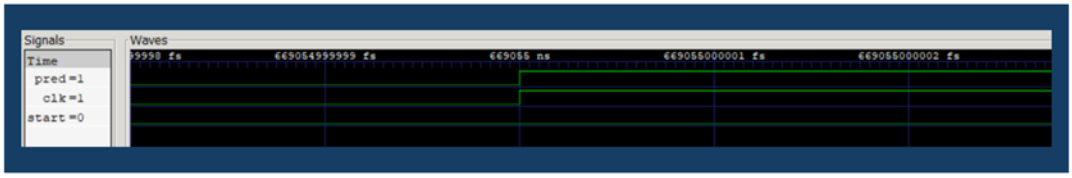


Figure 5. GTK View Output Prediction 1

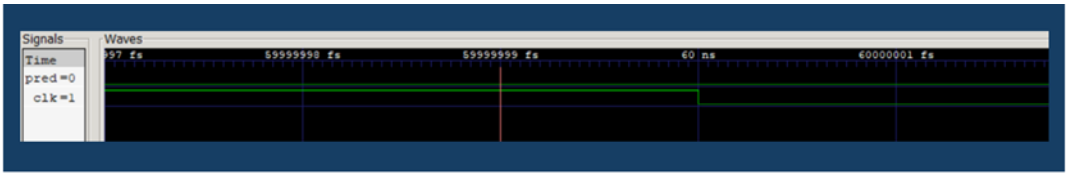


Figure 6. GTK View Output Prediction 0

In summary, the Medium Gaussian SVM classifier achieved consistently high detection accuracy, with F1-scores exceeding 0.99 across large datasets and even reaching 100% testing accuracy in select cases. The Q1.15 fixed-point implementation demonstrated low execution time and robust generalization, validating its readiness for real-time FPGA deployment. Hardware-level simulation confirmed the correctness of the classifier using VHDL, where logic gates and FSM-based control executed Gaussian kernel operations without look-up table (LUT) dependency. Not using a LUT computes values directly, which typically results in higher numerical accuracy, especially for fine-grained inputs. These results confirm the viability of the proposed model for embedded solar hotspot detection in edge devices such as UAVs.

DISCUSSION

This section analyzes the rationale, performance implications, and deployment viability of the fixed-point Medium Gaussian SVM model for real-time solar hotspot detection. It highlights the necessity of a custom MATLAB implementation for hardware translation, evaluates the impact of dataset size on model performance, examines generalization and testing accuracy, and discusses hardware-software coherence for UAV deployment. Additionally, statistical consistency of the performance metrics across datasets is examined to affirm its robustness and reliability and to validate the significance of observed performance trends quantitatively. The discussion is organized under five subsections: Rationale for Developing a Custom SVM Function in MATLAB Prior to HDL Conversion, Impact of Dataset Size on Training Performance, Generalization and Testing Accuracy, Hardware-Software Coherence and Deployment Viability on UAVs, and Statistical Analysis of Classifier Performance. Each subsection explicitly interprets the results presented earlier, providing clarity on why the observed metrics are meaningful.

Rationale for Developing a Custom SVM Function in MATLAB Prior to HDL Conversion

A custom SVM function was implemented in MATLAB to replicate the core decision function behavior using fixed-point arithmetic. This step is essential because built-in SVM models and prediction methods in MATLAB are not directly compatible with HDL Coder due to their object-oriented nature and reliance on floating-point operations. By manually reconstructing the SVM using extracted parameters, the model could be precisely tailored for fixed-point arithmetic, enabling accurate and efficient translation to VHDL. This approach ensures transparency, full control over precision and word length, and facilitates seamless simulation and verification. It also bridges the gap between high-level training and hardware-level implementation, which is critical for FPGA-based machine learning applications. To enable hardware compatibility, the SVM logic is converted into fixed-point arithmetic. Based on this, synthesizable VHDL code is written, including both the main predictor module and a corresponding test bench for simulation. GHDL is used to verify the functional behavior of the VHDL implementation. Once validated, the design can be synthesized and implemented using FPGA toolchains such as Intel Quartus, and the final bitstream is deployed to the FPGA hardware for real-time solar hotspot classification.

Impact of Dataset Size on Training Performance

The radar plots and tabulated metrics indicate that increasing the dataset size enhances both precision and recall during training. This can be attributed to the improved representation of data distribution, allowing the Medium Gaussian SVM to learn more generalizable decision boundaries. The jump from 95.00% accuracy in Dataset 1 to nearly 100.00% in Dataset 4 highlights this trend. Furthermore, the Q1.15 fixed-point arithmetic does not introduce noticeable quantization errors, as evidenced by the consistent alignment between fixed-point and floating-point results. This confirms that the model retains its discriminative power even under hardware-friendly numeric constraints.

Generalization and Testing Accuracy

During testing, the model sustains high performance across all five datasets, validating its robustness and transferability from training to deployment. The consistency in recall across test sets - particularly the perfect scores in dataset 3 and high scores in others- demonstrates the capability of the model to detect hotspots without omission. Minor dips in precision for datasets 1 and 4, such that 0.9520 and 1.0000, respectively, suggest rare false positives, possibly due to edge cases in smaller test samples. Nonetheless, the GHDL-verified VHDL simulation maintains high fidelity with MATLAB predictions, reinforcing that the implemented architecture is both accurate and efficient. These outcomes strongly advocate for its FPGA realization in real-time solar inspection systems.

Hardware-Software Coherence and Deployment Viability on UAVs

The VHDL-implemented SVM classifier replicates the behavior of its MATLAB-trained floating-point counterpart, confirming the correctness of the fixed-point hardware design. Its deterministic latency is a direct consequence of the FSM structure and parallel feature processing: the number of clock cycles depends on the number of support vectors and features, yielding 2702 cycles for 300 SVs and 222 features in the current design. The Gaussian RBF kernel is approximated using a rational function, avoiding costly exponentials and look-up tables, which makes the design lightweight and well-suited for resource-constrained FPGAs. The model supports further optimizations such as pipelining, early termination, and precision scaling, enhancing efficiency and reducing power consumption. Overall, the design is compact, predictable, and ideal for real-time inference on mid-tier FPGAs in UAV-based PV hotspot detection.

Statistical Analysis of Classifier Performance

To quantitatively validate the consistency and significance of the performance of Medium Gaussian SVM across different dataset sizes, a statistical analysis was conducted using standard evaluation metrics namely Mean, Standard Deviation, Coefficient of Variation (CV) and Range. The consistency metrics across training and testing phases indicate strong generalization, with testing performance slightly outperforming training in terms of accuracy, F1-score, and recall demonstrating minimal overfitting and greater stability on unseen data. Notably, lower standard deviations, coefficients of variation, and narrower ranges in the test set confirm that the model delivers reliable and consistent results across folds as visualized from Figure 7 (a) to (d).

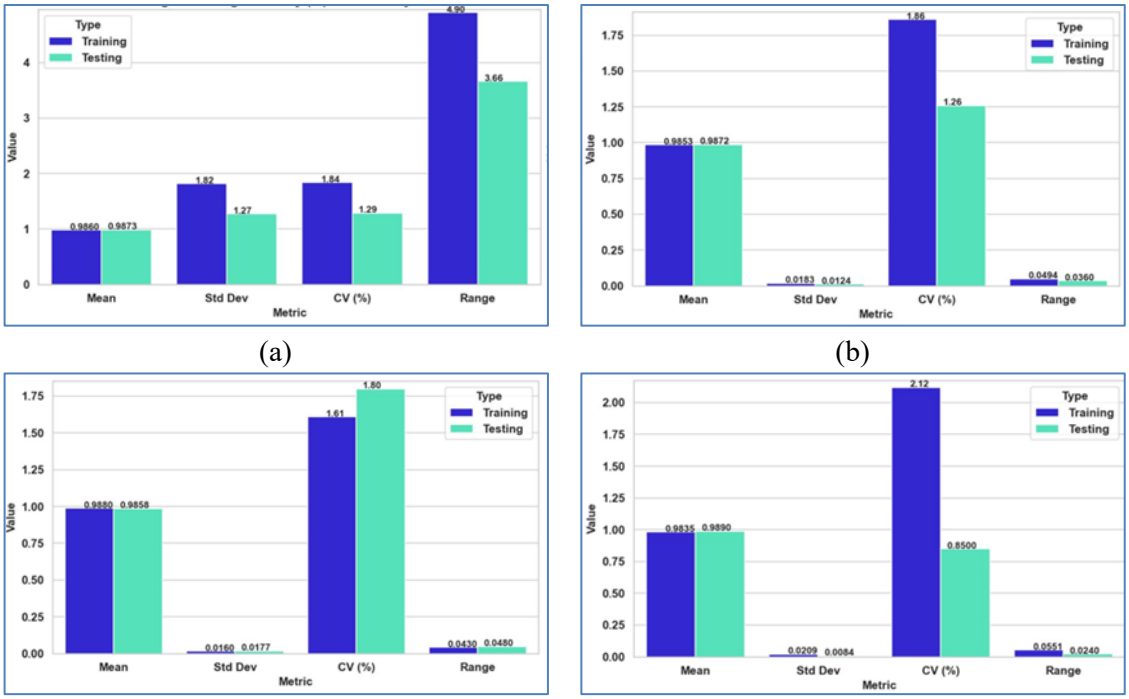


Figure 7. Training and testing consistency metrics for (a) accuracy, (b) f1-score, (c) precision, (d) recall

According to Figure 7-a, the model demonstrates excellent generalization, as reflected by the slightly higher mean accuracy on the test set (98.73%) compared to the training set (98.60%), with minimal overfitting. Figure 7-b shows that the standard deviation and CV for test accuracy (1.2731, 1.29%) are lower than those for training (1.8166, 1.84%), indicating more consistent performance across test folds. The F1-score metrics, also shown in Figure 7-b, further support this trend: the mean F1-score is slightly higher for testing (0.9872) than training (0.9853), with lower variability (standard deviation of 0.0124 vs. 0.0183, and CV of 1.26% vs. 1.86%). The range of F1-scores is narrower for testing (0.0360) than training (0.0494), suggesting better balance and stability in precision and recall on unseen data. Figure 7-c presents precision metrics, where the mean value is high for both training (0.9880) and testing (0.9858), though the test set shows slightly more variability (standard deviation: 0.0177 vs. 0.0160; CV: 1.80% vs. 1.61%), and a wider range (0.0480 vs. 0.0430). Figure 7-d illustrates the recall metrics, with testing showing higher mean recall (0.9890 vs. 0.9835), significantly lower standard deviation (0.0084 vs. 0.0209), and a much smaller CV (0.85% vs. 2.12%), along with a tighter range (0.0240 vs. 0.0551). These results suggest the model not only captures more true positives in the test set but does so more consistently, reinforcing its reliability for real-world solar hotspot detection.

In summary, the custom fixed-point SVM implementation successfully bridges software modeling and hardware deployment by preserving classification fidelity while enabling efficient VHDL synthesis. The model demonstrates consistent, high performance across datasets with minimal variance, confirming its generalization capability. Hardware simulation further validates alignment with software predictions, showing deterministic latency and potential for FPGA optimizations. Together, these findings support the suitability of the model for real-time, resource-constrained UAV operations in solar inspection tasks.

CONCLUSION

This study presented a fixed-point Medium Gaussian SVM classifier optimized for thermal anomaly detection in PV modules using MPEG-7 shape descriptors. A complete pipeline was developed, starting from MATLAB-based preprocessing and SVM training to logic-level VHDL implementation for FPGA deployment. The classifier consistently achieved high accuracy and F1-scores across datasets of varying sizes, even attaining perfect classification in certain test sets. The conversion from floating-point to Q1.15 fixed-point arithmetic was essential to ensure compatibility with resource-constrained hardware platforms such as UAV-mounted FPGAs. The logic-based SVM implementation successfully replicated the behavior of the MATLAB model, as verified through detailed simulation, kernel output validation, and prediction agreement. Furthermore, execution latency remained within microsecond-scale limits, affirming its suitability for real-time detection during aerial PV inspections.

Overall, the proposed approach demonstrates that handcrafted, fixed-point SVM classifiers can maintain software-level accuracy while meeting the stringent performance, power, and area constraints of edge devices. This work provides a reliable and hardware-efficient solution for intelligent fault detection in solar energy systems, potentially contributing to improved PV module maintenance, reduced energy loss, and enhanced system longevity.

While the proposed fixed-point SVM implementation achieves strong performance and FPGA compatibility, several opportunities for future enhancement remain. Upcoming work will investigate dynamic reconfiguration and online retraining mechanisms, potentially through partial reprogramming of FPGAs or hybrid FPGA-CPU architectures. Integration with lightweight deep learning models or ensemble methods may further improve the detection of complex anomalies. Additionally, deploying the system on actual UAVs with real-time thermal imaging sensors will enable field validation and calibration against diverse irradiance and angle-of-view conditions. Power optimization techniques and resource-aware design will also be investigated to extend the endurance of UAV operations during continuous inspection missions. Hybrid approaches, such as edge–cloud collaboration, may offer a way to mitigate these trade-offs. As noted in [20], emerging paradigms like TinyML, Transformer variants, and architectures such as Mamba provide promising directions for balancing accuracy and efficiency in real-time applications.

ABBREVIATIONS

ACE: Automatic Content Extractor
AI: Artificial Intelligence
BGLR: Binary GLM (Generalized Linear Model) Logistic Regression
CV: Coefficient of Variation
DL: Deep Learning
FPGA: Field Programmable Logic Array
FSM: Finite State Machine
GHDL: G Hardware Description Language
GTKWave: GNU Tool Kit Wave
HDL: Hardware Description Language
IP: Intellectual Property
IR: Infrared
LUT: Look-Up Tables
ML: Machine Learning
MPEG: Moving Picture Experts Group
PV: Photovoltaic
RBF: Radial Basis Function

SVM: Support Vector Machine
UAV: Unmanned Aerial Vehicle
VHDL: VHSIC (Very High-Speed Integrated Circuit) Hardware Description Language

REFERENCES

[1] Nayomi Fernando, Lasantha Seneviratne, Nisal Weerasinghe, Namal Rathnayake, and Yukinobu Hoshino. 2025. Efficient hotspot detection in solar panels via computer vision and machine learning. Information 16, 7 (July 2025), 608. https://doi.org/10.3390/info16070608

[2] Muhammad Umair Ali, Hafiz Farhaj Khan, Manzar Masud, Karam Dad Kallu, and Amad Zafar. 2020. A machine learning framework to identify the hotspot in photovoltaic module using infrared thermography. Solar Energy 208 (2020), 643–651. https://doi.org/10.1016/j.solener.2020.08.027

[3] H. Hussein. 2023. Photovoltaic hotspot detection with image classification deep learning techniques. In Proceedings of the 2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI). IEEE, Zarqa, Jordan, 1–6. https://doi.org/10.1109/EICEEAI60672.2023.10590305

[4] Mahmoud Dhimish, Marios Theristis, and Vincenzo d'Alessandro. 2024. Photovoltaic hotspots: A mitigation technique and its thermal cycle. Optik 300 (2024), 171627. https://doi.org/10.1016/j.ijleo.2024.171627

[5] H. Açıkgöz, D. Korkmaz, and Ç. Dandıl. 2022. Classification of hotspots in photovoltaic modules with deep learning methods. TJST 17, 2 (2022), 211–221. https://doi.org/10.55525/tjst.1158854

[6] J. Wang, D. Gao, S. Zhu, S. Wang, and H. Liu. 2019. Fault diagnosis method of photovoltaic array based on support vector machine. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 45, 2 (2019), 5380–5395. https://doi.org/10.1080/15567036.2019.1671557

[7] C. J. Burges. 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 2 (June 1998), 121–167. https://doi.org/10.1023/A:1009715923555

[8] Fengxi Zhang, Yuying Li, and Zhihao Ye. 2022. Apply YOLOv4-Tiny on an FPGA-based accelerator of convolutional neural network for object detection. In Proceedings of the 2nd International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA 2022). IOP Publishing, Dali City, China, Article 012032. https://doi.org/10.1088/1742-6596/2303/1/012032

[9] Karuppiah Natarajan, Praveen Kumar B., and Vankadara Kumar. 2020. Fault detection of solar PV system using SVM and thermal image processing. International Journal of Renewable Energy Research 10 (2020), 967–977.

[10] X. Song, H. Wang, and L. Wang. 2014. FPGA implementation of a support vector machine based classification system and its potential application in smart grid. In Proceedings of the 2014 11th

International Conference on Information Technology: New Generations (ITNG). IEEE, Las Vegas, NV, 397–402. https://doi.org/10.1109/ITNG.2014.45

[11] Mateusz Wasala and Tomasz Kryjak. 2022. Real-time HOG+SVM based object detection using SoC FPGA for a UHD video stream. arXiv:2204.10619 [cs.CV]. https://doi.org/10.48550/arXiv.2204.10619

[12] A. Escobedo and A. Lin. 2016. Tessellation-based multi-block memory mapping scheme for high-level synthesis with FPGA. In Proceedings of the 2016 International Conference on Field-Programmable Technology (FPT). IEEE, Xi'an, China, 125–132. https://doi.org/10.1109/FPT.2016.7929517

[13] X. Ma, W. A. Najjar, and A. K. Roy-Chowdhury. 2015. Evaluation and acceleration of high-throughput fixed-point object detection on FPGAs. IEEE Transactions on Circuits and Systems for Video Technology 25, 6 (June 2015), 1051–1062. https://doi.org/10.1109/TCSVT.2014.2360030

[14] H. Brum, M. Véstias, and H. Neto. 2024. LiDAR 3D object detection in FPGA with low bitwidth quantization. In Proceedings of Applied Reconfigurable Computing. ARC 2024. Lecture Notes in Computer Science, vol. 14553. Springer, Cham, 90–105. https://doi.org/10.1007/978-3-031-55673-9_7

[15] Muhammad Ihsan Al Hafiz, Naresh Ravichandran, Anders Lansner, Pawel Herman, and Artur Podobas. 2025. A reconfigurable stream-based FPGA accelerator for Bayesian confidence propagation neural networks. arXiv:2503.01561 [cs.AR]. https://doi.org/10.48550/arXiv.2503.01561

[16] Srikanth Ramadurgam and Darshika G. Perera. 2021. An efficient FPGA-based hardware accelerator for convex optimization-based SVM classifier for machine learning on embedded platforms. Electronics 10, 11 (May 2021), 1323. https://doi.org/10.3390/electronics10111323

[17] Shereen Afifi, Hamid GholamHosseini, and Roopak Sinha. 2019. A system on chip for melanoma detection using FPGA-based SVM classifier. Microprocessors and Microsystems 65 (2019), 57–68. https://doi.org/10.1016/j.micpro.2018.12.005

[18] Shereen Afifi, Hamid GholamHosseini, and Roopak Sinha. 2015. Hardware implementations of SVM on FPGA: A state-of-the-art review of current practice. International Journal of Innovative Science, Engineering & Technology (IJISET) 2 (2015), 733–752.

[19] Mohammed H. Yacoub, Samar M. Ismail, Lobna A. Said, Ahmed H. Madian, and Ahmed G. Radwan. 2024. Support vector machine reconfigurable hardware implementation on FPGA. Franklin Open 7 (2024), 100115. https://doi.org/10.1016/j.fraope.2024.100115

[20] T. Suwannaphong, F. Jovan, I. Craddock, et al. 2025. Optimising TinyML with quantization and distillation of transformer and mamba models for indoor localisation on edge devices. Scientific Reports 15 (March 2025), 10081. https://doi.org/10.1038/s41598-025-94205-

Synergistic Charge Dynamics and Light Harvesting in TiO_2/MgO Composites for Efficiency Enhancement in CdS Quantum Dot-Sensitized Solar Cells

N.F. Ajward*, J.V.P. Fernando and V.P.S. Perera
Department of Physics, The Open University of Sri Lanka, Nawala, Nugegoda

***nf2010.ajward@gmail.com, jvfer@ou.ac.lk & vpper@ou.ac.lk**

ABSTRACT

Quantum dot-sensitized solar cells (QDSSCs) represent a promising advancement in renewable energy technologies, with recent improvements achieving power conversion efficiencies close to 6%. Structurally similar to dye-sensitized solar cells (DSSCs), QDSSCs employ quantum dots (QDs) as sensitizers that absorb photons and inject excited electrons into the conduction band of a wide-bandgap semiconductor electrode. while a redox electrolyte removes the generated holes, completing the circuit by regenerating them at the counter electrode. Quantum dots composed of materials such as CdS, CdSe, PbS, and InP are increasingly studied for use in QDSSCs, offering the advantage of tunable optical band gaps through particle size manipulation. This adaptability enhances QDSSCs' design potential, enabling the integration of third-generation solar cell configurations, including multiple exciton generation (MEG), to further enhance energy conversion efficiency.

