Pertanika Journal of

SCIENCE & TECHNOLOGY

JST

VOL. 34 (5) OCT. 2026

PERTANIKA JOURNALS

A scientific journal published by Universiti Putra Malaysia Press

Pertanika Journal of Science & Technology
Pertanika Journal of Science & Technology

PERTANIKA JOURNAL OF SCIENCE & TECHNOLOGY

About the Journal

Overview

Pertanika Journal of Science & Technology is an official journal of Universiti Putra Malaysia. It is an open-access online scientific journal. It publishes original scientific outputs. It neither accepts nor commissions third party content.

Recognised internationally as the leading peer-reviewed interdisciplinary journal devoted to the publication of original papers, it serves as a forum for practical approaches in improve quality on issues pertaining to science and engineering and its related fields.

Pertanika Journal of Science & Technology currently publishes 6 issues a year (*February, April*, *June, August, October, and December*). It is considered for publication of original articles as per its scope. The journal publishes in **English** and it is open for submission by authors from all over the world.

The journal is available world-wide.

Aims and Scope

Pertanika Journal of Science & Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: bioinformatics, bioscience, biotechnology and bio-molecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management, and related fields of study.

History

Pertanika Journal of Science & Technology was founded in 1993 and focusses on research in science and engineering and its related fields.

Vision

To publish a journal of international repute.

Mission

Our goal is to bring the highest quality research to the widest possible audience.

Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions can expect to receive a decision within 90 days. The elapsed time from submission to publication for the articles averages 180 days. We are working towards decreasing the processing time with the help of our editors and the reviewers.

Abstracting and Indexing of Pertanika

Pertanika Journal of Science & Technology is now over 33 years old; this accumulated knowledge and experience has resulted the journal being abstracted and indexed in SCOPUS (Elsevier), Journal Citation Reports (JCR-Clarivate), EBSCO, ASEAN CITATION INDEX, Microsoft Academic, Google Scholar, and MyCite.

Citing Journal Articles

The abbreviation for Pertanika Journal of Science & Technology is *Pertanika J. Sci. & Technol.*

Publication Policy

Pertanika policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. It prohibits as well publication of any manuscript that has already been published either in whole or substantial part elsewhere. It also does not permit publication of manuscript that has been published in full in proceedings.

Code of Ethics

The *Pertanika* journals and Universiti Putra Malaysia take seriously the responsibility of all of its journal publications to reflect the highest in publication ethics. Thus, all journals and journal editors are expected to abide by the journal's codes of ethics. Refer to *Pertanika*'s **Code of Ethics** for full details, or visit the journal's web link at http://www.pertanika.upm.edu.my/code_of_ethics.php

Originality

The author must ensure that when a manuscript is submitted to *Pertanika*, the manuscript must be an original work. The author should check the manuscript for any possible plagiarism using any programme such as Turn-It-In or any other software before submitting the manuscripts to the Editorial Office, *Pertanika* Journal Publication Section.

All submitted manuscripts must be in the journal's acceptable similarity index range:

≤ 20% – PASS; > 20% – REJECT.

International Standard Serial Number (ISSN)

An ISSN is an 8-digit code used to identify periodicals such as journals of all kinds and on all media–print and electronic.

Pertanika Journal of Science & Technology: e-ISSN 2231-8526 (Online).

Lag Time

A decision on acceptance or rejection of a manuscript is reached in 90 days (average). The elapsed time from submission to publication for the articles averages 180 days.

Authorship

Authors are not permitted to add or remove any names from the authorship provided at the time of initial submission without the consent of the journal's Chief Executive Editor.

Manuscript Preparation

For manuscript preparation, authors may refer to Pertanika's **INSTRUCTION TO AUTHORS**, available on the official website of *Pertanika*.

Editorial Process

Authors who complete any submission are notified with an acknowledgement containing a manuscript ID on receipt of a manuscript, and upon the editorial decision regarding publication.

Pertanika follows a **double-blind peer-review** process. Manuscripts deemed suitable for publication are sent to reviewers. Authors are encouraged to suggest names of at least 3 potential reviewers at the time of submission of their manuscripts to *Pertanika*, but the editors will make the final selection and are not, however, bound by these suggestions.

Notification of the editorial decision is usually provided within 90 days from the receipt of manuscript. Publication of solicited manuscripts is not guaranteed. In most cases, manuscripts are accepted conditionally, pending an author's revision of the material.

The Journal's Peer Review

In the peer-review process, 2 to 3 referees independently evaluate the scientific quality of the submitted manuscripts. At least 2 referee reports are required to help make a decision.

Peer reviewers are experts chosen by journal editors to provide written assessment of the **strengths** and **weaknesses** of the written research, with the aim of improving the reporting of research and identifying the most appropriate and highest quality material for the journal.

Operating and Review Process

What happens to a manuscript once it is submitted to *Pertanika*? Typically, there are 7 steps to the editorial review process:

1. The journal's Chief Executive Editor and Editor-in-Chief examine the paper to determine whether it is relevance to the journal's needs in terms of novelty, impact, design, procedure, language as well as presentation and allow it to proceed to the reviewing process. If it is not appropriate, the manuscript is rejected outright and the author is informed.

2. The Chief Executive Editor sends the article-identifying information having been removed, to 2 to 3 reviewers. They are specialists in the subject matter of the article. The Chief Executive Editor requests that they complete the review within 3 weeks.

   Comments to authors are about the appropriateness and adequacy of the theoretical or conceptual framework, literature review, method, results and discussion, and conclusions. Reviewers often include suggestions to strengthening of the manuscript. Comments to the editor are in the nature of the significance of the work and its potential contribution to the research field.

Pertanika Journal of Science & Technology
Pertanika Journal of Science & Technology

Pertanika Journal of Science & Technology
Pertanika Journal of Science & Technology

3. The Editor-in-Chief examines the review reports and decides whether to accept or reject the manuscript, invite the authors to revise and resubmit the manuscript, or seek additional review reports. In rare instances, the manuscript is accepted with almost no revision. Almost without exception, reviewers' comments (to the authors) are forwarded to the authors. If a revision is indicated, the editor provides guidelines to attend to the reviewers' suggestions and perhaps additional advice on how to revise the manuscript.

4. The authors decide whether and how to address the reviewers' comments and criticisms and the editor's concerns. The authors return a revised version of the paper to the Chief Executive Editor along with specific information describing how they have addressed' the concerns of the reviewers and editor, usually in a tabular form. The authors may also submit a rebuttal if there is a need especially when the authors disagree with certain comments provided by the reviewers.

5. The Chief Executive Editor sends the revised manuscript out for re-review. Typically, at least 1 of the original reviewers will be asked to examine the article.

6. When the reviewers have completed their work, the Editor-in-Chief examines their comments and decides whether the manuscript is ready to be published, needs another round of revisions, or should be rejected. If the decision is to accept, the Chief Executive Editor is notified.

7. The Chief Executive Editor reserves the final right to accept or reject any material for publication, if the processing of a particular manuscript is deemed not to be in compliance with the S.O.P. of *Pertanika*. An acceptance letter is sent to all the authors.

   The editorial office ensures that the manuscript adheres to the correct style (in-text citations, the reference list, and tables are typical areas of concern, clarity, and grammar). The authors are asked to respond to any minor queries by the editorial office. Following these corrections, page proofs are mailed to the corresponding authors for their final approval. At this point, **only essential changes are accepted**. Finally, the manuscript appears in the pages of the journal and is posted online.

Pertanika Journal of

SCIENCE & TECHNOLOGY

Vol. 34 (5) Oct. 2026

A scientific journal published by Universiti Putra Malaysia Press

PJST
Pertanika Journal of Science & Technology
AN INTERNATIONAL PEER-REVIEWED JOURNAL

UNIVERSITY PUBLICATIONS COMMITTEE

CHAIRMAN
Zamberi Sekawi

EDITOR-IN-CHIEF
Mohamed Thariq Hameed Sultan
SEARCA Chair of Agriculture Technology and Innovation

JOURNAL SECTION
Head of Section:
Navaneetha Krishna Chandran

Journal Officers:
Ain Haziqah Ab Razak
Nur Izni Shafiqah Mohd Hamidi
Saundarya Saran

Editorial Assistants:
Zulinaardawati Kamarudin

Pre-press Officers:
Aisyah Umairah Jasmin
Engku Erina Zahirah Engku Mohd Suhaimi
Ku Ida Mastura Ku Baharom

IT Officer:
Kiran Raj Kaneswaran

Journal Section Office
Putra Science Park
1st Floor, IDEA Tower II
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM Serdang
Selangor, Malaysia

General Enquiry
Tel. No: +603 9769 1622
E-mail:
pertanika.proceeding@upm.edu.my
URL:
http://www.pertanika.upm.edu.my/

PUBLISHER
UPM Press Centre
Universiti Putra Malaysia
43400 UPM Serdang
Selangor, Malaysia
Tel: +603 9769 3606
E-mail: dir.penerbit@upm.edu.my
URL: http://penerbit.upm.edu.my

ASSOCIATE EDITOR
2024-2026

Miss Laiha Mat Kiah
Security Services Sn: Digital Forensic, Steganography, Network Security, Information Security, Communication Protocols, Security Protocols
Universiti Malaya, Malaysia

Saidur Rahman
Renewable Energy, Nanofluids, Energy Efficiency, Heat Transfer, Energy Policy
Sunway University, Malaysia

EDITORIAL BOARD
2024-2026

Abdul Latif Ahmad
Chemical Engineering
Universiti Sains Malaysia, Malaysia

Hsiu-Po Kuo
Chemical Engineering
National Taiwan University, Taiwan

Mohd. Ali Hassan
Bioprocess Engineering, Environmental Biotehnology
Universiti Putra Malaysia, Malaysia

Ahmad Zaharin Aris
Hydrochemistry, Environmental Chemistry, Environmental Forensics, Heavy Metals
Universiti Putra Malaysia, Malaysia

Ivan D. Rukhlenko
Nonliner Optics, Silicon Photonics, Plasmonics and Nanotechnology
The University of Sydney, Australia

Nor Azah Yusof
Biosensors, Chemical Sensor, Functional Material
Universiti Putra Malaysia, Malaysia

Azlina Harun@ Kamaruddin
Enzyme Technology, Fermentation Technology
Universiti Sains Malaysia, Malaysia

Lee Keat Teong
Energy Environment, Reaction Engineering, Waste Utilisation, Renewable Energy
Universiti Sains Malaysia, Malaysia

Norbahiah Misran
Communication Engineering
Universiti Kebangsaan Malaysia, Malaysia

Bassim H. Hameed
Chemical Engineering: Reaction Engineering, Environmental Catalysis & Adsorption
Qatar University, Qatar

Mohamed Othman
Communication Technology and Network, Scientific Computing
Universiti Putra Malaysia, Malaysia

Roslan Abd-Shukor
Physics & Materials Physics, Superconducting Materials
Universiti Kebangsaan Malaysia, Malaysia

Biswajeet Pradhan
Digital image processing, Geographical Information System (GIS), Remote Sensing
University of Technology Sydney, Australia

Mohd Shukry Abdul Majid
Polymer Composites, Composite Pipes, Natural Fibre Composites, Biodegradable Composites, Bio-Composites
Universiti Malaysia Perlis, Malaysia

Sodeifian Gholamhossein
Supercritical technology, Optimisation, nanoparticles, Polymer nanocomposites
University of Kashan, Iran

Ho Yuh-Shan
Water research, Chemical Engineering and Environmental Studies
Asia University, Taiwan

Mohd Zulkifly Abdullah
Fluid Mechanics, Heat Transfer, Computational Fluid Dynamics (CFD)
Universiti Sains Malaysia, Malaysia

Wing Keong Ng
Aquaculture, Aquatic Animal Nutrition, Aqua Feed Technology
Universiti Sains Malaysia, Malaysia

INTERNATIONAL ADVISORY BOARD
2024-2027

Hiroshi Uyama
Polymer Chemistry, Organic Compounds, Coating, Chemical Engineering
Osaka University, Japan

Mohini Sain
Material Science, Biocomposites,Biomaterials
University of Toronto, Canada

Mohamed Pourkashanian
Mechanical Engineering, Energy, CFD and Combustion Processes
Sheffield University, United Kingdom

Yulong Ding
Particle Science & Thermal Engineering
University of Birmingham, United Kingdom

ABSTRACTING AND INDEXING OF PERTANIKA JOURNALS

Pertanika Journal of Science & Technology is indexed in Journal Citation Reports (JCR-Clarivate), SCOPUS (Elsevier), EBSCO, ASEAN Citation Index, Microsoft Academic, Google Scholar and MyCite.

The publisher of Pertanika will not be responsible for the statements made by the authors in any articles published in the journal. Under no circumstances will the publisher of this publication be liable for any loss or damage caused by your reliance on the advice, opinion or information obtained either explicitly or implied through the contents of this publication.
All rights of reproduction are reserved in respect of all papers, articles, illustrations, etc., published in Pertanika. Pertanika provides free access to the full text of research articles for anyone, web-wide. It does not charge either its authors or author-institution for refereeing/publishing outgoing articles or user-institution for accessing incoming articles.
No material published in Pertanika may be reproduced or stored on microfilm or in electronic, optical or magnetic form without the written authorisation of the Publisher.
Copyright ©2026 Universiti Putra Malaysia Press. All Rights Reserved.

Pertanika Journal of Science & Technology
Vol. 34 (5) Oct. 2026

Contents

Foreword xiii
Mohamed Thariq Hameed Sultan

Data-Driven System Identification Modelling for the Reactor TRIGA PUSPATI Using Real Operational Data 2545
Nor Arymaswati Abdullah, Azura Che Soh, Siti Marhainis Othman, Ribhan Zafira Abdul Rahman, Samsul Bahari Mohd Noor, and Julia Abdul Karim

Pneumonia Screening and Prediction on Images using a Convolutional Neural Network Deep Learning Model 2569
Shuria Saaidin, Murizah Kassim, I Gede Pasek Suta Wijaya, and Nur Amirah Kamaluddin

Artificial Neural Network-Based UHF Partial Discharge Localisation in Mineral Oil with Datasets based on TDOA 2597
Zhi Yang Lim, Norhafiz Azis, Hafiz Rashidi Ramli, Mohd Amran Mohd Radzi, Nor Mohd Haziq Norsahperi, and Azrul Mohd Ariffin

Investigation of Assistive Grouser Traction Force in Toyoura Sand: Effects of Grouser Geometry and Contact Angle 2619
Mochamad Edoward Ramadhan, Ahmad Najmuddin Ibrahim, and Yasuhiro Fukuoka

Formulation of Regulatory Sandbox for Nanotechnology Innovation 2649
Nor Akhmal Hasmin, Juan Matmin, and Ainul Hafiza Zainudin

The Use of FSR-based Plantar Pressure Patterns for Deception Detection 2675
Gerelbat Batgerel, and Chunki Kwon

Urban Flooding as a Hydrological Connectivity Problem: An Entropy-Based GIS Assessment in a Low-Relief Tropical City 2691
Fachruddin Abdul Chaidir, Masha Nur Salsabiela Menhat, Mohd Azhafiz Abdullah, Muhammad Azraie Abdul Kadir, Effi Helmy Ariffin, Khairul Nizam Abdul Maulud, Ahmad Shakir Mohd Saudi, and Hasdinar Umar

Capacity-Limited Cooling Explains Simultaneous Energy Waste and Thermal Discomfort in Two University Buildings in the Southern Philippines 2715
Fejaycris O. Pillodar, Dioniela Mae C. Ellos, Jiza Gay P. Jarabe, John Russel A. Dalogdog, Monaliza Joy Z. Magsayo, Charlene Mae Z. Gaab, Manuel Anthony P. Jondonero, Melgie A. Alas, Rovick Pelos P. Tarife, Noel R. Estoperez, Mark June L. Aporador, Napoleon A. Enteria, Kim Emissary C. Magarin, Ryan Lloyd R. Carpentero, Hernando P. Bacosa, and Peter D. Suson

Regenerative Braking: A Range Extender for Electric Motorcycles in Hilly Terrain 2739
Mawada Ahmed, Siti Fauziah Toha, Khairul Affendy, Muhammad Abdullah, Salmiah Ahmad, Mohammad Osman Tokhi, and Ahmad Syahrin Idris

An Integrated Multi-Source Framework for Explainable Android Malware Detection with SOC Integration 2765
Pallavi M. Shimpi and Nitin Namdeo Pise

Time-dependent Polyethylene Terephthalate (PET) Biodegradation by *Aspergillus niger* and *Aspergillus oryzae* under Solid State Fermentation 2791
Latha Shankar, Norazlina Idris, Nurhafizah Ibrahim, Norakma, M. N, Rodiah, M. H, Wan Zarina, W. M, Noor Fazreen, D, and Nur Ain Izzati Mohd Zainudin

Fractional SIRM Epidemic Model with Microplastic-Mediated Environmental Transmission 2811
R. Sengothai, P. Balaganesan, D. Adhil Basha, and J. Renuka

Review Article 2835
Wireless Sensor Network and Leveraging IoT on Agro Farming: A Review
Muhammad Azwan Nasirudin, Mohammad Nishat Akhtar, and Farzad Ismail

Grey-Fuzzy Based Optimization of Machining Parameters in Turning of AA 7068/TiC MMCs 2887
Syed Altaf Hussain, Sudhamadhuri K, Upendra Rajak, and Abhishek Dasore

Review Article 2907
A Systematic Review of Hybrid Deep Learning and Optimization Techniques to Address Class Imbalance and Malware Detection in IoT Network
Ernest Ezema, and Julia Juremi

Context-Aware RFID-IoT Counterfeit Prevention: A Simulation-Based Proof of Concept Using Layered RBAC, MAC, ABAC, and ECA 2929
Ravichandran and Irene Anastasia Patric

Privacy-Preserving Federated Learning with Explainable AI for Multi-Site Autism Spectrum Disorder Biomarker Discovery 2947
Anusha Ravankol, and Nandula Anuradha

Diseases of Vegetable Plant Recognition Based on Enhanced SqueezeNet 2981
Kifaa Hadi Thanoon, Akram A. Abdulrahman, Ali Ibrahim Ahmed, Omar Ibrahim Alsaif, and Ibrahim Ahmed Saleh

A Novel Fusion of Computer Vision Capabilities and Deep Learning Techniques for OCT Retinal Disease Diagnosis 3003
Nagham Salim Mohammed, Omar S. Almolaa, and Omar Ibrahim Alsaif

Effects of Elevated CO_2-Grown Napier Grass and Cassava Starch Binder Ratios on Biomass Briquette Properties 3025
Muhammad Zulhilmi Mohd Nasirudin and Siti Zaharah Sakimin

Automated News Categorisation Using Transformer-Based Deep Learning Models: A Step toward Intelligent Media Classification 3047
B. Rupa Devi, Malyala Gayatri, Rajesh Lingam, Penubaka Balaji, S. Beatrice, and S. Athinarayanan

Review Article 3075
Nitrate-reducing Sulphide-Oxidising Bacteria (NR-SOB) in Corrosion Control: A Review
Aishah Zafre, Norinsan Kamil Othman, Najmiddin Yaakob, Azman Jalar, Maria Abu Bakar, and Rabiahtul Zulkafli

Mobile-Based Inventory Counting Using 2D Object Detection and Layer-Based Inference 3109
Gan Heng Lai, Goh Eg Su, and Johanna Ahmad

Bridging the Ocean Literacy Gap: Teacher Needs and Design Requirements for Virtual Reality-Based Sustainable Maritime Education for Coastal High Schools in Tanjungpinang, Indonesia 3131
Nurul Asikin, Arinda Eka Lidiastuti, Adam Fernando, and Rahmania Pamungkas

Antigenicity of Hydrolysed Casein in Goat's Milk Using Bromelain 3163
Nurul Izza Nabilah Mohd Samsudin, Chai Jie Ying, Rashidah Sukor, Aliah Zannierah Mohsin, and Siva Raseetha

Foreword

Welcome to the fifth issue of 2026 of the Pertanika Journal of Science and Technology (PJST)!

PJST is an open-access journal for studies in Science and Technology published by Universiti Putra Malaysia Press. It is independently owned and managed by the university for the benefit of the world-wide science community.

This issue contains 25 articles: three review articles; and the rest are regular articles. The authors of these articles come from different countries namely India, Indonesia, Iraq, Japan, Malaysia, Philippines, Qatar, South Korea, United Kingdom, and the United States of America.

The article titled "Pneumonia Screening and Prediction on Images using a Convolutional Neural Network Deep Learning Model" investigates the application of deep learning to support pneumonia screening by automating the classification of chest radiographs. The study compares two pretrained convolutional neural network architectures, VGG16 and Inception-v3, using a publicly available dataset comprising 5,216 training images and 624 independent test images classified as normal or pneumonia. Both extraction and fine-tuning approaches were evaluated under consistent testing conditions to assess their performance. The findings demonstrate that the fine-tuned Inception-v3 model achieved the strongest overall performance, with an accuracy of 93.6%, sensitivity of 95.1%, specificity of 91.0%, and an F1-score of 94.9%, while requiring approximately 60% of the training time per epoch of VGG16. The study also emphasises the importance of reporting performance on independent test data rather than relying on training-phase results, thereby encouraging more reliable assessment of medical image classification systems. By demonstrating the potential of transfer learning for chest radiograph analysis, this research contributes to the development of computer-assisted screening tools, particularly in healthcare settings where access to specialist radiological expertise may be limited. Full details of this study are available on page 2569.

Another noteworthy contribution is the article entitled "Time-dependent Polyethene Terephthalate (PET) Biodegradation by *Aspergillus niger* and *Aspergillus oryzae* under Solid State Fermentation", which addresses the environmental challenge of persistent plastic waste through fungal-mediated biodegradation. The study examines the ability of *Aspergillus niger* and *Aspergillus oryzae* to modify and reduce polyethene terephthalate (PET) under solid-state fermentation conditions over a ten-week period. PET samples

obtained from mineral water bottles were assessed through weight-loss measurements, xylanase activity, Fourier-transform infrared spectroscopy, and scanning electron microscopy to evaluate enzymatic activity and changes in polymer structure and surface morphology. The results show that *A. oryzae* produced the greatest PET weight reduction of 1.91% at week four, accompanied by xylanase activity of 467.8 U/mL, whereas *A. niger* exhibited a higher peak xylanase activity of 641.9 U/mL at week two but a lower maximum PET weight reduction of 0.76%. The analyses also revealed chemical and surface changes following fungal treatment, indicating the potential for biological modification of PET. Although the observed mass reductions were modest, the findings provide useful insights into the time-dependent behaviour of the two fungal species and support further investigation of low-cost, biologically based strategies for plastic-waste management. Further details of this study can be found on page 2791.

The issue also features the article titled "Regenerative Braking-Based Energy Recovery for Electric Motorcycles in Hilly Terrain", which presents an energy-management approach to improving the efficiency of electric motorcycles operating on roads with varying gradients. The research develops a MATLAB/Simulink model that incorporates vehicle dynamics, road-gradient effects, an electric drive system, a lithium-ion battery, and feedforward and proportional–integral–derivative (PID) control strategies. Using the Worldwide Motorcycle Test Cycle as a reference speed profile, the study evaluates speed tracking, regenerative braking behaviour, energy recovery, and changes in battery state of charge. The findings demonstrate that the PID-controlled system improved the simulated braking response compared with feedforward control, reducing the root-mean-square error from 14.10 m/s to 11.546 m/s. Recovered energy increased from 1.00×10^5 J to 2.20×10^5 J, while the battery state-of-charge variation improved from −13.0% to −10.2%. The simulation further confirmed that downhill braking enabled the electric machine to operate as a generator, returning part of the vehicle's mechanical energy to the battery. By highlighting the effectiveness of regenerative braking and appropriate control strategies in hilly driving conditions, this work advances energy-efficient electric mobility and more sustainable transportation technologies. The study's full contents are available on page 2739.

We anticipate that you will find the evidence presented in this issue to be intriguing, thought-provoking and useful in reaching new milestones in your own research. Please recommend the journal to your colleagues and students to make this endeavour meaningful.

All the papers published in this edition underwent Pertanika's stringent peer-review process involving a minimum of two reviewers comprising internal as well as external

referees. This was to ensure that the quality of the papers justified the high ranking of the journal, which is renowned as a heavily-cited journal not only by authors and researchers in Malaysia but by those in other countries around the world as well.

We would also like to express our gratitude to all the contributors, namely the authors, reviewers and Editorial Board Members of PJST, who have made this issue possible.

PJST is currently accepting manuscripts for upcoming issues based on original qualitative or quantitative research that opens new areas of inquiry and investigation.

Editor-in-Chief
Mohamed Thariq Hameed Sultan

Pertanika J. Sci. & Technol. 34 (5): 2545 - 2567 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Data-Driven System Identification Modelling for the Reactor TRIGA PUSPATI Using Real Operational Data

Nor Arymaswati Abdullah[1,2], Azura Che Soh[1]*, Siti Marhainis Othman[3], Ribhan Zafira Abdul Rahman[1], Samsul Bahari Mohd Noor[1], and Julia Abdul Karim[2]

[1]Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

[2]Reactor Instrumentation and Control Engineering, Reactor Technology Centre, Malaysian Nuclear Agency, Bangi, 43000 Kajang, Selangor, Malaysia

[3]Faculty of Electrical Engineering & Technology, Universiti Malaysia Perlis (UniMAP), Pauh Putra Campus, 02600 Arau, Perlis, Malaysia

ABSTRACT

Accurate reactor modelling is essential to ensure reliable system performance, particularly when equipment degradation, disturbances, and parameter variations reduce model accuracy over time. The existing model of the Reactor TRIGA PUSPATI (RTP), which is based on the First Principle (FP) approach, has experienced reduced accuracy due to ageing effects, maintenance activities, and system modifications. Therefore, this study proposes a data-driven System Identification (SI) approach using actual reactor operational input and output data to develop a more accurate reactor core subsystem model. The RTP system was divided into three subsystems, namely thermal-hydraulics, reactivity, and neutronics. SI-based models were developed and validated at three reactor power levels: 500 kW, 750 kW, and 1000 kW. The results demonstrated that the developed SI models achieved acceptable performance and accurately tracked the real reactor plant behaviour. The reactivity subsystem produced the closest representation of the actual plant, with Residual Mean Percentage (RMP) values ranging from 1.09% to 1.84%, while the thermal-hydraulics and neutronics subsystems recorded RMP values ranging from 0.69% to 27.45% and 1.28% to 31.54%, respectively. Comparative evaluation between the SI and FP models showed that the SI approach consistently produced lower residual errors and better dynamic representation of the RTP system. These findings demonstrate

ARTICLE INFO

Article history:
Received: 28 June 2026
Accepted: 05 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.01

E-mail addresses:
n.arymas@gmail.com (Nor Arymaswati Abdullah)
azuracs@upm.edu.my (Azura Che Soh)
marhainis@unimap.edu.my (Siti Marhainis Othman)
ribhan@upm.edu.my (Ribhan Zafira Abdul Rahman)
samsul@upm.edu.my (Samsul Bahari Mohd Noor)
julia@nm.gov.my (Julia Abdul Karim)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

the effectiveness of the SI approach in improving the accuracy of reactor subsystem modelling using real operational data.

Keywords: Data-driven, modelling, reactor TRIGA PUSPATI (RTP), real operational data, system identification

INTRODUCTION

The ageing of research reactors has become a significant issue in many countries due to both physical and non-physical degradation, which may reduce reactor performance, efficiency, and operational reliability (Leong et al., 2015; Martón et al., 2023; Nagai et al., 2011; Ratiko et al., 2020). Ageing effects generally occur gradually over time and may lead to performance deterioration, while radiation effects can introduce rapid and unpredictable changes caused by environmental conditions (Leong et al., 2015). In research reactors, these issues may also affect operational capability and radioisotope production (Nagai et al., 2011). Consequently, accurate modelling is essential to ensure reliable system operation and precise dynamic representation of the reactor behaviour.

In industrial and reactor systems, equipment degradation, disturbances, and parameter variations can increase the mismatch between the actual plant and the developed mathematical model. Conventional FP models are commonly developed based on fundamental physical laws to represent system dynamics. Although FP models are widely used and can provide accurate results, they are often computationally complex, require extensive system knowledge, and involve many parameters. In nuclear reactor applications, FP and physics-based models such as MCNP6, REBUS MCNP6, SRAC MCNP, and RELAP have been widely used for reactor analysis and simulation (Dim & Aghara, 2019; Nguyen et al., 2020; Safaei Arshi et al., 2021). However, these models require a detailed understanding of reactor physics and system behaviour, making the modelling process difficult and time-consuming. In many cases, the mathematical model is developed during the initial reactor installation and is not updated after equipment replacement, system modifications, or ageing effects. Consequently, the existing model may no longer accurately represent the plant's actual behaviour, leading to reduced modelling accuracy and system performance.

Artificial Intelligence (AI) approaches, especially machine learning techniques such as Artificial Neural Networks (ANNs), have become increasingly popular due to their ability to learn system behaviour and model nonlinear systems (Mao et al., 2022; Melo et al., 2024; Wu et al., 2025). ANN methods have been widely used in system modelling to improve prediction and control performance. Previous studies have applied ANN to heat exchangers and closed-loop systems to improve modelling accuracy and reduce model mismatch (Aguel et al., 2019; Son et al., 2022). However, ANN models usually require large amounts of data and high computational effort. Their performance may also be affected by random

parameter initialisation, which can reduce estimation accuracy and cause convergence to local optima (Mao et al., 2022). In addition, machine learning models may have difficulty adapting to changes in system behaviour after training (Kim & Kim, 2020).

Another widely used artificial intelligence approach is fuzzy logic, which is effective for modelling systems with uncertainty, nonlinearity, and incomplete information (Tang & Ahmad, 2024). Fuzzy logic has been widely applied in process modelling and control, including chemical processes, energy systems, robotics, and industrial fault diagnosis (Mohindru, 2024). In addition, the integration of fuzzy logic with optimisation and machine learning techniques has demonstrated promising performance in soft sensing, predictive maintenance, and real-time fault detection applications, particularly in industrial environments subject to sensor noise and operational disturbances (Khalid & Khan, 2023). However, fuzzy logic models may suffer from limitations such as sensitivity to membership function design, increased computational complexity in high-dimensional systems, and reduced interpretability when combined with advanced learning algorithms. Therefore, although fuzzy logic offers strong capabilities for handling uncertain and nonlinear systems, challenges related to scalability, robustness, and computational efficiency remain.

To overcome these limitations, data-driven modelling approaches such as SI have gained considerable attention in industrial applications. SI modelling develops mathematical representations directly from measured input and output data without requiring extensive prior knowledge of the system. This approach reduces modelling complexity while maintaining satisfactory accuracy in representing dynamic system behaviour. Recent studies have shown that SI-based modelling is highly effective for complex nonlinear industrial systems due to its strong adaptability, robustness, and ability to represent dynamic processes with high accuracy (Liu et al., 2024; Wang et al., 2024; Zhang et al., 2025). Compared with AI approaches such as ANNs and fuzzy logic, SI modelling generally requires fewer parameters and does not rely heavily on extensive training or self-learning. In addition, AI-based methods may produce inconsistent or uncertain decision-making results when system operating conditions change significantly. Therefore, SI modelling remains a practical, reliable, and computationally efficient approach for dynamic system modelling in complex industrial applications.

Currently, the RTP model is developed using the FP modelling approach. Throughout its operational lifetime, the reactor has undergone various modifications, maintenance activities, and equipment replacements, which may contribute to parameter variations and reduce the accuracy of the existing reactor model. In addition to ageing effects, preventive and corrective maintenance, as well as system modifications, have been implemented. However, model accuracy has not shown significant improvement, with the reported steady-state error remaining between 0.235 and 0.7 (Ghazali et al., 2017; Minhat et al., 2018). Therefore, this study proposes developing a new reactor core subsystem model using a

data-driven SI approach based on actual operational input and output data. The proposed approach is expected to account for ageing effects and time-varying characteristics while reducing modelling complexity and improving the accuracy and dynamic representation of the RTP system.

REACTOR CORE SYSTEM

The block diagram of the reactor core system is illustrated in Figure 1. The reactor core system consists of three integrated subsystems, namely the thermal-hydraulic model, reactivity model, and neutronics model. Each subsystem generates its own corresponding model, which collectively represents the overall RTP reactor system. The inputs and outputs of these models are interdependent, except for the insertion reactivity and inlet coolant temperature. The inlet coolant temperature is assumed constant because real plant data for this parameter are unavailable.

Thermal-hydraulics Model

The thermal-hydraulics modelling for the RTP system comprises a few parameters, as shown in Figure 2. Inputs and outputs are neutron power and inlet coolant temperature, and coolant temperature and fuel temperature, respectively. The real output data from the thermal-hydraulics model comes from sensors in the RTP system. The function of the inlet

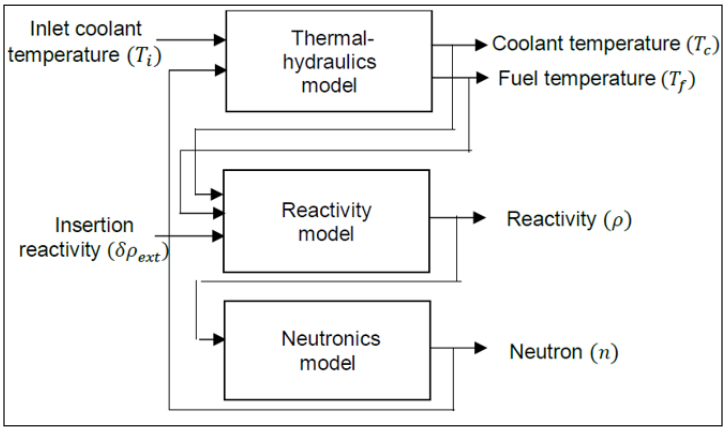


Figure 1. Block diagram of the reactor core system

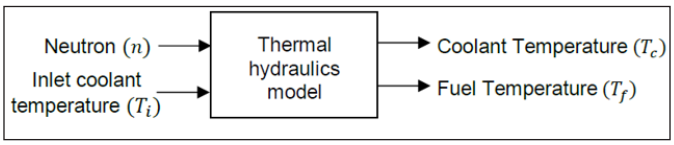


Figure 2. Thermal-hydraulics subsystem model for RTP

coolant temperature, Ti, is to dissipate heat in the reactor core, especially from the clad surface. This parameter is defined as a constant value at 32°C since there is no sensor for this variable. The neutron input is a nominal power from the neutronics model.

The temperature drop between the fuel and the coolant is compared with the temperature rise of the coolant, as expressed in Equation 1:

$$\frac{\bar{T}_f - \bar{T}_c}{\bar{T}_o - \bar{T}_i} = R_f w c_p \quad [1]$$

where c_p is the specific heat, $\bar{T}_o$ is the outlet coolant temperature, $\bar{T}_i$ is the inlet coolant temperature, R_f is the reactor core thermal resistance, and w is the coolant mass flow rate.

The dynamic variation of the coolant temperature is described by Equation 2 (Rafiei et al., 2019).

$$\frac{d}{dt}\bar{T}_c(t) = \frac{1}{\tau'}\left[\bar{T}_f(t) - \bar{T}_c(t)\right] - \frac{1}{\tau''}\left[\bar{T}_c(t) - \bar{T}_i\right]$$
$$\tau' = \frac{M_c c_p}{M_f c_f} \quad \text{and} \quad \tau'' = \frac{M_c c_p}{2\, w c_f} \quad [2]$$

where M_c is the coolant mass within the core. Nuclear fission produces heat that needs to be transferred. The temperature drop in the power reactor from fuel to coolant, which is proportional to the linear heat rate, can be described by Equation 3 (Arda & Holbert, 2016).

$$T_f(r,z) - T_c(r,z) = R_f P f_r(r) f_z(z), \quad [3]$$

where,

$$P = \frac{\bar{T}_f - \bar{T}_c}{R_f}$$

With consideration of cooling being cut off entirely, all the heat produced would result in the adiabatic heating of the fuel, as expressed in Equation 4:

$$M_f c_f \frac{d}{dt}\bar{T}_f(t) = P(t) \quad [4]$$

Where M_f and c_f are the total fuel mass and the specific heat. The approximate lumped-parameter model for thermal transients can be described by Equation 5 (Minhat et al., 2020).

$$M_f c_f \frac{d}{dt}\bar{T}_f(t) = P(t) - \frac{1}{R_f}\left[\bar{T}_f(t) - \bar{T}_c(t)\right]$$
$$\frac{d}{dt}\bar{T}_f(t) = \frac{1}{M_f c_f}P(t) - \frac{1}{\tau}\left[\bar{T}_f(t) - \bar{T}_c(t)\right] \quad [5]$$

The summarised thermal-hydraulic modelling is given by Equations 6 and 7 for the coolant and fuel temperatures, respectively (Cammi et al., 2013; Kazeminejad, 2012).

$$\frac{dT_C}{dt} = \frac{P(1-f)}{\tau_c K} + \frac{T_f - T_C}{\tau_c} - \frac{c_c}{\tau_c K w}\Gamma(T_C - T_i) \quad [6]$$

$$\frac{dT_f}{dt} = \frac{Pf}{\tau_f K} + \frac{T_f - T_C}{\tau_f} \quad [7]$$

where,

$T_C = wT_o + (1-w)T_i$ and

$$\Gamma = \sqrt{\frac{\delta_{in} g L v}{\alpha_2 w}}(T_C - T_i)$$

Where L is core height, α_2 is a factor for friction along the core channel, δ_{in} is the stationary density of the inlet water, K is the overall heat transfer coefficient between fuel and coolant, P is thermal power generated (also called actual core power) within the core volume by fission, τ_c is the time constant associated with the moderator and τ_f. The time constant is associated with the fuel.

Reactivity Model

The reactivity subsystem of the RTP consists of fuel temperature, coolant temperature, and insertion reactivity as input variables, with total reactivity as the output variable. The structure of the reactivity subsystem model is illustrated in Figure 3.

The fission chain reaction is essential for controlling the neutron population in the reactor. This behaviour is represented by the effective neutron multiplication factor, k_{eff}, which is defined in Equation 8 (Oka & Suzuki, 2013).

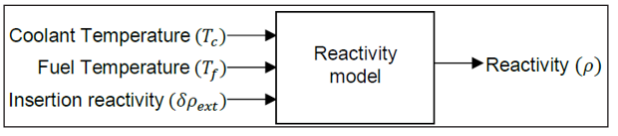


Figure 3. Reactivity subsystem model for RTP

$$k_{eff} = \frac{\text{Neutron production}}{\text{Neutron loss rate}} \quad [8]$$

where the neutron loss rate is the sum of neutron absorption and leakage rates. When $k_{eff} > 1$. The reactor is in a supercritical state, and the chain reaction increases. When $k_{eff} = 1$. The reactor is in a critical state, where the neutron population and reactor power remain constant. Meanwhile, when $k_{eff} < 1$. The reactor is subcritical, and the fission reaction decreases.

The reactor reactivity changes with the movement of the control rods and temperature variations within the reactor core. Therefore, the total reactivity can be expressed by Equation 9 (Ribeiro et al., 2020):

$$\rho(t) = \alpha_h \Delta h_{cr} + \alpha_c (T_c - T_c^0) + \alpha_f (T_f - T_f^0) \quad [9]$$

where Δh_{cr} is the extraction height of the control rod measured from the critical position, α_h is the rod worth coefficient, T_c^0 and T_f^0 are the initial coolant and fuel temperatures, respectively, while α_c and α_f represent the coolant and fuel temperature coefficients.

The reactivity data cannot be directly obtained from the Wide Range Nuclear Monitoring System (WRNMS). However, the insertion reactivity can be determined from the control rod position and converted to reactivity values using the rod-worth table. Since each control rod has a different rod worth value, the movement of the four control rods is used to regulate the reactor power level at the desired operating condition. The current reactivity value is computed based on the current rod height and is subsequently converted into a true reactivity value using specific rod worth tables. In other words, each control rod has a different differential reactivity rod worth, and these individual calibration curves were applied prior to determining the total insertion reactivity. These four rods cannot be summed because each rod will hold a different fuel rod.

For the RTP, the coolant temperature coefficient is very small ($\sim 1 \times 10^{-8}$ mΔk/k/°C) and can therefore be neglected (Ribeiro et al., 2020). In addition, the differences between the initial and current coolant and fuel temperatures are assumed to be insignificant. Hence, Equation 9 can be simplified to Equation 10, with the insertion reactivity defined by Equation 11.

$$\rho(t) \approx \delta\rho_{ext}(t) \quad [10]$$

where,

$$\delta\rho_{ext}(t) = \rho_r(t) - \rho_0(t) \quad [11]$$

Here, $\delta\rho_{ext}(t)$ represents the insertion reactivity, defined as the difference between the current reactivity value, $\rho_r(t)$, and the initial reactivity value, $\rho_0(t)$. This simplification indicates that the total reactivity is mainly influenced by changes in the control rod position.

Neutronics Model

The input and output parameters for the neutronics model are the reactivity and neutron values, respectively, as shown in Figure 4. The reactivity is produced by the reactivity model and fed into the neutronics model to generate neutrons. In RTP, the neutrons are detected by a fission chamber and fed into the Neutron Measurement System (NMS). Later, the signals are processed and converted into core power, and the power transformation rate is calculated as the Log Rate. This Log Rate will be utilised to measure the error of the Feedback Control Algorithm (FCA) in controlling the movement of the control rods.

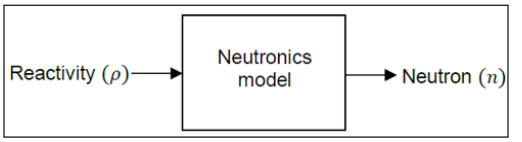


Figure 4. Neutronics subsystem model for RTP

The number of neutrons produced depends on numerous factors, such as the rate of neutron absorption, the fission reaction, the rate of neutron leakage, and the reactor's structural materials. The nuclear fission chain reaction generates neutrons, which will be extinct once they leak outside the reactor or are captured by nuclei. At a certain neutron generation rate, the generated neutrons are absorbed by nuclei, which will trigger fission and produce new neutrons. This is called a delayed neutron. Therefore, the numbers of delayed-neutron precursors and neutrons are used to describe reactor kinetics. The point reactor kinetics equation with six groups of delayed neutron precursors is defined by Equations 12 and 13 (Ribeiro et al., 2020).

$$\frac{dn(t)}{dt} = \left[k_{eff}(1-\beta) - 1\right]\frac{n(t)}{\Lambda} + \sum_{i=1}^{6}\lambda_i C_i(t)\,; \qquad [12]$$

where,

$$i = 1..6,;\ \ \beta = \sum_{i=1}^{6}\beta_i$$

$$\frac{dC(t)}{dt} = k_{eff}\beta_i\frac{n(t)}{\Lambda} - \lambda_i C_i(t) \qquad [13]$$

By referring to Equations 12 and 13, the reactor kinetics can be expressed in simplified forms given by Equations 14 and 15.

$$\frac{dn(t)}{dt} = \frac{\rho(t) - \beta}{\Lambda} n(t) + \sum_{i=1}^{6} \frac{\beta_i}{\Lambda} C_i(t) \quad [14]$$

$$\frac{dC_i(t)}{dt} = C_i n(t) - C_i n_i(t); \;\; i = 1, \ldots, 6 \quad [15]$$

Where n is the number of neutrons, C_i is the number of delayed neutrons precursors in group i, β_i is the delayed neutron fraction of group i, λ_i Is the decay constant for the delayed neutron precursor of the group i, k_{eff} is the effective multiplication factor and Λ is prompt neutron lifetime also called mean neutron generation time. Hence, the actual power can also be known as core power, which is relevant to the neutron density, and the actual core power can be expressed by Equation 16 (Wang et al., 2017).

$$P_a = \psi P_0 \quad [16]$$

where P_a is the actual core power, P_0 is nominal core power and ψ. It is a relative core power.

METHODOLOGY

Data Collection

The real input and output data were acquired at four power levels, i.e., 500kW, 750kW, and 1000kW. The data was recorded for up to 10,000 seconds only because it was sufficient for the RTP system to be completed within that period. Normally, at the beginning, the control rods of RTP were set to an initial height of 220 mm, except for the Regulator rod. When the power reached 15W to get the core excess value, the rods' movement was paused. The rod movement relies strictly on the natural operational and standard control rod manoeuvres required to achieve steady-state power demands (250kW, 500kW, 750kW, and 1000kW). This includes initialising rod heights to 220mm, pausing at 15W, and utilising the auto control system to reach the final power level. No intentional pseudo-random binary sequences or step tests were injected during the operational and data collection.

Before the data were tested, the real initial height of each control rod was identified. Even though the Reactor Digital Instrumentation Control System (ReDICS) in the RTP system sets the initial values for the Shim, Transient, and Safety rods at 220mm, the absolute rod heights differ due to adjustments to control rod movements. The actual initial height of each rod at all power levels is shown in Table 1.

All parameters of the core reactor system were acquired by WRNMS and stored in a computer. According to ReDICS, there are 342 parameters in the RTP system's data acquisition system. The parameters for each subsystem reactor system were intermingled

Table 1
Initial height for each rod

Power level	Initial height level (mm)			
	Shim	Safety	Regulator	Transient
500kW	200.6	200.0	294.5	200.4
750kW	199.6	200.3	233.9	199.5
1000kW	197.6	198.0	197.6	197.8

in the same document. For safety purposes, the data for each variable were recorded redundantly and alternately in the two systems in the same document. 10000 data samples were recorded for each parameter, with a sampling interval of 0.5 seconds. During the preprocessing phase, the data were resampled to a 1-second period, resulting in a refined dataset of 5000 data points for the identification modelling.

System Identification Modelling

Since the model is a black-box model structure, no previous knowledge is required. The Transfer Function (TF) model has been employed in this research to represent all subsystem models using the SI approach. The output of the SI modelling is a transfer function, also called a model, for each subsystem, representing the relationship between real input and output data and is unique to each model. The subsystem model of the RTP reactor system is depicted in Figure 5.

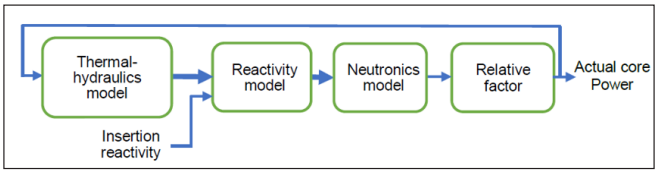


Figure 5. The subsystem model of the RTP reactor system

Since the parameters were in blended data sources, parameter modularisation was needed to categorise each parameter involved in each subsystem model by filtering the data. Then, the data was resampled for a 1-sec period. Therefore, a total of 5000 data points were used in this study, which is adequate to show that the RTP system could be completed within this period. The resampling was done by filtering out duplicate data. However, before analysing, the neutronics model data had to be normalised due to a large gap between the input and output data. The Min-Max Normalisation is expressed in Equation 17 (Chen, 2019).

$$u_i = \frac{[y_i - y_{min}]}{[y_{max} - y_{min}]} \qquad i = 1,2,3,..n \qquad [17]$$

Where u is the output after normalisation, y. Is it a real plant data set or an actual value, *i* is the number of data points until n. The total number of samples. Both datasets were normalised in the range of [0, 1]. The preprocessing flow in SI modelling is shown in Figure 6.

All data were in the time domain and were analysed after the spectra and time were plotted. Subsequently, the means for each data set in the preprocessing operations were removed. The mean value was removed by subtracting the mean value of all the input and the output data, as the Equation 18.

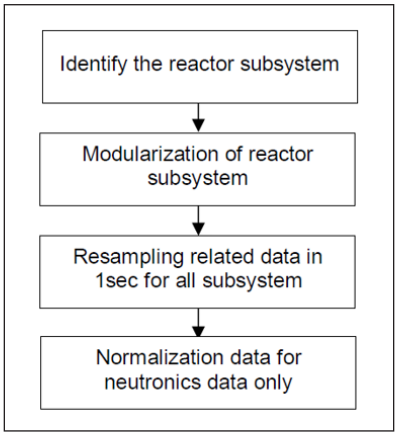


Figure 6. Preprocessing in SI modelling

$$y(t) = u(t) - \bar{u}(t) \quad [18]$$

where $u(t)$ is the input of the system and $\bar{u}(t)$ is the mean input of the system.

The aim of the operations is to obtain more accurate models, since arbitrary variations between input and output data cannot be captured by linear models. In other words, it is a method of data filtering. The data sets ranged between [1 2000].

In the estimation process, the poles and zeros were evaluated up to the 10th order, and the best-fit result was chosen. The best fit (%) quantifies the accuracy of the simulated output in tracking the measured real plat data. The metric represents the Normalised Root Mean Square Error (NRMSE) fit percentage, which normalises the prediction error by the variance of the true data. A value of 100% indicates a perfect match between the model and the actual system. The lower percentages indicate the presence of residual mismatches. The NRMSE (Best fit (%)) is defined in Equation 19 (Ljung, 1999);

$$\text{Best fit (\%)} = 100 \times \left[1 - \frac{\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}}\right] \quad [19]$$

Where y_i, is the real plant data (actual value), $\hat{y}_i$ is the predicted model output, $\bar{y}_i$ is the mean of the real plant data, n is the total number of the sample.

For parameter estimation, the covariance matrix of the estimated parameter vector was used to quantify uncertainty in properties such as poles and zeros, as well as in Bode

plots. For the validation process, Akaike's Final Prediction Error (FPE) was used to select the model order. The best fit produced the smallest FPE, as defined in Equation 20 (Roy et al., 2021).

$$\mathrm{FPE} = \frac{1 + d_\gamma/n}{1 - d_\gamma/n} \frac{1}{N} \sum_{t=1}^{n} \frac{1}{2} \varepsilon^2\left(t, \hat{\vartheta}\right) \quad [20]$$

where d_γ is the dimension of the model parameter vector, $\varepsilon(t)$ is the prediction error or residual, ϑ is the parameter vector, and n is the total amount of data.

The smallest FPE evaluated is defined as the estimated model that best matches the real plant, i.e., the most accurate model. The whole process, from identifying each reactor subsystem to modelling validation of the data, has been performed for all reactor subsystems. To provide further justification, the validation process included residual analysis to assess the estimated model's accuracy. The flow of the SI modelling process is depicted in Figure 7.

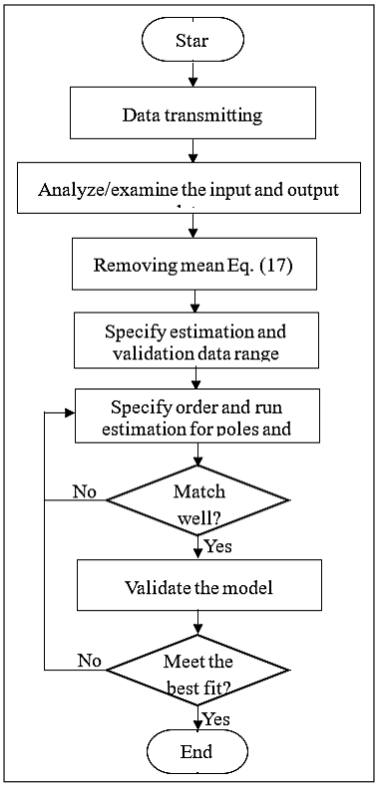


Figure 7. The flow process of SI modelling

The SI model for three subsystems, i.e., the thermal-hydraulics model, reactivity model, and neutronics model, was examined, with all results and plots, including best-fit, poles and zeros, FPE, residual, and RMP between Real Plant (RP) and SI, along with the FP model. The accuracy of the models was also evaluated through Mean Squared Error (MSE). The RMP and MSE are defined in Equations 21 and 22, respectively (Li et al., 2021).

$$\mathrm{RMP} = \left(\frac{y_i - \hat{y}_i}{y_i}\right) \times 100\% \quad [21]$$

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad [22]$$

Where y. Is it a real plant data set or an actual value, and $\hat{y}$ is a prediction data set, and i is the number of data points until n. The total number of samples.

RESULTS OF SYSTEM IDENTIFICATION MODEL

There were three subsystems of the reactor core system after modularisation and identification. As mentioned, these subsystems were tested at the specified power level and initial height after resampling for a 1-sec period. Generally, in SI modelling, after analysing the input and output data, the mean is removed, and then the ranges of the estimation and validation values are determined. Since removing means had been done, the pre-processing utilised zeroth-order detrending to eliminate the steady-state direct current (DC) offset, allowing the model to capture the transient variations. This work did not apply the first-order linear detrending. Therefore, any slow drifts due to xenon poisoning or sensor calibration shifts were not mathematically removed. Instead, these variations were left naturally embedded within the operational dataset.

Then, after specifying the order and running pole-and-zero estimation, model validation continues until the best-fit FPE and MSE are met. The figure of the real plant and the SI model was plotted and showed based on FPE, MSE and the best fit of the process. These figures are exposed in Figure 8.

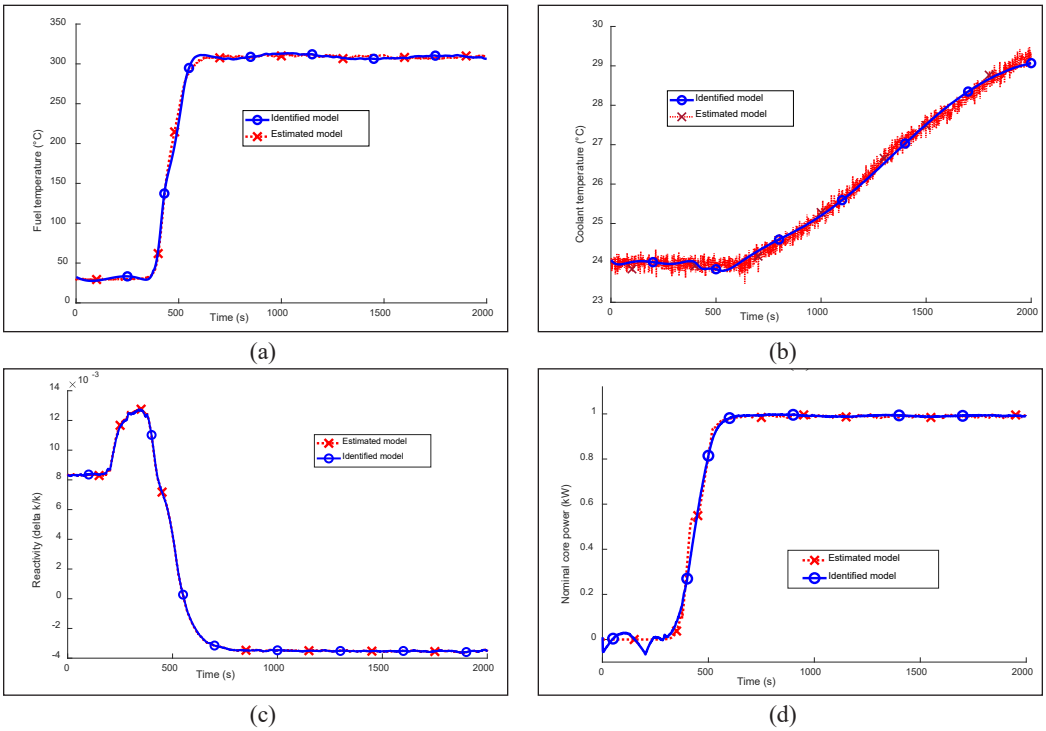


Figure 8. SI and real plant model for (a) fuel temperature of the thermal-hydraulics model, (b) coolant temperature of the thermal-hydraulics model, (c) reactivity model and (d) neutronics model

Generally, the SI model of fuel temperature in thermal-hydraulics, reactivity, and neutronics closely follows the real plant, except during the first hundred seconds in the neutronics model. Figure 8(a) The SI model for fuel temperature of the thermal-hydraulics model can follow the real plant model. The slightest difficulty in tracking from 500sec to 600sec due to strong negative temperature feedback from the reactor core. Thus, core power control performance at RTP shows non-smooth control around 500kW. As shown in Figure 8(b), the SI model of coolant temperature struggles to track the real plant model due to fluctuations in the real plant coolant temperature throughout the process. Furthermore, instead of the actual inlet temperature, a value of 32˚C was utilised due to the absence of inlet water temperature in the WRNMS data acquisition. Compared to others, the SI model of reactivity is the closest to the real plant, as shown in Figure 8(c). Figure 8(d) depicts the SI model of neutronics towards a real plant. For the first 400 seconds, the SI model appears to struggle to track the real plant. This is because the fission reaction is insignificant (the power level is around 25W) at this moment, and the coolant temperature affects it.