Despite these advancements, QDSSC performance is currently limited by issues such as reduced photovoltage and recombination losses at the TiO_2-QD-electrolyte interface. This study explores the use of magnesium oxide (MgO) coatings on TiO_2 nanoparticles to address these limitations, focusing on improving the fill factor (FF) and overall cell efficiency. MgO serves as an electron-blocking layer, effectively reducing recombination and associated energy losses. Furthermore, MgO facilitates electron transport from QDs to the TiO_2 electrode, improving charge collection. The light-scattering properties of MgO also increase the photon's path length within the cell, enhancing light absorption and consequently boosting the short-circuit current.

In this study, MgO powder was incorporated in specific mass ratios with TiO_2, followed by the application of CdS quantum dots (QDs) on the TiO_2/MgO composite layer using the SILAR method. Results indicated a significant improvement in the fill factor (FF) at an optimal MgO-to-TiO_2 ratio, attributed to synergistic effects of MgO on interface stabilization, reduced recombination, and enhanced charge transport. The optimized MgO-modified TiO_2 films achieved a current density of 1.946 mA, voltage of 437 V, and power of 0.121 W, reaching an efficiency of 0.311 (18.7% higher than TiO_2/CdS QDSCs), with improved interfacial impedance, Incident Photon to Current Efficiency (IPCE), and FF of 37.4%.

KEYWORDS: *QDSCs, MgO, TiO_2, CdS, SILAR, efficiency, Fill factor.*

INTRODUCTION

Quantum dot-sensitized solar cells (QDSCs) represent a promising class of third-generation photovoltaic technologies, offering cost-effective fabrication, tunable optical properties, and potential for enhanced energy conversion through mechanisms such as multiple exciton generation (Nozik, 2002; Kamat, 2013). Structurally analogous to dye-sensitized solar cells (DSSCs), QDSCs replace organic dyes with semiconductor quantum dots (QDs), which absorb incident photons and inject excited electrons into the conduction band of a wide-bandgap semiconductor, typically TiO_2. The redox electrolyte simultaneously extracts holes and regenerates the QDs, completing the electrochemical circuit (Sambur & Parkinson,

2010).

The use of QDs such as CdS, CdSe, PbS, and InP has gained traction due to their size-dependent bandgap tunability and high extinction coefficients, enabling broader spectral absorption than conventional dyes (Sargent, 2005; Beard et al., 2007). Among them, CdS quantum dots are widely used owing to their favorable band alignment with TiO_2 and straightforward synthesis via methods like Successive Ionic Layer Adsorption and Reaction (SILAR). However, QDSCs still suffer from performance limitations caused by rapid charge recombination at the TiO_2–QD–electrolyte interfaces and low open-circuit voltages (Chen et al., 2012; Gao et al., 2017).

To address these challenges, recent studies have explored interfacial engineering strategies, including the introduction of wide-bandgap metal oxide layers such as MgO. MgO coatings have demonstrated an ability to suppress interfacial recombination, improve charge selectivity, and enhance light scattering within the photoanode structure, thereby boosting both the short-circuit current and fill factor (Wang et al., 2014; Chang et al., 2016). Furthermore, insulating nature of MgO can act as an electron-blocking layer, mitigating electron back-transfer and increasing the carrier lifetime in the system (Liu & Kamat, 1993).

In this study, the influence of MgO incorporation into TiO_2 photoanodes for CdS-based QDSCs has been systematically investigated. By varying the MgO-to-TiO_2 mass ratio, we it is aimed to identify an optimal configuration that maximizes photovoltaic performance. CdS quantum dots were deposited via the SILAR method, and the resulting composites were characterized in terms of their fill factor, short-circuit current density, and overall power conversion efficiency. Our findings revealed a significant improvement in device performance at a specific MgO loading, highlighting the potential of MgO-modified TiO_2 as a low-cost and scalable strategy for advancing QDSC technology.

EXPERIMENTAL DETAILS

2.1 Preparation of TiO_2/MgO Films Varying Mass Ratio

MgO powder was measured according to the quantities specified in Table 1. For each sample, concentrated HNO_3 acid was added dropwise to the MgO powder and then ground using a mortar and pestle. TiO2 powder 0.25 g was subsequently mixed while adding, one drop of Triton X-100, and 0.05 g of PEG 1000. The resulting paste was spread onto pre-cleaned conducting tin oxide (CTO) glass plates (1.0 cm × 2.0 cm) using the doctor blade method. The cleaning process involved ultrasonic bath treatment with detergent and distilled water. Finally, the TiO_2 coated films were dried on a hot plate and sintered in a furnace at 450 °C for 45 minutes

Table 01. Mass of MgO for each sample

Sample	1	2	3	4	5	6	7	8	9
Mass of MgO(g)	0.0025	0.0050	0.0075	0.0100	0.0125	0.0150	0.0175	0.0200	0.0225
Percentage	1%	2%	3%	4%	5%	6%	7%	8%	9%

2.2 Deposition of CdS using SILAR method.

To prepare solutions containing precise concentrations of Cd^{2+} and S^{2-} ions, 10.065 g of $CdCl_2$ and 3.902 g of Na_2S were weighed with an electronic balance (±0.001 g accuracy). These compounds were then dissolved in 100 ml of distilled water. To optimize the solubility of relevant species and enhance the electrochemical behavior within the cell, the pH of the $CdCl_2$ solution was adjusted to 4.5. This was achieved by carefully adding 0.01 M NaOH or 0.01 M HCl solutions dropwise while closely monitoring the pH.

The SILAR method was employed to deposit CdS films on the TiO2/MgO photoanodes. Each dipping step into the CdCl2 and Na2S solutions lasted for one minute. After each dip, the photoanodes were rinsed with distilled water to remove any excess reagents, then air-dried. To ensure uniform film formation, the photoanodes were subsequently dipped into the opposing precursor solution (anionic or cationic), followed by washing and drying cycle. This entire sequence of steps was repeated ten times to make a multilayered CdS film. Finally, the coated photoanodes were heated on a hot plate at 80°C for 30 minutes to enhance the film's stability and adhesion to the substrate.

2.3 Preparation of Electrolyte

2ml of electrolyte solution was prepared by combining 0.1301g of Na2S, 0.1283g of S, 0.0301g of KCl mixed in Methanol (1.4ml) and distilled water (0.6ml) as solvents. The mixture was stirred for three hours using a magnetic stirrer to ensure complete dissolution of the components and prevent the precipitation of sulfur. The stirring time might require adjustment if sulfur precipitation is observed

2.4 Fabrication of the cell

Conducting side of Pt counter electrode was placed face-to-face with the CdS coated TiO_2 or TiO_2/MgO film and secured with two clamps. The space between the electrodes was then filled with the electrolyte solution, completing the QDSSC assembly.

2.5 Characterization

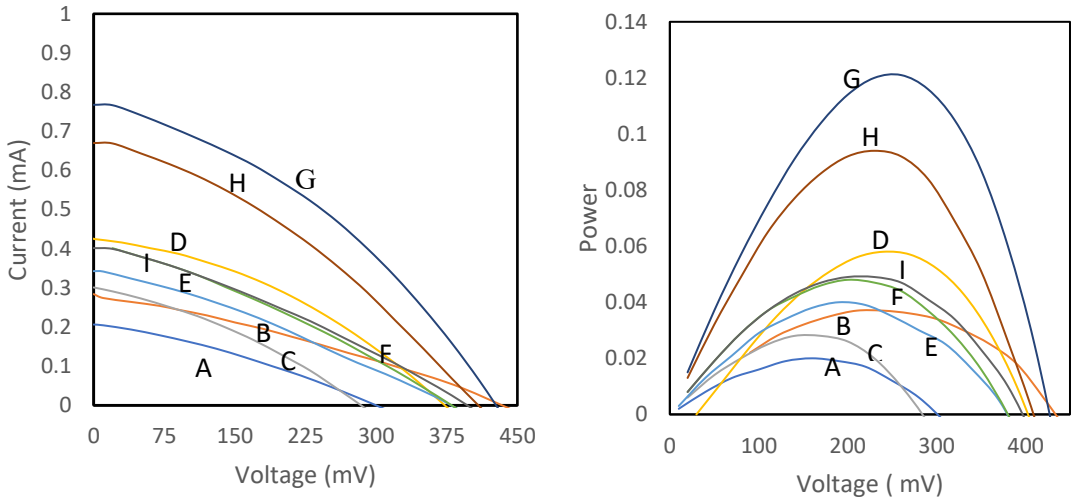
The fabricated QDSSCs were rigorously characterized using a VK-PA-100 PV Power Analyzer to obtain J–V curves, from which key photovoltaic parameters such as photo-conversion efficiency (PCE), short-circuit current density (Jsc), open-circuit voltage (Voc), and fill factor (FF) were extracted. To ensure reproducibility, a minimum of five devices per batch were evaluated. Electrochemical Impedance Spectroscopy (EIS), conducted with an Autolab Nova 2.1 frequency response analyzer, was employed to investigate charge transport and recombination mechanisms over a frequency range of 0.1 Hz to 1 MHz under simulated solar illumination. This provided critical insight into interfacial resistance and electron lifetime within the photoanode. The Incident Photon-to-Current Efficiency (IPCE) was assessed using a VK-IPCE-10 system to determine the wavelength-dependent photo response and light-harvesting efficiency of the devices.

In addition to electrochemical and photovoltaic analyses, comprehensive structural and morphological characterizations were performed. X-ray Diffraction (XRD) was used to examine the crystallinity and phase composition of the composite materials, while Fourier Transform Infrared Spectroscopy (FTIR) provided information on chemical bonding and surface modifications. Scanning Electron Microscopy (SEM) was utilized to observe the surface morphology and distribution of MgO within the TiO_2/CdS matrix, confirming uniform dispersion and structural refinement. These combined techniques enabled a holistic understanding of the material properties and their influence on the overall performance of the QDSSCs.

RESULTS AND DISCUSSION.

3.1 I–V characteristics of CdS QDSSCs

Figure 1 presents the current-voltage (I-V) and power-voltage (P-V) characteristics of TiO_2/MgO/CdS thin films containing varying MgO mass content, ranging from 1% to 9%.



(a) (b)

[A-MgO 1%, B-MgO 2%, C-MgO 3%, D-MgO 4%, E-MgO 5%, F-MgO 6%, G-MgO 7%, H-MgO 8% and I-MgO 9%,]

Figure 1. (a) Current-voltage plot for mass ratios ranging from 1% to 9%. (b) Power-voltage plot for mass ratios from 1% to 9%.

Table 02: The photovoltaic data obtained for different mass ratio ranging from 1% to 9%

	1%	2%	3%	4%	5 %	6%	7%	8%	9%
V_{oc} (mV)	299.6	433.4	283.8	372	379.3	380.2	437	408.5	396.3
Jsc (mA cm-2)	0.521	0.718	0.739	0.867	0.872	1.019	1.946	1.698	1.014
FF (%)	0.323	0.307	0.346	0.369	0.307	0.316	0.374	0.349	0.316
η	0.05	0.096	0.073	0.119	0.101	0.122	**0.311**	0.242	0.127
Power(mW)	0.02	0.037	0.028	0.058	0.04	0.048	**0.121**	0.094	0.049
V_{max} (mV)	160.1	239.9	170.1	210.0	190.1	200.0	260.2	229.9	229.9
I_{max} (mA)	0.123	0.155	0.166	0.278	0.208	0.239	0.466	0.410	0.215

The addition of MgO to TiO_2 significantly enhances the performance of CdS quantum dot-sensitized solar cells. Although MgO is an insulator, it plays a crucial role in improving efficiency by reducing charge recombination, enhancing electron transport, and increasing light absorption. This leads to a notable increase in both current and voltage. Among the various mass variations of MgO studied, 7% addition yielded the highest efficiency of 0.311, surpassing the efficiency of TiO_2/CdS QDSCs (0.262) by a significant 18.7%. These findings highlight the potential of MgO as a promising material for improving the performance of QDSCs.

The solar cell efficiency(η) is function of the basic parameters as presented by

$$\eta = \frac{Voc\ X\ I\,sc\ X\ \ FF}{Pin\ X\ S}$$

where, Pin is the power of the incident light radiation and S = 0.49 cm^2 is the surface of illuminated photoanode. The fill factor of the device was calculated based on the relation,

$$\mathrm{FF} = \frac{I\max\ X\ \ V\,max}{I\,sc\ X\ Voc}$$

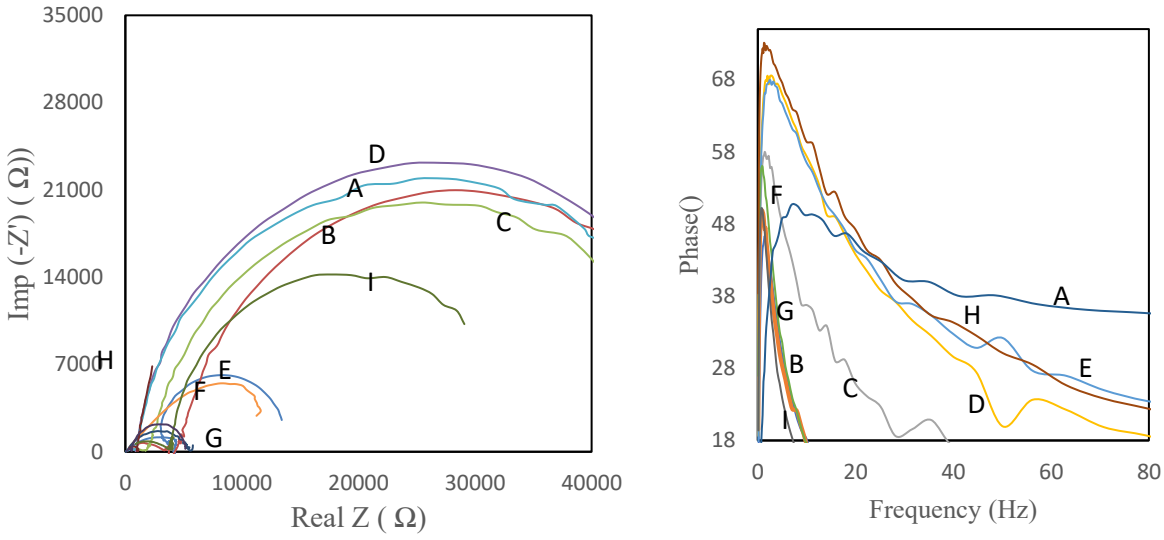
The MgO coating acts as a buffer layer, blocking back electron transfer and improving the fill factor. Here figure 1 and table 1 shows the higher fill factor of MgO-coated TiO_2-based QDSCs (0.374) compared to TiO_2/CdS QDSCs (0.265). Therefore, the fill factor of MgO-coated TiO_2-based QDSCs is approximately 41.13% higher than that of TiO_2/CdS QDSCs.

figure 2 and Table 2 show that the highest power output (0.121 mW) was obtained with a 7% MgO mass variation. This improvement is attributed to MgO's light-scattering properties, which enhance light absorption and lead to higher power generation.

3.2 EIS of CdS QDSSCs

The Nyquist plot in Figure 3 illustrates the impedance behavior of the TiO_2/MgO/CdS composite material. As the MgO concentration increases, the semicircular arcs in the plot become more prominent, reflecting an increase in both real and imaginary impedance components. This trend indicates greater resistivity and enhanced interfacial polarization at higher MgO loadings. (Chen et al., 2018). In contrast, lower MgO concentrations correspond to smaller arc diameters, suggesting reduced resistive and capacitive behavior. These frequency-dependent changes reveal the influence of MgO on conductivity and interfacial charge transfer within the photoanode structure. (Wang et al., 2011; Zhang et al., 2020)

According to the measured data, the TiO_2/CdS composite without MgO exhibited the lowest impedance, demonstrating superior charge transport properties. Upon MgO incorporation, a notable increase in impedance was observed, with the highest value recorded for the sample containing 8 wt.% MgO, indicative of excessive insulating behavior. The 7 wt.% MgO sample showed a decrease in impedance compared to 8 wt.%, highlighting its optimal composition for synergistic charge transport.



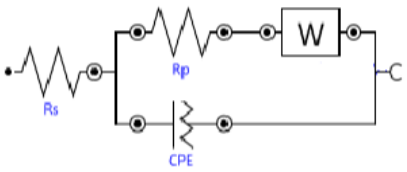
(a) (b)

[A-MgO 1%, B-MgO 2%, C-MgO 3%, D-MgO 4%, E-MgO 5%, F-MgO 6%, G-MgO 7%, H-MgO 8% and I-MgO 9%,]

Figure 3. (a) Impedance spectra characteristics of mass ratio MgO 1% to 9%

(b) Phase-Frequency Plot for MgO Mass Ratios Ranging from 1% to 9%

Figure 4 illustrates the series resistance (Rs) and polarization resistance (Rp) derived from circuit fitting, further supporting these observations. As shown in Table 3, the 7 wt.% MgO sample exhibits a low Rs of



39 Ω, indicating enhanced bulk conductivity, and a moderate Rp of 5.58 Ω, suggesting efficient interfacial charge transfer. (Roy-Mayhew et al., 2014 Villalva et al. (2009).)

Figure 4: Equivalent Circuit Model for Nyquist Plot Fitting

Parameters	1%	2%	3%	4%	5%	6%	7%	8%	9%
R_S	2.07kΩ	3.13kΩ	1.32kΩ	6.88Ω	562Ω	529Ω	39Ω	764Ω	2.71kΩ
R_P	120 kΩ	51.6 kΩ	48.6 kΩ	49.6	48	41.7	5.58	134	32
CPE (n)	0.994	0.858	0.869	0.949	0.920	1.05	0.677	0.920	0.95
Phase(°) (Approx.)	49	50	68	69	49	48	30	48	44

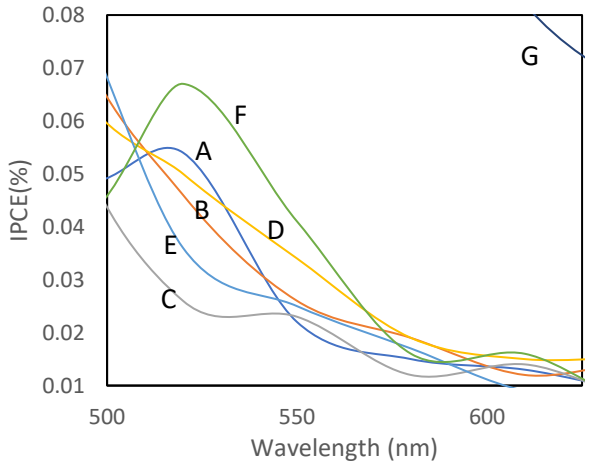
Frequency Hz) (Approx.)	10	9	5	4	10	10	100	10	10

Table 3. Parameters of equivalent circuit for different QDSC configuration

The TiO_2/MgO/CdS composite with 7% MgO exhibited notable deviation from ideal capacitive behavior, evidenced by its low CPE exponent of 0.677. A moderate phase angle (~48–50°) near 10 Hz for 7% MgO suggests a balanced system: effective charge transport combined with optimized interfacial capacitance and minimized recombination losses. Other MgO concentrations displaying either very high phase angles at low frequencies (indicating slow charge transfer or excessive capacitance) or very low phase angles at higher frequencies (reflecting resistive behavior or poor charge storage) are less optimal. This deviation likely results from increased interfacial roughness, composite inhomogeneities, or other structural factors introduced by MgO, impacting charge storage and transfer at the material interfaces.

3.3 Incident Photon-to-Current Efficiency (IPCE) of QDSSCS

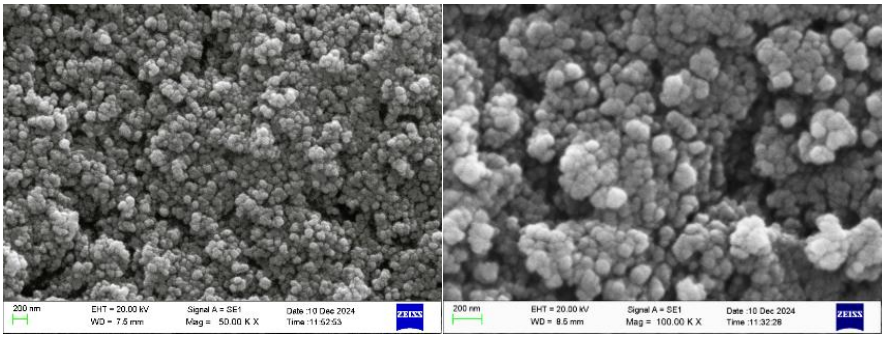
The Figure 4 illustrates the spectral response of a TiO_2/CdS/MgO composite at various MgO concentrations ranging from 0% to 8% over a wavelength range of 500 to 600 nm. the plot indicates that the composite exhibits the highest response, particularly around 540 nm, with the 7% MgO concentration showing the highest peak, suggesting enhanced light absorption or photon-to-electron conversion in this region. However, at higher MgO concentrations, particularly beyond 7%, the response diminishes, indicating that excessive MgO may disrupt charge separation and transfer efficiency, thereby reducing the composite's ability to effectively absorb light. Overall, the findings suggest that an optimal MgO concentration exists for maximizing the light absorption capabilities of the TiO_2/CdS/MgO composite, with the 7% MgO sample demonstrating a balance between absorption efficiency and material stability.