In a research reactor, when the coolant temperature is lower, the temperature gradient between the fuel and the coolant is higher. This increased temperature difference enhances the convective heat transfer, allowing more heat to be extracted from the fuel. As a result, the coolant temperature rises as it absorbs the heat from the fuel, while the fuel temperature decreases. This relation is shown in Figure 8, as coolant temperature is increasing in parallel with fuel temperature for both SI and the real plant model and interrelated as also depicted in Equation 6and Equation 7. The best fit, FPE and MSE for each model of the reactor core system are depicted in Table 2 below.

From Table 2, all SI models achieved a best fit above 95%, except for the coolant temperature model in the thermal-hydraulics system, which achieved 92.79%. Models with a best fit exceeding 95% indicate that the SI models are acceptable and can be considered nearly identical to the real plant behaviour. Achieving a better fit for the coolant temperature SI model was challenging due to fluctuations in the actual plant coolant temperature data throughout the process. Nevertheless, the SI models for both fuel and coolant temperatures in the thermal hydraulics system produced relatively similar results, with MSE values of approximately 6.773 and 6.873 and FPE values of 0.257 and 0.267, respectively. These values are larger compared to those obtained for the reactivity and neutronics SI models.

Table 2
SI modelling evaluation

Model	Best fit (%)	FPE	MSE
Fuel temperature for thermal hydraulics	96.62	0.257	6.773
Coolant temperature for thermal hydraulics	92.79	0.267	6.873
Reactivity	99.31	4.502E-9	1.892E-9
Neutronics	95.66	1.657E-4	1.618E-4

The SI models of reactivity and neutronics exhibit insignificant FPE and MSE values, with the reactivity SI model achieving the highest best fit of 99.31%. This demonstrates that the reactivity model most accurately represents the behaviour of the real plant. In addition, both the fuel-temperature and coolant-temperature models were developed using second-order transfer functions. The SI approach successfully generated a transfer function for each model. The resulting transfer functions for the thermal-hydraulics models of fuel temperature and coolant temperature are presented in Equations 23 and 24, respectively.

$$T_f = \frac{2.514s^2 + 0.1122s + 8.076e^{-5}}{s^2 + 2.3831e^{-2}s + 2.973e^{-6}} \quad [23]$$

$$T_c = \frac{377.1s + 189.6}{s^2 + 10.48s + 0.5157} \quad [24]$$

The reactivity model calculates the total reactivity of the reactor system by summing the insertion and feedback reactivities arising from fuel and coolant temperatures. After the SI modelling stage, the best-fit reactivity model achieved 99.31% with 4.502E-9 FPE, as shown in Table 2. The transfer function results for the reactivity model represent the insertion reactivity due to control rod movement, feedback reactivity due to fuel temperature, and feedback reactivity due to coolant temperature, as presented in Equations 25, 26, and 27, respectively.

Insertion reactivity due to control rod motion

$$\rho_{ext}(t) = \frac{2.409s + 8.816e^{-4}}{s^2 + 2.449s + 8.97e^{-4}} \quad [25]$$

Feedback reactivity due to fuel

$$\rho_f(t) = \frac{-8.08e^{-5}s - 5.092e^{-4}}{s + 5.908} \quad [26]$$

Feedback reactivity due to the coolant

$$\rho_c(t) = \frac{3.895e^{-5}s + 1.714e^{-8}}{s + 3.164e^{-4}} \quad [27]$$

For nominal core power, the input-output relationship was between reactivity and nominal core power. Based on the rods' movement, the reactivity model produced a reactivity value, which was utilised in the neutronics model to determine the nominal core

power, before being converted to the actual core power. The model for nominal core power from neutrons in the SI technique is given by Equation 28.

$$P_0 = \frac{1485.45^2 + 64.667075s + 21.75}{1537s^2 + 1542s + 5} \quad [28]$$

Analysis Results of the Thermal-Hydraulics Model

Figure 9 illustrates the fuel and coolant temperature responses of the thermal-hydraulics model at different power levels, namely 500kW, 750kW, and 1000kW. Starting from

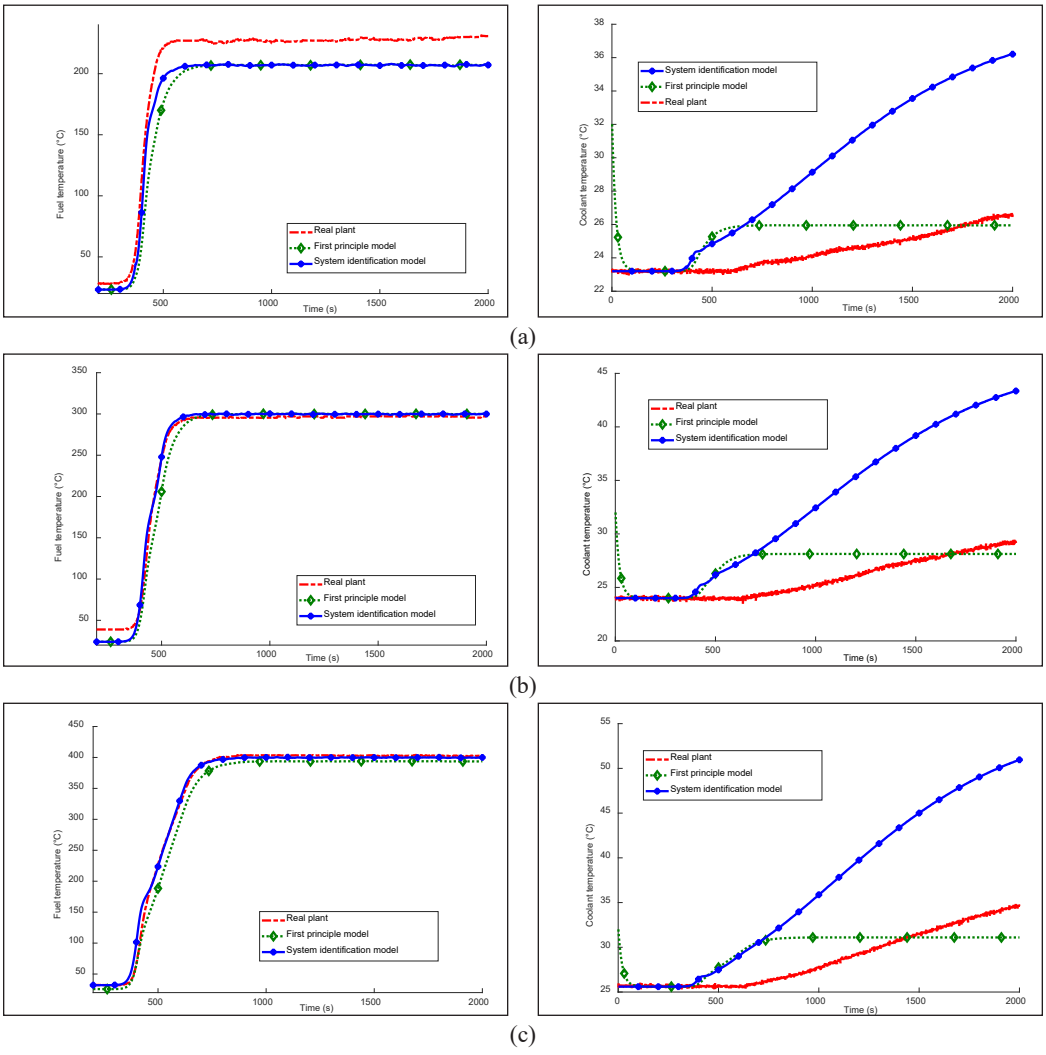


Figure 9. Fuel and coolant temperature responses of the thermal-hydraulics model at (a) 500kW, (b) 750kW, and (c) 1000kW

approximately 26°C, the fuel temperature increases with increasing power level and reaches nearly 300°C. The figures show that both the SI and FP models for the 500kW data set have difficulty accurately tracking the actual plant fuel temperature, exhibiting large steady-state deviations. However, as the power level increases, both models become better at following the real plant fuel temperature response. In particular, the fuel temperature responses at 750kW and 1000kW demonstrate that the SI model provides the best representation of the reactor system fuel temperature.

In contrast, the coolant temperature responses for all data sets and power levels do not closely follow the real plant coolant temperature. The figures indicate that the SI model predicts a gradual increase in coolant temperature to nearly twice the real plant response. Meanwhile, the FP model can control the initial increase in coolant temperature before maintaining it at a relatively constant value.

It is evident that the coolant temperature responses of both the SI and FP models deviate from the real plant behaviour. In this study, a fixed inlet coolant temperature of 32°C was used in the thermal-hydraulics model due to the lack of real-time inlet coolant temperature data. As a result, modelling coolant temperature becomes more challenging and may lead to inaccuracies in the model's response. Therefore, the inlet coolant temperature model requires further investigation and improvement. In addition, the coolant temperature coefficient ($\sim 1 \times 10^{-8}$ mΔk/k/°C) is insignificant compared to the fuel temperature coefficient (-0.1 mΔk/k/°C) (Minhat et al., 2018). Hence, the deviation in coolant temperature does not significantly affect the reactor core system power level.

Table 3 summarises the mean and RMP values of fuel temperature for the reactor system, using real plant data, the FP model, and the SI model. The results indicate that at the 500kW power level, both the SI and FP models have difficulty accurately tracking the real plant fuel temperature, as shown by the considerable RMP differences relative to other power levels. However, as the power level increases, the RMP values decrease, indicating that both the SI and FP models are better able to follow the real plant behaviour.

Table 3
Mean and RMP based on real plant data for fuel and coolant temperature

Power Level (kW)	Mean temperature, (ºC)			RMP (%)	
	Real plant	FP model	SI model	FP	SI
Fuel Temperature					
500	204.216	181.902	184.596	10.93	9.61
750	259.625	257.388	261.438	0.86	-0.69
1000	338.852	326.399	331.809	3.68	2.08
Coolant Temperature					
500	24.347	25.401	29.207	-4.33	-19.96
750	25.796	27.184	32.712	-5.38	-26.81
1000	28.702	29.662	36.580	-3.35	-27.45

Several negative RMP values were also observed, indicating that the predicted values from the FP and SI models exceeded the corresponding real plant values. As illustrated in Figure 9 for the 750kW fuel temperature response, the SI model is slightly lower than the real plant during the first 300sec but subsequently becomes noticeably higher than the real plant response. This deviation may be attributed to the increasing reactor power level, which affects the thermal behaviour of the reactor core and introduces nonlinear characteristics that the SI model finds difficult to capture accurately. In addition, uncertainties in heat transfer parameters and coolant flow behaviour may contribute to the observed differences between the model and the plant's actual response.

Analysis Results of the Reactivity Model

Plotted graphs comparing the SI, FP, and real plant reactivity models are shown in Figure 10 for all power levels. In all cases, the FP model closely follows the real plant response up to approximately 300sec; however, it subsequently decreases and stabilises at near-zero reactivity, deviating significantly from the real plant's total reactivity. In contrast, the SI model can track the real plant behaviour more consistently across all power levels.

From the figures, the SI reactivity model at all three power levels exhibits a trend similar to the real plant: both initially increase to approximately 4×10^{-3} delta k/k, then decrease

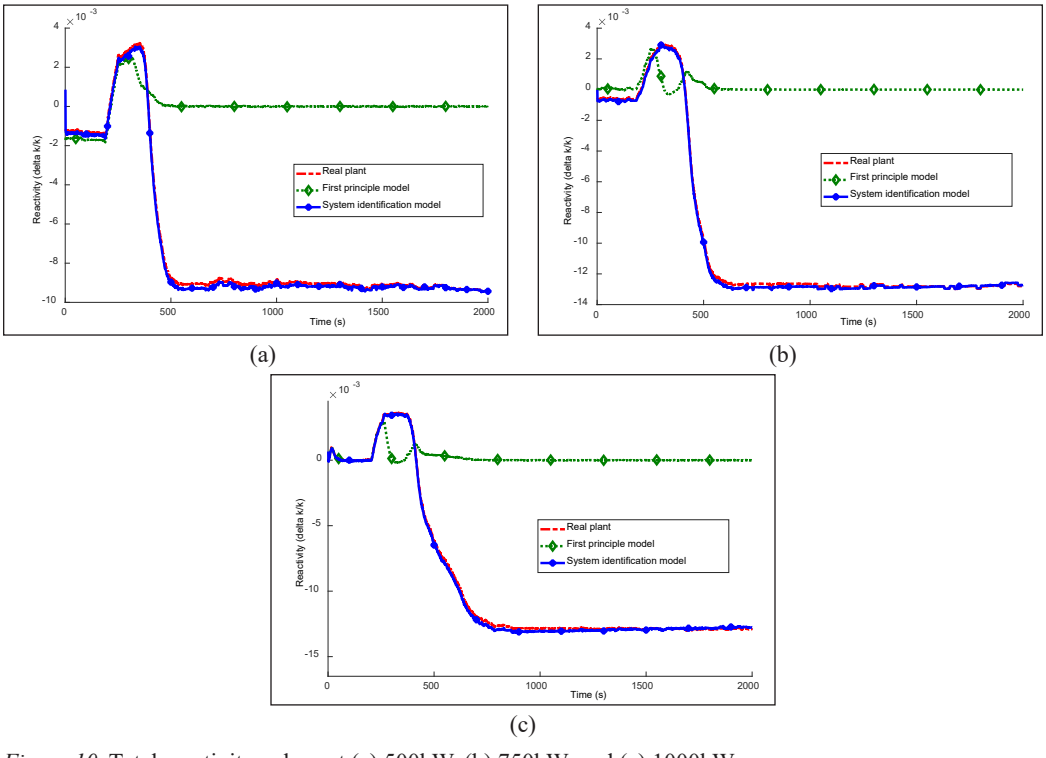


Figure 10. Total reactivity values at (a) 500kW, (b) 750kW, and (c) 1000kW

and stabilise at a negative reactivity level. For the 500kW case, the system stabilises at approximately -9×10^{-3} delta k/k. Similarly, for 750kW and 1000kW, the total reactivity stabilises at around -13×10^{-3} delta k/k and -14×10^{-3} delta k/k, respectively. However, at 500kW, the SI model shows a slight deviation from the real plant between 500sec and 900sec. This variation may be attributed to temperature feedback effects on reactivity, in which temperature changes influence neutron absorption behaviour and introduce modelling uncertainties under transitional operating conditions.

As shown in Table 4, the comparison of mean values and RMP for the reactivity models at each power level is presented. Although some RMP values are negative, the results indicate that the SI model consistently provides closer agreement with the real plant than the FP model, as reflected in its lower RMP values. The SI model RMP ranges from |1.09%| to |1.84%|, which is significantly smaller than that of the FP model, which ranges from 101.37% to 102.31%.

Table 4
Mean and RMP values of total reactivity based on real plant data

Power Level (kW)	Reactivity, ρ(Δk/k)			RMP (%)	
	Real plant	FP model	SI Model	FP model	SI Model
500	-7.685E-03	1.779E-04	-7.826E-03	102.31	-1.84
750	-1.062E-02	1.455E-04	-1.073E-02	101.37	-1.09
1000	-1.013E-02	1.601Ee-04	-1.025E-02	101.58	-1.18

Analysis Results of the Neutronics Model

Figure 11 illustrates the plotted responses of the neutronic model at each power level. Overall, the SI model demonstrates closer agreement with the real plant behaviour of the neutronic response across all power levels compared with the FP model. The results show that during the initial 200 seconds, the SI model closely tracks the real plant response. However, a temporary deviation is observed between approximately 200sec and 350sec. After 350 seconds, the SI model gradually converges towards the real plant trend, although it does not fully match the exact values.

As the power level increases from 500kW to 750kW, the SI model shows improved agreement with the real plant response. However, at 1000kW, a more noticeable discrepancy is observed, indicating reduced modelling accuracy at higher operating conditions. This suggests that the SI model faces greater challenges in capturing the neutronic behaviour at elevated power levels.

The input to the neutronics model is based on the reactor reactivity output. It is observed that when the reactivity increases and subsequently decreases between 250sec and 500sec, the neutronic response exhibits an inverse behaviour. This behaviour may be associated with xenon poisoning effects within the reactor core, which influence neutron absorption

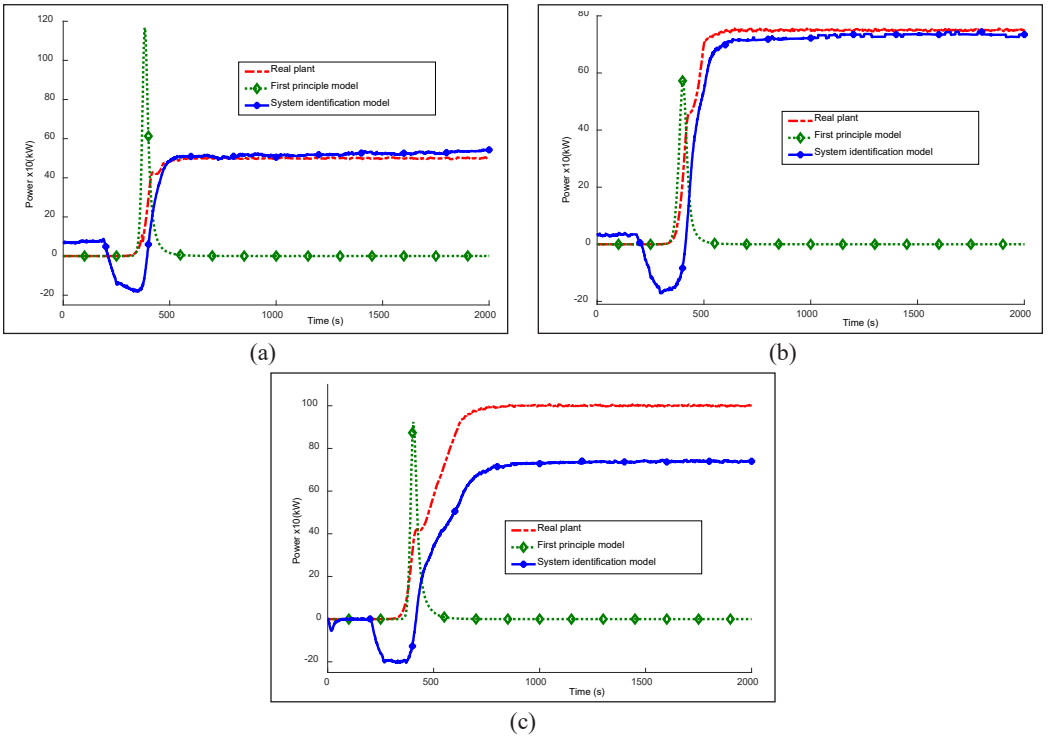


Figure 11. Actual core power response of the neutronics model at (a) 500kW, (b) 750kW, and (c) 1000kW power levels

and reactor kinetics. In contrast, the FP model does not accurately reproduce the real plant behaviour at all power levels. It shows an initial rise, peaking around 500sec, followed by a rapid decline back towards the initial level, and subsequently diverges from the real plant response for the remainder of the simulation.

The analysis for each power level was conducted by calculating the mean and RMP of the SI and FP models using real-plant data. The analysis results presented in Table 5 indicate that the RMP values for the SI models were significantly smaller, remaining below 8% for the 500kW and 750kW power levels. The neutronics model depends on the total reactivity value to produce the nominal core power. However, at the 1000kW power

Table 5
Mean and RMP based on real plant data for the neutronics model

Power Level (kW)	Power (kW)			RMP (%)	
	Real Plant	FP Model	SI Model	FP Model	SI Model
500	444.107	2.633	438.440	99.41	1.28
750	653.497	1.606	605.619	99.75	7.33
1000	843.109	2.265	577.233	99.73	31.54

level, the residual percentage increased to 31.54%. In contrast, the RMP values for all FP models deviated considerably from the real plant data, reaching approximately 99%. Therefore, the SI models outperformed the FP models, producing lower RMP values across all power levels.

CONCLUSION

The modelling approach for all subsystems of the reactor core system involved developing thermal-hydraulic, reactivity, and neutronics models. The evaluation used the highest best-fit value and the lowest FPE and MSE values to assess model accuracy, ensuring that the selected model was like the real plant. The results showed that the fuel temperature of the thermal-hydraulics, reactivity, and neutronics models provided the best fit, with 95% higher than that of the coolant temperature model. The lowest FPE and MSE values came from neutronics and reactivity models. In general, both coolant and fuel temperature models of thermal-hydraulics, as well as insertion reactivity due to control rod motion and neutronics models, produced a second-order model. However, both feedback reactivity due to fuel and coolant had been modelled as first-order. The MSE analysis showed that the reactivity and fuel temperature of the thermal hydraulics of FP models were lower than those of the SI model at all power levels. Meanwhile, the coolant temperature in the SI model's thermal-hydraulics and neutronics sections is lower at all power levels. This indicates that the neutronics and fuel temperature in the SI model's thermal hydraulics are closer to the real plant than those in FP.

ACKNOWLEDGEMENTS

The authors also express their sincere gratitude to the Malaysian Nuclear Agency for providing the data and technical advice that contributed to the successful completion of this research.

REFERENCES

Aguel, S., Meddeb, Z., & Jeday, M. R. (2019). Parametric study and modeling of cross-flow heat exchanger fouling in phosphoric acid concentration plant using artificial neural network. *Journal of Process Control, 84*, 133-145. https://doi.org/10.1016/j.jprocont.2019.10.001

Arda, S. E., & Holbert, K. E. (2016). Nonlinear dynamic modeling and simulation of a passively cooled small modular reactor. *Progress in Nuclear Energy, 91*, 116-131. https://doi.org/10.1016/j.pnucene.2016.03.033

Cammi, A., Ponciroli, R., Borio di Tigliole, A., Magrotti, G., Prata, M., Chiesa, D., & Previtali, E. (2013). A zero dimensional model for simulation of TRIGA Mark II dynamic response. *Progress in Nuclear Energy, 68*, 43-54. https://doi.org/10.1016/j.pnucene.2013.04.002

Chen, P. (2019). Effects of normalization on the entropy-based TOPSIS method. *Expert Systems with Applications, 136*, 33-41. https://doi.org/10.1016/j.eswa.2019.06.035

Dim, O. U., & Aghara, S. K. (2019). Modeling and evaluation of fissile material utilization of the UMLRR using Monte Carlo MCNP6 code. *Annals of Nuclear Energy, 130*, 365–376. https://doi.org/10.1016/j.anucene.2019.03.007

Ghazali, A. K., Minhat, M. S., & Hassan, M. K. (2017). Automated power control system for reactor TRIGA PUSPATI. *AIP Conference Proceedings, 1802*, Article 020010. https://doi.org/10.1063/1.4972908

Kazeminejad, H. (2012). Thermal–hydraulic modeling of reactivity insertion in a research reactor. *Annals of Nuclear Energy, 45*, 59-67. https://doi.org/10.1016/j.anucene.2012.02.017

Khalid, M., & Khan, M. T. (2023). Intelligent fuzzy based fault diagnosis and control techniques for industrial process systems: A review. *Engineering Applications of Artificial Intelligence, 122*, Article 106130.

Kim, B. S., & Kim, T. G. (2020). Modeling and simulation using artificial neural network-embedded cellular automata. *IEEE Access, 8*, 24056–24061. https://doi.org/10.1109/access.2020.2970547

Leong, C., Semiao, J., Santos, M. B., Teixeira, I. C., Teixeira, J. P., Batista, A. J. N., Goncalves, B., & Marques, J. G. (2015). Fast radiation monitoring in FPGA-based designs. In *2015 Conference on Design of Circuits and Integrated Systems (DCIS)* (pp. 1-6). IEEE. https://doi.org/10.1109/dcis.2015.7388590

Li, Q., Ma, Q., Li, Y., Chen, P., Ma, C., Zhao, B., & Chen, H. (2021). Numerical investigation of conjugated heat transfer of the plate-type fuel assembly in the research reactor. *Frontiers in Energy Research, 9*, Article 663533. https://doi.org/10.3389/fenrg.2021.663533

Liu, H., Chen, T., & Qin, S. J. (2024). State space system identification and fault detection for nonlinear industrial processes. *ISA Transactions, 146*, 120-132.

Ljung, L. (1999). *System identification: Theory for the user* (2nd ed.). Prentice Hall.

Mao, X., Song, S., & Ding, F. (2022). Optimal BP neural network algorithm for state of charge estimation of lithium-ion battery using PSO with Levy flight. *Journal of Energy Storage, 49*, Article 104139. https://doi.org/10.1016/j.est.2022.104139

Martón, I., Sánchez, A. I., Carlos, S., Mullor, R., & Martorell, S. (2023). Prognosis of wear-out effect on safety equipment reliability for nuclear power plants long-term safe operation. *Reliability Engineering & System Safety, 233*, Article 109121. https://doi.org/10.1016/j.ress.2023.109121

Melo, A., Câmara, M. M., & Pinto, J. C. (2024). Data-driven process monitoring and fault diagnosis: A comprehensive survey. *Processes, 12*(2), Article 251. https://doi.org/10.3390/pr12020251

Minhat, M. S., Mohd Subha, N. A., Hassan, F., & Mohamad Nordin, N. (2020). An improved control rod selection algorithm for core power control at TRIGA PUSPATI reactor. *Journal of Mechanical Engineering and Sciences, 14*(1), 6362-6379. https://doi.org/10.15282/jmes.14.1.2020.13.0498

Minhat, M. S., Selamat, H., & Mohd Subha, N. A. (2018). Adaptive control method for core power control in TRIGA Mark II reactor. *IOP Conference Series: Materials Science and Engineering, 298*, Article 012028. https://doi.org/10.1088/1757-899x/298/1/012028

Mohindru, P. (2024). Review on PID, fuzzy and hybrid fuzzy PID controllers for controlling non-linear dynamic behaviour of chemical plants. *Artificial Intelligence Review, 57*(4), Article 97. https://doi.org/10.1007/s10462-024-10743-0

Nagai, Y., Hatsukawa, Y., Kin, T., Hashimoto, K., Motoishi, S., Konno, C., Ochiai, K., Takakura, K., Sato, Y., Kawauchi, Y., Sato, N., Ohta, A., Yamabayashi, H., Tanase, M., Fujisaki, S., Teranaka, T., Takeuchi, N., & Igarashi, T. (2011). Successful labeling of 99mTc-MDP using 99mTc separated from 99Mo produced by 100Mo(n,2n)99Mo. *Journal of the Physical Society of Japan, 80*(8), Article 083201. https://doi.org/10.1143/jpsj.80.083201

Nguyen, N.-D., Nguyen, K.-C., Huynh, T.-N., Vo, D.-H.-D., & Tran, H.-N. (2020). Conceptual design of a 10 MW multipurpose research reactor using VVR-KN fuel. *Science and Technology of Nuclear Installations, 2020*, Article 7972827. https://doi.org/10.1155/2020/7972827

Oka, Y., & Suzuki, K. (Eds.). (2013). *Nuclear reactor kinetics and plant control*. Springer. https://doi.org/10.1007/978-4-431-54195-0

Ratiko, R., Wisnubroto, D. S., Nasruddin, N., & Mahlia, T. M. I. (2020). Current and future strategies for spent nuclear fuel management in Indonesia. *Energy Strategy Reviews, 32*, Article 100575. https://doi.org/10.1016/j.esr.2020.100575

Ribeiro, E. R. G. D., Palma, D. A. P., Henrice, E., Jr., & Su, J. (2020). Transient analysis of the Brazilian Multipurpose Reactor by the coupled neutronics and thermal hydraulics code NTHC1. *Annals of Nuclear Energy, 143*, Article 107449. https://doi.org/10.1016/j.anucene.2020.107449

Roy, D. K., Biswas, S. K., Saha, K. K., & Murad, K. F. I. (2021). Groundwater level forecast via a discrete space-state modelling approach as a surrogate to complex groundwater simulation modelling. *Water Resources Management, 35*(6), 1653-1672. https://doi.org/10.1007/s11269-021-02787-6

Safaei Arshi, S., Amin Mozafari, M., Jozvaziri, A., & Mirvakili, S. M. (2021). Investigation of safety aspects during steady state operation of Tehran research reactor fuel test loop. *Progress in Nuclear Energy, 140*, Article 103895. https://doi.org/10.1016/j.pnucene.2021.103895

Son, S. H., Kim, J. W., Oh, T. H., Jeong, D. H., & Lee, J. M. (2022). Learning of model-plant mismatch map via neural network modeling and its application to offset-free model predictive control. *Journal of Process Control, 115*, 112-122. https://doi.org/10.1016/j.jprocont.2022.04.014

Tang, H. H., & Ahmad, N. S. (2024). Fuzzy logic approach for controlling uncertain and nonlinear systems: A comprehensive review of applications and advances. *Systems Science & Control Engineering, 12*(1), Article 2394429. https://doi.org/10.1080/21642583.2024.2394429

Wang, G., Wu, J., Zeng, B., Xu, Z., Wu, W., & Ma, X. (2017). State-space model predictive control method for core power control in pressurized water reactor nuclear power stations. *Nuclear Engineering and Technology, 49*(1), 134-140. https://doi.org/10.1016/j.net.2016.07.008

Wang, Y., Zhao, C., & Huang, B. (2024). Data-driven dynamic modelling and system identification for nonlinear industrial process monitoring. *Chemical Engineering Research and Design, 204*, 215-228.

Wu, Z., Christofides, P. D., Wu, W., Wang, Y., Abdullah, F., Alnajdi, A., & Kadakia, Y. (2025). A tutorial review of machine learning-based model predictive control methods. *Reviews in Chemical Engineering, 41*(4), 359-400. https://doi.org/10.1515/revce-2024-0055

Zhang, X., Gao, Z., & Jiang, Q. (2025). Dynamic system identification and predictive modelling for complex industrial systems using data-driven approaches. *Computers & Chemical Engineering, 188*, Article 108771.

Pertanika J. Sci. & Technol. 34 (5): 2569 - 2596 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Pneumonia Screening and Prediction on Images Using a Convolutional Neural Network Deep Learning Model

Shuria Saaidin[1], Murizah Kassim[1,2]*, I Gede Pasek Suta Wijaya[3], and Nur Amirah Kamaluddin[1]

[1]Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[2]Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Kompleks Al-Khawarizmi, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[3]Department of Informatics Engineering, Faculty of Engineering, University of Mataram, Jalan Majapahit No. 62, Mataram 83125, West Nusa Tenggara, Indonesia

ABSTRACT

Pneumonia remains a leading cause of morbidity and mortality worldwide, and its radiographic diagnosis depends on expert readers who are unavailable in many settings. This study reports a comparative evaluation of two ImageNet-pretrained convolutional architectures, VGG16 and Inception-v3, for binary pneumonia classification on chest radiographs. The publicly available Kaggle chest radiograph collection was used, restricted to the Normal and Pneumonia classes: 5,216 training images (1,341 normal, 3,875 pneumonia) and an independent test partition of 624 images (234 normal, 390 pneumonia). Each architecture was evaluated under two regimes: feature extraction with a frozen convolutional base and fine-tuning of the upper convolutional blocks. All performance figures are reported on the held-out test partition with 95% Wilson confidence intervals and are accompanied by specificity, balanced accuracy, and the Matthews correlation coefficient, as the test partition is imbalanced. The fine-tuned VGG16 model achieved an accuracy of 0.905 (0.880-0.926), sensitivity of 0.956, specificity of 0.821, F1-score of 0.927, and MCC of 0.797. The fine-tuned Inception-v3 model achieved an accuracy of 0.936 (0.914-0.953), sensitivity 0.951, specificity 0.910, F1-score 0.949 and MCC 0.863, at approximately 60% of the training time per epoch required by VGG16. Fine-tuning improved VGG16 by raising specificity at the cost of sensitivity but produced no measurable change in Inception-v3. These held-out figures are substantially below the near-perfect accuracies commonly reported

ARTICLE INFO

Article history:
Received: 26 May 2025
Accepted: 17 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.02

E-mail addresses:
shuria6809@uitm.edu.my (Shuria Saaidin)
murizah@uitm.edu.my (Murizah Kassim)
gpsutawijaya@unram.ac.id (I Gede Pasek Suta Wijaya)
nuramirahkamaluddin97@gmail.com (Nur Amirah Kamaluddin)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

for these architectures on this benchmark, a discrepancy attributable to the reporting of training-phase rather than test-phase performance.

Keywords: Chest radiography, convolutional neural network, model evaluation, pneumonia screening, transfer learning

INTRODUCTION

Pneumonia continues to be a significant health challenge worldwide, particularly in developing countries (WHO, 2025). In 2021 alone, it was responsible for approximately 2.5 million fatalities, including about 700,000 children under five years old (UNICEF, 2024). Prompt and precise diagnosis of pneumonia is crucial to improve patient outcomes and reduce mortality (Lim, 2022). Chest X-ray imaging is a standard diagnostic tool for pneumonia; however, interpreting these radiographs relies on expert radiologists who may not be readily available, especially in resource-limited settings. The global shortfall in healthcare workers, projected to reach 18 million by 2030, further exacerbates the challenge of timely and reliable pneumonia diagnosis (Boniol et al., 2022). Recently, deep learning techniques, especially CNNs, have been applied to automate medical image analysis tasks. CNN-based models have demonstrated strong performance in detecting pneumonia from chest radiographs (Aljuaid et al., 2026). For example, a CNN using the VGG16 architecture achieved high diagnostic accuracy on a paediatric chest X-ray dataset, outperforming earlier methods (Kermany et al., 2018). Many researchers employ transfer learning by fine-tuning CNN architectures pre-trained on large datasets like ImageNet using pneumonia-specific datasets, such as the Kaggle chest X-ray collection, which contains thousands of images labelled as either normal or pneumonia, to enhance classification performance (Mahalle et al., 2024).

Despite this volume of work, three problems recur across the literature on this benchmark. First, most studies evaluate a single CNN architecture, so architectural differences are confounded with differences in preprocessing, optimiser settings and training schedule, which limits the generalisability of their conclusions. Second, the validation partition distributed with the dataset contains only sixteen images; studies that adopt it unmodified perform early stopping and model selection against a signal in which a single misclassification shifts the reported figure by 6.25 percentage points. Third, and therefore, reported accuracies are frequently training- or validation-phase values presented as held-out performance, and evaluation is often restricted to accuracy alone, with specificity, balanced accuracy, and the Matthews correlation coefficient omitted despite the class imbalance in the test partition. When confusion matrices are published alongside such claims, the stated accuracy is, in some cases, not recoverable from them.

The gap addressed by this study is therefore methodological rather than architectural. The aim is to establish a well-reported benchmark for two widely used CNN architectures,

VGG16 and Inception-v3, under an identical experimental protocol, and to quantify the margin by which held-out performance falls short of the figures commonly claimed. Both models are trained on the same Kaggle chest X-ray dataset with consistent preprocessing and augmentation, and each is examined under two training regimes: an initial phase with the convolutional base frozen, and a subsequent fine-tuning phase in which the upper blocks are unfrozen for domain-specific retraining. The primary focus is the held-out test partition, evaluated for accuracy, sensitivity, specificity, F1-score, balanced accuracy, the Matthews correlation coefficient, ROC-AUC, and average precision, each reported with a 95% confidence interval, together with training and inference times. The contributions are as follows. (i) A comparative evaluation of the two architectures reported entirely on held-out data with confidence intervals and statistical testing of the pairwise differences. (ii) Evidence that fine-tuning does not improve both operating characteristics uniformly but redistributes error between sensitivity and specificity, which has direct consequences for a screening application. (iii) An explicit account of the reporting practices by which accuracy on this benchmark becomes overstated. No claim is made to architectural or theoretical novelty.

RELATED WORK

This section reviews prior work on CNN-based pneumonia screening, organised by architecture, training strategy, frequency-domain methods, and evaluation practices, and concludes by identifying the limitations of the existing literature that motivate the present study.

CNN Architectures for Chest Radiography

Several studies have been reviewed in this research to gain insight into the standard deep learning architectures used in settings like ours. It was observed that most of them use CNNs and combine the method with a variety of deep learning architectures for pneumonia detection. For example, models such as VGG16, ResNet, DenseNet, and Inception-v3 have been shown to identify pneumonia in chest X-ray images, achieving high accuracy (Shin et al., 2016; Tan & Le, 2019). However, these studies agree that the common challenge they face is the scarcity of large, annotated datasets. Most studies have used transfer learning as a powerful technique to mitigate this issue, and this approach is also applied in this research. By training CNNs on large-scale general image datasets and then fine-tuning them for the target task, such as pneumonia X-ray classification, researchers can improve model accuracy while reducing training time (Abd Rahman et al., 2022; Raghu et al., 2019; Yamashita et al., 2018). In addition, ensemble methods, which combine predictions from multiple models, have been shown to further improve diagnostic accuracy by leveraging the complementary strengths of different architectures, resulting in more robust and reliable predictions. (Azis et al., 2022; Gulshan et al., 2019; Hu et al., 2024).

Data Augmentation and Preprocessing

Similar to transfer learning, data augmentation techniques are often employed to boost the performance of deep learning models, particularly when training data is limited. Common augmentations include geometric transformations such as rotations, translations, zooms, and flips, which artificially expand the diversity of the training set. (Sharma et al., 2024). Proper preprocessing of chest X-ray images, like intensity normalisation and contrast enhancement, is also important for improving model performance. Together, these augmentation and preprocessing steps help models generalise better when used with new, unseen data. The effectiveness of deep learning models is typically evaluated using various metrics; for instance, accuracy quantifies the overall correctness of predictions. (El Houby, 2024). The following section describes the performance metrics typically used to evaluate deep learning performance.

Frequency-Domain and Hybrid Architectures

A parallel line of work departs from purely spatial-domain feature extraction by embedding frequency-domain representations directly within the network. Fourier Analysis Network formulations replace conventional dense activations with sine and cosine components, so that a quasi-periodic structure is modelled explicitly rather than approximated by successive spatial convolutions. This principle has been applied to medical image classification in the CONV-FAN-POX architecture, which combines an EfficientNetV2 backbone with Fourier analysis blocks and reports a mean accuracy of 0.9881 and an F1-score of 0.9856 across five-fold cross-validation on the six-class augmented MSLD2.0 dataset, using approximately seven million trainable parameters, substantially fewer than comparable spatial-domain networks (Tülü & Kaya, 2026). Similarly, hybrid attention-based architectures have integrated hierarchical Swin Transformer feature extraction with convolutional neural networks, leveraging self-attention and cross-attention mechanisms to capture both local and global representations for pneumonia classification from chest X-ray images, achieving an accuracy of 98.56% on the publicly available Chest X-Ray dataset and demonstrating improved robustness over earlier deep learning methods (Kunjumon et al., 2026). Notably, the same architecture attains a mean accuracy of 0.8176 on the original, unaugmented partition of that dataset, and the authors report both figures explicitly rather than the higher one alone. Such approaches are relevant to chest radiography because pulmonary texture carries diagnostic information across multiple spatial frequencies that a fixed convolutional receptive field captures only indirectly. They are not evaluated here, since the present study is deliberately confined to the two architectures whose reported performance on this benchmark it seeks to correct. Still, they represent a more promising direction for architectural contribution than further replication of standard transfer-learning pipelines.

Performance Metrics

There are several performance metrics to assess pneumonia detection models, including accuracy, which represents the overall proportion of correct predictions. The second metric is sensitivity or recall, which is the proportion of actual positives, such as pneumonia cases, that were correctly identified as true positives. Lastly, there is specificity, which is the ability to correctly identify normal or non-pneumonia cases among all actual normal cases, such as the true negative rate. The AUC-ROC is a single statistic that summarises, over all classification thresholds, sensitivity as a function of 1-specificity (Müller et al., 2022; Namdar et al., 2021). In the present study, all these indices were calculated simultaneously to provide a comprehensive and detailed assessment of the model's diagnostic performance. With the development of deep learning techniques, their application in clinical practice has great potential to change the pattern of pneumonia screening by enabling fast, accurate diagnosis. There is a need for continued research and collaboration between AI and the medical community to address the remaining challenges and realise the technology's full potential (Abdulahi et al., 2024; Basubrin, 2025). As the field develops, integrating these models into clinical practice could significantly improve patient outcomes and alleviate the strain on the healthcare system.

Limitations of Existing Work

Four (4) weaknesses recur across the studies surveyed above. First, the validation partition distributed with this benchmark contains sixteen images and is frequently adopted unmodified, so that model selection is performed against a signal with no useful resolution, and a reported validation accuracy of 100% carries almost no information. Second, several studies report training-phase or validation-phase accuracy in place of held-out test accuracy; where confusion matrices are published, the accuracy stated in the text is, in some cases, not recoverable from them. Third, the test partition of this collection is imbalanced at approximately 1.7:1 in favour of pneumonia, yet specificity, balanced accuracy, and the Matthews correlation coefficient are frequently omitted, allowing a model that performs poorly on the minority class to still present a high headline accuracy. Fourth, interpretability evidence is rarely provided, leaving unresolved whether classification is driven by pulmonary findings or by acquisition and positioning artefacts that correlate with class membership in this collection. The present study is directly opposed to these four points.

METHODOLOGY

Figure 1 illustrates the research tasks, which were divided into seven phases. The phases included dataset preparation, model development, training, evaluation, and prediction time analysis.

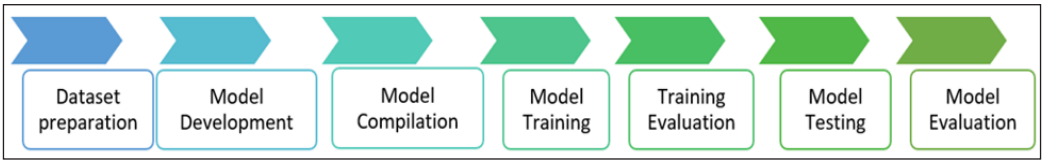
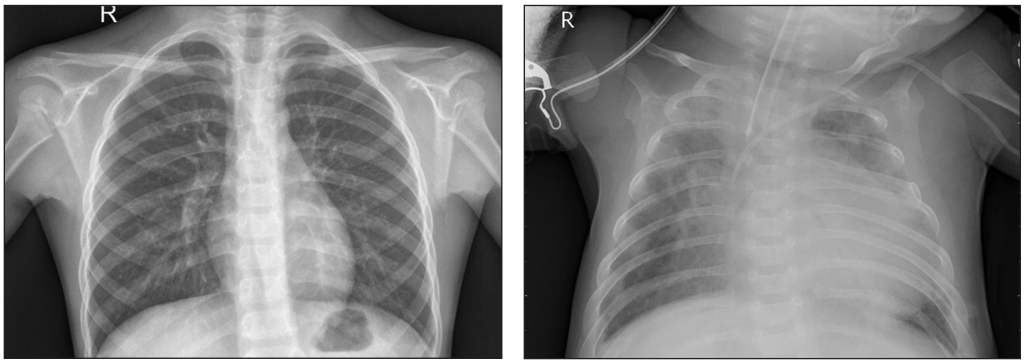


Figure 1. Research activity

Data Collection

In this study, data from the Chest X-ray dataset, including Pneumonia, COVID-19, and Tuberculosis, were utilised. The dataset is available on Kaggle, curated by (JTIPTJ, 2021). Chest X-ray images were divided into three folders: train, test, and validation. Each folder contains subfolders for the image categories: Normal, Pneumonia, COVID-19, and Tuberculosis. This dataset contains 7,135 X-ray images, providing a comprehensive set of labelled images for analysis. However, in this study, only Normal and Pneumonia images were utilised. A sample of datasets is shown in Figures 2(a) and 2(b).



(a) Normal chest X-ray (b) Chest x-ray with pneumonia

Figure 2. Sample *d*ata (a) Normal chest X-ray; (b) Chest X-ray with pneumonia

Dataset Preparation

The data were partitioned into three subsets: a training set, a validation set, and a test set. The training set contained 5,216 images, of which 1,341 were normal, and 3,875 were pneumonia, resulting in an imbalance in favour of pneumonia cases. The validation partition distributed with the dataset consists of only 16 images, eight normal and eight pneumonia, and the test partition comprises 624 images, 234 normal and 390 pneumonia. The test partition is therefore imbalanced at approximately 1.7:1 in favour of pneumonia, and accuracy alone is not an adequate summary of performance on it; specificity, balanced accuracy and the Matthews correlation coefficient are consequently reported alongside

accuracy throughout Section 4. The distributed validation partition is acknowledged to be too small to support reliable model selection or early stopping, since a single misclassification can alter the reported validation accuracy by 6.25 percentage points; therefore, validation-phase figures are reported in this work for completeness only, and no conclusion is drawn from them. The test partition was held out entirely and used once, after all training and model-selection decisions were final. For data preprocessing, Keras was used to normalise pixel values by rescaling image intensities to the range [0, 1] and to perform on-the-fly data augmentation. The augmentations applied were random rotations, shear transformations, zooming, horizontal flips, and brightness adjustments, as described in Shorten and Khoshgoftaar (2019). These transformations were applied during training to overcome data sparsity. The augmentation will also improve the models' robustness by exposing them to a wider range of images. Such normalisation and augmentation strategies are known to facilitate faster convergence and better generalisation in CNNs (Srivastava et al., 2014).

Experimental Environment

All experiments were conducted on a single workstation laptop equipped with an Intel Core i7 processor and 16 GB of system RAM. Models were implemented in TensorFlow using the Keras 3.12 API under Python 3.12, with NumPy 2.2.6. The convolutional backbones were obtained from tf.keras.applications with ImageNet weights, and each network's corresponding preprocessing input transformation was applied as an input layer to ensure preprocessing matched the convention used to learn the pretrained weights. Input images were resized to 224 × 224 × 3. The timing measurements reported in Section 4.5 were obtained on this configuration.

Model Development

Four (4) models were created to analyse the effect of initial training versus fine-tuning: VGG16 Initial Training, VGG16 Fine-Tuning, Inception v3 Initial Training, and Inception v3 Fine-Tuning. All the models were trained on top of a pre-trained base model (ImageNet weights), as in Deng et al. (2009). The architecture for VGG16 and Inception v3 consists of the Base Model, which is a pre-trained VGG16 or Inception v3 without the top classification layer (Simonyan & Zisserman, 2014; Szegedy et al., 2016). A Flatten layer was utilised to transform the 2D feature maps to 1D feature vectors. The flattened vector was passed to a fully connected dense layer, which learned non-linear combinations of the extracted convolutional features. A dropout layer was added to avoid overfitting by randomly setting a fraction of the units to zero at each update. (Srivastava et al., 2014). Finally, the output layer was a single neuron with a sigmoid as the activation for binary classification.

Training Process

The models were optimised using Adam with a learning rate of 1×10^{-4} during initial training, reduced to 1×10^{-5} during fine-tuning, and binary cross-entropy as the loss function. (Kingma & Ba, 2014). The training was completed by compiling a model. Models were built using the Adam optimiser and the binary cross-entropy loss function Equation 1.

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] \quad [1]$$

Where:
y_i is the actual label
p_i is the predicted probability
N is the number of samples.

Callbacks were used with the model checkpoint to write model weights after each epoch if the validation loss decreased. CSV Logger was extended to log training and validation scores per epoch. We used early stopping to halt training if the validation loss had not decreased in 5 epochs, and restored the best weights (Prechelt, 1998). Training metrics, including loss and accuracy, were plotted over epochs in TensorBoard. The models were configured to be trained for 20 epochs. Nevertheless, early stopping was used to terminate the training if the validation loss did not decrease for five consecutive epochs, thereby avoiding overtraining and limiting resource expenditure.

Performance Evaluation

The models were evaluated during the training and testing phases to comprehensively assess their performance.

Training Phase

The training phase was done to check training accuracy. The ratio of correctly predicted instances to the total instances in the training set. Accuracy was calculated using Equation 2. The loss value calculated on the training set indicates how well the model is learning. The ratio of correctly predicted instances to the total instances in the validation set. The loss value calculated on the validation set is used to monitor overfitting.

$$Accuracy = \frac{Number\ \ of\ correct\ \ predictions}{Total\ \ number\ \ of\ predictions} \quad [2]$$

Testing Phase

The testing phase is done to identify the ROC AUC Curve. The Receiver Operating Characteristic curve and the area under it indicate the model's ability to distinguish between classes (Bradley, 1997) as Equation 3.

$$AUC = \int_0^1 TPR(FPR)d(FPR) \quad [3]$$

Where:
TPR is the true positive rate
FPR is the false positive rate

A Precision-Recall Curve was plotted to compare precision versus recall at various threshold settings, providing insights into the trade-off between precision and recall for different decision thresholds (Davis & Goadrich, 2006).

$$Precision = \frac{True\ positives}{True\ positives + False\ positives} \quad [4]$$

$$Recall = \frac{True\ positives}{True\ positives + False\ negatives} \quad [5]$$

Hyperparameters were identified, including the Learning Rate. The learning rate was set to 1×10^{-4} while training the custom top layers, with the convolutional base frozen, and was reduced by an order of magnitude to 1×10^{-5} during the fine-tuning phase to avoid disrupting the pretrained representations with large gradient updates. A batch size of 32 was used for both initial training and fine-tuning. Initially, the epoch was set to 20. However, each training phase consisted of multiple epochs, with early stopping implemented to terminate training if validation performance did not improve. The prediction time for each model was measured to assess its efficiency in real-time applications. The inference times for each model on a batch of images were measured and compared. In the present work, two existing CNN architectures, VGG16 and Inception, were compared in their ability to detect pneumonia in chest X-ray images. The evaluation metrics included training accuracy, validation accuracy, training loss, validation loss, precision, recall, F1 Score, and AUC-ROC. Training and validation accuracy evaluate the performance of the models in classifying pneumonia cases from the training data and from previously unseen data points, respectively. On the other hand, training and validation loss signal the error rates during these periods. Precision evaluates true-positive predictions among all positive predictions but reduces the overall number of false positives. Recall assesses the model's performance in detecting positive cases, which reduces false negatives. The F1-score (the harmonic mean of precision and recall) is a balanced measure for evaluating the model's

performance. AUC-ROC curve describes how well the models differentiate between positive and negative cases at various thresholds. The overall evaluation values indicate the predictive ability, reliability, and potential effectiveness of the models in clinical practice, which can enhance pneumonia detection.

Formal Specification and Additional Evaluation Metrics

The classification pipeline is specified formally as follows. Let x denote an input radiograph after resizing, and let r(x) = x/255 denote the rescaling operator mapping intensities to the interval [0, 1]. During training, a stochastic augmentation operator, A (·), is applied, composed of random rotation, shear, zoom, horizontal flip and brightness adjustment; A is the identity at validation and test time. Let f(·) denote the pretrained convolutional feature extractor with parameters θ, and g (·) the classification head with parameters φ, comprising a flattening operation, a dense layer, a dropout layer and a single sigmoid unit. The predicted probability of the pneumonia class is given by Equation 6.

$$p = g_{\varphi}(f_{\theta}(A(r(x)))),\ p \in (0, 1) \quad [6]$$

In the initial training phase, the feature-extractor parameters θ are held fixed at their ImageNet-pretrained values and only φ is updated, as in Equation 7. In the fine-tuning phase, the upper convolutional blocks θ_top □ θ are unfrozen and updated jointly with φ at a reduced learning rate, as in Equation 8. In both phases, the objective L is the binary cross-entropy given in Equation 1.

$$\varphi^* = \arg\min_{\varphi} L (\varphi; \theta \text{ fixed}) \quad [7]$$

$$(\theta_{top}^*, \varphi^*) = \arg\min_{(\theta_{top}, \varphi)} L(\theta_{top}, \varphi) \quad [8]$$

In inference, the decision rule assigns the pneumonia class when the predicted probability reaches a threshold τ, fixed at 0.5 throughout this study, as in Equation 9.

$$\hat{y} = 1 \text{ if } p \geq \tau, \text{ and } \hat{y} = 0 \text{ otherwise, with } \tau = 0.5 \quad [9]$$

Because the test partition is imbalanced, three additional measures are reported alongside the accuracy, precision, and recall defined in Equations 2, 4, and 5. Specificity, given in Equation 10, quantifies the proportion of normal cases correctly identified. The F1-score, given in Equation 11, is the harmonic mean of precision and recall. Balanced accuracy, given in Equation 12, is the unweighted mean of sensitivity and specificity and is insensitive to class prevalence. The Matthews correlation coefficient, given in Equation 13, summarises all four cells of the confusion matrix in a single value bounded by −1

and +1 and is the most conservative of the summary measures used here. In Equations 10 to 13, TP, TN, FP, and FN denote true positives, true negatives, false positives and false negatives, respectively.

$$\text{Specificity} = TN / (TN + FP) \quad [10]$$

$$F1 = 2 \cdot (\text{Precision} \cdot \text{Recall}) / (\text{Precision} + \text{Recall}) \quad [11]$$

$$\text{Balanced Accuracy} = (\text{Sensitivity} + \text{Specificity}) / 2 \quad [12]$$

$$MCC = (TP \cdot TN - FP \cdot FN) / \sqrt{[\ (TP+FP)(TP+FN)(TN+FP)(TN+FN)\ ]} \quad [13]$$

RESULTS AND DISCUSSIONS

Model Accuracy

Figure 3 visualises the training and validation accuracy trends for the four trained models. Evaluating the VGG16 and Inception-v3 models before and after fine-tuning reveals notable insights into their classification capabilities. The VGG16 model exhibited a promising learning trajectory in its initial training phase. Starting with a training accuracy of 86.73%, it improved steadily to about 97.48% by the final epoch. The validation accuracy for VGG16 showed greater fluctuation, starting at around 75% in the first epoch, dipping to 68.75% in the second, and reaching 100% by the seventh epoch. Because the distributed validation partition contains only 16 images, a validation accuracy of 100% corresponds to sixteen correct classifications and carries very little information; these validation figures are reported to describe the course of training and are not estimates of generalisation.

This variability suggests that while the model was learning effectively on the training data, it struggled to consistently generalise to the validation data early on, possibly due to the limited size of the validation set. Upon fine-tuning, the performance of VGG16 improved significantly on both the training and validation sets. The fine-tuned VGG16 model's training accuracy started at around 86.61% when we resumed training with the pre-trained weights and quickly climbed to 99.59%. More importantly, its validation accuracy reached 100% by the fifth epoch of fine-tuning and remained perfect across several subsequent epochs. This reflects performance on the 16-image validation partition and does not indicate generalisation; the corresponding held-out test accuracy is 0.9054, as reported in Table 5.

The Inception-v3 model demonstrated robust performance from the outset of initial training. Its training accuracy began at 92.67% and improved to 99.11% throughout the training epochs. The validation accuracy for Inception was also high initially, around 96.36%, with a slight dip in some later epochs, but generally remaining strong and stable. The peak validation accuracy during initial training was 97.31%, which was observed around the 11th epoch, indicating that Inception-v3 was able to generalise well early on.

Fine-tuning further slightly boosted Inception's performance. The fine-tuned Inception model's training accuracy improved from an already high 93.72% to 99.59%. Its validation accuracy also increased, peaking at 99.52% by about the 17th epoch. This suggests that fine-tuning helped Inception to solidify its already high accuracy, achieving near-perfect validation accuracy. Overall, both models demonstrated substantial accuracy improvements through training, with fine-tuning providing an extra performance boost. VGG16 benefited more dramatically from the fine-tuning phase (achieving perfect validation accuracy only after fine-tuning), whereas Inception-v3 showed strong performance throughout and modest gains from fine-tuning. Table 1 compares the initial and final training and validation accuracies for all models, summarising these observations.