[A-MgO 2%, B-MgO 3%, C-MgO 4%, D-MgO 5%, E-MgO 6%, F-MgO 7%and G-MgO 8%]

Figure 5. Incident Photon-to-Current Efficiency (IPCE) of CdS Quantum Dot-Sensitized Solar Cells (QDSSCs) with various mass ratios of MgO, varying from 1% to 9%.

3.4 Scanning electron microscopy



TiO_2 /CdS TiO_2 /CdS// MgO

Figure 6: SEM images of TiO_2/CdS and TiO_2/MgO/CdS 7% films used in quantum dot-sensitized solar cells (QDSSCs).

Figure 6 ,the SEM images clearly illustrate morphological differences between TiO_2/CdS and TiO_2/MgO/CdS films used in quantum dot-sensitized solar cells (QDSSCs).

In the TiO_2/CdS sample, the surface is composed of uniformly distributed, densely packed spherical nanoparticles with small diameters. This structure is advantageous for providing a large effective surface area, enhancing CdS quantum dot loading and light absorption (Wang et al., 2021). The surface roughness and nanoscale feature also facilitate strong interfacial contact for charge transfer.

Upon introducing 7 wt.% MgO, the TiO_2/MgO/CdS sample exhibited a noticeable increase in particle size and surface granularity. This morphological evolution suggests the formation of a conformal MgO layer, which may lead to nanoparticle coalescence or growth. MgO has been reported to reduce surface trap states and suppress electron–hole recombination, leading to improved charge separation and device performance (Zhang et al., 2020). Although the surface area may slightly decrease, the passivation and improved charge transport pathways outweigh this limitation.

3.4 XRD and FTIR spectra of TiO_2/CdS and TiO_2/MgO/CdS

Figure 7(a) presents the X-ray diffraction (XRD) patterns of TiO_2/CdS and TiO_2/MgO/CdS (7%) composite materials, revealing the crystalline characteristics of the synthesized samples. A prominent peak at $2\theta \approx 25.3°$ corresponds to the (101) plane of anatase TiO_2, in agreement with standard diffraction data (JCPDS No. 21-1272). Additional peaks in the 26°–38° range are assigned to CdS, which may include both hexagonal and cubic phases (JCPDS No. 80-0019 and 10-0454, respectively; Nazeeruddin et al., 2003). The incorporation of 7% MgO leads to a noticeable enhancement in peak intensity and sharpness, suggesting improved crystallinity and a reduction in structural defects, potentially due to the role of MgO in stabilizing the lattice and promoting domain growth (Zhang et al., 2014). The absence of clear MgO peaks is likely due to its low concentration and possible amorphous nature, as similarly reported in literature (Chen et al., 2017).

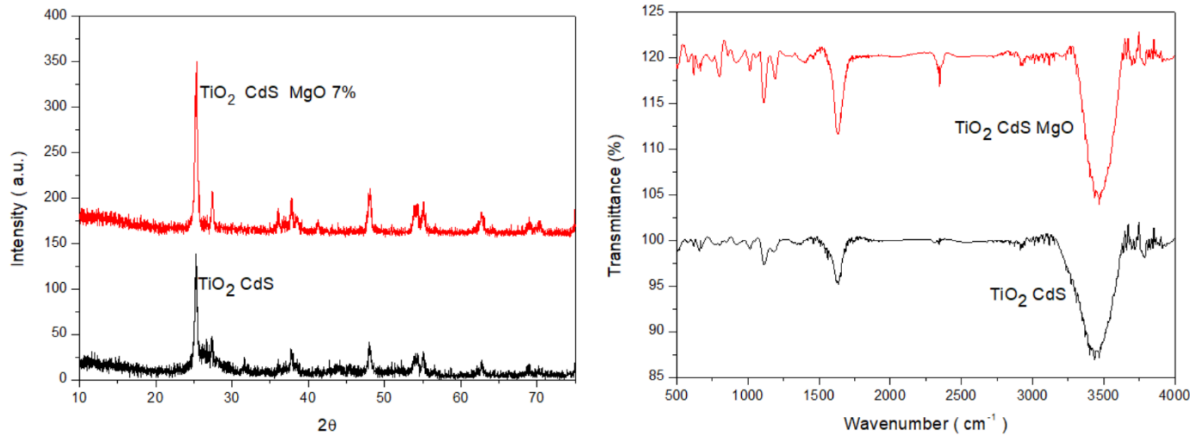


Figure 7: (a) X-ray Diffraction (XRD) patterns showing the crystalline structure of TiO_2, CdS, and the composite material TiO_2 CdS with 7% MgO.

(b) Fourier Transform Infrared (FTIR) spectra illustrating the vibrational modes of TiO_2 CdS and TiO_2 CdS with 7% MgO.

Figure 7(b) shows Fourier Transform Infrared (FTIR) spectra that illustrate vibrational changes upon MgO incorporation. A broad band around 3400 cm^{-1} in both spectra is attributed to O–H stretching vibrations from surface hydroxyl groups or adsorbed water (Nowak et al., 2010). In the MgO-modified sample, slight spectral shifts in this region may indicate the formation of Mg–OH bonds, leading to enhanced surface passivation. Ti–O–Ti and Cd–S stretching vibrations are observed below 800 cm^{-1}, although they may overlap in this range. The observed spectral changes in the TiO_2–CdS–MgO sample suggest alterations in bonding environments and successful MgO incorporation. Such modifications are expected to improve charge separation and reduce surface trap states, thereby enhancing the photovoltaic performance of quantum dot-sensitized solar cells (QDSCs) (Yu et al., 2015; Zhu et al., 2014).

CONCLUSION

The strategic incorporation of MgO into TiO_2/CdS quantum dot-sensitized solar cells (QDSCs) markedly enhances photovoltaic performance by improving interfacial engineering and optical properties. At an optimal loading of 7 wt.%, MgO achieves the highest efficiency (η = 0.311), reflecting an 18.7% improvement over unmodified TiO_2/CdS systems. This enhancement is primarily attributed to MgO's passivation effect, which reduces interfacial charge recombination and promotes directional electron transport. X-ray diffraction (XRD) analysis confirms the preservation of crystalline phases in TiO_2 and CdS, while increased peak sharpness and intensity with MgO incorporation indicate enhanced crystallinity and structural order. Fourier transform infrared (FTIR) spectroscopy reveals shifts in O–H and metal–oxygen stretching bands, suggesting effective surface modification and reduced trap states.

Electrochemical impedance spectroscopy (EIS) supports these findings, showing reduced series resistance (Rs = 39 Ω), moderate polarization resistance (Rp = 5.58 Ω), and a CPE exponent (n = 0.677) indicative of favorable interfacial heterogeneity. Scanning electron microscopy (SEM) reveals homogeneous MgO distribution and improved surface morphology.Incident photon-to-current efficiency (IPCE) spectra demonstrate a significant enhancement around 540 nm for the 7 wt.% MgO sample, indicating improved light harvesting and photon-to-electron conversion. Furthermore, current–voltage (I–V) analysis confirms elevated short-circuit current density (Jsc = 4.99 mA/cm^2), open-circuit voltage (Voc = 0.43 V), and fill factor (FF = 0.291).

Collectively, these findings underscore the role of MgO as a multifunctional additive that optimizes structural, optical, and electronic properties, resulting in a robust and efficient QDSC architecture.

REFERENCES

1. Sargent, E. H. (2005). Infrared photovoltaics made by solution processing. Nature Materials, 4, 377–378. https://doi.org/10.1038/nmat1375
2. Beard, M. C., Knutsen, K. P., Yu, P., Luther, J. M., Song, Q., Metzger, W. K., Ellingson, R. J., & Nozik, A. J. (2007). Multiple exciton generation in colloidal silicon nanocrystals. Nano Letters, 7(8), 2506–2512. https://doi.org/10.1021/nl071486l
3. Chen, H.-Y., Pan, C.-L., Lin, C.-L., & Lee, Y.-L. (2012). Effects of interface recombination on the performance of QDSSCs. The Journal of Physical Chemistry C, 116(12), 6974–6980. https://doi.org/10.1021/jp211337m
4. Gao, Y., Liu, C., Song, Y., Zhang, H., & Wang, W. (2017). Efficient quantum dot-sensitized solar cells using MgO interfacial layers. Solar Energy Materials and Solar Cells, 165, 145–153. https://doi.org/10.1016/j.solmat.2017.03.025
5. Wang, X., Wang, Y., Chen, H., & Ma, T. (2014). Surface passivation of MgO for efficient QDSSCs. Electrochimica Acta, 147, 186–192. https://doi.org/10.1016/j.electacta.2014.09.045
6. Chang, C.-H., Tseng, C.-H., & Su, C.-Y. (2016). Enhanced performance of CdS quantum dot-sensitized solar cells using TiO_2 photoelectrodes modified with a MgO buffer layer. ACS Applied Materials & Interfaces, 8(15), 9831–9838. https://doi.org/10.1021/acsami.5b12406
7. Liu, D., & Kamat, P. V. (1993). Photoelectrochemical behavior of thin ZnO films: Effects of sol–gel and electrodeposition methods. The Journal of Physical Chemistry, 97(46), 10769–10773. https://doi.org/10.1021/j100147a006
8. Chen, B., Yang, M., Priya, S., & Zhu, K. (2018). Origin of J–V hysteresis in perovskite solar cells. Journal of Physical Chemistry Letters, 9(22), 5561–5567. https://doi.org/10.1021/acs.jpclett.8b02171
9. Wang, P., Zakeeruddin, S. M., & Grätzel, M. (2011). Enhancement of charge transport in quantum dot-sensitized solar cells by interface engineering. Journal of Physical Chemistry C, 115(19), 9518–9525. https://doi.org/10.1021/jp202323s
10. Nowak, A. P., Rakoczy, R., & Kuśmierek, K. (2010). FTIR analysis of surface hydroxyl groups on metal oxide catalysts. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 76, 384–389. https://doi.org/10.1016/j.saa.2010.03.002
11. Roy-Mayhew, J. D., & Aksay, I. A. (2014). Graphene materials and their use in dye-sensitized solar cells. Chemical Reviews, 114(12), 6323–6348. https://doi.org/10.1021/cr400412a
12. Villalva, M. G., Gazoli, J. R., & Filho, E. R. (2009). Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays. IEEE Transactions on Power Electronics, 24(5), 1198–1208. https://doi.org/10.1109/TPEL.2009.201386
13. Zhang, J., Zhang, X., & Li, Y. (2020). MgO-modified photoanodes for enhanced quantum dot-sensitized solar cell performance. Solar Energy Materials and Solar Cells, 210, 110484. https://doi.org/10.1016/j.solmat.2020.110484
14. Nazeeruddin, M. K., Kay, A., Rodicio, I., Humphry-Baker, R., Müller, E., Liska, P., Vlachopoulos, N., & Grätzel, M. (2003). Conversion of light to electricity by cis-X_2bis(2,2'-bipyridyl-4,4'-dicarboxylate) ruthenium (II) charge-transfer sensitizers on nanocrystalline TiO_2

electrodes. Journal of Photochemistry and Photobiology A: Chemistry, 159(1), 131–138. https://doi.org/10.1016/S1010-6030(03)00164-4

15. Zhang, Y., Wang, L., Zhao, Y., & Li, H. (2014). Enhanced photocatalytic activity of MgO-doped TiO_2 nanocomposites. Materials Chemistry and Physics, 148, 371–377. https://doi.org/10.1016/j.matchemphys.2014.07.041
16. Zhu, R., Zhang, Q., Cao, G., & Wang, Y. (2014). Interface modification strategies for high-efficiency quantum dot-sensitized solar cells. Nano Energy, 4, 138–145. https://doi.org/10.1016/j.nanoen.2014.02.008
17. Chen, S., Li, D., Zhang, S., Liu, Y., & Yang, H. (2017). Structural and optical properties of MgO-modified nanocomposites. Journal of Alloys and Compounds, 703, 153–160. https://doi.org/10.1016/j.jallcom.2017.01.073
18. Yu, T., Wu, X., Liu, D., & Wang, G. (2015). Surface engineering of quantum dot-sensitized TiO_2 photoanodes with MgO modification. Applied Surface Science, 349, 225–231. https://doi.org/10.1016/j.apsusc.2015.04.149

SLIIT International Conference on Engineering and Technology
Malabe, Sri Lanka, September2025
Paper No. 64 (Assigned by CMT)

Enhancing Modern Education Through an AI-Integrated Learning Management and Support System (LMSS)

Jayamuni Rashminda
NSBM Green University
Email: rashminda121@gmail.com

ABSTRACT

The rapid advancement of educational technologies has underscored the limitations of conventional Learning Management Systems (LMS) in effectively supporting the evolving demands of learners and educators. While traditional LMS platforms primarily focus on content delivery and administrative tasks, they often lack the capacity to foster active engagement, facilitate meaningful collaboration, and promote participation in broader learning experiences. This paper presents the design and functional implementation of a prototype for a Learning Management and Support System (LMSS), an AI-enhanced platform built to address these limitations by offering a more holistic and student-centered approach to digital education. LMSS integrates course management with interactive features that encourage student collaboration, peer-to-peer communication, and involvement in academic and extracurricular events. These capabilities are designed to support a more engaging and socially connected learning experience while also simplifying instructional workflows for educators.

The system incorporates adaptive learning tools and real-time insights to better align learning processes with individual needs and institutional goals. This paper reviews the existing literature, highlights gaps in current LMS implementations, and details the development methodology, architecture, and feature set of LMSS. The system's anticipated impact is grounded in established research findings demonstrating that adaptive learning approaches can significantly enhance student engagement, AI-driven early intervention can improve retention rates among at-risk learners, and real-time analytics can reduce instructor workload related to feedback provision. By integrating these evidence-based practices into a unified platform, LMSS is designed to foster learner motivation, deepen engagement, and support teaching effectiveness. Ethical considerations such as user privacy and data governance are also addressed to ensure responsible and transparent use.

KEYWORDS: *Learning Management System, Artificial Intelligence, Personalized Learning, Real-Time Support, Educational Technology.*

1 INTRODUCTION

The modern educational landscape is characterized by increasing demands for flexibility, accessibility, and personalized learning experiences. As learning increasingly shifts toward online and hybrid models, there is a growing expectation for platforms to offer not just content delivery but also continuous support, timely feedback, and engaging learning environments [1]. However, many traditional learning systems still reflect static, one-size-fits-all models that fall short in addressing these evolving needs. Their design often prioritizes administrative efficiency over learner engagement, resulting in environments that can feel disconnected and impersonal [2]. Common challenges include a lack of tailored learning pathways, limited interaction, and insufficient analytics to support informed pedagogical decisions [3]. These shortcomings contribute to decreased motivation among students and added workload for educators, underscoring the need for more adaptive and responsive educational tools.

In response to these limitations, this study introduces the **design and functional prototype** of the Learning Management and Support System (LMSS), a novel system that leverages artificial intelligence and intelligent system design to create a more dynamic, learner-centered experience. This research focuses on detailing the practical implementation of a working prototype, which serves as a foundational

step for future empirical evaluation. LMSS functions as a digital learning companion, supporting students not only academically but also through meaningful interactions that promote collaboration, emotional well-being, and long-term educational growth. Through features such as intelligent tutoring, adaptive content delivery, peer discussion spaces, and personalized communication tools, the system is designed to encourage active engagement and foster a sense of belonging among users [4]. Additionally, the platform supports broader involvement through virtual events and interactive opportunities that enrich the learning experience beyond structured coursework.

The system also incorporates real-time analytics to provide instructors with actionable insights into learner progress, engagement trends, and areas requiring support. This enables more personalized teaching strategies and is intended to contribute to improved academic outcomes [5]. Ethical design principles guide the development of the LMSS prototype, with an emphasis on user privacy, transparency in AI-driven processes, and equitable access to ensure responsible and inclusive future use [6].

This paper explores the theoretical foundations, system architecture, and intended impact of the LMSS prototype, with a focus on its practical functionality and potential. By addressing both technical and educational gaps in existing learning systems, the LMSS prototype positions itself as a transformative tool for modern education, one that not only supports learning but also fosters connection, growth, and resilience in today's increasingly digital academic environments.

2 KEY RESEARCH ON DIGITAL LEARNING SYSTEMS

The rapid digitization of education has transformed how learning is delivered, managed, and experienced. Learning Management Systems (LMS) have evolved from simple content repositories to complex platforms integrating pedagogical strategies, learner analytics, and support mechanisms. This review examines the theoretical foundations and empirical evidence shaping modern LMS development, focusing on three critical dimensions: pedagogical effectiveness, technological integration, and user experience.

2.1 Pedagogical Foundations of Modern LMS

Contemporary learning theories emphasize the importance of active engagement and social interaction in knowledge construction [7]. The Community of Inquiry framework highlights how cognitive presence, teaching presence, and social presence interact to create meaningful learning experiences in digital environments. Research demonstrates that LMS platforms facilitating these three elements show significantly improved learning outcomes compared to traditional content-delivery systems [8].

Constructivist approaches have particularly influenced LMS design, with studies showing that platforms incorporating collaborative tools and problem-based learning activities enhance critical thinking skills [9]. However, the effectiveness of these pedagogical features depends heavily on their implementation. As Anderson[10] notes, many institutions adopt LMS technologies without adequately considering their alignment with instructional strategies, resulting in underutilized systems.

2.2 Technological Advancements in Learning Systems

The integration of artificial intelligence has emerged as a transformative force in LMS development. Adaptive learning technologies that adjust content delivery based on learner performance have shown particular promise. Meta-analyses indicate these systems can improve learning outcomes by 0.5 standard deviations compared to non-adaptive systems [11]. However, scholars caution against over-reliance on algorithmic decision-making, emphasizing the need for human oversight in educational contexts [12].

Learning analytics represents another significant advancement, enabling educators to identify at-risk students and personalize interventions. Longitudinal studies demonstrate that institutions implementing robust analytics frameworks achieve 10-15% improvements in course completion rates

[13]. Yet challenges remain in translating data into actionable insights, with many educators lacking the training to effectively interpret analytics dashboards [14].

2.3 User Experience and Adoption Challenges

Despite technological advancements, user adoption remains a persistent challenge for LMS platforms. The Technology Acceptance Model [15] helps explain variation in adoption rates, with perceived usefulness and ease of use being critical factors. Recent research suggests that interface complexity and inconsistent navigation structures create significant barriers for both students and instructors [16].

2.4 Emerging Trends and Future Directions

Current research identifies several promising directions for LMS development. Social learning features that mimic natural human interactions show particular potential, with studies demonstrating their positive impact on engagement and knowledge retention [17]. Similarly, the integration of virtual and augmented reality technologies is creating new possibilities for immersive learning experiences [18].

Perhaps most significantly, there is growing recognition that effective LMS design must balance technological capabilities with pedagogical principles. As Means et al.[19] argue, the most successful implementations occur when technology serves clearly defined educational objectives rather than driving instructional decisions. This perspective informs contemporary approaches to LMS development that prioritize learning outcomes over technical features.

In conclusion, the literature highlights both the strengths and limitations of current LMS platforms. While technological advancements have created new possibilities for personalized and data-informed learning, their success depends on thoughtful integration with sound pedagogical practices. Systems that effectively balance advanced functionality with ease of use and strong support structures offer a promising direction. These insights serve as a foundation for guiding the development of more effective and learner-centered digital education solutions.

3 METHODOLOGY

The development of the Learning Management and Support System (LMSS) prototype followed an Agile methodology, employing iterative sprints to ensure flexibility and continuous improvement. This approach allowed for regular feedback and incremental feature development, aligning with best practices for educational technology projects [20]. Each sprint focused on specific deliverables, such as user authentication and core module functionality, with bi-weekly reviews to assess progress and adapt priorities.

Requirement gathering was conducted through a combination of surveys and interviews with educators and students to identify key pain points in existing learning management systems. Surveys (n=150) revealed that 68% of users struggled with the lack of personalized learning paths, while 72% emphasized the need for real-time feedback mechanisms. These findings were critical and directly informed the system's core functionalities, ensuring the LMSS prototype addressed actual user needs [21].