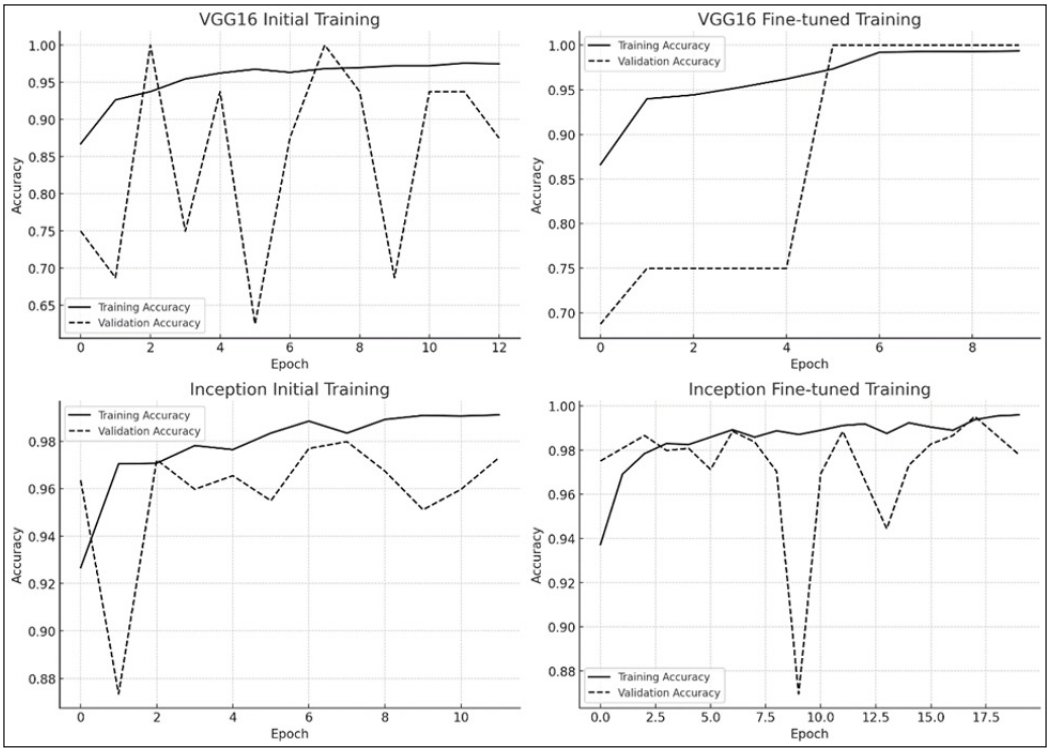


Figure 3. The evaluation of VGG16 and inception models

Table 1 shows that both VGG16 and Inception models demonstrated significant improvements through the training phases, with fine-tuning substantially enhancing their performance. While VGG16 showed considerable fluctuations in validation accuracy during initial training, fine-tuning enabled it to consistently achieve perfect validation accuracy. Conversely, Inception displayed strong performance throughout, with minimal volatility and substantial improvements post-finetuning. The overall analysis highlights the robustness

of Inception and the significant enhancement of VGG16 through fine-tuning, making both models highly effective for classification tasks.

Table 1
Comparison of training and validation accuracies for all models

Model	Training Phase	Initial Training Accuracy	Final Training Accuracy	Initial Validation Accuracy	Final Validation Accuracy
VGG16	Initial	86.73%	97.48%	75.00%	87.50%
VGG16	Fine-tuned	86.61%	99.59%	68.75%	100.00%
Inception	Initial	92.67%	99.11%	96.36%	97.31%
Inception	Fine-tuned	93.72%	99.59%	97.51%	99.52%

Loss Performance for All Models

Figure 4 visualises the training loss and validation loss for all models. An Initial Training was identified for the VGG16 model, and the loss performance exhibited a distinctive pattern. The training loss decreased significantly, from 0.30 to 0.0659, indicating effective learning. However, the validation loss started high at 0.7976, peaked at 0.9322 in the second epoch, and then dropped sharply to 0.1235 by the third epoch. These fluctuations indicate periods of overfitting and recovery. The lowest validation loss was recorded at 0.0835 in epoch 7, suggesting moments of improved generalisation despite initial challenges. Fine-tuning the VGG16 model resulted in a substantial decrease in training loss, from 0.2863 to 0.0114, demonstrating improved learning efficiency. The validation loss showed a general downward trend, starting at 0.4904, with minor fluctuations. Notable peaks in validation loss were observed early in the training, indicating initial overfitting. However, these were followed by a significant drop, with the lowest validation loss of 0.0495 recorded in epoch 7. This reflects that fine-tuning improved the model's generalisation ability, effectively reducing errors on unseen data.

The initial training of the Inception model displayed a different pattern. The training loss decreased significantly from 0.1651 to 0.0297, highlighting efficient learning. The validation loss started low at 0.0907, then increased to 0.3605 in the first epoch, and fluctuated throughout training. These fluctuations suggest intermittent overfitting. Despite these variations, the lowest validation loss was 0.0361 at epoch 11, demonstrating strong generalisation during specific periods. In the fine-tuned training phase, the Inception model's training loss decreased from 0.1636 to 0.0114, indicating continued improvement in learning efficiency. The validation loss showed a general downward trend, starting at 0.0744, but experienced more significant fluctuations. Notable peaks in validation loss were observed at epochs 5 (0.0794), 9 (0.6852), and 13 (0.3019), indicating periods of

overfitting or model adjustments. The lowest validation loss was 0.0205 in epoch 17, reflecting substantial improvements in minimising errors on unseen data.

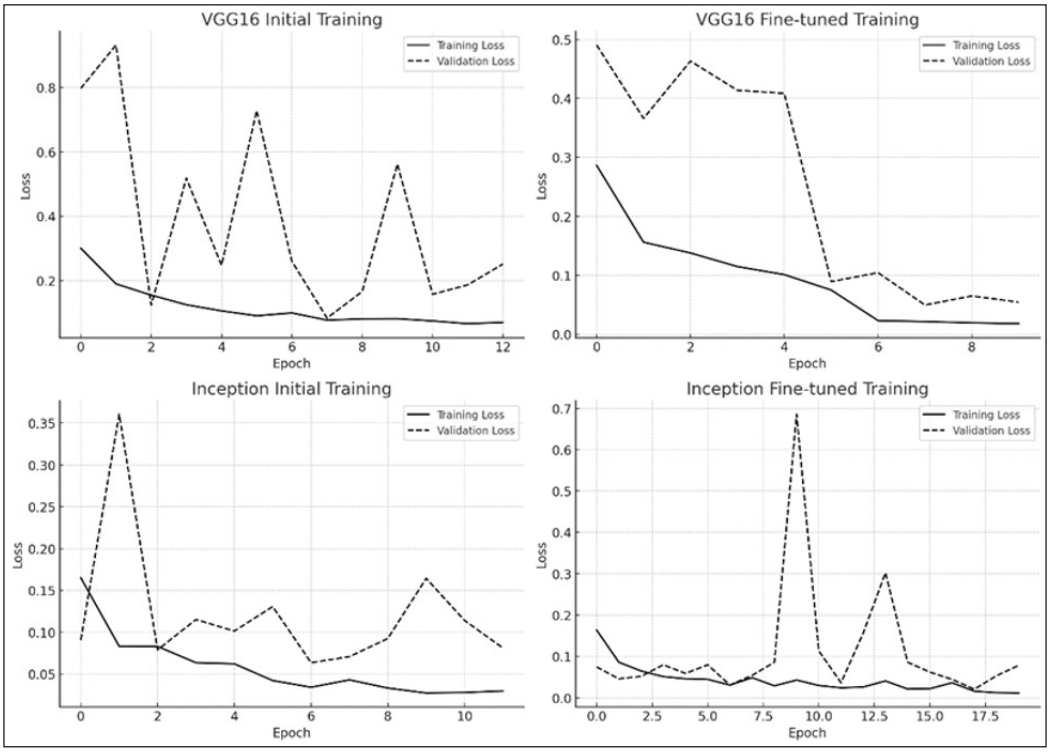


Figure 4. Training loss and validation loss for 4 models

Table 2
Comparison of training and validation loss for all models

Model	Training Phase	Initial Training Loss	Final Training Loss	Initial Validation Loss	Final Validation Loss
VGG16	Initial	0.3000	0.0659	0.7976	0.0835
VGG16	Fine-tuned	0.2863	0.0114	0.4904	0.0495
Inception	Initial	0.1651	0.0297	0.0907	0.0361
Inception	Fine-tuned	0.1636	0.0114	0.0744	0.0205

Table 2 presents the analysis of loss performance for both VGG16 and Inception models, highlighting their learning and generalisation dynamics. VGG16 exhibited substantial improvement post-finetuning, significantly decreasing training and validation losses, indicating enhanced learning and generalisation capabilities. Inception demonstrated robust learning with consistently low training and improved validation losses, particularly after fine-tuning. Despite occasional fluctuations, both models exhibited strong performance

in minimising errors and improving accuracy, underscoring their effectiveness in classification tasks.

Model Performance Evaluation

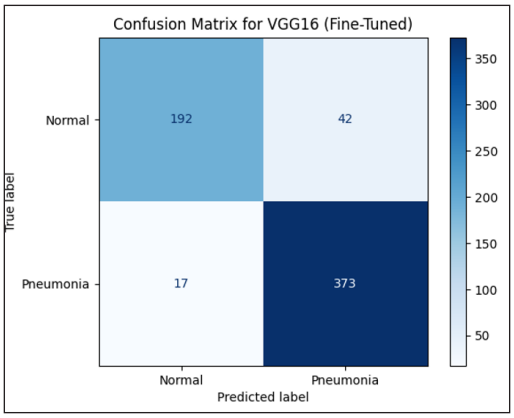
In this study, we trained and evaluated two deep learning models, VGG16 and Inception, to classify chest X-ray images as normal or pneumonia. The models were trained on a large dataset, ensuring comprehensive learning and robust performance. An independent test partition of 624 images was used for evaluation, comprising 234 normal and 390 pneumonia cases. Because this partition is imbalanced at approximately 1.7:1 in favour of pneumonia, accuracy alone would be biased towards the majority class; therefore, specificity, balanced accuracy, and the Matthews correlation coefficient are reported alongside it in Table 5. The provided figures present the confusion matrices for two deep learning models, VGG16 and Inception, evaluated before and after fine-tuning. A confusion matrix is a crucial tool for analysing classification models, as it provides a detailed breakdown of the model's performance in terms of true positives, true negatives, false positives, and false negatives. These matrices offer insights into the models' efficacy in distinguishing between normal and pneumonia cases. Figure 5 (a) to (d) illustrates the confusion matrices for all four models trained.

VGG16 Model Performance

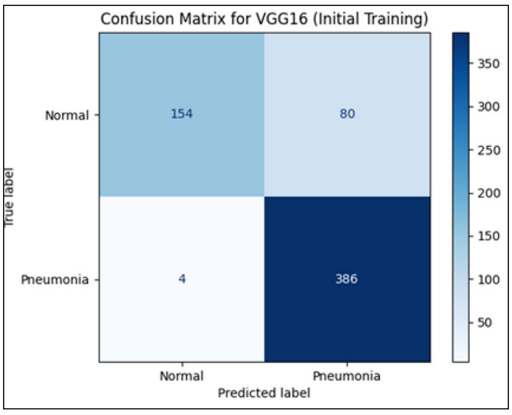
Table 3 presents the confusion matrix for the VGG-16 model at both the initial and fine-tuned stages. The initial training reveals a strong performance in identifying pneumonia cases, with a very high number of true positives and a low number of false negatives. This indicates that the model is susceptible to pneumonia and correctly identifies most pneumonia cases. However, the model struggles more with correctly classifying normal cases, as the higher number of false positives is 80, and the lower number of true negatives is 154. This suggests that while VGG16 is highly sensitive to pneumonia cases, it tends to misclassify normal cases as pneumonia. This performance imbalance suggests an opportunity to improve by increasing the model's specificity without compromising its sensitivity. After fine-tuning, the VGG16 model shows a notable improvement in correctly classifying normal cases, with a significant increase in true negatives from 154 to 192 and a decrease in false positives from 80 to 42. This improvement suggests that the fine-tuning process effectively enhanced the model's ability to identify

Table 3
Comparison of VGG16 at initial training and VGG16 at fine-tuning

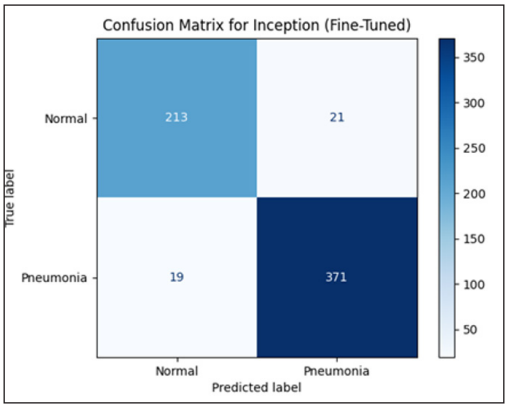
VGG16	VGG16 at Initial Training	VGG16 at Fine-Tuned
True Positives (TP)	386	373
True Negatives (TN)	154	192
False Positives (FP)	80	42
False Negatives (FN)	4	17



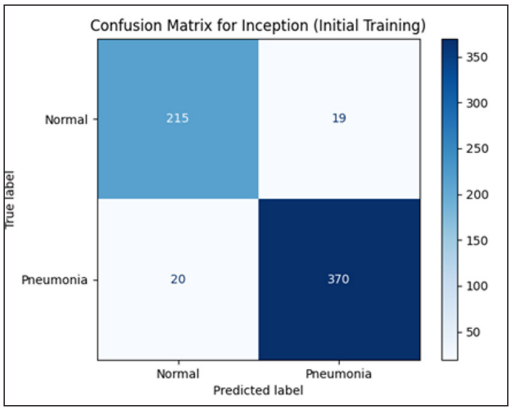
(a) VGG16 Fine-Tuned Model



(b) VGG16 Initial Training Model



(c) Inception Fine-Tuned Model



(d) Inception Initial Training Model

Figure 5. (a) to (d) Confusion matrices for four models

normal cases, thereby increasing its specificity correctly. However, this improvement in specificity comes at a cost: the number of true positives decreases from 386 to 373, while the number of false negatives increases from 4 to 17. This indicates that while the fine-tuning process improved the model's ability to classify normal cases correctly, it slightly compromised its ability to accurately identify pneumonia cases. The increase in false negatives is particularly concerning in medical diagnosis contexts, where missing a positive case can have serious implications. Comparing the two models, it is evident that the Inception model generally performs better than VGG16 in terms of balanced accuracy for both normal and pneumonia cases. The initial Inception training shows fewer misclassifications than VGG16, with fewer false positives and false negatives. Fine-tuning minimally impacts Inception's performance, maintaining its high accuracy, whereas VGG16

shows a significant shift in performance metrics post-finetuning, with improved specificity but reduced sensitivity.

ROC Analysis

Figure 6 displays the ROC curves for four different models evaluated in the study: VGG-16 at both initial and fine-tuned stages, and Inception (both initial and fine-tuned). The ROC curve is a graphical representation of a model's diagnostic performance, plotting the true positive rate against the false positive rate at a fixed specificity.

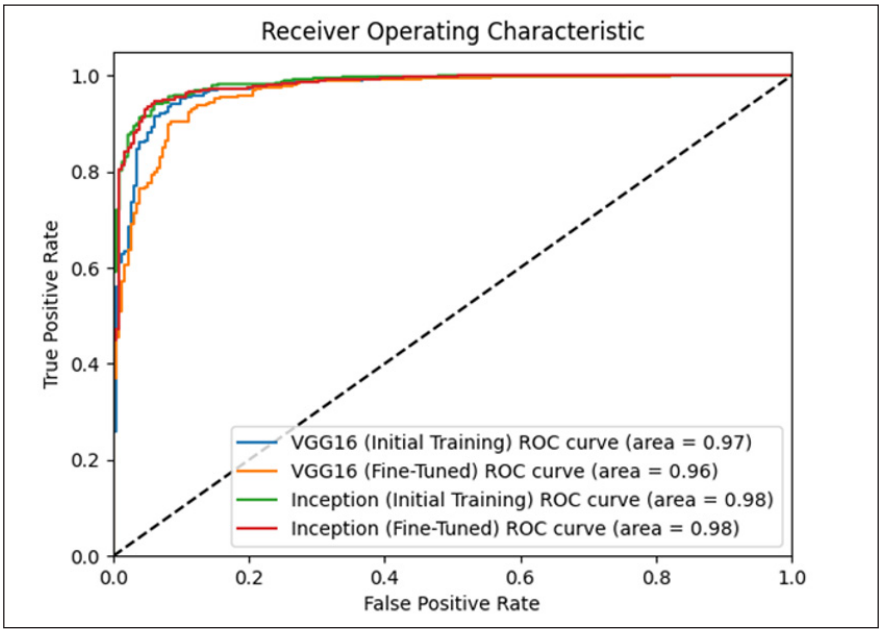


Figure 6. ROC curves for all models

The receiver operating characteristic (ROC) curves for two advanced deep learning models, VGG16 and Inception, are illustrated above after fine-tuning and regularisation. The ROC curve is a key tool for examining the performance of classification models as a function of the trade-off between true positive rate (sensitivity) and false positive rate. The analysis of these curves provides a deep insight into the model's efficiency in class discrimination. ROC curve was plotted with the x-axis representing the false positive rate and the y-axis representing the true positive rate. The false positive rate (FPR) is the ratio of false positives to the total number of positives, N/(false positives + true negatives). On the other hand, the true positive rate (TPR), also known as sensitivity or recall, computes the ratio of the true positives to the sum of true positives and false negatives. The diagonal line represents the performance of a random classifier, with an AUC of 0.50. For any

model that's better than random classification, its curve will be above this line. Analysis of the ROC curves provides valuable insights into the efficacy of the VGG16 and Inception architectures.

The initial training achieves an AUC of 0.97, with high sensitivity and specificity for VGG16. However, with fine-tuning, the AUC drops slightly to 0.96. This slight drop was due to overfitting, and there is a hint that fine-tuning on top of augmentation made the model aware of some repeated images. However, VGG16 remains an excellent model, with an AUC significantly higher than the baseline. On the contrary, the Inception model shows remarkable stability and robustness. With VGG16 as an example, Inception can initially obtain an AUC of 0.98. This noteworthy result demonstrates that the Inception architecture is well-suited for classification. It is also notable that, even after fine-tuning, Inception still achieves an AUC of 0.98, indicating that the model's performance remains very high and is not significantly affected by the fine-tuning. This stable result indicates that Inception is robust and suitable for the data with few performance degradations.

The inception, which is pretrained and fine-tuned, has higher AUC values, which is an obvious indication that it has an a priori advantage in the model architecture better suited for this classification. Although VGG16 performed well, the slight performance drop after fine-tuning suggests it may be more prone to overfitting or other fine-tuning issues than Inception. In summary, the ROC curves of VGG16 and Inception-based models demonstrate that they are effective classifiers with excellent performance. Being passable in terms of AUC, and equally withstanding fine-tuning, Inception would likely be a marginally better option than VGG16. These observations are useful for applications that require a robust classifier, and for which Inception’s robustness and slightly better accuracy may make it an attractive choice. Yet, VGG16 achieves good performance as well, indicating that both architectures have high potential in many classification tasks.

Average Precision

Figure 7 shows the Precision–Recall (PR) curves for the models. PR curves are especially useful for evaluating models on imbalanced datasets, as they focus on performance for the positive class (pneumonia) and can be more informative than ROC curves when one class is rare. The test partition used here is imbalanced at approximately 1.7:1 in favour of pneumonia, so PR analysis is directly applicable and complements the ROC results. Both models exhibit high precision across most of the recall range. For both VGG16 and Inception, precision remains near 1.0 when recall is moderate, and then precision declines as recall approaches 1.0. This is a typical pattern: to capture the very last few per cent of pneumonia cases, which pushes recall toward 100%, the models must tolerate more false positives, which lowers precision. The drop in precision at the extremely high-recall end indicates that some predictions made to capture the remaining pneumonia cases were normal

cases, like false positives. We computed the Average Precision (AP) for each PR curve, which summarises the precision-recall trade-off as a single number, which is essentially the area under the PR curve.

The VGG16 model achieved an AP of approximately 0.98 in its initial state, and this AP remained around 0.98 after fine-tuning. This consistency suggests that fine-tuning did not significantly change VGG16's precision-recall characteristics – the model was already performing near-optimally in terms of balancing precision and recall. The Inception-v3 model achieved a slightly higher AP, about 0.99, both before and after fine-tuning. This indicates that Inception had a marginal advantage in precision-recall performance, maintaining extremely high precision across all recall levels. The fact that both models' AP scores are very high and that neither model's score decreased after fine-tuning underscores that both the initial and fine-tuned models make very few mistakes in a precision-recall sense. For VGG16, an AP of 0.98 indicates that, on average, it achieves 98% precision across all recall levels, which translates to a strong result. Inceptions around 0.99 AP means it performed slightly better, which aligns with its higher precision and the lower false-positive rate observed in the confusion matrix. In summary, the precision–recall analysis reinforces the earlier findings: both models are highly effective at identifying pneumonia, with Inception being a touch better, and the fine-tuning stage preserved the models' excellent precision, recall balance which improves VGG16's recall slightly at the expense of some precision, as seen in the confusion matrix discussion, but overall AP remained very high.

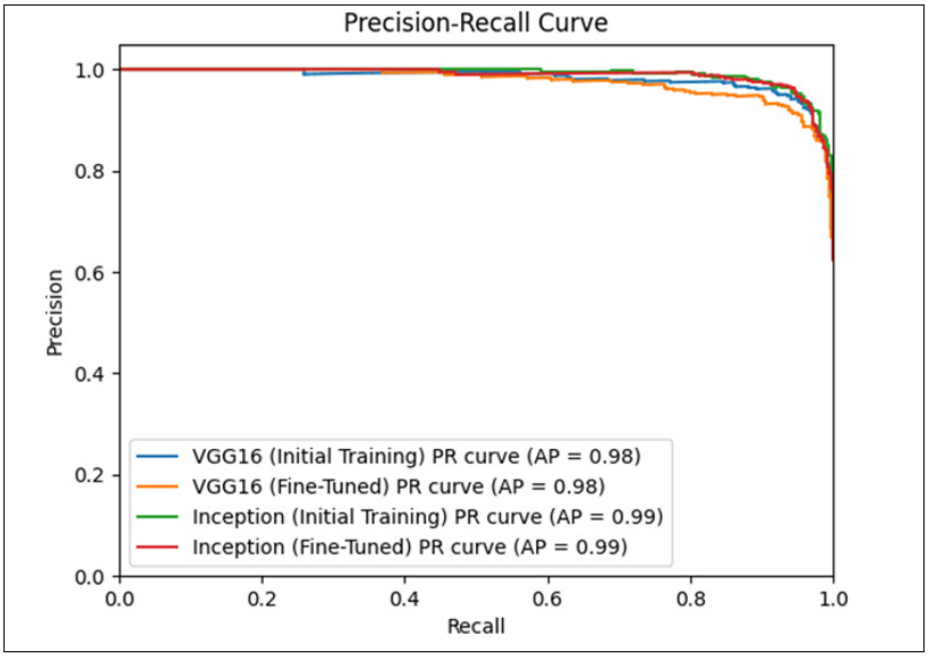


Figure 7. Precision-Recall (PR) curves for VGG16 and Inception deep learning models

F1 Score Analysis

We examined the F1-score for each model to further quantify the balance between precision and recall. The F1-score is the harmonic mean of precision and recall, providing a single metric that balances the two (with an F1 of 1 indicating perfect precision and recall, and 0 indicating the worst). For the VGG16 model, the initial training phase yielded an F1-score of approximately 0.90 on the test set. This indicates that VGG16 initially achieved a strong recall of almost 99% but a somewhat lower precision of around 83%, as reflected in its relatively low F1 score. After fine-tuning, VGG16's F1-score improved to about 0.93. This improvement is consistent with what we observed earlier: fine-tuning significantly increased VGG16's precision (fewer false positives) while only slightly reducing recall, resulting in a better overall balance and thus a higher F1-score.

The Inception-v3 model had a higher F1-score to start with, roughly 0.95 in the initial phase. Both high precision and high recall, over 95% for each metric. After fine-tuning, Inception's F1-score remained essentially the same, at about 0.95. The slight changes in precision and recall for Inception due to fine-tuning were negligible and cancelled out in terms of F1, which is one might expect at most a 0.001 drop, as seen, which is insignificant at this scale. This reiterates that fine-tuning did not have a notable impact on Inception's already well-optimised performance. Comparing the two models: initially, Inception's F1-score is around 0.95, which is higher than VGG16's, which is around 0.90, meaning Inception offered a better precision-recall trade-off without any adjustments. After fine-tuning, VGG16's F1-score rose to 0.93, closing much of the gap, while Inception stayed around 0.95. These values indicate that both models achieve an excellent balance between avoiding false positives and false negatives on the test set, with Inception maintaining a slight edge. Table 4 summarises the precision, recall, and F1-score for both models under initial and fine-tuned conditions, based on test-set outcomes. It provides a quick comparison of how fine-tuning shifted these key performance metrics.

Table 4
Comparative performance metrics for VGG16 and inception models: Initial vs. Fine-Tuned Training

	VGG16Initial Training	VGG16 Fine-Tuned Training	Inception Model Initial Training	Inception Model Fine-Tuned Training
Precision	0.8283	0.8988	0.9512	0.9464
Recall	0.9897	0.9564	0.9487	0.9513
F1 Score	0.9019	0.9267	0.9499	0.9488

Test-Set Performance and Statistical Analysis

The accuracy previously discussed in Figure 4 is for the training and validation phases and must not be read as an estimate of held-out performance. Table 5, therefore, reports

performance computed directly from the test-set confusion matrices of Figure 5, on the independent test partition of 624 images (234 normal, 390 pneumonia). Confidence intervals are Wilson score intervals at the 95% level, which are preferred to the normal approximation for proportions close to unity.

Table 5
Test-Set Performance for All Models (n = 624; 234 normals, 390 pneumonia).

Model	Accuracy	95% CI	Sensitivity	Specificity	F1	MCC
VGG16 (initial)	0.8654	0.8364-0.8899	0.9897	0.6581	0.9019	0.7213
VGG16 (fine-tuned)	0.9054	0.8799-0.9260	0.9564	0.8205	0.9267	0.7969
Inception-v3 (initial)	0.9375	0.9157-0.9539	0.9487	0.9188	0.9499	0.8668
Inception-v3 (fine-tuned)	0.9359	0.9139-0.9526	0.9513	0.9103	0.9488	0.8630

Fine-tuning altered VGG16 substantially, raising test accuracy from 0.8654 to 0.9054. The mechanism is a redistribution of error rather than a uniform improvement: specificity rose from 0.6581 to 0.8205 while sensitivity fell from 0.9897 to 0.9564, and the number of false negatives increased from 4 to 17. In a screening application, where a missed case carries a greater cost than a false alarm, this is a consequential trade rather than an unambiguous gain. For Inception-v3, the effect of fine-tuning was negligible: accuracy moved from 0.9375 to 0.9359 and the Matthews correlation coefficient from 0.8668 to 0.8630, both differences lying well within the confidence intervals. Inception-v3 is therefore best characterised as already converged after the feature-extraction phase alone, and the earlier statement that fine-tuning boosted its performance is not supported by the test-set data.

Comparing the two fine-tuned architectures, Inception-v3 classified 584 of 624 test images correctly, compared with 565 for VGG16, a difference of 19 images, or 3.0 percentage points. Under a two-proportion z-test, this corresponds to $z = 1.99$ and $p = 0.047$. Because both models were evaluated on the same images, their errors are positively correlated, so this unpaired test overestimates the variance of the difference. Consequently, the reported p-value is a conservative upper bound, and a paired analysis of the per-image predictions would be expected to yield a smaller value. The magnitude of the difference is nonetheless small: the 95% confidence intervals of the two fine-tuned models overlap across the range 0.914 to 0.926, and their Matthews correlation coefficients differ by 0.066. The advantage of Inception-v3 is accordingly described here as modest rather than decisive, and the case for preferring it rests as much on its lower computational cost, reported in Section 4.5, as on its classification performance.

Prediction Duration

Apart from accuracy, an important consideration for deploying deep learning models is their computational efficiency, particularly the time required to make predictions or to train, which indirectly affects how quickly updates can be made. Figure 8 presents the time required by each model to make predictions on 624 X-ray images, which represent the size of our test set, as well as the models' training times per epoch. The VGG16 model, known for its deeper architecture and larger number of parameters, exhibited higher computational demands. In the initial training phase, VGG16 required approximately 94 seconds to complete 20 epochs, averaging roughly 4.7 seconds per epoch. Upon fine-tuning, VGG16's computational load increased: it took around 110 seconds to complete 20 epochs, which takes about 5.5 seconds per epoch. This increase is expected because fine-tuning involves unfreezing more layers of the network, resulting in more computations per epoch. In terms of inference time, processing the entire test set would scale similarly with VGG16, which is relatively slow due to its complexity. In contrast, the Inception-v3 model is architecturally optimised for efficiency. During initial training, Inception completed 20 epochs in approximately 60 seconds, which is about 3.0 seconds per epoch, substantially faster than VGG16. With fine-tuning, Inception's time per epoch increased only slightly, taking about 66 seconds for 20 epochs, which takes about 3.3 seconds per epoch. This modest increase reflects the additional computations required for fine-tuning some layers; however, overall, Inception remained significantly faster than VGG16. Consequently, Inception can process a batch of 624 test images in significantly less time than VGG16. The comparison clearly shows that Inception-v3 is more computationally efficient than VGG16. VGG16's deeper network and larger fully connected layers contribute to its longer training and inference times. In a real-world deployment scenario, this means that a system using Inception could handle more images per unit time or provide results to clinicians more quickly than one using VGG16, especially on hardware with limited processing power. On the other hand, VGG16's improved accuracy after fine-tuning might be beneficial in scenarios where maximum accuracy is needed, and the infrastructure can support the heavier model. In summary, if computational resources or time are constrained, for instance, deploying an automated pneumonia detection system in a busy clinic or on edge devices such as mobile X-ray units, the Inception-v3 model is the preferable choice due to its shorter prediction times. VGG16 remains a viable model for scenarios where slightly longer processing times are acceptable in exchange for high accuracy (especially after fine-tuning). Figure 8 summarises the observed prediction durations for each model.

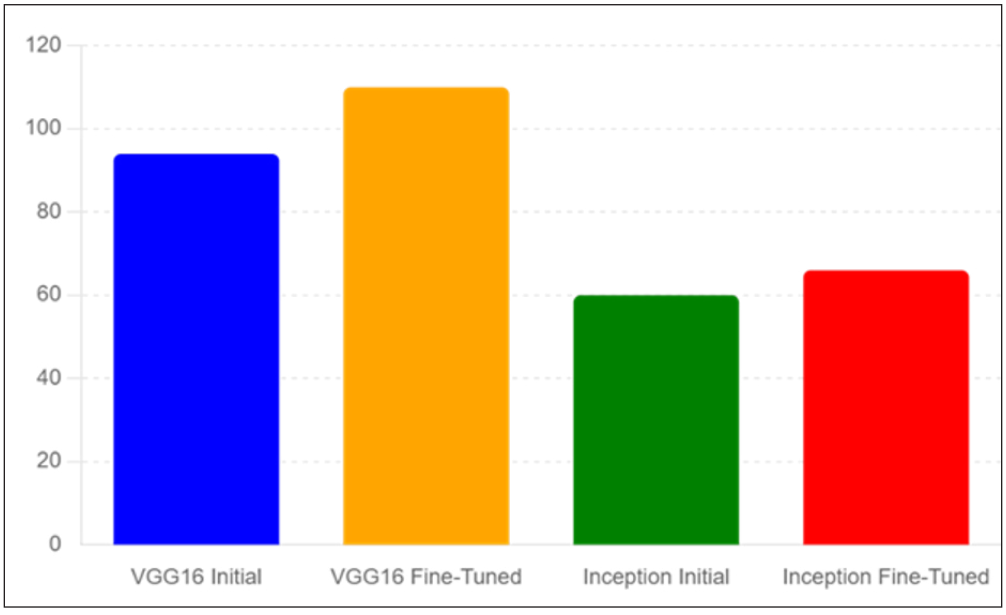


Figure 8. Prediction time duration for each model

DISCUSSION

Interpretation of the Corrected Results

The corrected test-set figures are approximately six to nine percentage points below the accuracies commonly reported for these architectures on this benchmark. The discrepancy is not attributable to the training procedure, which follows established practice, but to the reporting convention. Final training-phase accuracy exceeds 0.99 for both fine-tuned models, whereas held-out accuracy does not exceed 0.936; where a study reports the former without distinguishing it from the latter, the resulting figure overstates deployable performance by roughly the margin observed here. The same mechanism affects validation-phase reporting even more severely: a validation accuracy of 100% on the distributed 16-image partition is consistent with a true held-out accuracy of 0.905, because the partition is too small to resolve a difference that large. This is not a defect specific to the present experiments but rather a property of the benchmark distribution that is inherited by any study that adopts it without modification.

Comparison with Prior Work

Against the hybrid CNN–XGBoost model of Hedhoud et al. (2023), which reported 87% accuracy with fast inference, the fine-tuned Inception-v3 model evaluated here achieves 93.6% accuracy on a held-out partition of comparable size; however, the two studies use

different partitions, so the comparison is indicative only. More generally, the corrected figures sit at the lower end of the published range for this dataset, which is the expected consequence of reporting held-out rather than training-phase performance. Direct numerical comparison with studies that do not state which partition their headline accuracy was computed on is not meaningful, and this irreproducibility is itself one of the findings of the present work.

The Limits of Discrimination Metrics as Evidence of Clinical Validity

A high AUC is not by itself sufficient evidence of clinical validity. A classifier may achieve strong discrimination using image features that correlate with the class label but carry no diagnostic meaning, such as differences in patient positioning, exposure settings, annotation markers, or scapular and diaphragmatic borders that systematically differ between the two classes in a single-source collection. Establishing that a model attends to pulmonary findings requires spatial evidence, most obtained through class-activation mapping; recent work in medical imaging uses Grad-CAM for precisely this purpose, demonstrating that activation concentrates on lesion-relevant rather than background regions (Tülü & Kaya, 2026). No such analysis was performed in the present study; therefore, the results reported here establish discrimination performance without demonstrating that the discrimination is anatomically grounded. This is a material constraint on the conclusions that can be drawn: the accuracy figures in Table 5 quantify how often each model is correct on this test partition, but they do not license any claim about why it is correct, nor about how the models would behave on radiographs acquired under different conditions. Both architectures are accordingly presented as benchmark results rather than as candidates for clinical deployment, and interpretability analysis is identified in Section 5.5 as the first requirement for any subsequent work.

Six limitations constrain the interpretation of these results. First, the validation partition distributed with this benchmark contains only sixteen images and cannot support reliable model selection or early stopping; validation-phase as in Figure 4 should accordingly be read as descriptive of the training process rather than as estimates of generalisation. Second, the data derive from a single institution and a single paediatric population, so performance on adult radiographs, on different acquisition hardware, or under different exposure protocols is unknown, and no external validation was performed. Third, no patient-level metadata accompany the images, which prevents subgroup analysis by age, sex or comorbidity and makes it impossible to verify that images are independent at the patient level. Fourth, the decision threshold was fixed at 0.5 throughout rather than optimised against a clinically specified minimum sensitivity, which is the operating point a screening application would require. Fifth, comparisons between models were assessed using unpaired two-proportion tests and non-overlapping confidence intervals rather than paired tests on

the per-image predictions, such as McNemar's test or DeLong's test for correlated ROC curves. The unpaired approach is conservative in this setting, so the reported differences are not overstated, but a paired analysis would give a more precise estimate of the difference between architectures. Sixth, no interpretability or class-activation analysis was performed, so there is no evidence that either model bases its predictions on pulmonary findings rather than on incidental image characteristics; this is the single largest obstacle to interpreting the reported accuracies as clinically meaningful.

Future Work

Four directions follow directly from these limitations. The first and most immediate is class-activation analysis of the trained models using Grad-CAM or a comparable method, applied to a stratified sample of images correctly and incorrectly classified for both classes, to establish whether the decision evidence is anatomically plausible. Until this is done, the accuracy figures reported here cannot support any claim of diagnostic validity. Second, threshold selection should be driven by a specified minimum sensitivity, with specificity and the resulting review burden reported at that operating point, rather than by the default value of 0.5. Third, external validation in an independent cohort is a precondition for any claim of clinical utility, since the present results characterise generalisation only within a single collection. Fourth, constraining the classifier to anatomically relevant regions via lung-field segmentation as a preprocessing step would directly test whether the performance reported here depends on extrapulmonary image content and could be evaluated by re-running the present protocol on masked inputs. Ensemble combination of the two architectures, model compression for deployment on lower-end hardware, and extension to the multi-class setting distinguishing bacterial from viral pneumonia or screening for tuberculosis and COVID-19 remain worthwhile but secondary to these four.

CONCLUSION

This study reports a comparative evaluation of VGG16 and Inception-v3 for pneumonia classification on chest radiographs, conducted and reported entirely on held-out data. The fine-tuned Inception-v3 model achieved a test accuracy of 0.936 (95% CI 0.914–0.953), compared with 0.905 (0.880–0.926) for the fine-tuned VGG16 model, and did so at approximately 60% of the training time per epoch, making it the preferable choice when computational resources or response time are constrained. Fine-tuning benefited VGG16 by raising specificity from 0.658 to 0.821 at the cost of a reduction in sensitivity from 0.990 to 0.956 and produced no measurable change in Inception-v3. These held-out figures are materially below the near-perfect accuracies frequently reported for the same architectures on the same benchmark; the difference is attributable to the reporting of training-phase rather than test-phase performance, compounded by a distributed validation partition too

small to support model selection. Neither architecture is therefore ready for deployment as a screening aid without external validation, calibration of the decision threshold against a clinical sensitivity requirement, and verification that classification is driven by pulmonary findings. The contribution of this work lies in establishing a correctly reported reference point on a benchmark where published performance has drifted above what the underlying data support.

ACKNOWLEDGEMENT

Authors acknowledge the Universiti Teknologi MARA in collaboration with Universitas Mataram, Fakultas Teknik, Program Studi Teknik Informatika, Indonesia, under direct grant no 600-TNCPI/PBT 5/3 INT (064/2024).

REFERENCES

Abd Rahman, N. N., Yusoff, M., & Kassim, M. (2022). HOG and Haralick feature extractions with machine learning methods for COVID-19 X-ray image classification. *International Journal of Engineering Trends and Technology, 70*(10), 8-17. https://doi.org/10.14445/22315381/IJETT-V70I10P202

Abdulahi, A. T., Ogundokun, R. O., Adenike, A. R., Shah, M. A., & Ahmed, Y. K. (2024). PulmoNet: A novel deep learning based pulmonary diseases detection model. *BMC Medical Imaging, 24*(1), Article 51. https://doi.org/10.1186/s12880-024-01227-2

Aljuaid, H., Adlan, H., Alkebsi, B., Alfurhood, B. S., Liotta, A., & Cavallaro, L. (2026). An experimental comparison of deep learning models for pneumonia classification from chest X-ray images. *Biomedical Signal Processing and Control, 112*, 108742. https://doi.org/10.1016/j.bspc.2025.108742

Azis, S. F. S., Shahbudin, S., Kassim, M., Mohamad, R., & Rahman, F. Y. A. (2022). Photovoltaic module defects classification analysis using DenseNet architecture. In *2022 IEEE Symposium on Industrial Electronics & Applications (ISIEA)* (pp. 1-5). IEEE. https://doi.org/10.1109/ISIEA54517.2022.9873681

Basubrin, O. (2025). Current status and future of artificial intelligence in medicine. *Cureus, 17*(1), Article e77561. https://doi.org/10.7759/cureus.77561

Boniol, M., Kunjumen, T., Nair, T. S., Siyam, A., Campbell, J., & Diallo, K. (2022). The global health workforce stock and distribution in 2020 and 2030: A threat to equity and 'universal' health coverage? *BMJ Global Health, 7*(6), e009316. https://doi.org/10.1136/bmjgh-2022-009316

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition, 30*(7), 1145-1159. https://doi.org/10.1016/S0031-3203(96)00142-2

Davis, J., & Goadrich, M. (2006). The relationship between precision–recall and ROC curves. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 233-240). Association for Computing Machinery. https://doi.org/10.1145/1143844.1143874

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248-255). IEEE. https://doi.org/10.1109/CVPR.2009.5206848

El Houby, E. M. F. (2024). COVID-19 detection from chest X-ray images using transfer learning. *Scientific Reports, 14*(1), Article 11639. https://doi.org/10.1038/s41598-024-61693-0

Gulshan, V., Rajan, R. P., Widner, K., Wu, D., Wubbels, P., Rhodes, T., Whitehouse, K., Coram, M., Corrado, G., Ramasamy, K., Raman, R., Peng, L., & Webster, D. R. (2019). Performance of a deep-learning algorithm vs manual grading for detecting diabetic retinopathy in India. *JAMA Ophthalmology, 137*(9), 987-993. https://doi.org/10.1001/jamaophthalmol.2019.2004

Hedhoud, Y., Mekhaznia, T., & Amroune, M. (2023). An improvement of the CNN-XGboost model for pneumonia disease classification. *Polish Journal of Radiology, 88*, e483-e493. https://doi.org/10.5114/pjr.2023.132533

Hu, Q., Jiang, J., Xu, H., & Kassim, M. (2024). IMNE: Maximizing influence through deep learning-based node embedding in social network. *Swarm and Evolutionary Computation, 88*, Article 101609. https://doi.org/10.1016/j.swevo.2024.101609

JTIPTJ. (2021). *Chest X-ray (pneumonia, COVID-19, tuberculosis)* [Data set]. Kaggle. https://www.kaggle.com/datasets/jtiptj/chest-xray-pneumoniacovid19tuberculosis

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., . . . Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell, 172*(5), 1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1412.6980

Kunjumon, A., V, S., & Nair, J. J. (2026). A hybrid deep learning framework for improved pneumonia detection in chest X-ray images. *Procedia Computer Science, 282*, 1043-1055. https://doi.org/10.1016/j.procs.2026.05.146

Lim, W. S. (2022). Pneumonia—overview. In S. M. Janes (Ed.), *Encyclopedia of respiratory medicine* (2nd ed., pp. 185–197). Academic Press. https://doi.org/10.1016/B978-0-12-801238-3.11636-8

Mahalle, V. S., Kandoi, N. M., & Patil, S. B. (2024). Transfer learning by fine-tuning pre-trained convolutional neural network architectures for image recognition. In D. Mishra, X. S. Yang, A. Unal, & D. S. Jat (Eds.), *Data science and big data analytics* (pp. 273-287). Springer. https://doi.org/10.1007/978-981-99-9179-2_21

Müller, D., Soto-Rey, I., & Kramer, F. (2022). Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes, 15*(1), Article 210. https://doi.org/10.1186/s13104-022-06096-y

Namdar, K., Haider, M. A., & Khalvati, F. (2021). A modified AUC for training convolutional neural networks: Taking confidence into account. *Frontiers in Artificial Intelligence, 4*, Article 582928. https://doi.org/10.3389/frai.2021.582928

Prechelt, L. (1998). Automatic early stopping using cross validation: Quantifying the criteria. *Neural Networks, 11*(4), 761–767. https://doi.org/10.1016/S0893-6080(98)00010-0

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). *Transfusion: Understanding transfer learning with applications to medical imaging* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1902.07208

Sharma, V., Sandhu, R., Kaur, J., Kaur, H., Chouhan, S., & Rakhra, M. (2024). A novel CNN-based approach for pneumonia detection using Keras and TensorFlow. In 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT). IEEE. https://doi.org/10.1109/ICEECT61758.2024.10739031

Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging, 35*(5), 1285-1298. https://doi.org/10.1109/TMI.2016.2528162

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*(1), Article 60. https://doi.org/10.1186/s40537-019-0197-0

Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1409.1556

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*(56), 1929–1958. https://jmlr.org/papers/v15/srivastava14a.html

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2818-2826). IEEE. https://doi.org/10.1109/CVPR.2016.308

Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In K. Chaudhuri & R. Salakhutdinov (Eds.), *Proceedings of the 36th International Conference on Machine Learning* (Vol. 97, pp. 6105-6114). PMLR. https://proceedings.mlr.press/v97/tan19a.html

Tülü, Ç. N., & Kaya, Y. (2026). Convolutional Fourier analysis network (CONV-FAN-POX): A novel time–frequency approach for medical image analysis. *Biomedical Signal Processing and Control, 117*, 109698. https://doi.org/10.1016/j.bspc.2026.109698

UNICEF, United Nations Children's Fund. (2024). *Pneumonia in children statistics.* https://data.unicef.org/topic/child-health/pneumonia/

WHO, World Health Organization. (2025). *Pneumonia.* https://www.who.int/health-topics/pneumonia/

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging, 9*(4), 611-629. https://doi.org/10.1007/s13244-018-0639-9

Pertanika J. Sci. & Technol. 34 (5): 2597 - 2618 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Artificial Neural Network-Based UHF Partial Discharge Localisation in Mineral Oil with Datasets Based on TDOA

Zhi Yang Lim[1], Norhafiz Azis[1]*, Hafiz Rashidi Ramli[2], Mohd Amran Mohd Radzi[1], Nor Mohd Haziq Norsahperi[2], and Azrul Mohd Ariffin[3]

[1]Advanced Lightning, Power and Energy Research Centre (ALPER), Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2]Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[3]Department of Electrical and Electronics Engineering, College of Engineering, Institute of Power Engineering, Universiti Tenaga Nasional (UNITEN), Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia

ABSTRACT

This study presents a dataset generation procedure to train artificial neural networks (ANNs) for ultra-high frequency (UHF) partial discharge (PD) localisation in an oil-filled tank. The datasets were generated using the Time Difference of Arrival (TDOA) equations and augmented with samples containing random deviations. The PD localisation performance of the ANNs trained on the datasets was compared with that of the artificial bee colony (ABC), bat algorithm (BA), genetic algorithm (GA), particle swarm optimisation (PSO), and a non-iterative algorithm, based on the TDOAs of experimental UHF PD signals. An experiment was conducted in which UHF PD signals were captured with 4 patch-type wideband antennas and denoised using the variational mode decomposition (VMD) method. TDOAs were computed from both the original and denoised UHF PD signals through the First Peak Method (FPM) and Cumulative Energy Method (CEM). For the original UHF PD signals, the FPM produces TDOAs with lower mean absolute errors (MAEs) than the CEM. In contrast, the CEM produces TDOAs with lower MAEs than the FPM for the denoised UHF PD signals. TDOAs with lower MAEs generally lead to lower PD localisation errors. The trained ANNs consistently achieve the highest PD localisation accuracy, with average computation times ranging from 36.75 ms to 43.17 ms.

Keywords: Artificial neural network (ANN), mineral oil, partial discharge (PD) localisation, Time Difference of Arrival (TDOA), ultra-high frequency (UHF)

ARTICLE INFO

Article history:
Received: 22 May 2026
Accepted: 17 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.03

gs68246@student.upm.edu.my (Zhi Yang Lim)
norhafiz@upm.edu.my (Norhafiz Azis)
hrhr@upm.edu.my (Hafiz Rashidi Ramli)
amranmr@upm.edu.my (Mohd Amran Mohd Radzi)
nmhaziq@upm.edu.my (Nor Mohd Haziq Norsahperi)
Azrula@uniten.edu.my (Azrul Mohd Ariffin)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

The continuous operation of a power transformer relies on the integrity of its insulation system. Statistics indicate that insulation breakdown is a primary cause of transformer failures (Pattanadech & Muhr, 2016). Prolonged and unaddressed partial discharge (PD) within the insulation is a major contributor to such breakdowns (Chan et al., 2023). Therefore, PD diagnostics are among the most common techniques for monitoring transformer insulation condition and supporting predictive maintenance (AJ et al., 2018). An important aspect of PD diagnostics is the localisation of the PD sources within transformers (Zaki et al., 2020). This information can facilitate the formulation of targeted maintenance strategies for the transformers. Generally, PD localisation involves PD signal acquisition, PD signal denoising, Time Difference of Arrival (TDOA) computation, and application of the PD localisation algorithm (Desai & Sarathi, 2018).

The ultra-high frequency (UHF) method is a common technique for PD signal acquisition that detects electromagnetic signals emitted by PD sources in the frequency range of 0.3 to 3 GHz (Boya-Lara et al., 2022). The capability for online PD detection and source localisation represents a major advantage of the UHF method as compared with conventional electrical techniques based on IEC 60270 (Jahangir et al., 2017). Although transformer tanks function as a shield that increases the signal-to-noise ratio (SNR) of UHF PD signals (Jahangir et al., 2017), field measurements can still be affected by various types of noises, such as white noise, impulsive noise, and periodic narrowband interference (Boya-Lara et al., 2022). Therefore, effective denoising techniques are essential for improving the accuracy of the UHF method and minimising the risk of false diagnosis. Multivariate wavelet (MW) denoising, empirical mode decomposition (EMD), and variational mode decomposition (VMD) are common denoising methods for UHF PD signals (Long et al., 2021). Studies report that VMD can achieve better denoising performance at lower computational cost than EMD and wavelet-based methods (Long et al., 2021; Zhang et al., 2019).

Typical techniques for TDOA computation include the First Peak Method (FPM), Cumulative Energy Method (CEM), and Cross-Correlation Method (CCM). FPM determines the arrival time of a PD signal based on the first local peak whose absolute magnitude exceeds a predefined threshold (Hashim et al., 2023; Sinaga et al., 2012). In CEM, the arrival time of a PD signal is determined from the knee point of its cumulative energy curve (Kundu et al., 2009). For both the FPM and the CEM, TDOAs are calculated from the differences between the arrival times at 2 sensors. CCM computes the TDOA from the maximum point of the cross-correlation curve between 2 PD signals, assuming that signals from the same PD source share similar waveform patterns (Sinaga et al., 2012). Different TDOA computation methods may yield different TDOAs for the same PD signal (Kakeeto et al., 2008; Sinaga et al., 2012), and the performance of each method varies

across cases (Long et al., 2021). Since variations in TDOAs can affect the accuracy of PD localisation (Desai & Sarathi, 2018), the TDOA computation technique serves as a crucial factor in the overall PD localisation framework.

Over the years, research on TDOA-based PD localisation algorithms progresses from iterative approaches, such as the least-squares and Newton's algorithms, to optimisation algorithms such as the genetic algorithm (GA), particle swarm optimisation (PSO), and quantum genetic algorithm (QGA) (Lim et al., 2025). These optimisation algorithms do not require initial guesses and generally achieve higher localisation accuracy than the iterative approaches (Rathod et al., 2022). A non-iterative algorithm that provides faster estimation is also proposed by Kundu et al. (2009). To improve localisation accuracy, recent research explores the use of artificial intelligence (AI) techniques for PD localisation (Chan et al., 2023). For example, an artificial neural network (ANN) method is proposed by Taha et al. (2020) to determine optimal sensor placements on the transformer tank wall to enhance PD localisation accuracy. Hashim et al. (2023) investigate the application of ANN and adaptive neuro-fuzzy inference system (ANFIS) models for acoustic PD localisation. The study by Khodaveisi et al. (2024) evaluates several AI techniques, such as support vector machine (SVM) and backpropagation neural network (BPNN), for UHF PD localisation using a time-reversal approach.

For PD localisation, the performance of ANN depends on the training dataset. The datasets can be prepared from PD measurements conducted on the transformer (Hashim et al., 2023). However, this requires repeated on-site PD measurements (Zaki et al., 2020), and the actual PD location remains unknown unless the transformer tank is dismantled. These constraints might limit the size of the generated dataset, thereby affecting the performance of the trained ANNs (Basheer & Hajmeer, 2000). An alternative approach to dataset generation involves simulating PD signal propagation within the transformer tank, as demonstrated by Khodaveisi et al. (2024). This approach demands considerable time, and access to in-service transformer parameters is sometimes restricted (Chan et al., 2023). Therefore, this study adopts a different method for dataset generation, in which dataset samples are computed using the TDOA equations. This method avoids the need for large-scale on-site PD data collection and requires less time than the simulation-based approach. TDOA samples with random deviations are included to augment the datasets, as practical TDOAs often differ from the theoretical TDOAs due to various intricate factors, such as the complex propagation paths within the transformer tank and noise at the measurement sites (Ghosh et al., 2017; Jangjoo et al., 2023). The objective of this study is to examine the UHF PD localisation performance of ANNs trained on datasets generated by the proposed method. An experiment is conducted with a steel tank filled with mineral oil. UHF PD signals are acquired at 2 distinct PD locations within the tank, and the VMD method is applied to denoise the signals. The FPM and CEM are employed to compute

the TDOAs from both the original and denoised UHF PD signals. PD localisation is then performed using the computed TDOAs. The localisation performance of the trained ANNs is compared with the artificial bee colony (ABC), bat algorithm (BA), genetic algorithm (GA), particle swarm optimisation (PSO), and a non-iterative algorithm in terms of distance error and computation time.

METHODOLOGY

Mathematical Model of TDOA-Based PD Localisation

The test tank used in this study measured 1.00 m in length, 0.30 m in width, and 0.25 m in height, as shown in Figure 1. A 3-dimensional (3D) coordinate system was defined for the test tank, with the lower-left corner at the back set as the origin, O (0, 0, 0). The coordinates of the 4 UHF sensors were denoted as (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), and (x_4, y_4, z_4), while the PD source location was represented by (x, y, z). Based on this setup, the TDOA-based PD localisation model employed in this study was mathematically described by Equations 1-4 (Chan et al., 2023; Desai & Sarathi, 2018). In these equations, v refers to the propagation velocity of UHF signals in mineral oil, which is 2×10^8 m/s (CIGRE Working Group D1.29, 2017; Sinaga et al., 2012). The term T_1 denotes the time of arrival (TOA) of the UHF PD signal at UHF Sensor 1 (S_1), which served as the reference sensor. The TDOAs, τ_{21}, τ_{31}, and τ_{41}, were computed as the differences between the TOAs at UHF Sensors 2, 3 and 4, and the TOA at the reference sensor.