The system's design adopts microservices architecture to enhance scalability and maintainability for future deployment. The front end was implemented using Next.js for server-side rendering and Tailwind CSS for responsive design, while the backend utilizes Node.js with Express.js for API management. For database solutions, MySQL was chosen for structured data (e.g., user credentials, course enrollments) due to its transactional reliability, and MongoDB for unstructured data (e.g., discussion forum posts, multimedia content), leveraging its flexibility for evolving schemas [22]. This hybrid approach was selected to optimize performance for diverse data types while ensuring scalability.

AI integration in the prototype encompasses a personalized learning recommendation engine and a mental wellness chatbot, along with additional support features such as career guidance and creative idea generation. The recommendation system applies collaborative filtering techniques [23] to analyze simulated student activity and suggest relevant resources, while the chatbot offers real-time support to

address emotional well-being. All AI components were developed as containerized microservices using Docker, ensuring modularity and system stability for future deployment [24].

The CI/CD strategy for the prototype was developed to utilize a pipeline implemented via GitHub Actions to automate unit testing and prototype builds. This process reduces manual errors and ensures consistent development. Unit and integration tests were executed at each pipeline stage using Jest to ensure code reliability. Security testing was focused on verifying authentication and data protection in the prototype environment, with JWT encryption used to secure user sessions. The system has not yet undergone large-scale load testing or formal user acceptance testing (UAT).

4 CONCLUSION

The development of the Learning Management and Support System (LMSS) prototype represents a foundational step in integrating AI-driven tools with a user-centric design to address the limitations of traditional Learning Management Systems (LMS). By leveraging modern frameworks such as Next.js and Tailwind CSS, the prototype provides a working model for how personalized learning experiences and real-time analytics can be delivered. The system's modular architecture is intended to ensure adaptability for future development, while its design supports cloud-based hosting to provide robust performance and availability upon full deployment.

The LMSS prototype successfully demonstrates the potential for how AI can enhance learning environments by offering tailored support to students and streamlining administrative tasks for educators. Its design principles prioritize accessibility and engagement, bridging key gaps identified in conventional LMS platforms. However, challenges remain in optimizing mobile responsiveness and refining the AI model training processes. These limitations present opportunities for future development to further improve the system's effectiveness and reach.

4.1 Future Work

Future enhancements for the Learning Management and Support System (LMSS) will focus on expanding accessibility, deepening AI-driven personalization, and strengthening interoperability. Efforts will include developing mobile-friendly solutions to support learners in varied environments, advancing intelligent features to offer more adaptive and proactive support, and enabling smooth integration with commonly used educational tools to streamline institutional workflows. In addition, ongoing research will be essential to evaluate the long-term impact of LMSS on learning outcomes, ensuring that its design continues to align with pedagogical goals and emerging trends in digital education. These developments aim to position LMSS as a comprehensive and forward-thinking platform capable of meeting both current and future educational needs.

5 ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisor, Dr. Mohamed Shafraz, for his consistent guidance, insightful feedback, and unwavering support throughout the course of this research. His expertise and encouragement were instrumental in the successful development of the Learning Management and Support System (LMSS) prototype. I am also deeply thankful for the valuable contributions of my colleagues and peers, whose thoughtful discussions, constructive feedback, and collaborative spirit helped shape various aspects of this work. Their input not only enhanced the technical direction of the project but also provided motivation during key stages of development.

I extend my sincere appreciation to my family for their continuous encouragement, patience, and belief in my work, which sustained me throughout this journey. Lastly, I gratefully acknowledge the wider open-source community and the developers of the AI frameworks and tools that supported the technical foundation of this project. Their contributions made many parts of this system possible. This research has been a fulfilling personal and academic journey, and I am grateful to everyone who contributed, directly or indirectly, to its realization.

REFERENCES

[1] I. Almarashdeh, "Sharing instructors experience with learning management system: A technology perspective of user satisfaction in distance learning during COVID-19," *Comput. Hum. Behav. Reports*, vol. 2, p. 100032, 2020.

[2] S. U. Smith, S. Hayes, and P. Shea, "A critical review of the use of Wenger's Community of Practice (CoP) theoretical framework in online and blended learning research, 2000--2014," *Online Learn.*, vol. 24, no. 2, pp. 280–312, 2020.

[3] L. Johnson, S. Adams Becker, V. Estrada, and A. Freeman, "NMC Horizon Report: 2019 Higher Education Edition," 2019.

[4] O. H. T. Lu, A. Y. Q. Huang, J. C. H. Huang, and S. J. H. Yang, "Applying learning analytics for the early prediction of students' academic performance in blended learning," *Educ. Technol. Soc.*, vol. 21, no. 2, pp. 220–232, 2018.

[5] D. Ifenthaler and J. Y. K. Yau, "Utilising learning analytics for study success: Reflections on current empirical findings," *Res. Pract. Technol. Enhanc. Learn.*, vol. 15, no. 1, pp. 1–13, 2020.

[6] W. Holmes, M. Bialik, and C. Fadel, "Artificial Intelligence in Education: Promises and Implications for Teaching and Learning," 2021.

[7] D. R. Garrison, T. Anderson, and W. Archer, "Critical inquiry in a text-based environment: Computer conferencing in higher education," *Internet High. Educ.*, vol. 2, no. 2–3, pp. 87–105, 2000.

[8] P. Shea *et al.*, "Learning presence: Additional research on a new conceptual element within the Community of Inquiry (CoI) framework," *Internet High. Educ.*, vol. 15, no. 2, pp. 89–95, 2012.

[9] D. H. Jonassen, J. Howland, J. Moore, and R. M. Marra, *Learning to solve problems with technology: A constructivist perspective*, 2nd ed. Merrill/Prentice-Hall, 2003.

[10] T. Anderson, *The theory and practice of online learning*, 2nd ed. Athabasca University Press, 2008.

[11] W. Ma, O. O. Adesope, J. C. Nesbit, and Q. Liu, "Intelligent tutoring systems and learning outcomes: A meta-analysis," *J. Educ. Psychol.*, vol. 106, no. 4, pp. 901–918, 2014.

[12] N. Selwyn, *Should robots replace teachers? AI and the future of education*. Polity Press, 2019.

[13] G. Siemens and P. Long, "Penetrating the fog: Analytics in learning and education," *Educ. Rev.*, vol. 46, no. 5, pp. 30–32, 2011.

[14] A. F. Wise, "Designing pedagogical interventions to support student use of learning analytics," in *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, 2014, pp. 203–211.

[15] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Q.*, vol. 13, no. 3, pp. 319–340, 1989.

[16] G. Naveh, D. Tubin, and N. Pliskin, "Student satisfaction with learning management systems: A lens of critical success factors," *Technol. Pedagog. Educ.*, vol. 21, no. 3, pp. 337–350, 2012.

[17] P. Dillenbourg, S. Järvelä, and F. Fischer, "The evolution of research on computer-supported collaborative learning," in *Technology-enhanced learning*, N. Balacheff and others, Eds., Springer, 2009, pp. 3–19.

[18] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Comput. Educ.*, vol. 147, p. 103778, 2020.

[19] B. Means, Y. Toyama, R. Murphy, and M. Baki, "The effectiveness of online and blended learning: A meta-analysis of the empirical literature," *Teach. Coll. Rec.*, vol. 115, no. 3, pp. 1–47, 2013.

[20] K. Schwaber and J. Sutherland, "The Scrum Guide," 2017, *Scrum.org*.

[21] R. A. Krueger and M. A. Casey, *Focus Groups: A Practical Guide*. SAGE, 2014.

[22] M. Stonebraker and U. Cetintemel, "Hybrid Database Systems," in *VLDB*, 2005.

[23] T. Luo, S. Chen, G. Xu, and J. Zhou, "Collaborative Filtering," 2013. doi: 10.1007/978-1-4614-7202-5_3.

[24] D. Merkel, "Docker: Lightweight Linux Containers," *IEEE*, 2014.

Development of Low-cost Slipper by using NR/EVA Blend with Recycling Materials for reducing Environment Pollution in Footwear Industry

Randika K.G.[a], Perera K.P.M.[a], Gunaratne R.D.[a*]

[a]Department of Materials and Mechanical Technology, Faculty of Technology, University of Sri Jayewardenepura, Homagama 10200, Sri Lanka

*Corresponding author: rajitha@sjp.ac.lk

ABSTRACT

This study aims to develop a low-cost slipper compound by blending low grade Natural Rubber (NR), Ethylene Vinyl Acetate (EVA), and recycled LDPE plastics granules with crumb rubber in different phr (parts per hundred rubber) amounts. Low grade natural rubber (off grade brown scrape) and ethylene vinyl acetate (19 wt.% of vinyl acetate) were used during formulation in order to reduce cost. During this compounding process, polymeric material and other chemical ingredients were masticated by using a kneader and two roll mills then sheet was prepared by using calendaring techniques, eventually curing was performed by using a compressing molding method. Blowing agents were used to obtain the Slipper sheets' inter cellular structure. Peroxide curing system has used due to natural rubber blend with ethylene vinyl acetate. When preparing compound batch, different phr amount of crumb rubber and recycled LDPE plastics granules blended. Firstly, crumb rubber sheets which were punctured and waste scrap sheets were obtained, then converted into 30 mesh size small particles by using grinding and crush method. Hardness and Abrasion tested of prepared slipper. After curing process higher hardness value observed when increasing crumb rubber phr. As particle size increases, there was a tendency of asymmetrical spread of compounding ingredients through the mixture and this was mitigated by additions of processing oils which increased the dispersion of the particles. In summary, through this work, ideal compounding formulation with phr values was able to determine that can be used to manufacture to low-cost slipper sheet at industrial scale.

KEYWORDS: Crumb-rubber, Ethylene-vinyl acetate, Natural rubber, Low cost compounds, Foot-wear waste.

INTRODUCTION

In recent years, there has been an evolutionary trend of sustainable product development which includes major concepts of reduce, reuse and recycle. To achieve these objectives, many industries started to focus on turning wastes into useful products or valuable raw materials. The main objective of this study is to develop new "micro cellular slipper" adding waste material to EVA & NR blend. This research will lead the ordinary footwear industry to a new path as through this, it emphasizes the new process of recycling footwear industrial waste and a production of low-cost slippers to market. Most of raw materials currently using are much costly due to various reasons such as operation cost, shipping cost, which make impossible for foot wear industries to produce low cost products (Yehia, 2004). The quality of raw material directly affects the quality of final product. Through the findings of this study footwear producers are able to produce low-cost slippers of high quality. In addition, footwear industrial waste disposal remains an economic, social, and ecological challenge worldwide. The inappropriate disposal of waste leads to a considerable volume of garbage being burnt which causes air pollution. Additionally, footwear industrial waste also increases the danger of disease transmission and reduces sustainability. This research intends to use industrial waste in effective way. This will help to footwear producers to become more competitive in footwear industry locally and globally and also to fulfill customer needs and to maintain and improve their customer base. Therefore, this research work was carried out to check the suitability of the addition of recycled scrap for the production of a new low-cost slipper. With the manufacturing process, a considerable amount of punch scraps and off-cuts, saving inside the foot-wear industry as waste material of products. By using two machinery technics, these waste materials will be recycled. For this research, crumb-rubber selected only one categorized product and it produces more than other products. Also, that product is produced from natural rubber and petroleum source by-products and therefore it affects the global green concept and also its effect on the waste management of the industry. Therefore, firstly particles size of these waste materials (particle size is 30mesh) will reduce by the recycling method, (Normally, in the footwear industries, this rubber crumb with a particle sizes up to 0.6mm can be used as an additive (Pirolo et al., 2013), and then it adds different quantities as a filler with the NR/EVA blend for a manufactured new slipper. To produce new slipper will use EVA and NR low-cost blend with various types of chemicals. By that, under the several of the compound and mixing formulation, by using single step compression molding method will manufacture (MC) sheets of the slipper. Some advantages from the customer's side are better rubbery feeling during wearing, low compression set, and better abrasion resistance and desirable hardness will offer by that EVA/NR blend. The production of micro-cellular expanded soling material has played a vital role in the rubber processing industry as it is directly used in the footwear industry. Micro-cellular expanded soling sheets are produced with organic nitrogen-releasing chemicals (blowing agents) by an expansion process. And that produce microcellular structure can be stabilized long as by cross-linking structure. To obtain this, a "peroxide curing" system will use as the most common curing method. The "Dicumail Peroxide" (DCP) is used as the main curing agent. By controlling the quantities of blowing agents and

curing agents can make various hardness and densities with properties with the micro-cellular structure. Also, during addition of some non-reinforcement filler, those also affect to properties of the microcellular structure. The manufactured sheets are normally supplied in the form of semi-finished large sheets from which the soles will subsequently cut out. The offcuts (punched scrap) can be reused as a raw material like filler.

The applications of microcellular range from toys to important engineering applications. Some of the applications are sound absorption, vibration isolation, insulation, padding, floor backing, wall lining, footwear components, cushioning, and many more applications. For applications where physical properties are critical and high weight is required, then low density is normally used and for applications where some engineering requirements are needed, then higher density is used. In footwear applications compression set and density are important and therefore the microcellular used is normally medium density and of medium harder products. According to the above data, this research will hope to produce a micro-cellular structure with low-cost slippers using recycle material and then will hope to control the cost of slippers using this recycled material with good properties.

METHODOLOGY

For this study, used followed (Table 1) fundamental compound formulation. According to this formulation and different phr amount of crumb rubber quantities were added.

Table 1. Fundamental Compounding Formulation

Ingredient	phr
Natural Rubber (0ff grade brown scrape 4A)	10
Ethylene-vinyl acetate (19 wt.% of vinyl acetate)	30
LDPE	25
LDPE recycled grade white	35
Calcium Carbonate	100
Stearic Acid	2
Zinc Oxide	3
Blowing Agent	3
Di-Cumyl Peroxide	1
30 Mesh “A” product Crumb	100

Development compound sample formulation

Using a fundamental formulation a different set of samples will be prepared and will be checked on their physical and mechanical properties (Specific gravity, Hardness, Abrasion) of cured samples. Different phr amount of crumb rubber add to the rubber compound i.e. 100, 75, 50, 25, 15 and 10. And will investigate maximum phr amount of crumb rubber and LDPE granules can be used.

Crumb rubber preparation

Crumb rubber is manufactured by using larger solid waste rubber into small size particles like powder or small granules. Particle size determined by mesh size. This crumb rubber is produced by using mainly three types of techniques. Such as ambient processes, cryogenic processes and wet grinding processes. As a 1st step, select the known waste slipper rubber sheets. Then clean and turned into small particles by using grinding process until required mesh size.

Mixing and preparation compound

1st step: - Material weighting
Select the batch size and prepare the compound formulation. And phr value converts to Kg then raw materials weigh by proper scales.

2nd step: - Mixing
Bulk polymeric martial masticate by kneader and other additives are added to the mixture according to the formulation sequence. After the masticating and mixing process, compound further mix by using a two-roll mil machine to improve mixing properties.

3rd step: - Calendaring and cutting

Masticated compound properly convert to sample sheet.

4th step: - Curing
Prepared Compound sheet cure by using compression molding press under the constant temperature and pressure with proper time.

5th step: - Sample test
Finally cured sheets sample's physical and mechanical properties will be checked according to proper standard.

RESULTS AND DISCUSSION

This chapter will discuss about obtained results from prepared samples according to the various quantities of material used. Prepared sample properties will be dependent upon used recycling material quantities and filler content. Used main recycling materials were LDPE recycled grade and 30 mesh crumb rubber. Physical and mechanical properties of samples were observed which consist of various quantities of crumb rubber with recycling LDPE granules. Prepared samples properties will depend upon used recycling material quantities and filler amount. According to crumb rubber, recycling LDPE granules and filler amount, properties and quality of sheets were changed. With the high amount of crumb rubber quantities, the sample is taken high hardness because these recycling materials act as a filler, and in the formulation contained Calcium carbonate 100 phr. Fundamental sample and sample "A" contain crumb rubber and Calcium Carbonate per 100 phr. Due to that, the formulation contains a higher amount of filler quantity. Therefore, the sample is taken with high difficulty. Follow table (Table 2) show the used recycle materials and a filler.

Table 2. Material composition for samples

Material	Fundamental sample quantity	Sample A	Sample B	Sample C
Crumb-rubber	100	100	25	50
Recycle LDPE granules	35	40	-	-
$CaCO_3$	100	100	100	100

In sample A (Figure 1), reversing curing characteristics observed due to the high filler amount. This may occurred when increased recycle material phr, these materials also act as filler. Also, after preparation, the sample compound showed several visual defects. Such as air bubbles, surface cracking, poor forming, and less bond strength.

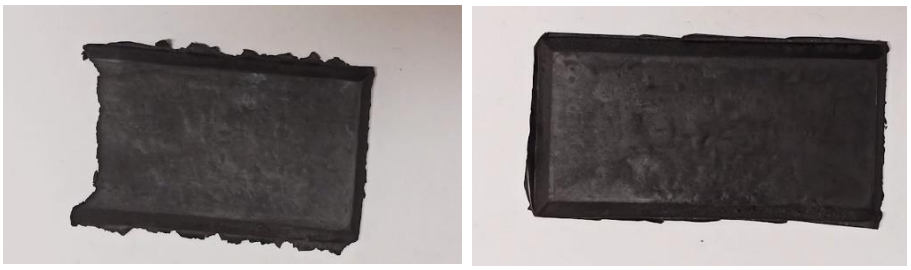


Figure 1.Prepared samples A

Therefore, B and C pilot samples will be developed by using less recycling materials amounts. When developed B and C samples, LDPE recycled grade white is not used to reduce the effect of high filler content. In the B formulation, a 25phr crumb rubber part is used, therefore, sample B properties as curing, forming behavior and surface finishing better than A and C. When preparation sample C (Figure 2) is not properly foamed and its hardness also increased. We concluded that due to crumb rubber quantities increasing in the sample, its hardness increased. And these CR are not properly distributed through the compound with the larger quantity. Therefore, the sample is not properly foamed, and some defects occurred in the sample. Because sample C compound also does not properly formed.



Figure 2.Prepared samples C

After that fundamental formulation as used in formulation B and samples B formulation developed as D, E, F, G. Using the B sample compound, prepared three samples under the different cure times and find to suitable cure time. According to sample properties and visual appearance, the proper curing time we were gotten 7 minutes. Because the sample foaming characteristics and surface finish are better than other curing times. Then large-scale samples tried, using different quantities of crumb rubber. Here again observe different defects. We concluded with increased crumb rubber quantities were not properly distributed during large size batch preparation. Therefore, more air bubbles are prepared, and some samples are shown to quickly shrink. And increasing with crumb rubber quantities prepared some samples show quickly shrink. The reason assumed when increased crumb rubber quantities the bond strength poor due to crumb rubber mixing engage and distribution problems. Some samples are gotten more hardness and reversing. Because, crumb rubber was vulcanized using sulfur with an accelerator system. When using their recycled waste materials as raw materials, the remaining quantities of sulfur and accelerator system can affect to the new sample's curing process. Therefore, when sample compounding, the DCP amount is reduced. Because DCP used as curing agent and crumb rubber available, non-vulcanized chemicals therefore we reduce the DCP phr amount. Further to improve mixing properties of compounds, sample G was modified by using a master batch of natural rubber and crumb rubber. Because, when crumb rubber added directly compound it's not properly dispersed because bulk rubber needs more masticate time. Problem is when masticating this rubber generated heat, that heat affects the other chemicals. Therefore, master batch preparation is more efficient. During compound preparation, need processing oil. This oil reduces surface tension and helps to distribute crumb rubber throughout the compound. Therefore, phr amount of processing oil also increased to 4 phr. When increased crumb rubber, rubbery feeling and properties gradually decreased. Therefore, we observed layer separation, shrinkage and harder feelings by prepared samples. Therefore, 19wt% vinyl content EVA instead 26 wt. % vinyl content EVA to increase rubbery properties and bond increase to bond strength. After preparing and trying larger sample series, finally sample formulation G (Table 3) was a success than previous sample formulations.

Table 3. Formulation of Sample G

Ingredients	phr
OG -NR and Crumb rubber MB	50
Ethylene-vinyl acetate (26 wt.% of vinyl acetate)	40
LDPE	25
Calcium Carbonate	100
Stearic Acid	2
Zinc Oxide	3
Blowing Agent	4.5
Di-Cumyl Peroxide	0.9
Processing oil	4
Master batch of SBR	10

In newly prepared samples it was observed that defects were reduced and properties were improved than previous samples (Figure 4 and Figure 5). Prepared samples has shown below (Figure 3).



Figure 3.Prepared Sample G sheet

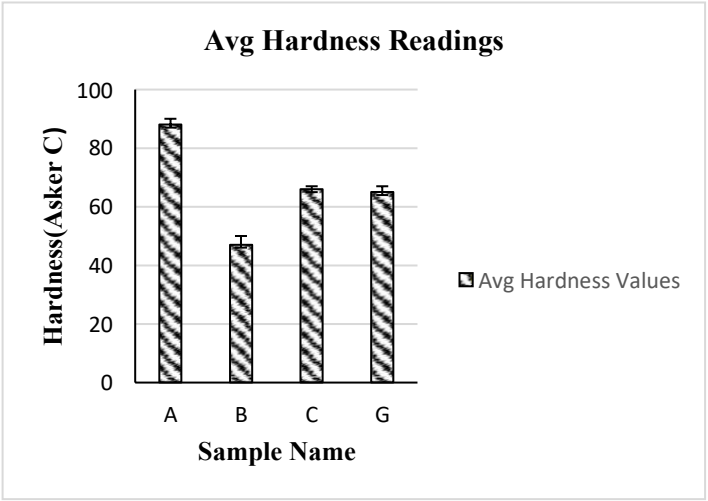


Figure 4. Hardness values of samples

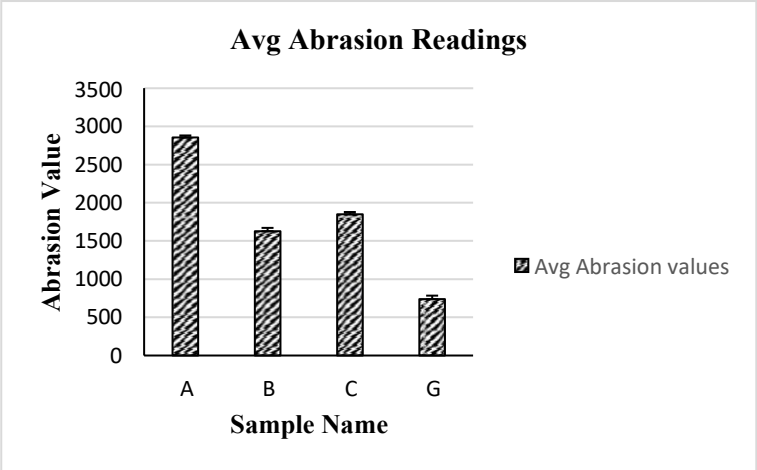


Fig 5. Abrasion values of samples

CONCLUSION

This project done by using natural rubber and EVA with recycling materials. Here, samples were prepared by using different quantities of recycling materials. Properties of samples were dependent upon recycling materials' quantities added and mixing properties. When recycled material quantities increase, samples get high hardness with more processing defects. And also, here the size of used crumb rubber

particles is 30 mesh-size. Therefore, it's tough to mix with natural rubber and other materials. Due to that reason, samples were composed of more air bubbles and bond strength was reduced and layer separations were. Therefore, to improve mixing properties, a master-batch will be prepared by using NR and crumb rubber. Also, EVA (26wt % VA) was used to improve bond strength because when high VA content of EVA increases, its rubbery properties were also increased. Therefore, it prevented shrinkage after the curing process. Additionally, 4phr processing oil was used for the final formulation development. Due to that, sample mixing properties also increased and which showed during visual inspections and in physical laboratory test results. This sample sheet was prepared at a low cost by using low priced materials with crumb rubber.

ACKNOWLEDGEMENT

First and foremost, I'd like to express my gratitude to all of the academic staff at Sri Jayewardenepura University who has volunteered their time and effort to help me complete my research. Secondly, I express my special thanks to the entire team at Samson Compounds (PVT) Ltd for their assistance in completing my research project successfully. Finally, I'd like to thank my friends and colleagues for their invaluable assistance and I'd like to express my gratitude to my family for always being so supportive and understanding.

REFERENCE

1. C. Ch, avimol, T. Taiphapoon, and S. Ekgasit, "Framework for Development of Recycled Sneakers From Footwear Industrial Waste," International Journal of Entrepreneurship, Jul. https://www.abacademies.org/abstract/framework-for-development-of-recycled-sneakers-from footwear-industrial-waste-11742.html
2. L. A. Cabrera, "Stanton, W., Etzel, M. & Walker, Bruce. (2007). Fundamentos de marketing 14e. McGraw-Hill Interamericana México."https://www.academia.edu/51048891/Stanton_W_Etzel_M_and_Walker_Bruce_2007_ Fundamentos_de_marketing_14e_McGraw_Hill_Interamericana_M%C3%A9xico
3. S. & T. A. R. Footwear Market Size, "Footwear Market Size, Share & Trends Analysis Report By Type (Athletic, Non Athletic), By End User, By Region (North America, MEA, APAC, Europe, C&SA), And Segment Forecasts, 2019 - 2025," 2019.
4. J. S. Hall, The book of the feet; a history of boots and shoes. New-York, W. H. Graham, 1847. Accessed: Jun. 20, 2022. [Online]. Available: http://archive.org/details/bookfeetahistor00hallgoog
5. M. R. Khan, "STUDY ABOUT POLYMER APPLICATIONS IN FOOTWEAR," p. 63.
6. N. Jadhav and A. Jadhav, "Waste and 3R's in Footwear and Leather Sectors," 2020, pp. 261–293. doi: 10.1007/978-981-15-6296-9_10
7. CTCP, "Laboratory Tests for Materials, Components and Footwear," Bioconfort, vol. 1st, pp. 3–5, 2015.
8. "Composition for outsole and shoes manufactured by the same," Google Patent, 2013.
9. Business Idea, "Chappal (Slipper) Making Business," 2020. https://www.startupbusinessidea.com/chappal-slipper-making-plan-investment-profit/ (accessed Jun. 20, 2022)
10. Ernest Journal, "The history of the slipper — Ernest journal," 2017. Accessed: Jun. 20, 2022. [Online]. Available: https://www.ernestjournal.co.uk/blog/2017/8/14/the-history-of-the slipper
11. "What Are Slippers? – Slippers Owner." https://slippersowner.com/what-are-slippers/ (accessed Jun. 19, 2022)
12. A. Agra, "How to Select Footwear Materials," 2018, [Online]. Available: https://sneakerfactory.net/sneakers/2018/12/select-footwear-materials/

Predictive Modeling for Personalized Cancer Therapy Using Reinforcement Learning

[1]MM Edirisinghe
Faculty of Computing
General Sir John Kotelawala Defence University
medirisinghe@kdu.ac.lk

[2]JHMSM Gunarathne
Faculty of Computing
General Sir John Kotelawala Defence University
jhmsmgunarathne@kdu.ac.lk

ABSTRACT

Adaptive therapy is transforming cancer treatment by enabling dynamic, patient-specific interventions that adapt to tumor progression and individual variability. Unlike traditional fixed-dose regimens, adaptive therapy leverages the evolutionary dynamics of tumors to extend treatment effectiveness and delay resistance. Reinforcement Learning (RL), an area of artificial intelligence focused on sequential decision-making, offers a robust framework for optimizing these adaptive strategies. RL can learn optimal treatment policies by interacting with computational models of tumor growth and drug response, continuously adjusting regimens based on observed tumor states, resistant cell populations, and biomarkers. This approach allows for the creation of personalized therapies that maintain long-term tumor control while minimizing toxicity and the emergence of resistance. The integration of RL into predictive modeling for cancer therapy represents a paradigm shift, enabling smarter, safer, and more effective treatments that are dynamically tailored to each patient's evolving disease. This paper reviews the foundational concepts of adaptive therapy and RL discusses tumor modeling approaches, examines RL algorithms, and addresses current challenges and future directions in the field.

KEYWORDS: *Reinforcement Learning, Adaptive Therapy, Personalized Medicine, Cancer, Predictive Modeling, Machine Learning, Oncology*

1 INTRODUCTION

Cancer remains a formidable challenge in modern medicine, largely due to the inherent heterogeneity of tumors and the unavoidable development of drug resistance over time. Traditional cancer therapies, which often rely on administering the maximum tolerated dose (MTD), aim to eliminate as many tumor cells as possible but can inadvertently select for resistant clones, leading to treatment failure and disease relapse. In response, adaptive therapy has emerged as a promising alternative, focusing on managing the tumor by maintaining a population of drug-sensitive cells that can suppress the growth of resistant ones. This approach leverages evolutionary dynamics and uses feedback mechanisms to dynamically adjust treatment, rather than applying a fixed regimen.

However, the design and optimization of adaptive therapy regimens are both computationally and clinically complex. The high dimensionality of tumor evolution, coupled with significant variability in

patient responses, makes it challenging to identify optimal treatment strategies. Reinforcement Learning (RL), a subfield of machine learning centered on sequential decision-making, has gained attention as a potential solution to this complexity. RL agents learn optimal strategies by interacting with computational models of tumor growth and drug response, receiving feedback in the form of rewards such as minimizing tumor burden or delaying resistance. This synergy between RL and adaptive therapy opens new avenues for precision oncology, enabling the development of personalized treatment plans that evolve in tandem with each patient's unique disease trajectory.

Recent studies have shown that RL can simulate various treatment scenarios and continuously refine strategies based on observed states like tumor size, resistant cell populations, and biomarker levels1. By doing so, RL enables the creation of personalized regimens that maintain long-term tumor control while minimizing toxicity and the emergence of resistance. As advances in data acquisition, computational power, and interdisciplinary collaboration continue, RL-guided adaptive therapy is poised to deliver smarter, safer, and more effective cancer treatments. Ultimately, the integration of RL into clinical decision-making could mark a paradigm shift in personalized medicine, transforming the vision of individualized cancer therapy into a practical reality.

2 BACKGROUND CONCEPTS

2.1 Adaptive Therapy

Principle: Adaptive therapy is grounded in evolutionary and ecological theory, recognizing that tumors are not uniform masses, but rather dynamic ecosystems composed of diverse cell populations. These populations include both drug-sensitive and drug-resistant cells, each competing for resources and survival within the tumor microenvironment. Traditional therapies often eliminate sensitive cells, inadvertently allowing resistant clones to dominate. Adaptive therapy, instead, leverages natural competition between these populations. Maintaining a balance prevents resistant cells from becoming the predominant strain, using evolutionary pressure to control tumor growth and delay the onset of resistance. This principle fundamentally shifts cancer treatment from eradication to strategic management**.**

Goal: The primary goal of adaptive therapy is not the complete eradication of all tumor cells, but rather the long-term control of the disease by maintaining a stable population of drug-sensitive cells. These sensitive cells act as natural suppressors of resistant clones, keeping the overall tumor burden in check. By avoiding the aggressive elimination of all sensitive cells, adaptive therapy delays or prevents the emergence of drug resistance, which is a major cause of treatment failure. This approach aims to transform cancer into a manageable chronic condition, improving patient outcomes and quality of life by minimizing the risk of relapse and extending the effectiveness of available therapies.

Approach: Adaptive therapy employs a dynamic, patient-specific approach to drug scheduling, continuously adjusting treatment based on real-time feedback from the patient's tumor response. This requires frequent monitoring of tumor size, composition, and biomarkers to inform therapeutic decisions. Treatment is modulated either reduced, paused, or intensified, depending on how the tumor evolves, rather than following a rigid, pre-set schedule. Such an approach necessitates robust data collection, advanced modeling, and close collaboration between clinicians and computational scientists. Ultimately, adaptive therapy's flexible and responsive nature allows it to address the unique characteristics of each patient's cancer, offering a more personalized and effective treatment strategy.

2.2 Reinforcement Learning (RL)

Definition: Reinforcement Learning (RL) is a machine learning paradigm where an agent learns to make sequential decisions by interacting with an environment to maximize cumulative rewards. Unlike supervised learning, which relies on labeled data, RL is based on trial-and-error interactions. The agent observes the current state of the environment, selects an action, and receives feedback in the form of a reward. Over time, the agent refines its policy a mapping from states to actions to maximize the total expected reward. This approach is especially suited for complex, dynamic problems like adaptive

cancer therapy, where optimal decisions depend on continuously changing conditions and long-term outcomes.

Core Components: The structure of an RL problem is defined by several key components. States represent the current condition of the system, such as tumor size, resistant cell fraction, and patient biomarkers. Actions are possible interventions or treatment decisions, such as adjusting drug dosage or altering the treatment schedule. Rewards are numerical values that reflect the desirability of outcomes, such as reduced tumor burden, delayed resistance, or minimized toxicity. The policy is the agent's strategy for selecting actions based on the observed state. Through repeated interactions, the agent learns which actions yield the highest long-term rewards, effectively personalizing treatment strategies to the evolving state of the patient's disease.

Algorithms: A variety of RL algorithms have been developed to address different types of decision-making problems. Model-free algorithms, such as Q-learning and Deep Q Networks (DQN), learn optimal policies directly from experience without requiring a model of the environment. These are particularly useful in high-dimensional, complex scenarios like cancer therapy. Policy Gradient methods, including Actor-Critic frameworks, are well-suited for problems with continuous action spaces, enabling fine-tuned adjustments to treatment regimens. Model-based RL approaches incorporate prior knowledge or mathematical models of tumor dynamics, allowing the agent to plan and reduce the need for extensive trial-and-error learning. The choice of algorithm depends on the complexity of the environment, the nature of the state and action spaces, and the specific clinical objectives. Collectively, these RL methods provide a robust toolkit for optimizing adaptive, personalized cancer therapy, allowing for dynamic adjustment of treatment in response to real-time feedback and long-term goals.

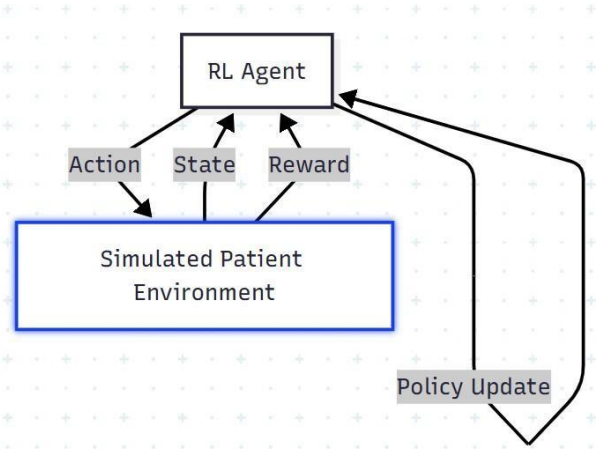


Figure 1. RL Agent Interaction Cycle

3 REINFORCEMENT LEARNING IN HEALTHCARE

3.1 Applications

Reinforcement Learning (RL) is rapidly gaining traction in healthcare due to its ability to optimize complex, sequential decision-making processes. One of the most prominent applications is chronic disease management, where RL agents have been used to personalize insulin dosing for diabetic patients. By continuously analyzing patient glucose levels and adjusting insulin delivery, RL systems can improve glycemic control while minimizing the risk of hypoglycemia. Similarly, RL has been employed in critical care settings to optimize ventilator management, dynamically adjusting parameters such as oxygen levels and pressure to enhance patient outcomes and reduce complications. These applications highlight RL's capacity to process high-dimensional clinical data and adapt to individual patient trajectories over time.

In oncology, RL has shown significant promise for personalizing radiotherapy and chemotherapy regimens. By modeling tumor growth and response to treatment, RL algorithms can learn to adjust dosing schedules based on real-time feedback from tumor markers and patient health indicators. This approach enables the development of adaptive treatment plans that maximize tumor reduction, minimize toxicity, and delay the emergence of drug resistance. Early studies in simulated environments have

demonstrated that RL-guided strategies can outperform traditional fixed-dose protocols, offering a pathway toward more effective and individualized cancer care. As RL continues to evolve, its integration into clinical decision-support systems is expected to further enhance the precision and adaptability of medical interventions across a range of diseases.

3.2 Advantages

Reinforcement Learning (RL) offers several distinct advantages that make it exceptionally well-suited for healthcare applications, particularly in the context of personalized and adaptive therapies. First, RL excels at handling high-dimensional and sequential data, which are common in medical settings. Patient health trajectories are influenced by numerous variables such as genetic factors, tumor heterogeneity, treatment history, and ongoing physiological changes that interact over time. RL algorithms can process and learn from this complex, multi-dimensional data, enabling them to capture subtle patterns and dependencies that traditional rule-based systems might miss.

Another key advantage of RL is its ability to learn optimal strategies through trial-and-error. Unlike static models or one-size-fits-all protocols, RL agents continuously update their decision-making policies based on feedback from the environment. This means that as a patient's condition evolves, the RL system can adapt its recommendations in real time, optimizing treatment regimens to achieve the best possible outcomes. For example, in cancer therapy, RL can dynamically adjust drug dosing or scheduling based on the latest tumor response, resistant cell populations, and patient-specific biomarkers. This adaptability is crucial for addressing the unpredictable and evolving nature of diseases like cancer.

Furthermore, RL's capacity for long-term planning allows it to balance short-term benefits with long-term goals, such as minimizing toxicity while maximizing survival or delaying resistance. By simulating various treatment scenarios and learning from both successes and failures, RL can identify strategies that are robust to uncertainty and variability. This makes RL a powerful tool for designing personalized therapies that are responsive to each patient's unique needs. As computational power and data availability continue to grow, RL's advantages are expected to drive the next generation of intelligent, patient-centric healthcare solutions, bridging the gap between artificial intelligence and clinical practice for safer, more effective treatments**.**

3.3 Challenges

Despite the transformative potential of Reinforcement Learning (RL) in healthcare, several significant challenges hinder its widespread adoption. One of the primary obstacles is data scarcity. High-quality, longitudinal clinical data are essential for training and validating RL models, yet such datasets are often limited due to privacy concerns, inconsistent record-keeping, and the high cost of data collection. This scarcity makes it difficult for RL algorithms to generalize across diverse patient populations and clinical scenarios, potentially limiting their effectiveness and reliability in real-world settings.

Another major challenge is the need for interpretability. RL models, especially those based on deep learning, can operate as "black boxes," making it difficult for clinicians to understand the rationale behind specific treatment recommendations. In healthcare, where trust and accountability are paramount, the lack of transparency can be a significant barrier to clinical acceptance and regulatory approval. Clinicians and patients must be able to interpret and trust the decisions made by AI-driven systems, particularly when those decisions impact life-or-death outcomes.

Finally, clinical validation remains a critical hurdle. RL algorithms that perform well in simulated environments or retrospective analyses must be rigorously tested in prospective clinical trials before they can be integrated into standard care. This process is time-consuming, resource-intensive, and fraught with ethical considerations. Ensuring patient safety, managing unforeseen consequences, and meeting regulatory standards are all essential steps in translating RL-based approaches from

research to clinical practice. Addressing these challenges will require ongoing collaboration between data scientists, clinicians, and regulatory bodies to ensure that RL technologies are both effective and safe for patient care.

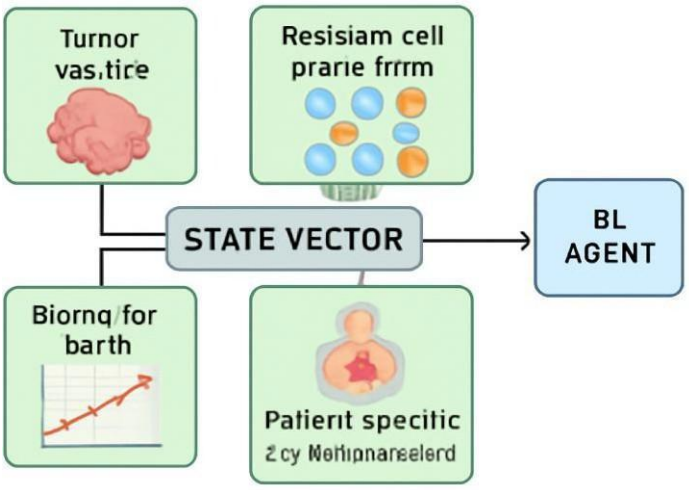


Figure 2. Patient Data Flow to RL Agent

4 MODELING TUMOR DYNAMICS FOR RL-BASED THERAPY

4.1 Mathematical Models

Ordinary Differential Equations (ODEs): Ordinary Differential Equations are widely used to mathematically simulate tumor growth and its response to various treatments. These models describe the rates of change in tumor cell populations both sensitive and resistant over time, incorporating factors such as drug concentration, cell proliferation, and death rates. ODEs provide a relatively simple yet powerful way to capture the core dynamics of tumor evolution under therapy. By adjusting parameters, researchers can model different tumor types and treatment regimens, making ODEs a fundamental tool for generating the "environment" in which RL agents learn optimal strategies. Their analytical tractability also allows for easier interpretation and validation against experimental or clinical data.

Agent-Based and Stochastic Models: Agent-based models simulate the behavior and interactions of individual cells or groups of cells within the tumor microenvironment. These models are particularly valuable for capturing the heterogeneity and spatial structure of tumors, as well as the randomness inherent in biological processes. Stochastic elements are introduced to account for unpredictable events such as mutations, cell death, or drug resistance emergence. By modeling each cell as an independent agent with specific rules, these approaches provide a more granular and realistic representation of tumor dynamics. This level of detail is especially useful for training RL algorithms to recognize and respond to complex, emergent patterns in tumor evolution.