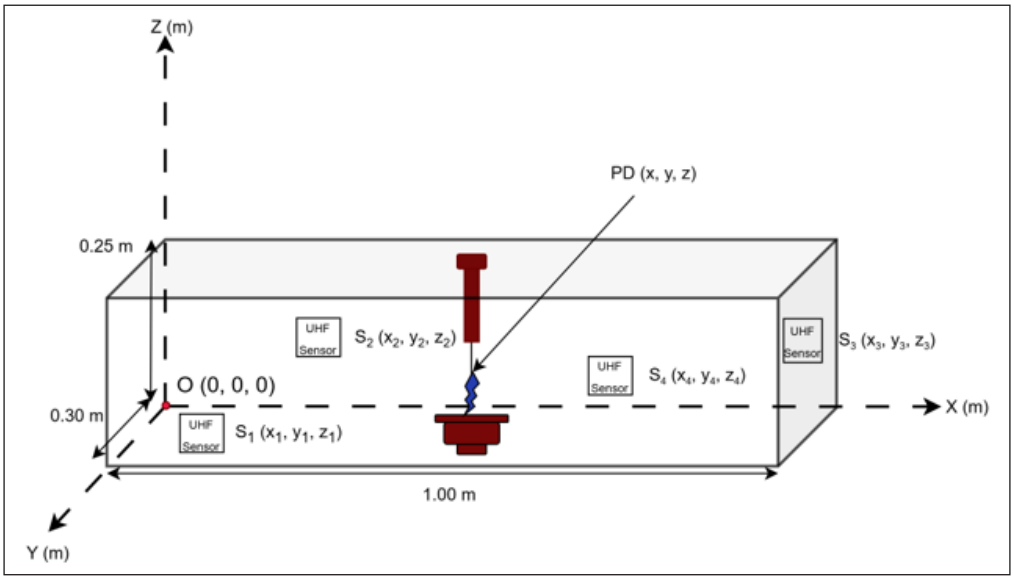


Figure 1. The test tank and the adopted 3D coordinate system

$$(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = (v \times T_1)^2 \quad [1]$$

$$(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = [v \times (T_1 + \tau_{21})]^2 \quad [2]$$

$$(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2 = [v \times (T_1 + \tau_{31})]^2 \quad [3]$$

$$(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2 = [v \times (T_1 + \tau_{41})]^2 \quad [4]$$

Experimental Setup

A laboratory-based experiment was set up for UHF PD signal acquisition in accordance with IEC 60270 and IEC TS 62478 (Hashim et al., 2023), as illustrated in Figure 2. The test tank used in the setup was made of steel, with 4 side walls 1 mm thick. A total of 65 litres of mineral oil, with a viscosity of 9.6 mm^2/s at 40°C and a density of 0.875 kg/dm^3 at 20°C, was filtered and dried based on the procedures outlined in (Hashim et al., 2023; Mohd Hashim et al., 2023). The treated oil was then transferred into the test tank. A needle with a tip radius of 3 μm and a copper plane electrode with a diameter of 50 mm and an edge radius of 3 mm were used to establish a needle-to-plane electrode configuration (Hashim et al., 2023). The electrode configuration was placed at 2 distinct locations within the test

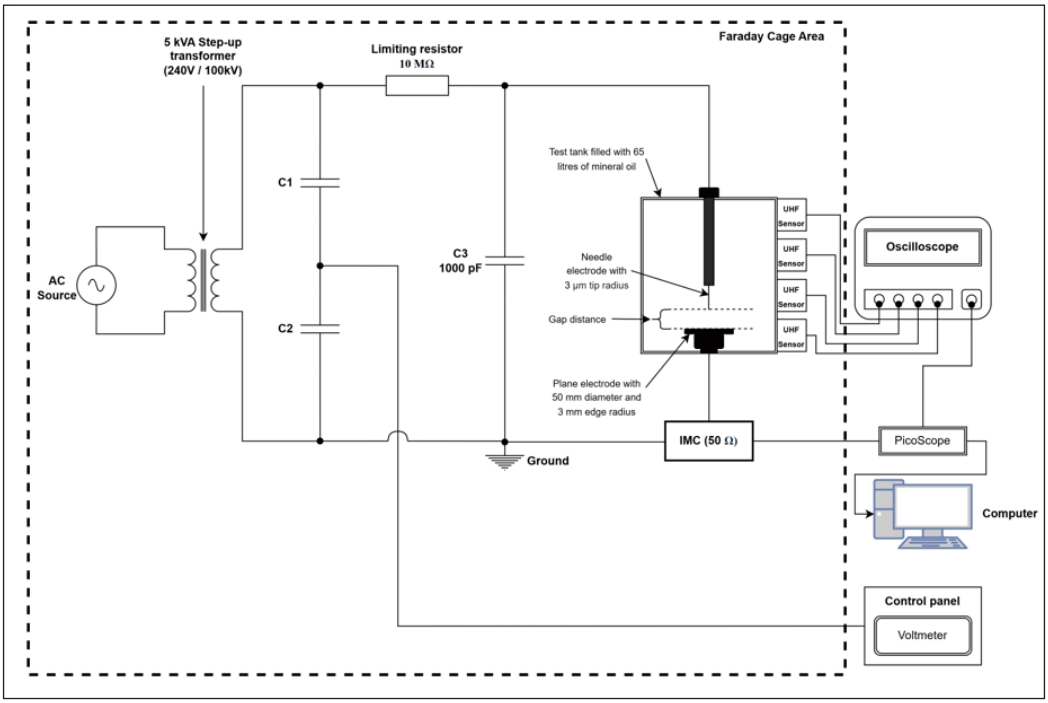


Figure 2. Experimental setup for UHF PD signal acquisition

tank, one at a time, to create 2 PD source locations. Based on the coordinate system depicted in Figure 1, the tip of the needle electrode was positioned at (0.500, 0.150, 0.080) m for PD Location 1 and (0.250, 0.100, 0.083) m for PD Location 2. The gap distances between the needle and plane electrodes for the 2 PD locations were fixed at 35 mm and 38 mm, respectively. In total, 4 patch-type wideband antennas with a frequency range of 0.7 to 2.7 GHz served as the UHF sensors and were connected to a LeCroy WaveMaster 830Zi oscilloscope to capture the UHF PD signals emitted by the PD source. The oscilloscope's measurement bandwidth and sampling rate were set to 3 GHz and 40 GS/s, respectively (Sinaga et al., 2012). At the same time, an impedance matching circuit (IMC) with 50 Ω input and output impedances, along with a PicoScope 6804E, was used to perform electrical PD measurements. These electrical PD signals were used to verify the occurrence of PD and were excluded from further analysis. A total of 20 random combinations of sensor coordinates were used to acquire UHF PD signals at each PD location. Each combination comprised the coordinates of the 4 UHF sensors positioned at different locations on the 4 side walls of the test tank.

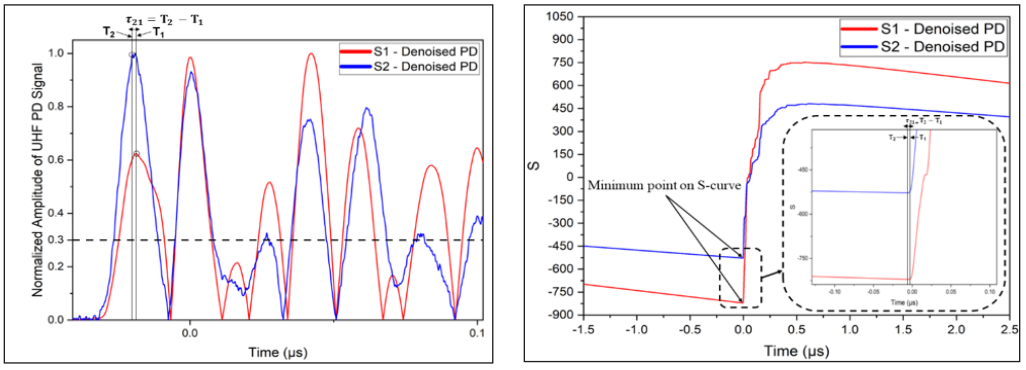
UHF PD Signal Denoising and TDOA Computation

The VMD algorithm was applied in accordance with Long et al. (2017) and Zhang et al. (2019) to denoise the UHF PD signals obtained from the experiment. The quadratic penalty term, α, and the time-step of the dual ascent, τ, were set to 2,000 and 0, respectively (Zhang et al., 2019). Each of the UHF PD signals was decomposed into several band-limited intrinsic mode functions (BLIMFs), which are amplitude-modulated and frequency-modulated signal components within specific frequency bands (Zhang et al., 2019). The total bandwidth of the BLIMFs was minimised while the sum of all BLIMFs exactly reproduced the original PD signal. The number of decomposition modes in the VMD, K, was adaptively selected based on the mode-mixing judgement algorithm proposed by Zhang et al. (2019). Initially, K was set to 2 and increased by 1 if the frequency spectrum of any BLIMF contained more than one active component. To avoid excessive computation time, the maximum K was limited to 10. The segmentation length of the frequency spectrum, L_1, and the threshold for amplitude criterion, g, in the mode-mixing judgement algorithm were set to 10,000 and 0.5, respectively (Zhang et al., 2019). Equation 5 was employed to calculate the kurtosis of each of the BLIMFs (Zhang et al., 2019), where X represents the BLIMF, $x(i)$ denotes the ith sample of the BLIMF, $\bar{x}$ is the mean of the BLIMF, and n is the total number of samples in the BLIMF. The calculated kurtoses were subsequently used to identify the effective BLIMFs that contain significant PD signal components. Since kurtosis is sensitive to sudden changes in the BLIMF waveform, BLIMFs associated with PD generally exhibit higher kurtoses than the noise-dominated BLIMFs (Zhang et al., 2019). This is because UHF PD signals are characterised by transient pulses with a very short rise

time followed by damped oscillations (Zhang et al., 2019). Therefore, based on the criterion proposed by Long et al. (2017) and Zhang et al. (2019), BLIMFs with kurtoses greater than 10 were considered effective and were used to reconstruct the denoised UHF PD signal.

$$Kurtosis\ (X) = \frac{\frac{1}{n}\sum_{i=1}^{n}(x(i) - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x(i) - \bar{x})^2\right)^2} \quad [5]$$

The FPM and CEM were subsequently employed to compute the TDOAs from both the original and denoised UHF PD signals. In the FPM, each UHF PD signal was first converted to its absolute amplitude and then normalised to its maximum value (Sinaga et al., 2012). A magnitude threshold of 0.3 was used in this study, and the TOA was extracted at the instant where a local peak first exceeded this threshold, as depicted in Figure 3(a). The TDOA was calculated as the TOA at the sensor of interest minus the TOA at the reference sensor, as shown in Figure 3(a). For the CEM, the cumulative energy curves of the original and denoised UHF PD signals were utilised to compute the variable *S* based on the equation provided by Kundu et al. (2009). The time at the minimum point of the resulting *S* curve was identified as the TOA of the respective UHF PD signal (Kundu et al., 2009), as illustrated in Figure 3(b). The TDOAs of the CEM were computed in the same manner as in the FPM.



(a) (b)

Figure 3. TDOA determination based on (a) FPM and (b) CEM

Artificial Neural Networks (ANNs) Training

The procedures used to generate the datasets for ANN training are summarised in Figure 4. First, the test tank was virtually divided into a grid with 5 mm, 2 mm, and 2 mm spacing along the *x*-, *y*-, and *z*- axes, respectively. This grid resulted in a total of 3,824,226 possible

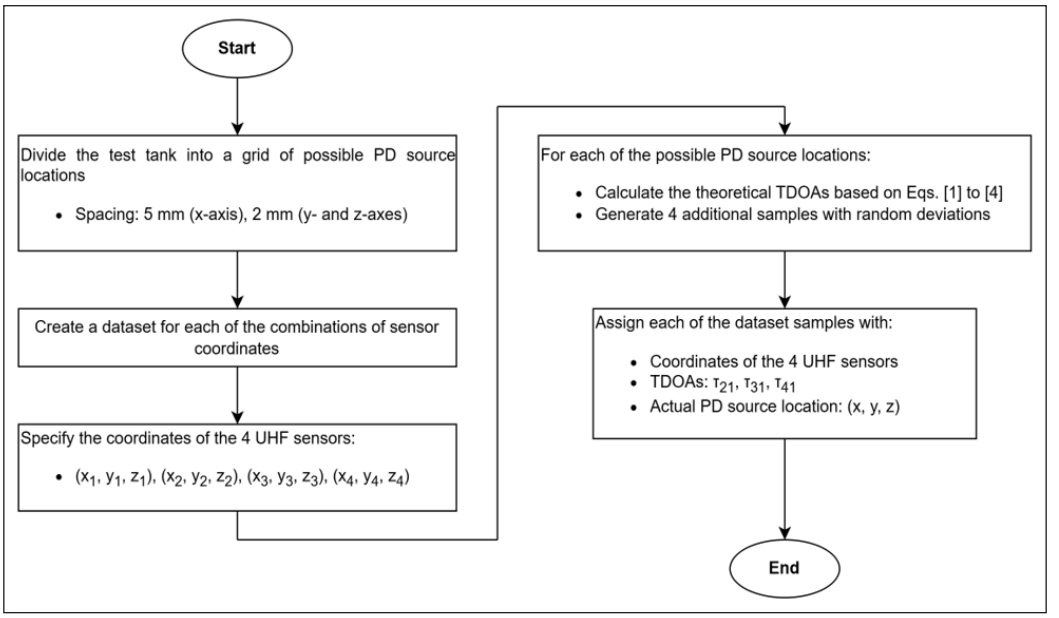


Figure 4. The dataset generation procedures based on the TDOA equations

PD source locations, as illustrated in Figure 5. For each possible PD source location, the theoretical TDOAs were calculated using Equations 1-4 to form 1 sample in the dataset. To account for factors that could affect the precise determination of TDOAs, such as the sophisticated propagation path within the transformer tank, noise at the measurement site, and other uncertainties (Ghosh et al., 2017; Jangjoo et al., 2023), 4 additional samples were created for each of the possible PD source locations. Each additional sample was generated by adding random deviations, drawn from a normal distribution with a mean of 0 ns and a standard deviation of 0.3 ns, to the theoretical TDOAs. This approach aligns with Zaki et al. (2020), who reported that TDOA determination errors can be modelled with a normal distribution. These steps were repeated for each sensor coordinate combination considered in this study, yielding a total of 20 datasets.

The adaptive moment estimation (Adam) algorithm was employed to train the ANNs (Kingma & Ba, 2014), with the maximum number of epochs capped at 300 and the mini-batch size set to 1,912,113 samples. The initial learning rate, η, the exponential decay rate for the first moment, β_1, and the exponential decay rate for the second moment, β_2, were set to 0.001, 0.9, and 0.999, respectively (Kingma & Ba, 2014). The architecture of the ANNs was determined by the steps shown in Figure 6. The training inputs were normalised by a constant of 5.37 ns, which represents the theoretical maximum TDOA within the test tank. This constant was obtained by dividing the tank diagonal length by the propagation velocity of UHF PD signals in oil. For each of the hidden layers, the smallest number of

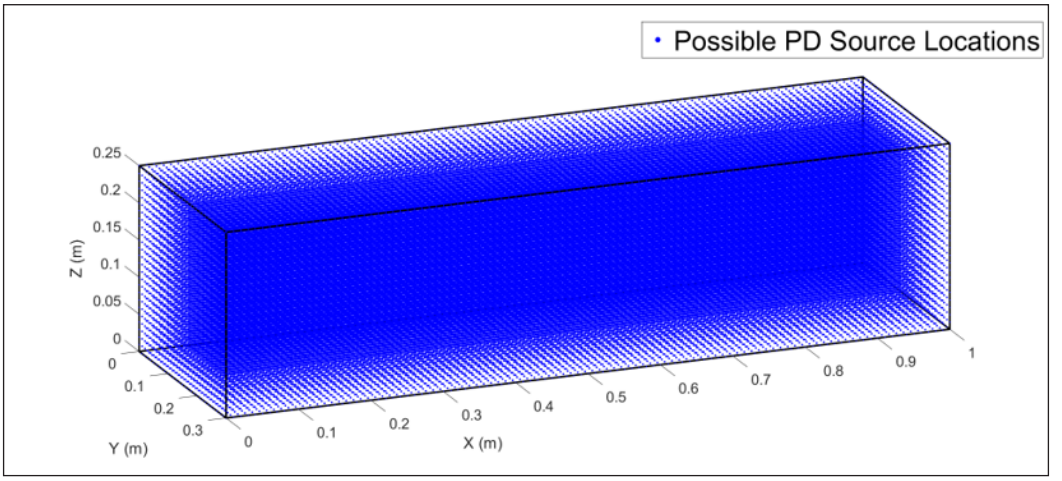


Figure 5. The virtual grid of possible PD source locations within the test tank

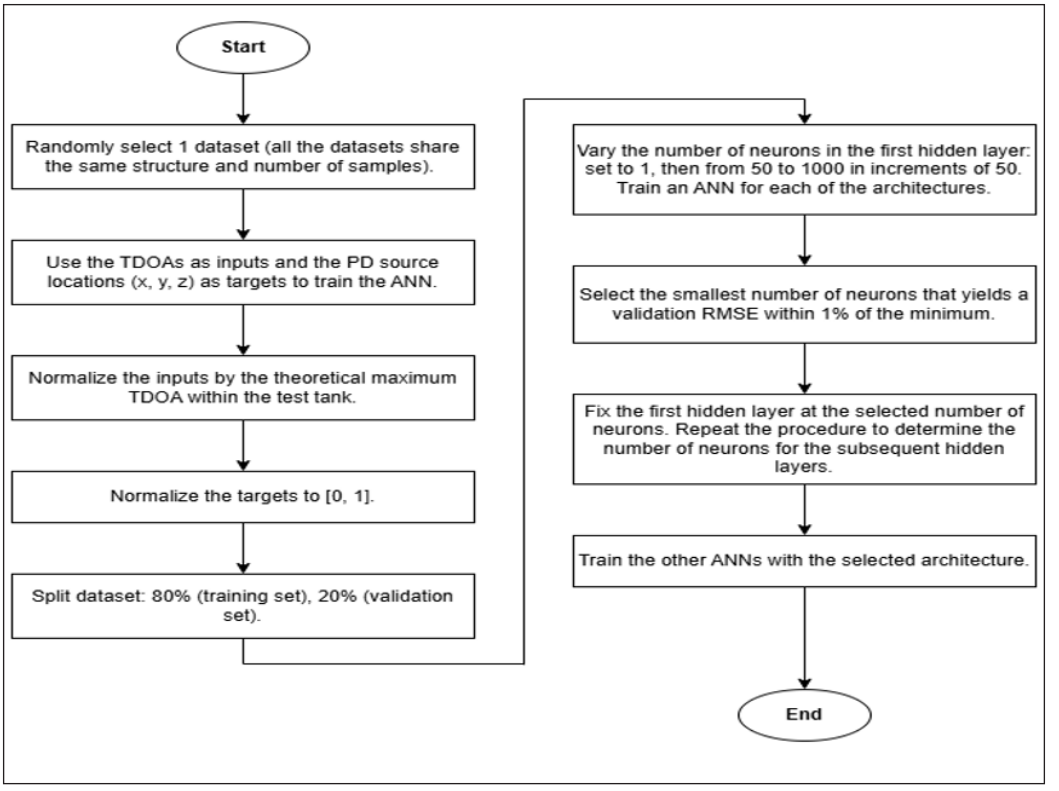
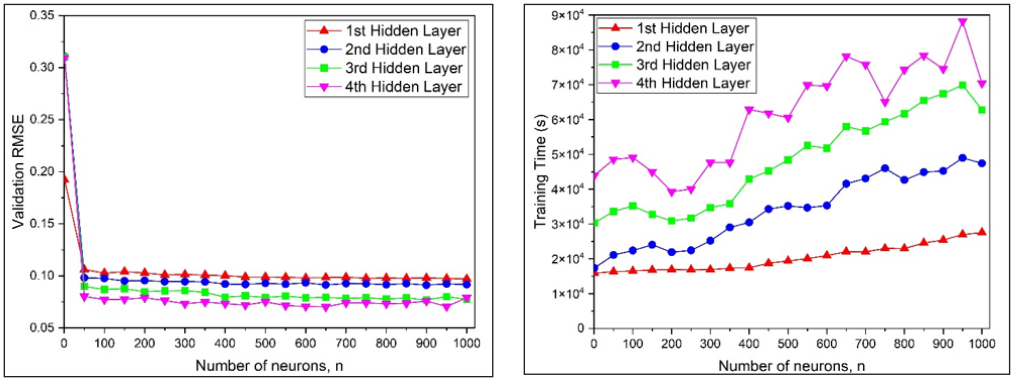


Figure 6. The steps used to determine the architecture for the ANNs in this study

neurons that produced a validation RMSE no more than 1% above the minimum RMSE was selected. For instance, the minimum RMSE was achieved by 1,000 neurons in the first hidden layer, at 0.10. However, the number of neurons in the first hidden layer was set to 600 because this configuration resulted in a 24.02% shorter training time with only a 0.99% higher RMSE, as depicted in Figure 7.



(a) (b)

Figure 7. (a) Validation RMSE and (b) training time for all the tested ANN architectures

According to the steps in Figure 6, the selected ANN architecture comprised 3 hidden layers with 600, 400, and 400 neurons, respectively, as illustrated in Figure 8. The rectified linear unit (ReLU) activation function was applied to all the hidden-layer neurons (Khodaveisi et al., 2024). For PD localisation, each of the ANNs was configured with 3 input neurons for the TDOAs and 3 output neurons for the *x*-, *y*-, and *z*- coordinates of the estimated PD location. As indicated in Figure 7, the architecture was chosen to balance ANN performance and training time. Thus, it may not represent the optimal design. The validation set performance of the trained ANNs was assessed by the root mean square error (RMSE) and Pearson correlation coefficients (R_x, R_y, R_z) given in Equations 6, 7a, 7b and 7c (Khodaveisi et al., 2024). In these equations, the predicted PD location for the ith validation sample was expressed as $(x_{p,i}, y_{p,i}, z_{p,i})$, while the corresponding actual PD location was denoted by $(x_{c,i}, y_{c,i}, z_{c,i})$. The terms $\overline{x_p}$, $\overline{y_p}$, and $\overline{z_p}$ represented the mean of the predicted PD source coordinates, whereas $\overline{x_c}$, $\overline{y_c}$, and $\overline{z_c}$ denoted the mean of the actual PD source coordinates, computed across *n* samples, where *n* is the total number of validation samples.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\left(x_{p,i} - x_{c,i}\right)^2 + \left(y_{p,i} - y_{c,i}\right)^2 + \left(z_{p,i} - z_{c,i}\right)^2\right]} \qquad [6]$$

$$R_x = \frac{\sum_{i=1}^{n}(x_{p,i}-\overline{x_p})(x_{c,i}-\overline{x_c})}{\sqrt{\left[\sum_{i=1}^{n}(x_{p,i}-\overline{x_p})^2\right]\left[\sum_{i=1}^{n}(x_{c,i}-\overline{x_c})^2\right]}} \quad [7a]$$

$$R_y = \frac{\sum_{i=1}^{n}(y_{p,i}-\overline{y_p})(y_{c,i}-\overline{y_c})}{\sqrt{\left[\sum_{i=1}^{n}(y_{p,i}-\overline{y_p})^2\right]\left[\sum_{i=1}^{n}(y_{c,i}-\overline{y_c})^2\right]}} \quad [7b]$$

$$R_z = \frac{\sum_{i=1}^{n}(z_{p,i}-\overline{z_p})(z_{c,i}-\overline{z_c})}{\sqrt{\left[\sum_{i=1}^{n}(z_{p,i}-\overline{z_p})^2\right]\left[\sum_{i=1}^{n}(z_{c,i}-\overline{z_c})^2\right]}} \quad [7c]$$

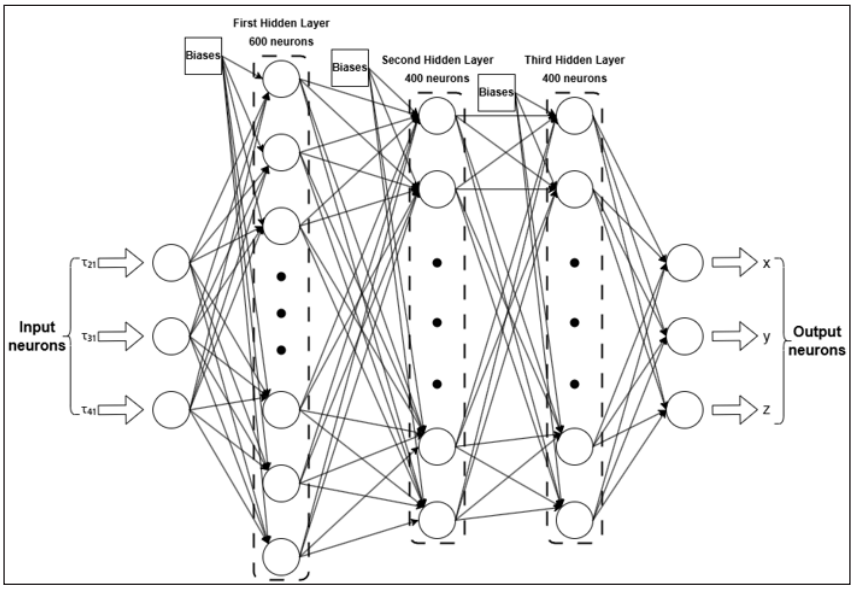


Figure 8. The ANN architecture used in this study

PD Localisation Performance Analysis

The 4 TDOA scenarios derived from the experimental UHF PD signals were used to evaluate the PD localisation performance of the trained ANNs. For simplicity, the TDOAs computed from the original UHF PD signals through the FPM and CEM were referred to as Scenarios 1 and 2, respectively, whereas the TDOAs computed from the denoised UHF PD signals through the FPM and CEM were designated as Scenarios 3 and 4, respectively. The artificial bee colony (ABC) (Lim et al., 2025), bat algorithm (BA) (Chakravarthi et al., 2017), genetic algorithm (GA) (Cintra Veloso et al., 2006), particle swarm optimisation (PSO) (Tang et al., 2008), which are optimisation algorithms, along with the non-iterative algorithm proposed by Kundu et al. (2009), were employed to benchmark the PD

localisation performance of the ANNs. The parameters and implementation procedures of the optimisation algorithms were adopted from Lim et al. (2025). Equation 8 was used to calculate the distance error between the estimated and actual PD locations (Desai & Sarathi, 2018; Lim et al., 2025). The average distance error was computed as the mean across all combinations of sensor coordinates for each PD location. The average computation time for each PD localisation algorithm was also recorded for comparison.

$$\Delta D = \sqrt{(x_{actual} - x_{estimated})^2 + (y_{actual} - y_{estimated})^2 + (z_{actual} - z_{estimated})^2} \quad [8]$$

RESULTS AND DISCUSSION

UHF PD Signal Acquisition and VMD Denoising

An example of the electrical and UHF PD signals obtained from the experiment is shown in Figure 9. A slight time delay of a few nanoseconds (ns) is observed between the first peaks of the UHF PD signals and the first peak of the electrical PD signal. This delay is likely due to the propagation of the PD signals as electromagnetic waves from the PD source to the UHF sensors. Electrical PD measurements are carried out with a calibrated IMC, where an apparent charge of 100 pC corresponds to an output voltage of 258 mV ± 20 mV. The electrical pulse in Figure 9 exceeds this voltage level and is therefore identified as a PD pulse. The close alignment between the electrical and UHF pulses suggests that both signals originate from the same source and therefore supports the identification of the UHF pulses as PD signals.

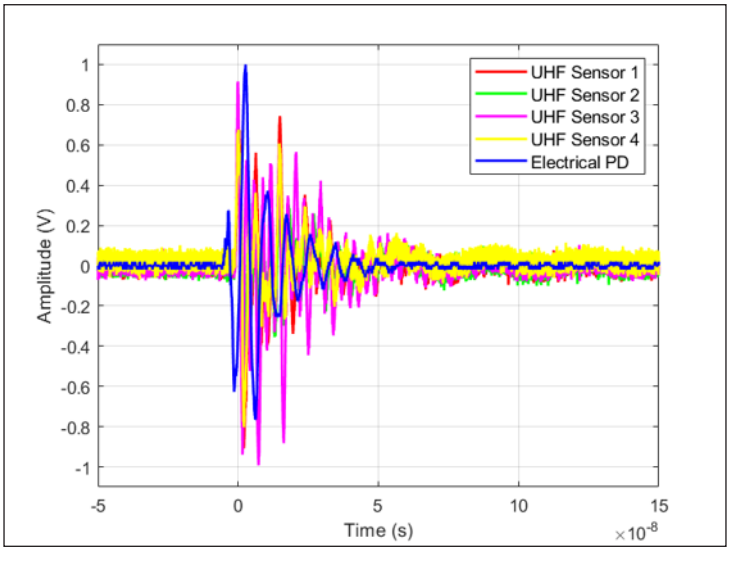
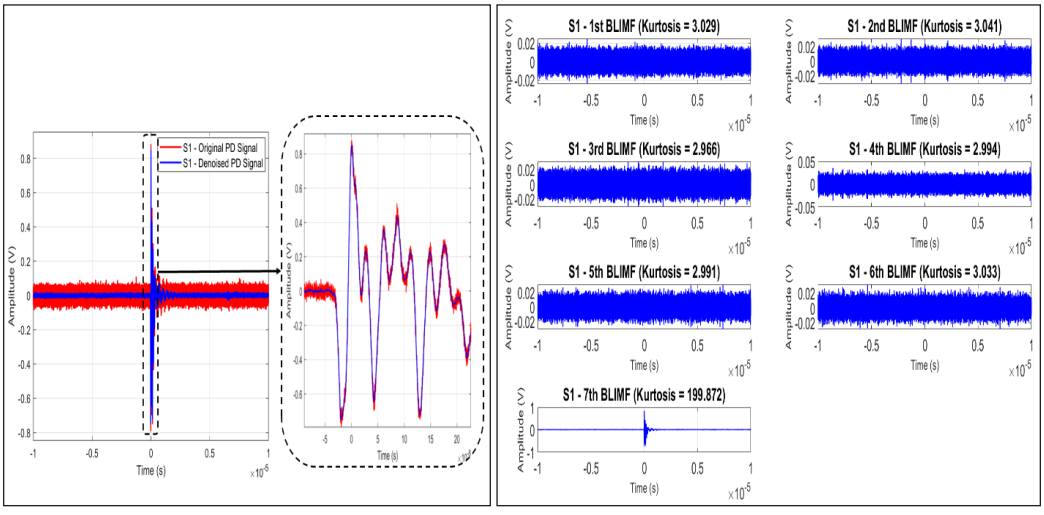


Figure 9. An example of the electrical and UHF PD signals obtained from the experiment

The effect of VMD denoising on the UHF PD signal is illustrated in Figure 10(a), where the magnitude of the PD signal is slightly reduced, while the general shape of the PD pulse is retained. VMD denoising also smooths the PD waveform and reduces ripples, as shown in Figure 10(a). Figure 10(b) depicts an example of the decomposition of an original UHF PD signal into 7 BLIMFs based on the VMD and mode-mixing judgment algorithm. It is observed that only the 7th BLIMF has a kurtosis greater than 10. Therefore, it is considered the effective BLIMF and is used to reconstruct the denoised UHF PD signal. Based on the equations in Mohd Hashim et al. (2023), the average normalised correlation coefficient (NCC) and the normalised root mean square error (NRMSE) between the original and denoised UHF PD signals at PD Location 1 are 0.91 and 0.40, respectively. For PD Location 2, the average NCC is 0.86, and the average NRMSE is 0.50. The average time required to denoise the UHF PD signals is 126.8 s for PD Location 1 and 96.1 s for PD Location 2.