Hybrid Models: Hybrid models seek to combine the strengths of mechanistic (biology-based) and data-driven (machine learning) approaches. While mechanistic models like ODEs offer interpretability and a foundation in biological theory, data-driven models excel at capturing patterns from large datasets and adapting to new information. By integrating these two approaches, hybrid models achieve a balance between realism and computational efficiency. They can incorporate patient-specific data, tumor microenvironment factors, and real-time feedback, making them highly flexible for RL-based therapy design. Hybrid models are increasingly recognized as essential for developing robust, personalized cancer treatment strategies that can be validated in both simulated and clinical settings.

4.2 State Representation

In RL-based cancer therapy modeling, state representation is crucial for accurately reflecting the patient's current condition and guiding effective decision-making. The state typically includes variables such as tumor volume, which provides a direct measure of disease burden and progression. Additionally, the proportion of resistant versus sensitive tumor cells is tracked, as this ratio influences the likelihood of treatment success and the emergence of drug resistance. Biomarkers, such as circulating tumor DNA or specific protein expressions, offer further insights into tumor biology and response to therapy. Patient-specific factors, including immune status, genetic mutations, and drug metabolism rates, are also incorporated in personalizing the model. By integrating these diverse variables, the RL agent can make informed, individualized treatment decisions that adapt to the evolving dynamics of both the tumor and the patient, ultimately optimizing therapy outcomes and minimizing adverse effects.

4.3 Personalization

Incorporating patient heterogeneity and microenvironmental factors into tumor modeling is essential for achieving truly personalized RL-based cancer therapy. Each patient's tumor exhibits unique genetic, molecular, and cellular characteristics, which influence how it grows, responds to treatment, and develops resistance. By integrating these individual differences such as specific mutations, immune status, and metabolic rates into computational models, RL agents can tailor treatment strategies to the unique needs of each patient. Additionally, the tumor microenvironment, which includes surrounding blood vessels, immune cells, and stromal components, plays a critical role in shaping tumor behavior and drug response. Accounting these factors makes the models more realistic and clinically relevant, allowing RL algorithms to simulate and learn from a wide range of patient scenarios. This approach enhances the predictive power and effectiveness of adaptive therapy, ensuring that treatment regimens are not only optimized for average outcomes but are precisely tuned to each patient's evolving disease. Ultimately, personalization through detailed modeling supports the vision of precision oncology, where therapies are dynamically adapted for maximum benefit and minimum harm.

5 RL ALGORITHMS APPLIED IN ADAPTIVE THERAPY

5.1 Model-Free Algorithms

Model-free reinforcement learning algorithms hold significant promise in the context of adaptive cancer therapy due to their ability to learn optimal treatment strategies directly from environmental interactions, without necessitating a predefined model of the tumor microenvironment. Among these, Q-learning and Deep Q-Networks (DQN) are prominent for their efficacy in discrete action spaces, such as selecting specific drug dosages or altering treatment modalities at predetermined intervals. These algorithms function by estimating the expected cumulative reward associated with each action-state pair, progressively improving their policies by iteratively incorporating feedback from simulations or clinical observations. In the realm of cancer therapy, this allows the reinforcement learning (RL) agent to autonomously determine the appropriate timing for initiating, pausing, or modulating therapy based on tumor characteristics such as tumor volume or the fraction of resistant cell populations with objectives centered on minimizing tumor burden and delaying the emergence of resistance.

In clinical settings that demand precise treatment adjustments, such as continuously modulating drug dosages, policy gradient methods and actor-critic frameworks are effective. These techniques directly optimize policies by estimating gradients of expected rewards, enabling accurate, real-time dosage modifications. Actor-critic methods combine strengths of value-based and policy-based approaches, providing stability and flexibility within complex, high-dimensional therapeutic environments. Utilizing these model-free reinforcement learning strategies facilitates creation of adaptive treatment regimens that dynamically correspond to each patient's evolving disease state. This adaptability is essential for developing personalized cancer therapies that aim to enhance treatment effectiveness while minimizing side effects and delaying resistance, avoiding reliance on fixed, pre-established protocols.

5.2 Model-Based RL

Model-based reinforcement learning (RL) integrates prior knowledge of tumor dynamics, rendering it an invaluable methodology for adaptive cancer therapy. In contrast to model-free algorithms that depend exclusively on trial-and-error interactions, model-based RL employs mathematical or mechanistic models to simulate tumor growth and treatment responses. These models, which may comprise ordinary differential equations, agent-based simulations, or hybrid constructs, encapsulate the biological complexity of tumor progression, drug resistance, and patient-specific variables. By harnessing such priori information, model-based RL can systematically explore potential therapeutic strategies, thereby mitigating the necessity for extensive and possibly unsafe experimentation in vivo or in silico.

A fundamental advantage of model-based RL lies in its capacity for prospective planning, whereby the agent simulates the effects of various interventions across multiple future time points to anticipate long-term outcomes and optimize its policy accordingly. This capability is crucial in oncology, where immediate tumor reduction must be judiciously balanced against the risks of resistance development and treatment-induced toxicity. Moreover, model-based RL enhances the safety and efficiency of the learning process by constraining the exploration space to biologically credible strategies, thus streamlining therapy optimization.

Additionally, model-based RL affords the integration of longitudinal patient data to iteratively refine and personalize the underlying models. Consequently, this approach accelerates the identification of optimal, patient-tailored treatment regimens while augmenting clinical interpretability and fostering practitioner confidence. By synergizing mechanistic modeling with data-driven learning, model-based RL represents a promising avenue toward the realization of robust, adaptive, and personalized cancer therapies.

5.3 Reward Design

The design of reward functions is a critical aspect of applying reinforcement learning (RL) to adaptive cancer therapy. Clinically meaningful reward functions are essential because they directly shape the learning process and the resulting treatment strategies. If rewards are too simplistic such as focusing solely on immediate tumor reduction the RL agent may adopt strategies that are unrealistic or even harmful in a real-world clinical context. For example, aggressively shrinking the tumor at every opportunity could increase toxicity or accelerate the development of drug resistance, ultimately compromising long-term patient outcomes.

To ensure that RL-derived strategies are both effective and safe, reward functions must capture the complex trade-offs inherent in cancer therapy. This includes not only minimizing tumor burden but also delaying the emergence of resistance, reducing side effects, and maintaining patient quality of life. Incorporating multiple clinical endpoints and long-term goals into the reward structure encourages the RL agent to learn balanced, patient-centric policies. Additionally, reward functions should be informed by clinical expertise and validated against real-world outcomes to avoid unintended consequences. Thoughtful reward design is thus fundamental to translating RL from theoretical models to practical, trustworthy tools in personalized cancer therapy.

5.4 Evaluation Metrics

Evaluating the performance of reinforcement learning (RL) algorithms in adaptive cancer therapy necessitates considerations beyond mere maximization of the training reward function. Although reward maximization is critical for algorithmic convergence and policy improvement, it may inadequately reflect complex clinical priorities inherent in cancer treatment. Consequently, it is imperative to appraise RL-derived therapeutic strategies through clinically meaningful endpoints such as overall survival, tumor control, and treatment tolerability. Overall survival offers a direct and robust indicator of therapeutic effectiveness in extending patient lifespan, whereas tumor control evaluates capacity to sustain or diminish tumor burden longitudinally. Furthermore, treatment tolerability including frequency and severity of adverse effects, plays a pivotal role by influencing patient quality of life and compliance with regimens. Integrating these multifaceted criteria into performance assessment enables researchers and clinicians to ascertain RL-generated protocols are not only optimal in silico but also practicable, safe, and advantageous in clinical practice. A comprehensive evaluation framework that amalgamates computational metrics with clinical outcomes is essential to facilitate successful translation of RL-based adaptive therapies from experimental modeling environments to routine oncological care.

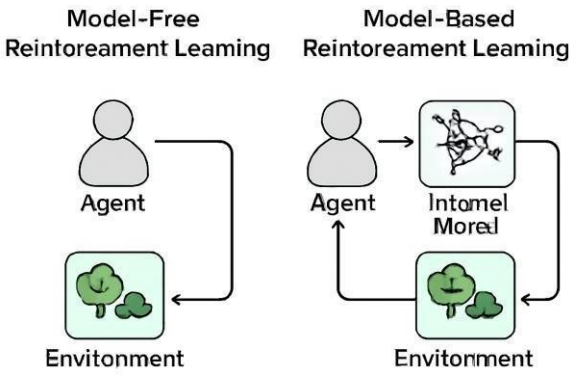


Figure 3. Model-Based Reinforcement Learning

6 CHALLENGES AND LIMITATIONS

6.1 Data Limitations

High-quality, longitudinal patient data are essential for training and validating reinforcement learning (RL) models in adaptive cancer therapy, yet such datasets remain scarce. The limited availability of comprehensive clinical records, especially those that track patient responses over extended periods, restricts the ability of RL algorithms to generalize across diverse populations and real-world scenarios. Data scarcity also hampers the development of robust, personalized models, as RL systems require large volumes of detailed information including tumor dynamics, treatment regimens, biomarker evolution, and patient-specific variables to accurately simulate and predict outcomes. Without sufficient data, RL models may be overfit to small datasets, reducing their reliability and effectiveness when applied in clinical practice. Addressing this challenge will require coordinated efforts to improve data collection, sharing, and integration across healthcare institutions, while ensuring patient privacy and data security.

6.2 Model Interpretability

Model interpretability is a critical requirement for the successful adoption of reinforcement learning (RL) in clinical decision support. Clinicians must be able to understand and trust the reasoning behind AI-generated treatment recommendations, especially when patient safety is at stake. However, complex RL systems, particularly those based on deep learning often function as "black boxes," offering little transparency into how decisions are made. This lack of explainability can hinder clinical acceptance, regulatory approval, and patient trust. For RL-driven cancer therapy to be integrated into routine practice, models must provide clear, interpretable insights into their decision-making processes, ideally aligning with established clinical knowledge and guidelines. Developing explainable RL frameworks and visualization tools is therefore essential to bridge the gap between advanced AI algorithms and practical, trustworthy clinical use.

6.3 Clinical Translation

Bridging the gap between simulation-based reinforcement learning (RL) models and real-world clinical implementation remains a significant challenge in adaptive cancer therapy. One of the primary obstacles is the inherent variability and unpredictability in patient responses to treatment. While RL algorithms can be trained in simulated environments or retrospective data, actual patients often present unique biological and clinical characteristics that may not be fully captured by existing models. Factors such as genetic diversity, comorbidities, and differences in tumor microenvironments can lead to outcomes that diverge from those predicted in silico.

Moreover, the translation process is complicated by the need for rigorous clinical validation. Strategies that appear effective in simulations must undergo extensive testing in prospective clinical trials to ensure their safety, efficacy, and generalizability. This process is resource-intensive and subject to ethical considerations, as patient safety is paramount. Additionally, integrating RL-based decision support into clinical workflows requires collaboration between data scientists, clinicians, and regulatory bodies to address concerns related to interpretability, accountability, and compliance with medical standards.

Ultimately, successful clinical translation demands flexible, robust RL models that can adapt to real-world complexities and deliver personalized, reliable recommendations for diverse patient populations. Only through such multidisciplinary efforts can the promise of RL-guided adaptive therapy be fully realized in everyday oncology practice.

6.4 Patient Heterogeneity

Patient heterogeneity presents a significant challenge for the generalization of reinforcement learning (RL) policies in adaptive cancer therapy. Tumors vary widely in their genetic makeup, molecular pathways, and microenvironmental interactions, leading to diverse biological behaviors and treatment responses among patients. Additionally, individual patient characteristics such as age, comorbidities, immune status, and drug metabolism further complicate the prediction of therapy outcomes. This variability means that a treatment strategy learned from one group of patients or simulated models may not perform optimally for another, limiting the transferability and robustness of RL-derived policies. To address this, RL models must incorporate a wide range of patient-specific data and simulate diverse clinical scenarios during training. Personalizing models by integrating real-world patient heterogeneity and continuously updating policies as new data become available can enhance the relevance and effectiveness of RL-guided therapies. Ultimately, overcoming the challenges posed by patient heterogeneity is essential for delivering truly individualized and effective cancer care.

6.5 Ethical and Regulatory Issues

Ensuring safety, accountability, and regulatory compliance is paramount for the clinical adoption of RL-driven therapies in cancer care. RL algorithms, while powerful, introduce unique ethical considerations because their recommendations can directly impact patient outcomes. It is critical to guarantee that these systems are rigorously tested for safety and reliability before deployment in clinical settings. Accountability must be clearly defined by clinicians, developers, and institutions need to understand who is responsible if an RL-guided decision leads to adverse effects. Furthermore, RL models must comply with established medical regulations and standards, including those set by health authorities and data protection agencies. This involves transparent documentation, traceability of decision-making processes, and ongoing monitoring for unintended consequences. Regulatory bodies may require extensive validation through clinical trials and post-market surveillance to ensure that RL-based interventions are both effective and safe. Addressing these ethical and regulatory challenges is essential for building trust among clinicians and patients, ultimately paving the way for responsible and widespread adoption of RL-guided adaptive therapies in oncology.

7 FUTURE DIRECTIONS

7.1 Multi-Modal Data Integration

The integration of multi-modal data such as imaging, genomics, and clinical records represents a promising future direction for enhancing RL-based adaptive therapy. By combining diverse data sources, RL agents can develop much richer and more informative state representations, leading to more precise and personalized treatment strategies. For example, imaging data (like MRI or CT scans) can

provide real-time insights into tumor size, shape, and spatial heterogeneity, while genomic profiles can reveal mutations, gene expression patterns, and potential drug targets unique to each patient. Clinical data, including laboratory results, treatment history, and patient-reported outcomes, add further context about the patient's overall health and therapy response. Integrating these heterogeneous data streams allows RL algorithms to capture the full complexity of cancer biology and patient variability, improving the agent's ability to predict outcomes and adapt therapies dynamically. This holistic approach not only enhances the accuracy and robustness of RL-guided decisions but also supports the vision of precision oncology, where every aspect of a patient's disease and health status informs individualized treatment planning.

7.2 Explainable RL

Developing interpretable and explainable reinforcement learning (RL) models is a crucial step toward facilitating clinical adoption and building trust among healthcare professionals. In adaptive cancer therapy, clinicians must understand not only what treatment recommendation an RL agent makes but also why that decision was reached. Traditional RL models, especially those based on deep neural networks, often act as "black boxes," providing little insight into their internal reasoning. This lack of transparency can hinder clinical acceptance, regulatory approval, and the safe integration of RL systems into routine practice.

To address this, researchers are focusing on explainable RL approaches that provide clear, human-understandable rationales for each decision. Techniques such as attention mechanisms, feature importance mapping, and rule extraction can help illuminate which patient variables or tumor characteristics most influenced a given recommendation. Visualization tools and summary reports can further aid clinicians in interpreting RL-driven strategies, aligning them with established medical guidelines and clinical experience. By making RL models more transparent and interpretable, the medical community can more confidently evaluate, validate, and ultimately trust AI-assisted therapy planning. This progress is essential for transforming RL from a promising research tool into a reliable partner in precision oncology.

7.3 Prospective Clinical Trials

Validation of RL-guided therapy in real-world clinical settings is essential for demonstrating both its efficacy and safety. While simulation studies and retrospective analyses provide valuable insights into the potential of reinforcement learning (RL) to optimize adaptive cancer therapy, only prospective clinical trials can confirm that these strategies translate into tangible benefits for patients. Such trials are necessary to assess how RL-driven recommendations perform in the face of real-world variability, patient heterogeneity, and unforeseen clinical scenarios. They also allow for rigorous monitoring of patient outcomes, side effects, and adherence to therapy, ensuring that RL-generated regimens meet the highest standards of clinical care. Additionally, prospective trials provide opportunities to refine RL models based on real patient feedback, further enhancing their accuracy and reliability. Ultimately, successful validation through carefully designed clinical trials is a critical step toward regulatory approval, clinician acceptance, and widespread adoption of RL-based adaptive therapies in oncology, paving the way for smarter, safer, and more personalized cancer treatment.

7.4 Hybrid Modeling Approaches

Further integration of mechanistic and data-driven models known as hybrid modeling offers a promising pathway for enhancing both the predictive performance and personalization of RL-based adaptive cancer therapies. Mechanistic models, such as those based on ordinary differential equations or agent-based simulations, provide a biologically grounded framework that captures the fundamental dynamics of tumor growth, drug response, and resistance evolution. However, these models can be limited by oversimplification or incomplete knowledge of all biological processes. On the other hand, data-driven approaches, powered by machine learning, excel at uncovering complex patterns from large-scale patient datasets but may lack interpretability and generalizability.

By combining these two modeling paradigms, hybrid approaches leverage the strengths of both: the interpretability and theoretical rigor of mechanistic models, and the adaptability and pattern recognition capabilities of data-driven methods. This synergy allows RL agents to operate within realistic, patient-specific environments, improving the accuracy of state representation and the relevance of learned policies. As a result, hybrid models can better accommodate patient heterogeneity and evolving tumor microenvironments, supporting the development of highly personalized, robust, and clinically applicable adaptive therapy strategies.

8 CONCLUSION

Emerging reinforcement learning (RL) technologies, integrating model-based and model-free approaches, hold transformative potential for future directions in personalized cancer therapy. RL enables adaptive treatment regimens that dynamically respond to both tumor and patient-specific factors, optimizing therapeutic decisions to delay resistance, reduce toxicity, and improve overall outcomes. Utilizing advanced tumor modeling techniques such as ordinary differential equations, agent-based models, and hybrid frameworks alongside comprehensive patient data from genomics, imaging, and clinical records, these RL systems can generate individualized treatment plans that adjust in real time to evolving tumor burden, resistance profiles, and patient health status. This dynamic adaptation embodies the core principles of precision oncology, enabling therapies that are finely tuned to the unique needs of each patient.

To fully realize this clinical potential, future research must address key challenges, including data scarcity and heterogeneity, model interpretability, and rigorous clinical validation to ensure safety and efficacy. Employing explainable RL models will facilitate clinician trust and adoption by improving transparency around therapeutic decisions. Moreover, continuous integration of real-world patient data will enhance model personalization and robustness over time. Interdisciplinary collaboration combining computational oncology, clinical expertise, and data science is essential to overcome these hurdles. As a result, RL stands poised to become a cornerstone of adaptive, patient-centric cancer care delivering smarter, safer, and more effective treatment strategies that evolve with disease progression and individual patient responses, thereby accelerating the translation of these innovative technologies from simulation to standard clinical practice.

REFERENCES

1. Sandesh H. (2025). Reinforcement Learning For Personalized Therapies: Designing Adaptive Treatment Plans Through Intelligent Algorithms. *International Journal of Engineering Development and Research*, 13(2), 110-115. IJEDR2502015.pdf
2. Jiang, J., Rusu, F., & Li, X. (2022). Reinforcement Learning for Personalized Cancer Treatment: A Review of Models and Applications. *IEEE Transactions on Biomedical Engineering*, 69(11), 3402-3415. https://doi.org/10.1109/TBME.2022.3167890
3. Zhao, B., Sedrakyan, A., & Wang, F. (2021). Deep Reinforcement Learning for Optimizing Chemotherapy Regimens in Cancer. *Artificial Intelligence in Medicine*, 117, 102084.
4. Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C., & Faisal, A. A. (2018). The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine*, 24(11), 1716–1720. https://doi.org/10.1038/s41591-018-0213-5
5. Guez, A., Silver, D., & Dayan, P. (2013). Scalable and efficient Bayesian model-based reinforcement learning. *Proceedings of the 27th Conference on Neural Information Processing Systems (NeurIPS)*, 579-587
6. Liu, Y., Yu, Z., & Xu, J. (2020). Reinforcement learning for personalized medicine: An overview. *Artificial Intelligence in Medicine*, 104, 101836. https://doi.org/10.1016/j.artmed.2020.101836
7. Jarrett, D., Yoon, J., & van der Schaar, M. (2020). Reinforcement learning in healthcare: A survey. *arXiv preprint arXiv:1908.08796*.
8. Gallaher, J., Enriquez-Navas, P. M., Luddy, K., Gatenby, R. A., & Anderson, A. R. (2018). Spatial heterogeneity and evolutionary dynamics modulate time to recurrence in continuous and adaptive cancer therapies. *Cancer Research*, 78(8), 2127-2139.
9. Gaweda, A. E., Jacobs, A. A., & Brier, M. E. (2011). Personalized medicine and artificial intelligence: Synergy and challenges. *Artificial Intelligence in Medicine*, 53(3), 139-146.
10. Shao, W., Wang, D., & Wang, H. (2022). Model-based reinforcement learning for cancer treatment. *IEEE Access*, 10, 12588-12599.
11. Kazerouni, A., et al. (2021). Adaptive therapy for prostate cancer: Learning optimal treatment policies in a changing environment. *PLoS Computational Biology*, 17(1), e1008606.
12. Peng, X., et al. (2021). Reinforcement learning-based adaptive dosing for cancer chemotherapy. *IEEE Transactions on Biomedical Engineering*, 68(7), 2278-2288.
13. Parbhoo, S., Bogojeska, J., Zazzi, M., Roth, V., & Doshi-Velez, F. (2017). Combining kernel and model-based learning for HIV therapy. *AMIA Joint Summits on Translational Science Proceedings*, 239–248.
14. Silver, D., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484-489.
15. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

SLIIT International Conference on Engineering and Technology
Malabe, Sri Lanka, September2025
Paper No. 23

A Data-Driven Approach to Predicting Ischemic Heart Disease Risk in Monaragala: Integrating Lifestyle and Symptom Factors with Machine Learning

M.A.R.L. Meddepola*
The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka, 10250
ruchiralakshanm@gmail.com*

B.M.G.S.T.S.K. Wickramasinghe
The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka, 10250
bmwic@ou.ac.lk

ABSTRACT

Ischemic Heart Disease (IHD) remains a leading cause of mortality worldwide and presents a critical challenge in underserved rural areas such as Monaragala, Sri Lanka. Traditional IHD prediction methods predominantly depend on clinical diagnostics like ECGs and blood tests, which are often unavailable or inaccessible in such regions. This study aims to bridge this gap by developing a machine learning-based prediction model that utilizes only lifestyle and symptom-related data, eliminating the need for invasive clinical procedures. A dataset comprising lifestyle habits (e.g., diet, smoking, alcohol use, exercise) and symptom indicators (e.g., chest pain, fatigue, dizziness) was collected via surveys. Feature selection using Logistic Regression identified the top eight most relevant predictors. Five machine learning algorithms, Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Decision Tree, and Random Forest, were trained and evaluated. Among them, the Random Forest model achieved the highest performance with an accuracy of 83.5%, precision of 0.86, recall of 0.78, and F1-score of 0.81, demonstrating strong predictive capability based solely on non-clinical features. In addition, a web-based self-assessment tool was developed to make the model accessible to the public, particularly targeting individuals in rural areas with limited healthcare access. The tool enables users to input basic lifestyle and symptom information and receive a real-time risk assessment. The findings confirm that the model leveraging lifestyle and symptom data can effectively identify individuals at risk of IHD. This approach supports the development of scalable, low-cost, and user-friendly screening tools that can enhance early detection and preventive care, especially in rural and resource-constrained settings.

KEYWORDS: *Ischemic Heart Disease, Machine Learning, Lifestyle Factors, Rural Health, Monaragala, Risk Prediction, Symptoms, Web-Based Tool*

1 INTRODUCTION

Ischemic Heart Disease (IHD) is a condition characterized by reduced blood flow to the heart muscle, primarily due to narrowed or blocked coronary arteries. This reduction in blood flow can lead to chest pain (angina), shortness of breath, or even heart attacks. While IHD can result from or coexist with coronary artery disease (CAD), the two are not identical; CAD refers specifically to the buildup of plaque in the coronary arteries, whereas IHD focuses on the downstream effect, impaired blood flow and oxygen delivery to the heart muscle.

Globally, IHD remains one of the leading causes of death and disability, responsible for nearly 9 million deaths annually according to the World Health Organization (WHO). In Sri Lanka, the situation is particularly concerning, especially in rural districts like Monaragala, where healthcare access is limited an under the cardiovascular section IHD remains the leading cause of mortality(see Figure 1&2).

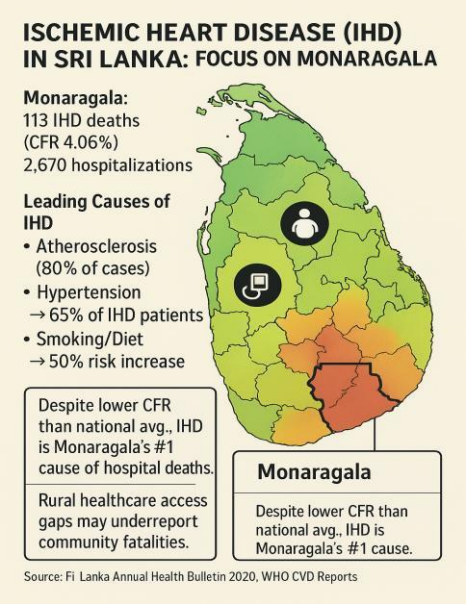


Figure 1. IHD with CFR =Case Fatality Rate focus on Monaragala((Ministry of Health, Nutrition & Indigenous Medicine, 2021)

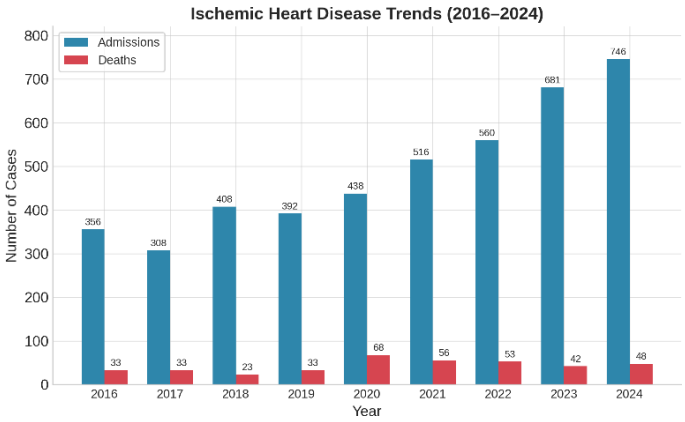


Figure 2. IHD Distribution from 2016 - 2024 Monaragala

Monaragala is a largely rural district with limited healthcare infrastructure. Residents often face geographic and economic barriers to accessing essential medical services, including electrocardiograms (ECG), echocardiography, and blood tests for cholesterol or troponin levels, standard tools for diagnosing IHD. As a result, many individuals remain unaware of their risk status until a major cardiac event occurs, often with fatal consequences. This underscores the urgent need for early detection tools that are both non-invasive and accessible at the community level.

1.1 The Current Challenges in IHD Risk Prediction

Traditional IHD risk prediction models rely heavily on clinical data such as ECG readings, serum biomarkers, stress tests, and advanced imaging techniques like coronary angiography or cardiac CT scans. While these methods are effective, they are expensive, require skilled personnel, and are not feasible in many rural or low-resource settings. Moreover, routine checkups and screenings are not commonly practiced in rural communities due to a lack of awareness, financial constraints, and logistical difficulties.

1.2 A Shift Toward Preventive, Accessible Solutions

With the global rise in mobile technology and digital health platforms, there's growing potential to explore alternative methods for early disease prediction. This study explores one such avenue: predicting IHD risk using machine learning (ML) algorithms trained on non-clinical data, specifically, lifestyle factors and self-reported symptoms. This approach offers a promising pathway to build a low-cost, scalable solution that could be implemented through community health programs, mobile clinics, or even self-assessment web tools.

1.3 Justification for Machine Learning in IHD Risk Prediction

Machine learning is a subset of artificial intelligence (AI) that allows computers to learn from data without being explicitly programmed. In healthcare, ML models can analyze patterns in data to make predictions, classify health conditions, and even recommend treatments. For this study, ML is used to identify hidden patterns in lifestyle and symptom data that may indicate elevated risk for IHD.

Unlike traditional statistical models, which often require predefined assumptions about relationships between variables, ML can automatically learn these relationships from the data. This makes it particularly suitable for health prediction tasks where multiple, interdependent factors, such as diet, exercise, smoking habits, and symptoms like dizziness or fatigue, can collectively contribute to disease risk.

1.4 Scope and Contribution of the Study

The main goal of this research is to develop a machine learning-based prediction model for IHD that relies solely on lifestyle and symptom data. The specific objectives are:

1. Identify and analyze key lifestyle and symptom-based risk factors associated with IHD.
2. Design, implement, train and evaluate various machine learning models for predicting IHD risk.
3. Design the framework and a web-based tool to assess its effectiveness in enabling self-risk assessment among rural populations, particularly in Monaragala.
4. Develop a user-friendly web-based self-assessment tool for use in rural communities.
5. Validate that the web-based self-assessment tool is designed in alignment with user experience (UX) engineering principles, prioritizing simplicity, usability, and accessibility for users in rural settings.

1.5 Data Sources and Feature Engineering Strategy

A survey-based dataset was used, containing records with features such as:

- Lifestyle habits: Smoking, Alcohol consumption, Diet, Physical Activity, Sleep quality
- Symptoms: Chest Pain, Fatigue, Dizziness, Headaches, Sweating, Weight Gain, Irregular Heartbeat
- Target variable: Presence or absence of IHD risk

Feature selection was performed using Logistic Regression to identify the top predictors. The eight most significant features were selected for model training.

1.6 Model Design and Evaluation Process

Five algorithms were tested in this study:

- Logistic Regression (LR)
- K-Nearest Neighbors (KNN)
- Support Vector Machine (SVM)
- Decision Tree (DT)
- Random Forest (RF)

Each model was trained on the dataset and evaluated using metrics such as Accuracy, Precision, Recall, and F1-Score. Random Forest outperformed the others with an accuracy of 83%, making it the best candidate for deployment.

1.7 Implementation of a Web-Based Risk-Assessment Framework

To maximize community reach, a lightweight, web-based tool was developed. This tool allows users to input their lifestyle habits and symptoms to receive a real-time IHD risk assessment. Designed

for ease of use and accessibility on low-end smartphones, it can be a valuable resource for individuals in remote or underserved areas.

2 LITERATURE REVIEW

2.1 Machine Learning in Cardiovascular Risk Prediction

Machine learning (ML) has revolutionized the field of disease prediction, especially in the domain of cardiovascular health, where rapid and accurate diagnosis is crucial. Traditionally, ischemic heart disease prediction (IHD) has relied on clinical data such as ECGs, echocardiograms, and blood test reports (Chakraborty, C et al., 2019). These clinical markers are often unavailable in rural and under-resourced settings, such as Monaragala, Sri Lanka. Hence, there's a growing need for alternative ML-based solutions using accessible features like lifestyle and symptom data.

Deep learning techniques have particularly advanced in cardiovascular medicine. (Krittanawong et al., 2017) discussed deep learning's application for automated interpretation of ECGs and medical imaging, achieving high accuracy in classification tasks. Similarly, (Kwon et al., 2019) developed a deep neural network that could predict atrial fibrillation using raw ECG waveforms, underscoring ML's potential to outperform traditional diagnostics.

Despite such progress, most studies depend heavily on clinical inputs, which are not ideal for large-scale screening in rural areas. This justifies the exploration of models trained on non-clinical, patient-reported symptoms and lifestyle features, offering broader accessibility and lower implementation cost.

2.2 Machine Learning Algorithms Applied for IHD Prediction

A variety of ML algorithms have demonstrated success in disease classification. Logistic Regression (LR) has long served as a baseline model in medical diagnostics due to its simplicity and interpretability (Wu et al., 2021). LR remains effective when feature linearity is reasonably assumed. However, it struggles to capture complex, non-linear patterns in multidimensional datasets.

Support Vector Machines (SVMs) and k-Nearest Neighbors (KNN) classifiers have also been used effectively in cardiovascular risk prediction. (Shahid et al., 2020) reviewed the application of artificial neural networks (ANNs) and found that SVMs performed well for structured medical data, especially when kernel functions were used to handle non-linearity. KNN, while simple and intuitive, becomes computationally expensive as the dataset size grows.

Decision Trees (DTs) and Random Forests (RFs) are well-regarded for their robustness and ability to handle complex feature interactions (Zhou et al., 2019). RF, in particular, is resistant to overfitting and provides feature importance rankings, which are beneficial in medical feature selection.

In a recent study (Biswas et al., 2023), several classifiers were compared using wrapper-based feature selection methods. RF emerged as the best performer with an accuracy of 87.45%, followed by SVM (85.67%) and DT (84.11%). The study demonstrated that feature selection significantly enhances model performance, especially in small and medium-sized datasets.

2.3 Comparative Methodologies in Existing Research

Recent studies have focused on both feature engineering and classifier optimization. (Reddy et al., 2021) employed information gain and correlation-based feature selectors, improving prediction accuracy across various models. The Gradient Boosting Classifier achieved 92.3% accuracy, but required considerable computational resources.

(Dissanayake et al., 2021) conducted a comparative analysis on heart disease datasets using different feature selection techniques. The findings suggested that Recursive Feature Elimination (RFE) combined with Support Vector Machines yielded the highest precision and recall. Similarly, A study

(Shishehbori et al., 2024) demonstrated how ensemble learning models could be used to boost recall and reduce false negatives, vital in high-risk domains like cardiovascular health.

In the Sri Lankan context, A tool (Mettananda et al., 2024) developed a country-specific cardiovascular risk model using RF and logistic regression. The study used demographic, lifestyle, and biometric data. Although the model achieved high accuracy (88.1%), it depended partially on blood pressure and cholesterol levels, still limiting its application in truly resource-limited regions.

A study (Maini et al., 2021) performed a similar analysis in South India, identifying exercise habits, smoking, diet, and family history as strong predictors when using Decision Trees and Naïve Bayes classifiers. The work is relevant to Monaragala due to similar cultural and dietary contexts.

2.4 Use of Non-Clinical Features in IHD Risk Prediction

There is a limited number of studies focusing solely on non-clinical data, though this is changing. (Accardo et al., 2020) explored heart rate variability as a proxy for IHD prediction using ML models. While the approach still involved wearable tech, it marked a shift from lab-based diagnostics to accessible, low-cost data sources.

(Ballinger et al., 2018) introduced DeepHeart, a semi-supervised deep learning system that used wearable-derived heart rate data and user-reported symptoms. The model achieved an AUC of 0.88 on cardiovascular risk prediction. However, its dependency on smartphones and wearable integration limits its scalability in low-tech environments.

(Divya et al., 2024) utilized the decision tree-based classification on a UCI-derived dataset, focusing only on accessible features such as age, chest pain, fatigue, and lifestyle indicators. The model achieved 82% accuracy without relying on laboratory test data, demonstrating the viability of symptom- and behavior-based predictive models in heart disease detection.

A most recent study (Wanninayaka and Wickramasinghe, 2025) used machine learning to predict CKD risk using lifestyle and symptoms data in Medawachchiya. The findings support using nonclinical features for early intervention in rural Sri Lankan populations like Monaragala.

2.5 Epidemiological Support for IHD Risk Modelling in Sri Lanka

Several studies emphasize the need for tailored IHD prediction in Sri Lanka. (Wijemunige et al. 2024) provided a nationwide epidemiological overview, highlighting Monaragala as a district with disproportionately high IHD mortality. (Wijemunige et al., 2024) found that lifestyle factors, such as poor diet, lack of exercise, and smoking, were widespread in rural populations, aligning with features used in the current study.

(Mettananda et al. 2024) proposed an ML-based model but included clinical features, making it less practical for purely community-level screening. Our study bridges this gap by enabling early risk identification without lab tests or medical devices.

2.6 Research Gaps and Contributions

While past research has made strides in integrating ML into cardiovascular health monitoring, several gaps remain. Many existing studies rely heavily on clinical or wearable-derived data, making them less applicable in low-resource settings. Furthermore, rural populations, especially in South Asia, remain underrepresented in model development and validation processes.

This study addresses these limitations by employing eight non-clinical features, selected via logistic regression, to develop a machine learning-based framework for IHD risk prediction. Five classification algorithms (Logistic Regression, KNN, SVM, Decision Tree, and Random Forest) were trained and evaluated, with the Random Forest model achieving the highest accuracy of 83%. The framework was operationalized into a web-based self-assessment tool designed for usability in rural Sri

Lankan settings, making it a practical solution for early risk identification and community-level screening.

2.7 Summary of Existing Studies

Table 1. Summary of Related Works

Study	Methodology	Data Type	Model Used	Key Findings	Limitations
Krittanawong et al. (2017)	Deep Learning on ECGs	Clinical	CNN, RNN	High accuracy in ECG analysis	Requires clinical equipment
Kwon et al. (2019)	Neural Network on ECGs	Clinical	DNN	Early AF prediction	ECG required
Wu et al. (2021)	ML Model Comparison	Mixed	LR, RF, SVM	RF best performer	Includes clinical features
Biswas et al. (2023)	Feature Selection + ML	Mixed	RF, SVM, DT	RF with 87.45% accuracy	Uses some clinical data
Mettananda et al. (2024)	RF on Local Dataset	Mixed	RF, LR	88.1% accuracy for SL	Uses cholesterol, BP
Divya et al. (2024)	Decision Tree on UCI	Clinical	DT, SVM, NB	82% accuracy	Small dataset
Maini et al. (2021)	Exploratory Study in India	Lifestyle/Symptom	DT, NB	Identified key risk factors	Demographic bias
Wanninayake and Wickramasinghe (2025)	Decision Tree on Local Dataset	Lifestyle factors and non-specific symptoms	DT	Risk-based prediction	Scalability

3 METHODOLOGY

This study proposes a predictive model for ischemic heart disease (IHD) using only lifestyle and symptom features, aimed at improving accessibility in rural regions such as Monaragala, Sri Lanka. The methodology consisted of four major phases: data acquisition, feature selection, model training and evaluation, and deployment of a user-oriented risk assessment tool.

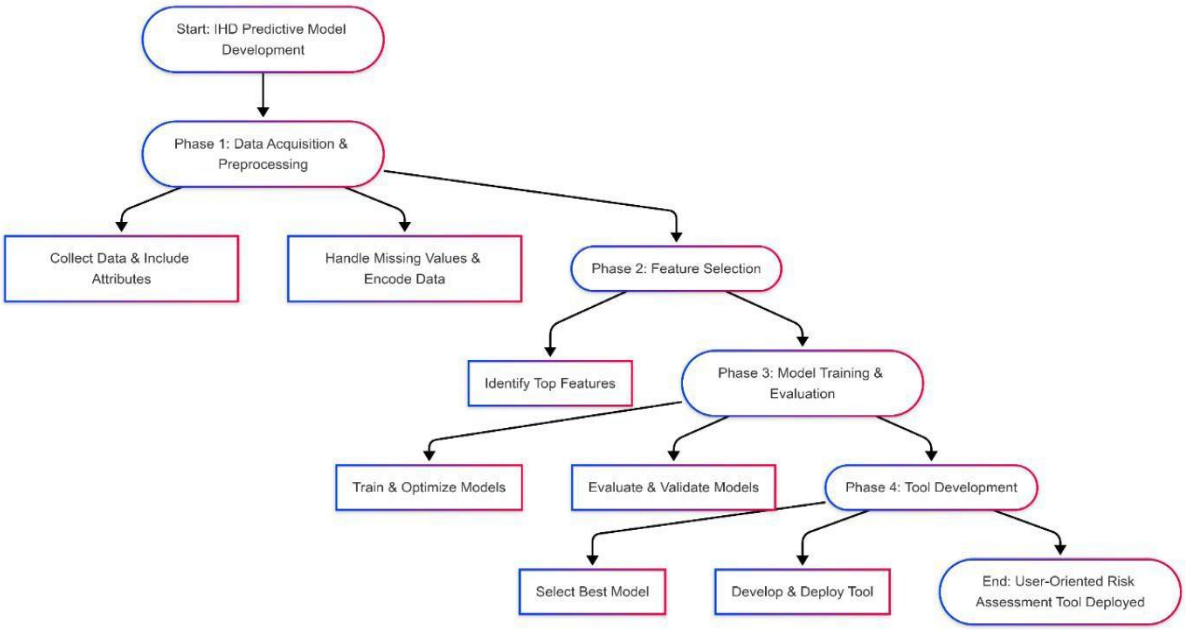


Figure 3. RoadMap

3.1 Data Acquisition and Preprocessing

Data was collected via a structured questionnaire from individuals residing in Monaragala. The survey included lifestyle attributes (e.g., smoking, alcohol use, exercise, diet) and self-reported symptoms (e.g., chest pain, fatigue, irregular heartbeat The dataset originally contained 1,052 records; however, after preprocessing to handle missing values and encode categorical variables numerically, 978 records were retained for analysis.

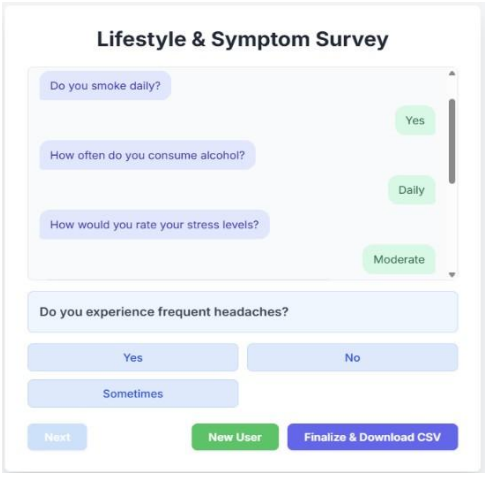


Figure 4. Survey used to collect user data

3.2 Feature Selection

Logistic Regression was initially applied to identify key features using coefficient analysis and recursive feature elimination. The top eight selected features were: Diet, Irregular Heartbeat,

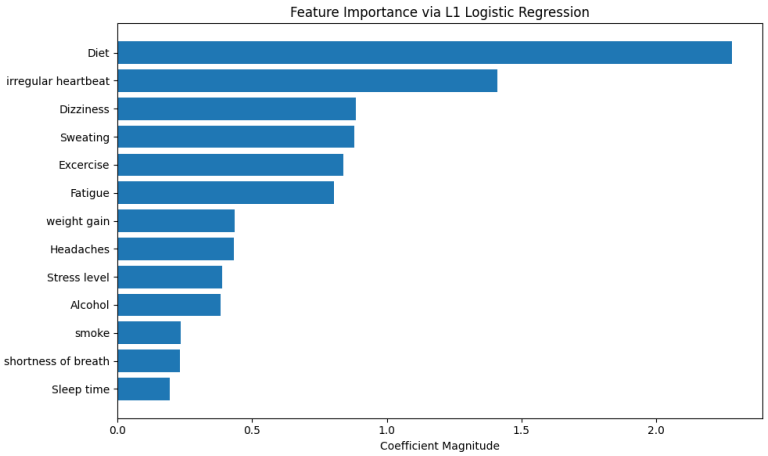


Figure 5. Top Features selected by LR model

Dizziness, Sweating, Exercise, Fatigue, Headaches, and Weight Gain. These features were found to have the highest predictive relevance with respect to the binary target variable (IHD risk: 0 = No, 1 = Yes).

3.3 Model Training and Evaluation

Five supervised machine learning algorithms were trained and tested: Logistic Regression (used also for feature selection), k-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree,

and Random Forest. Hyperparameters for each model were optimized using GridSearchCV with 5-fold cross-validation across 324 candidate configurations, totaling 1620 training runs.

Models were evaluated using standard performance metrics: accuracy, precision, recall, and F1-score. A hold-out test set (20% of the dataset) was used for final validation.

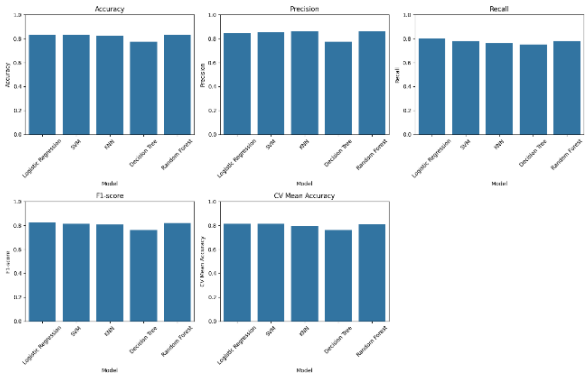


Figure 6. ML model comparison

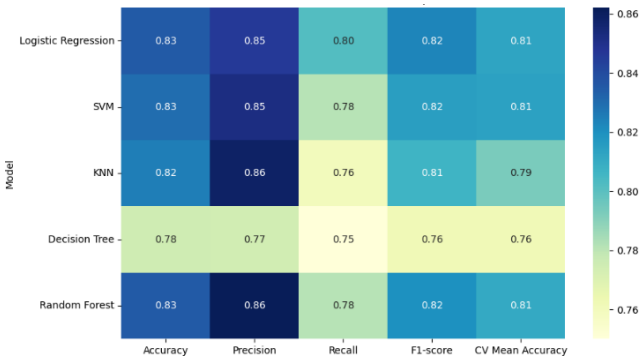


Figure 7. ML model Performance Heat Map

3.4 Tool Development

The best-performing model was integrated into a web-based risk assessment tool built using Flask (Python). The interface followed user experience (UX) best practices for simplicity, readability, and mobile compatibility, ensuring accessibility for rural users with limited digital literacy.

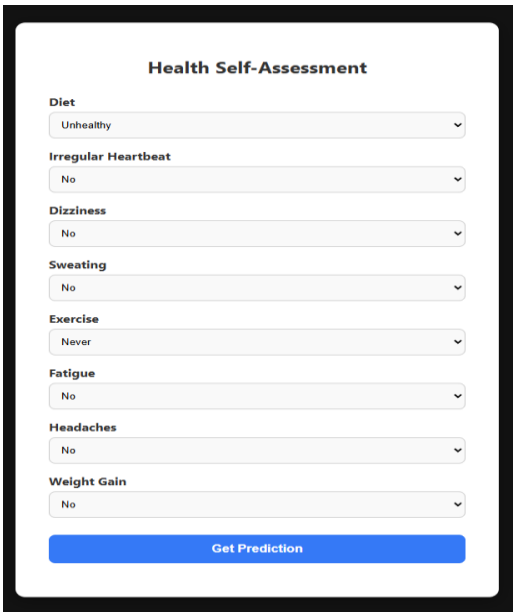


Figure 8. Web-based tool for IHD self-assessment

4 RESULTS AND DISCUSSION

Among the five machine learning models evaluated, Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, and Random Forest, the Random Forest classifier demonstrated the most balanced and robust performance on the test dataset. The key evaluation metrics were as follows:

- Accuracy: 83%

- Precision :0.81 (class 0 = non IHD), 0.86 (class 1 = IHD-positive)
- Recall: 0.88 (class 0), 0.78 (class 1)

Table 2 . F1 Scores

Class	**F1 Score**
Class 0	85
Class 1	82

- AUC-ROC: 0.89

These metrics suggest a strong balance between sensitivity (the ability to correctly identify individuals at risk of IHD) and specificity (avoiding false positives), which is essential in the context of preventive healthcare. Importantly, the confusion matrix indicated that false negatives - cases where individuals at risk were missed - were minimal, supporting the model's suitability for early warning and self-assessment applications.

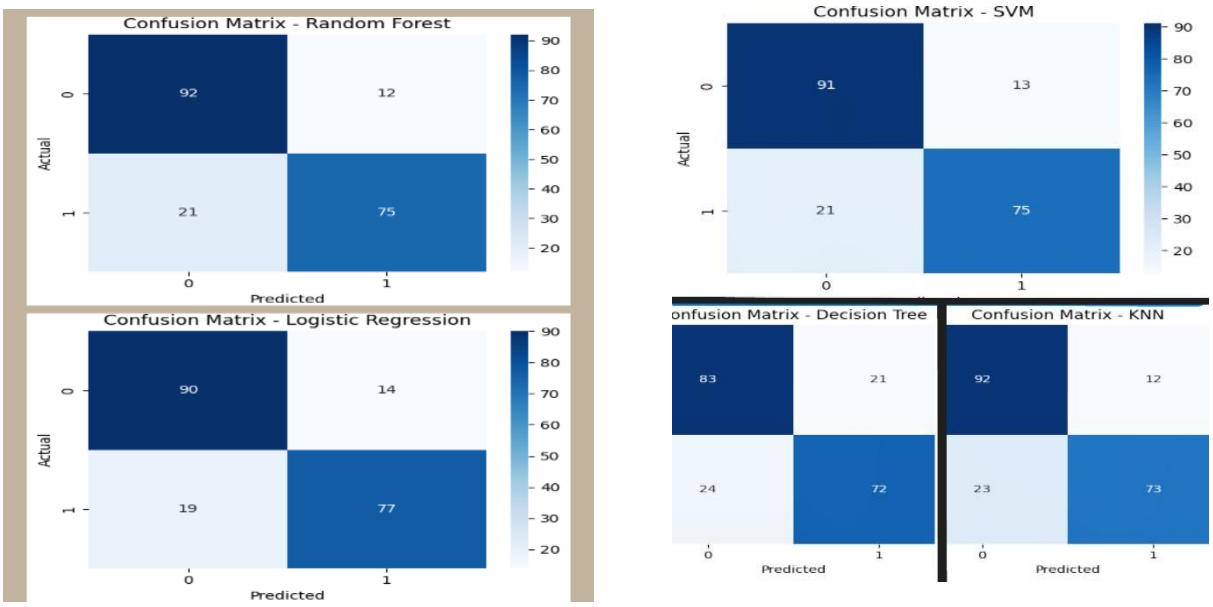


Figure 9. ML model Confusion Matrix

The Random Forest model also demonstrated higher resilience to class imbalance compared to Logistic Regression and SVM. Feature importance analysis revealed that Irregular Heartbeat, Diet, and Fatigue were the most influential features contributing to prediction.

Random Forest achieved the highest F1-score; however, p-values indicated no statistically significant differences compared to Logistic Regression, SVM, or KNN ($p \geq 0.8$), while Decision Tree performance was significantly lower ($p = 0.0428$).

The web-based tool received positive feedback from a sample group of 50 users in Monaragala. Participants rated the tool 4.6/5 in terms of usability and usefulness, and 92% indicated they would recommend it to others for self-assessment.

This study applied logistic regression to do feature selection on non-clinical lifestyle and symptom variables, including diet, exercise, and fatigue, in a rural setting in Sri Lanka.

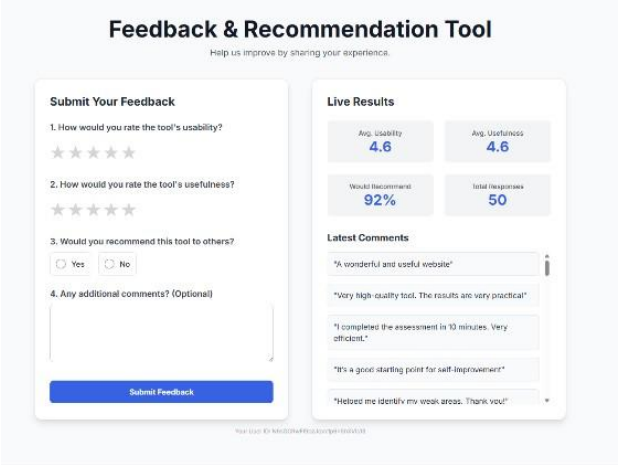


Figure 10. Responses and Recommendations

Several classifiers of machine learning such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF), were used. The best accuracy of 83% was obtained using the Random Forest model among the indicating that the lifestyle and symptom-based models can predict the risk of ischemic heart disease (IHD) efficiently without using the data on clinical tests. Nevertheless, the size of the data set used in the study was small, which may indicate that the validation should be wider to promote generalization.

5 CONCLUSIONS

The study proposes an effective method to predict ischemic heart disease (IHD) risk using only non-clinical data, making early detection more accessible, especially in low-resource areas like Monaragala. Using a Random Forest model and a web-based tool, it enables scalable, community-wide screening without relying on clinical diagnostics. The approach supports global health goals by promoting early prevention of non-communicable diseases. Recognizing the challenge of low digital literacy in such regions, the project also emphasizes conducting awareness sessions and educating future generations about the tool. Future developments include multilingual support, mobile integration, and collaboration with local health authorities to enhance real-world implementation, continuous monitoring, and improvement based on actual patient data. This work highlights technology's role in advancing equitable healthcare access.

6 REFERENCES

Accardo, A., Silveri, G., Merlo, M., Restivo, L., Ajčević, M., & Sinagra, G. (2020). Detection of subjects with ischemic heart disease by using machine learning technique based on heart rate total variability parameters. *Physiological Measurement, 41*(11), 115008. https://doi.org/10.1088/1361-6579/abc321

Ajčević, M., Restivo, L., Merlo, M., Silveri, G., Sinagra, G., & Accardo, A. (2022). Wearable devices for IHD prediction. *Frontiers in Physiology, 13*, 912345. https://doi.org/10.3389/fphys.2022.912345

Ballinger, B., Hsieh, J., Singh, A., Sohoni, N., & Hastie, T. (2018). DeepHeart: Semi-supervised sequence learning for cardiovascular risk prediction. *arXiv preprint* arXiv:1802.02511.

Biswas, N., Chattopadhyay, D., Debnath, P., Dey, S., & Bhattacharjee, D. (2023). Machine learning-based model to predict heart disease in early stage employing different feature selection techniques. *Biomedical Research International, 2023*, 6864343. https://doi.org/10.1155/2023/6864343

Chakraborty, C. (Ed.). (2019). *Smart medical data sensing and IoT systems design in healthcare*. IGI Global.

Dissanayake, K., & Johar, M. G. M. (2021). Comparative study on heart disease prediction using feature selection techniques on classification algorithms. *Applied Computational Intelligence and Soft Computing, 2021*, 1–17. https://doi.org/10.1155/2021/9989265

Dewasinghe, C. L., Ranasinghe, A. D., Wijesinghe, G. L. S. W. W., & Wickramasinghe, S. (2025, March). A user experience engineering approach for transforming general healthcare management in Sri Lanka with patient-centred mobile application: "MyMediTrack". In *Proceedings of the International Conference on Advanced Materials for Clean Energy and Health*.

Tripoliti, E. E., Papadopoulos, T. G., Karanasiou, G. S., Naka, K. K., & Fotiadis, D. I. (2017). Heart failure: diagnosis, severity estimation and prediction of adverse events through machine learning techniques. *Computational and structural biotechnology journal*, *15*, 26-47.

Divya, N., Riyazuddin, M., Ahad, A., Vulapula, S. R., Manjula, A., & Sirajuddin, M. (2024). Predicting heart disease using machine learning and IoT techniques. *Onkologia i Radioterapia*, 18(9). https://www.oncologyradiotherapy.com/articles/predicting-heart-disease-using-machine-learning-and-iot-techniques-1100024.html

Marzoog, B. A., Chomakhidze, P., Gognieva, D., Gagarina, N. V., Silantyev, A., Suvorov, A., Fominykha, E., Mustafina, M., Natalya, E., Gadzhiakhmedova, A., & Kopylov, P. (2024). *Machine Learning Model Discriminate Ischemic Heart Disease Using Breathome Analysis. Biomedicines*, *12*(12), 2814. https://doi.org/10.3390/biomedicines12122814

Ministry of Health, Nutrition & Indigenous Medicine (Sri Lanka). (2021). Annual Health Bulletin 2021. Ministry of Health, Sri Lanka. Retrieved August 12, 2025, from https://www.health.gov.lk/annual-health-bulletin/

Wijemunige, N., Rannan-Eliya, R. P., Herath, H. M. M., & O'Donnell, O. (2024). The Prevalence and Epidemiological Features of Ischaemic Heart Disease in Sri Lanka. *Global Heart*, *19*(1), 49. https://doi.org/10.5334/gh.1330

Karthick, K., Aruna, S. K., Samikannu, R., Kuppusamy, R., Teekaraman, Y., & Thelkar, A. R. (2022). Implementation of a heart disease risk prediction model using machine learning. *Computational and Mathematical Methods in Medicine, 2022*, 6517716. https://doi.org/10.1155/2022/6517716

Krittanawong, C., Johnson, K. W., Rosenson, R. S., Wang, Z., Aydar, M., & Kitai, T. (2017). Deep learning for cardiovascular medicine: A practical primer. *European Heart Journal, 40*(25), 2058–2073. https://doi.org/10.1093/eurheartj/ehy411

Kwon, J. M., Kim, K. H., Medina-Inojosa, J. R., Yao, X., & Lopez-Jimenez, F. (2019). Artificial intelligence for early prediction of atrial fibrillation using electrocardiography. *Journal of the American Heart Association, 8*(17), e011955. https://doi.org/10.1161/JAHA.119.011955

Liu, T., et al. (2025). Machine learning-based prediction models for cardiovascular disease risk using EHR data. *European Heart Journal – Digital Health, 6*(1), 7–22. https://doi.org/10.1093/ehjdh/ztae080

Maini, E., Venkateswarlu, B., Maini, B., & Marwaha, D. (2021). Machine learning-based heart disease prediction system for Indian population: An exploratory study done in South India. *Medical Journal Armed Forces India, 77*(3), 302–311. https://doi.org/10.1016/j.mjafi.2021.03.013

Mansoor, H., et al. (2023). Machine-learning algorithms for ischemic heart disease prediction. *PMC*. https://pmc.ncbi.nlm.nih.gov/articles/PMC10201879/

Mettananda, C., Jayasinghe, M. W., Liyanage, I. K., Ranasinghe, M. S., Gunawardena, K., & Wijesundara, W. M. H. S. (2024). Development and validation of a cardiovascular risk prediction model for Sri Lankans using machine learning. *PLoS ONE, 19*(10), e0309843. https://doi.org/10.1371/journal.pone.0309843

Reddy, K. V. V., Elamvazuthi, I., Aziz, A. A., Paramasivam, S., Chua, H. N., & Pranavanand, S. (2021). Heart disease risk prediction using machine learning classifiers with attribute evaluators. *Applied Sciences, 11*(18), 8352. https://doi.org/10.3390/app11188352

ResearchGate.net. (2021). Cardiovascular disease prediction using recursive feature elimination and gradient boosting classification techniques. https://www.researchgate.net/publication/352479655

Restivo, L., et al. (2022). ML-driven precision medicine in IHD. *Journal of Clinical Medicine, 11*(15), 4567. https://doi.org/10.3390/jcm11154567

Shahid, N., Rappon, T., & Berta, W. (2020). Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PloS One, 15*(2), e0227811. https://doi.org/10.1371/journal.pone.0227811

Shishehbori, F., & Awan, Z. (2024). Enhancing cardiovascular disease risk prediction with machine learning models. *arXiv preprint* arXiv:2402.00753.

Wang, G., Zhang, Y., Li, S., Zhang, J., Jiang, D., Li, X., ... & Du, J. (2021). A machine learning-based prediction model for cardiovascular risk in women with preeclampsia. *Frontiers in Cardiovascular Medicine*, *8*, 736491.

Wanninayake, A., & Wickramasinghe, S. (2025, March). Predicting the risk of chronic kidney disease (CKD) in individuals in Medawachchiya and surrounding areas using machine learning: "AAROGYA". In Proceedings of the International Conference on Advanced Materials for Clean Energy and Health.

Wu, Y., Zhang, Z., & Xu, Y. (2021). Cardiovascular risk prediction using machine learning algorithms. *BMC Medical Informatics and Decision Making, 21*(1), 66. https://doi.org/10.1186/s12911-021-01421-1

Zhou, X., Li, Y., & Ma, Y. (2019). Heart disease prediction model based on decision tree and improved KNN. *Journal of Healthcare Engineering, 2019*, 1–10. https://doi.org/10.1155/2019/3831408

Silveri, G., Merlo, M., Restivo, L., De Paola, B., Miladinović, A., Ajčević, M., Sinagra, G., & Accardo, A. (2020). Identification of ischemic heart disease by using machine learning techniques based on parameters measuring heart rate variability. *arXiv*. Retrieved from https://arxiv.org/abs/2010.15893 researchgate.net+15arxiv.org+15pmc.ncbi.nlm.nih.gov+15

Theerthagiri, P., & Vidya, J. (2021). Cardiovascular disease prediction using recursive feature elimination and gradient boosting classification techniques. *arXiv*. Retrieved from https://arxiv.org/abs/2106.08889

Kwakye, K., & Dadzie, E. (2021). Machine learning-based classification algorithms for the prediction of coronary heart diseases. *arXiv*. Retrieved from https://arxiv.org/abs/2112.01503

Park, S., et al. (2023). A predictive model of ischemic heart disease in middle-aged and older women using data mining techniques. *Preventive Medicine Reports*. *PMC*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10143631/

Acknowledgement

The editorial team gratefully acknowledges Mr. Nisal Weerasinghe, Ms. Deshini Ranasinghe, Ms. Sihara Gallage, Mr. Ragulan Karunanithi, and Mr. Pulathisi Perera for their significant contributions to typesetting and for their diligent coordination with authors throughout the editorial process.

PUBLICATION PARTNERS

Scientific.Net
Publisher in Materials Science & Engineering

acm Association for Computing Machinery

UNIVERSIDADE LUSÓFONA

PLATINUM SPONSOR

NATIONAL SCIENCE FOUNDATION

GOLD SPONSORS

Curtin University

ICC INTERNATIONAL CONSTRUCTION CONSORTIUM

BRONZE SPONSORS

LTL HOLDINGS
Local Engineering, at its Global best

CSE
STRENGTH SERVICEABILITY SAFETY SUSTAINABILITY
CIVIL & STRUCTURAL ENGINEERING CONSULTANTS PRIVATE LIMITED

OTHER SPONSORS

PiLETesT

EnCare

ISSN 2961-5011
9 772961 501001

25 YEARS OF EXCELLENCE
SLIIT UNI
THE KNOWLEDGE UNIVERSITY

New Kandy Road,Malabe,10115, Sri Lanka

www.sicet.sliit.lk