(a) (b)

Figure 10. (a) Example of original and denoised UHF PD signals. (b) BLIMFs obtained from the decomposition of the original UHF PD signal in (a)

TDOA Computation Based on FPM and CEM

Figure 11(a) shows the mean absolute errors (MAEs) of the TDOAs derived from the UHF PD signals of PD Location 1. The MAEs are computed in 2 steps. First, the absolute differences between the TDOAs derived from the original or denoised UHF PD signals and the theoretical TDOAs computed from Equations 1-4 are calculated. Second, the absolute differences across all combinations of sensor coordinates are averaged to obtain the MAEs. For PD Location 1, the τ_{21}, τ_{31}, and τ_{41} in Scenario 1 have MAEs that are 26.01%, 24.45%, and 30.01% lower than those in Scenario 2, while the MAEs in Scenario

4 are 47.11%, 49.15%, and 52.68% lower than those in Scenario 3. As illustrated in Figure 11(b), a similar trend is observed for PD Location 2, where Scenario 1 yields 30.29%, 29.26%, and 23.69% lower MAEs than Scenario 2, and Scenario 4 produces 44.58%, 44.25%, and 46.39% lower MAEs than Scenario 3. These findings indicate that neither the FPM nor the CEM consistently outperforms the other. The FPM yields lower MAEs for the original UHF PD signals, whereas the CEM produces lower MAEs for the denoised UHF PD signals. In addition, VMD denoising improves the UHF PD signal waveforms and enables both the FPM and the CEM to achieve lower MAEs in TDOA computation.

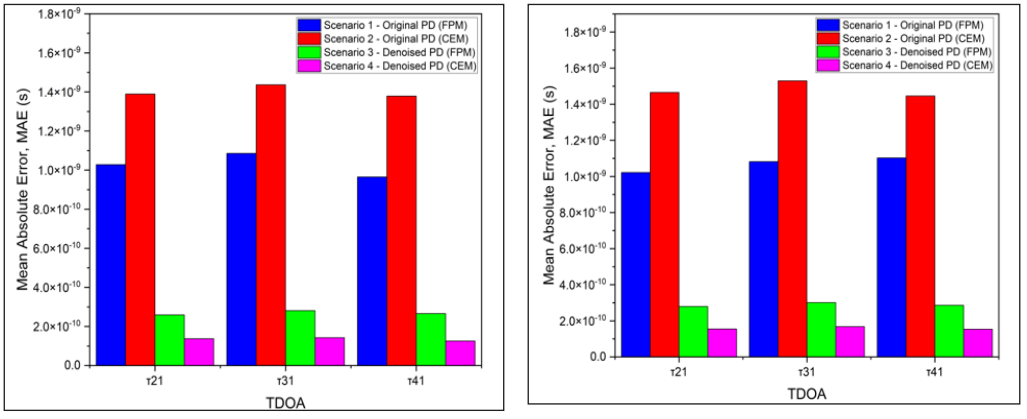


Figure 11. Mean absolute errors of the TDOAs obtained by the FPM and CEM from the original and denoised UHF PD signals of (a) PD Location 1 and (b) PD Location 2

Validation Set Performance of the Trained ANNs

In this study, validation sets are used instead of independent test sets because the trained ANNs are subsequently evaluated using the TDOAs derived from the experimental UHF PD signals. Figure 12 depicts the validation RMSE and the Pearson correlation coefficients of all 20 trained ANNs. It is found that the correlation R_x is the highest, ranging from 0.96 to 0.99. This indicates that the x-coordinates predicted by the ANNs are very close to the target x-coordinates. The correlations R_y and R_z are lower, with an average of 0.88 and 0.74, respectively. The weak performance along the y- and z-axes is likely due to the shorter dimensions of the test tank in these directions (Khodaveisi et al., 2024), which leads to less variation in the corresponding coordinates among the samples in the datasets. The trained ANNs achieve an average RMSE of 7.23 cm, which is comparable to the performance reported for the backpropagation neural network in Khodaveisi et al. (2024).

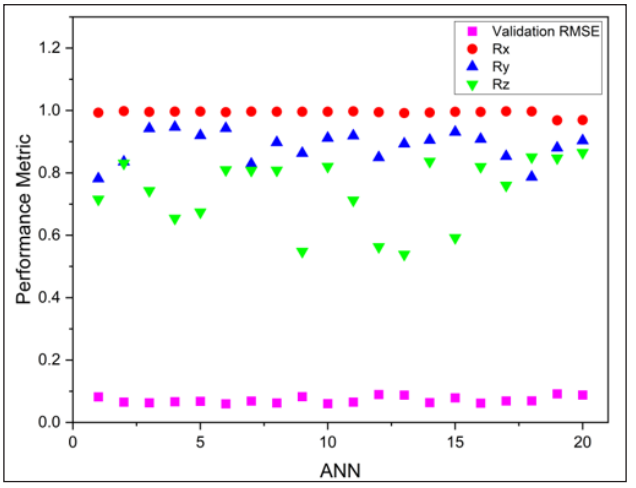


Figure 12. Validation RMSE and Pearson correlation coefficients of the trained ANNs

PD Localisation Performance Analysis

The estimated PD locations produced by the trained ANNs, ABC, BA, GA, PSO, and the non-iterative algorithm for PD Location 1 are presented in Figure 13. Figure 14(a) shows that the PD localisation algorithms yield lower average distance errors in Scenario 1 than in Scenario 2. The distance errors decrease when the UHF PD signals are denoised, as indicated by Scenarios 3 and 4. Scenario 4 produces lower distance errors than Scenario 3. These results demonstrate that PD localisation accuracy is affected by the precision of TDOA computation (Chan et al., 2023; Desai & Sarathi, 2018), where TDOAs with low MAEs lead to high PD localisation accuracy. As shown in Figure 14(a), the trained ANNs achieve the highest PD localisation accuracy across all 4 scenarios. For instance, in Scenario 4, the trained ANNs record the lowest average distance error of 5.74 cm, which is 46.34%, 46.59%, 49.83%, 49.77%, and 56.60% lower than the errors of the ABC, BA, GA, PSO, and the non-iterative algorithm, respectively. In contrast, the non-iterative algorithm consistently produces the lowest localisation accuracy because a few of its estimated PD locations fall outside the tank, as illustrated in Figure 13. When both solutions from the non-iterative algorithm are located outside the tank, the one closest to the actual PD location is taken as the estimated PD location. However, some of the out-of-tank estimated PD locations are too far and are thus not displayed in Figure 13. The ABC, BA, GA, and PSO algorithms exhibit similar PD localisation accuracy across the 4 scenarios. However, as depicted in Figure 14(b), the ABC algorithm has the shortest average computation time. The trained ANNs require 37.97 to 43.17 ms to estimate the PD locations, whereas all the optimisation algorithms have average computation times above 638 ms. The non-iterative algorithm provides the fastest PD localisation, with an average computation time of 0.34 to 0.57 ms.

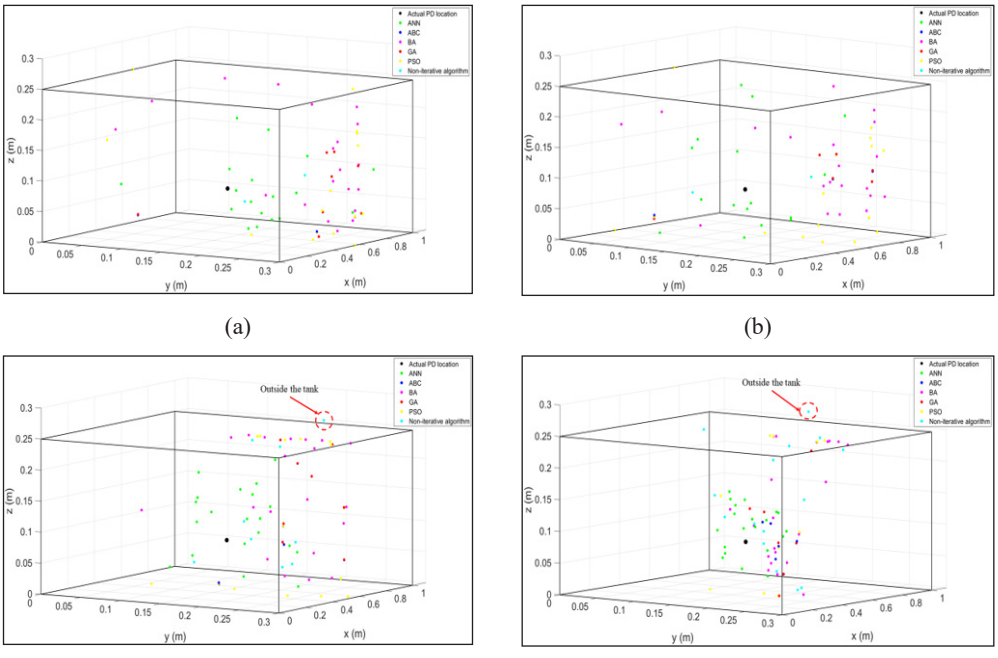


Figure 13. PD localisation results of the trained ANNs (green), ABC (blue), BA (magenta), GA (red), PSO (yellow), and the non-iterative algorithm (cyan) for PD Location 1 under (a) Scenario 1, (b) Scenario 2, (c) Scenario 3, and (d) Scenario 4

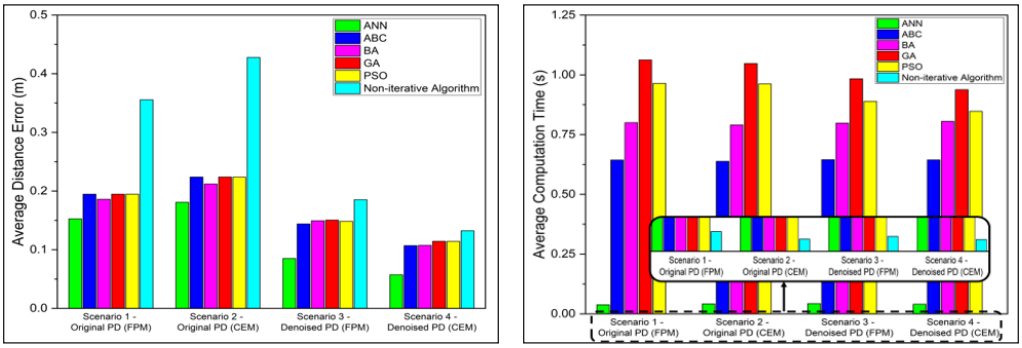


Figure 14. (a) Average distance error and (b) average computation time of the trained ANNs, ABC, BA, GA, PSO, and non-iterative algorithm for PD localisation at PD Location 1

Figure 15 illustrates the estimated PD locations obtained by the trained ANNs, ABC, BA, GA, PSO, and the non-iterative algorithm for PD Location 2. The average distance errors and average computation times for each PD localisation algorithm are presented in Figures 16(a) and 16(b), respectively. A similar trend to PD Location 1 is observed

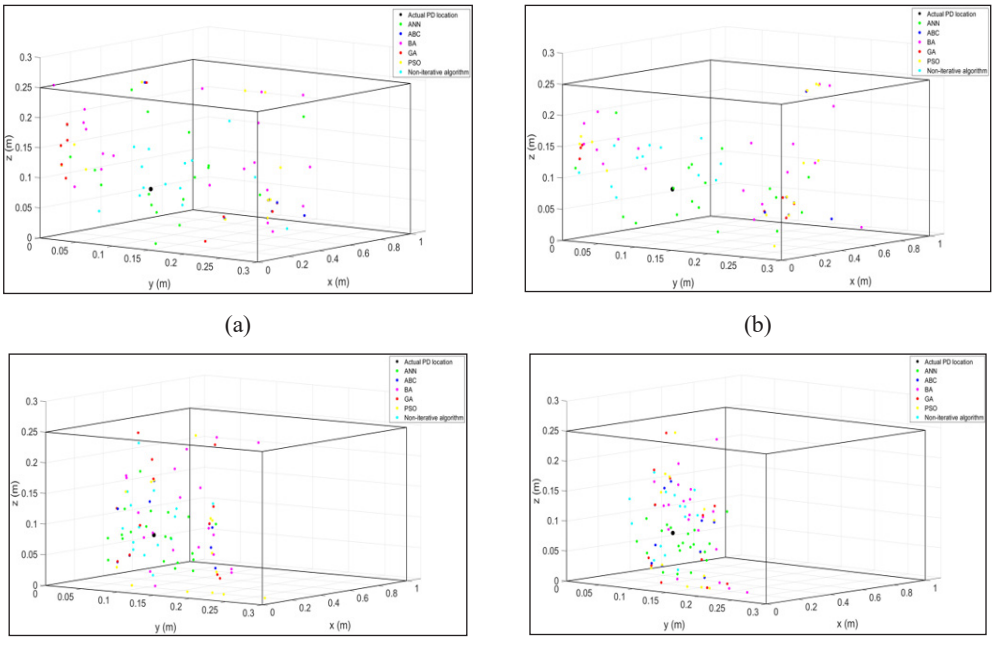


Figure 15. PD localisation results of the trained ANNs (green), ABC (blue), BA (magenta), GA (red), PSO (yellow), and the non-iterative algorithm (cyan) for PD Location 2 under (a) Scenario 1, (b) Scenario 2, (c) Scenario 3, and (d) Scenario 4

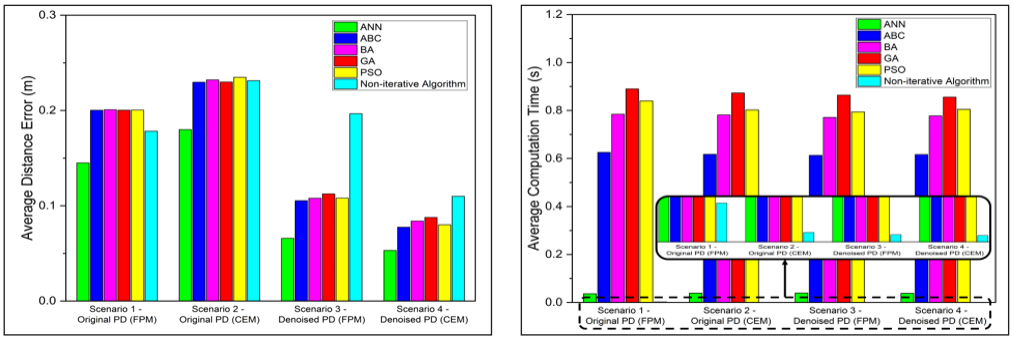


Figure 16. (a) Average distance error and (b) average computation time of the trained ANNs, ABC, BA, GA, PSO, and non-iterative algorithm for PD localisation at PD Location 2

in Figure 16(a), where Scenario 1 results in lower distance errors than Scenario 2, and Scenario 4 yields lower errors than Scenario 3. The trained ANNs produce the lowest average distance errors across all scenarios, with a minimum error of 5.32 cm in Scenario 4. This is 31.46%, 36.71%, 39.44%, 33.63%, and 51.69% lower than the errors of the ABC, BA, GA, PSO, and the non-iterative algorithm, respectively. The ABC consistently

yields higher PD localisation accuracy than the BA, GA, and PSO, as shown in Figure 16(a). For example, in Scenario 3, the distance error of the ABC is 10.54 cm, which is 2.36%, 6.33%, and 2.52% lower than the errors of the BA, GA, and PSO, respectively. In Scenario 1, the non-iterative algorithm achieves higher PD localisation accuracy than the optimisation algorithms, with an 11% lower distance error than ABC, as shown in Figure 16(a). In contrast, the non-iterative algorithm yields the largest distance errors in Scenarios 3 and 4. As depicted in Figure 16(b), the trained ANNs require an average of 36.75 to 39.71 ms to perform PD localisation for PD Location 2. This is shorter than the computation times of the optimisation algorithms, which generally exceed 613 ms. As in the case of PD Location 1, the non-iterative algorithm achieves the shortest average computation time, ranging from 0.36 ms to 2.18 ms.

Overall, the accuracy of PD localisation depends on the precision of TDOA computation. For both PD locations, TDOAs with low MAEs, obtained from the FPM for the original UHF PD signals and from the CEM for the denoised UHF PD signals, generally result in low average distance errors. The trained ANNs achieve the highest PD localisation accuracy, with minimum average distance errors of 5.74 cm and 5.32 cm for PD Locations 1 and 2, respectively. These errors correspond to 5.35% and 4.96% of the tank diagonal length, respectively. The good PD localisation accuracy of the ANNs is likely due to the datasets used for training. Random deviations are added to the theoretical TDOAs during dataset generation. As a result, the ANNs are exposed to non-ideal TDOAs during training. In contrast, the optimisation algorithms and the non-iterative algorithm do not undergo a training process prior to PD localisation. Therefore, these algorithms are more sensitive to TDOA errors compared to the trained ANNs. In terms of computation time, the trained ANNs provide faster estimations than the ABC, BA, GA, and PSO, but are slower than the non-iterative algorithm.

CONCLUSION

In conclusion, this study investigated the PD localisation performance of ANNs trained with TDOA datasets generated from the TDOA equations. The trained ANNs exhibit higher correlation for the x-coordinates as compared to the y- and z-coordinates. For the original UHF PD signals obtained from the experiment, the FPM computes the TDOAs with lower MAEs than the CEM, whereas for the UHF PD signals denoised by the VMD method, the CEM records lower MAEs than the FPM. The PD localisation results indicate that TDOAs with low MAEs generally yield lower average distance errors. For all PD locations and TDOA scenarios in this study, PD localisation with the trained ANNs consistently yields the lowest average distance errors. This suggests that the dataset generation procedures adopted in this study offer an alternative to extensive field data collection or simulation-based approaches. The average computation time of the trained ANNs for UHF PD

localisation is shorter than that of the ABC, BA, GA, and PSO, but longer than that of the non-iterative algorithm.

ACKNOWLEDGEMENT

This study was funded by Universiti Putra Malaysia Inisiatif Putra (Fast Track) (GP-FASTRACK/2024/9791100) funding scheme and Inisiatif Putra Berkumpulan (GP-IPB/2022/9717000).

LIST OF ABBREVIATIONS

ABC	Artificial bee colony
Adam	Adaptive moment estimation
AI	Artificial intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial neural network
BA	Bat algorithm
BLIMF	Band-limited intrinsic mode function
BPNN	Backpropagation Neural Network
CCM	Cross-Correlation Method
CEM	Cumulative Energy Method
EMD	Empirical mode decomposition
FPM	First Peak Method
GA	Genetic algorithm
IMC	Impedance matching circuit
MAE	Mean absolute error
MW	Multivariate wavelet denoising
NCC	Normalised correlation coefficient
NRMSE	Normalised root mean square error
PD	Partial discharge
PSO	Particle swarm optimisation
QGA	Quantum genetic algorithm
ReLU	Rectified linear unit
RMSE	Root mean square error
SNR	Signal-to-noise ratio
SVM	Support vector machine
TDOA	Time Difference of Arrival
TOA	Time of arrival

UHF Ultra-high frequency
VMD Variational mode decomposition

REFERENCES

AJ, C., Salam, M. A., Rahman, Q. M., Wen, F., Ang, S. P., & Voon, W. (2018). Causes of transformer failures and diagnostic methods – a review. *Renewable and Sustainable Energy Reviews*, *82*, 1442-1456. https://doi.org/10.1016/j.rser.2017.05.165

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods*, *43*(1), 3-31. https://doi.org/10.1016/S0167-7012(00)00201-3

Boya-Lara, C., Rivera-Caballero, O., & Ardila-Rey, J. A. (2022). A comparative study of denoising techniques for UHF signals from partial discharge. In *2022 8th International Engineering, Sciences and Technology Conference (IESTEC)* (pp. 595-601). IEEE. https://doi.org/10.1109/IESTEC54539.2022.00099

Chakravarthi, M. K., Giridhar, A. V., & Sarma, D. V. S. S. S. (2017). Localization of partial discharge source in power transformer using bat algorithm. In *2017 7th International Conference on Power Systems (ICPS)* (pp. 382-385). IEEE. https://doi.org/10.1109/ICPES.2017.8387324

Chan, J. Q., Raymond, W. J. K., Illias, H. A., & Othman, M. (2023). Partial discharge localization techniques: A review of recent progress. *Energies*, *16*(6), Article 2863. https://doi.org/10.3390/en16062863

CIGRE Working Group D1.29 (2017). *Partial discharges in transformers (CIGRE technical brochure no. 676).* CIGRE.

Cintra Veloso, G., Borges Da Silva, L., Lambert-Torres, G., & Pinto, J. O. P. (2006). Localization of partial discharges in transformers by the analysis of the acoustic emission. In *2006 IEEE International Symposium on Industrial Electronics* (pp. 537-541). IEEE. https://doi.org/10.1109/ISIE.2006.295515

Desai, B. M. A., & Sarathi, R. (2018). Identification and localisation of incipient discharges in transformer insulation adopting UHF technique. *IEEE Transactions on Dielectrics and Electrical Insulation*, *25*(5), 1924-1931. https://doi.org/10.1109/TDEI.2018.007294

Ghosh, R., Chatterjee, B., & Dalai, S. (2017). A method for the localization of partial discharge sources using partial discharge pulse information from acoustic emissions. *IEEE Transactions on Dielectrics and Electrical Insulation*, *24*(1), 237-245. https://doi.org/10.1109/TDEI.2016.006080

Hashim, A. H. M., Azis, N., Jasni, J., Radzi, M. A. M., Kozako, M., Jamil, M. K. M., & Yaakub, Z. (2023). Application of ANFIS and ANN for partial discharge localization in oil through acoustic emission. *IEEE Transactions on Dielectrics and Electrical Insulation*, *30*(3), 1247-1254. https://doi.org/10.1109/TDEI.2023.3264958

Jahangir, H., Akbari, A., Werle, P., & Szczechowski, J. (2017). UHF PD measurements on power transformers-advantages and limitations. *IEEE Transactions on Dielectrics and Electrical Insulation*, *24*(6), 3933-3940. https://doi.org/10.1109/TDEI.2017.006149

Jangjoo, M. A., Allahbakhshi, M., & Mirzaei, H. R. (2023). UHF sensors positioning on the power transformer tank to enhance the partial discharge localization accuracy. *Electric Power Systems Research*, *218*, Article 109174. https://doi.org/10.1016/j.epsr.2023.109174

Kakeeto, P., Judd, M., Pearson, J., & Templeton, D. (2008). Experimental investigation of positional accuracy for UHF partial discharge location. In *2008 International Conference on Condition Monitoring and Diagnosis* (pp. 1070-1073). IEEE. https://doi.org/10.1109/CMD.2008.4580467

Khodaveisi, F., Karami, H., Karimpour, M. Z., Rubinstein, M., & Rachidi, F. (2024). Partial discharge localization in power transformer tanks using machine learning methods. *Scientific Reports*, *14*(1), Article 11785. https://doi.org/10.1038/s41598-024-62527-9

Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization.* arXiv. https://doi.org/10.48550/arXiv.1412.6980

Kundu, P., Kishore, N. K., & Sinha, A. K. (2009). A non-iterative partial discharge source location method for transformers employing acoustic emission techniques. *Applied Acoustics*, *70*(11-12), 1378-1383. https://doi.org/10.1016/j.apacoust.2009.07.001

Lim, Z. Y., Azis, N., Mohd Hashim, A. H., Mohd Radzi, M. A., Norsahperi, N. M. H., & Mohd Ariffin, A. (2025). Acoustic emission partial discharge localization in oil based on artificial bee colony. *Pertanika Journal of Science and Technology*, *33*(1), 241-259. https://doi.org/10.47836/pjst.33.1.11

Long, J., Wang, X., Dai, D., Tian, M., Zhu, G., & Zhang, J. (2017). Denoising of UHF PD signals based on optimised VMD and wavelet transform. *IET Science, Measurement & Technology*, *11*(6), 753-760. https://doi.org/10.1049/iet-smt.2016.0510

Long, J., Wang, X., Zhou, W., Zhang, J., Dai, D., & Zhu, G. (2021). A comprehensive review of signal processing and machine learning technologies for UHF PD detection and diagnosis (i): Preprocessing and localization approaches. *IEEE Access*, *9*, 69876-69904. https://doi.org/10.1109/ACCESS.2021.3077483

Mohd Hashim, A. H., Azis, N., Jasni, J., Mohd Radzi, M. A., Kozako, M., Mohd Jamil, M. K., & Yaakub, Z. (2023). Examination on the denoising methods for electrical and acoustic emission partial discharge signals in oil. *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, *11*(3), 791-804. https://doi.org/10.52549/ijeei.v11i3.4463

Pattanadech, N., & Muhr, M. (2016). Partial discharge inception voltage investigation of mineral oil: Effect of electrode configurations and oil conditions. *IEEE Transactions on Dielectrics and Electrical Insulation*, *23*(5), 2917-2924. https://doi.org/10.1109/TDEI.2016.7736853

Rathod, V. B., Kumbhar, G. B., & Bhalja, B. R. (2022). Partial discharge detection and localization in power transformers based on acoustic emission: Theory, methods, and recent trends. *IETE Technical Review*, *39*(3), 540-552. https://doi.org/10.1080/02564602.2021.1871672

Sinaga, H. H., Phung, B. T., & Blackburn, T. R. (2012). Partial discharge localization in transformers using UHF detection method. *IEEE Transactions on Dielectrics and Electrical Insulation*, *19*(6), 1891-1900. https://doi.org/10.1109/TDEI.2012.6396945

Taha, I. B. M., Dessouky, S. S., Ghaly, R. N. R., & Ghoneim, S. S. M. (2020). Enhanced partial discharge location determination for transformer insulating oils considering allocations and uncertainties of

acoustic measurements. *Alexandria Engineering Journal*, *59*(6), 4759-4769. https://doi.org/10.1016/j.aej.2020.08.041

Tang, L., Luo, R., Deng, M., & Su, J. (2008). Study of partial discharge localization using ultrasonics in power transformer based on particle swarm optimization. *IEEE Transactions on Dielectrics and Electrical Insulation*, *15*(2), 492-495. https://doi.org/10.1109/TDEI.2008.4483469

Zaki, A. U. M., Hu, Y., & Jiang, X. (2020). Partial discharge localization in 3-D with a multi-dnn model based on a virtual measurement method. *IEEE Access*, *8*, 87434-87445. https://doi.org/10.1109/ACCESS.2020.2993190

Zhang, J., He, J., Long, J., Yao, M., & Zhou, W. (2019). A new denoising method for UHF PD signals using adaptive VMD and SSA-Based shrinkage method. *Sensors*, *19*(7), Article 1594. https://doi.org/10.3390/s19071594

Pertanika J. Sci. & Technol. 34 (5): 2619 - 2647 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Investigation of Assistive Grouser Traction Force in Toyoura Sand: Effects of Grouser Geometry and Contact Angle

Mochamad Edoward Ramadhan[1,2], Ahmad Najmuddin Ibrahim[2]*, and Yasuhiro Fukuoka[3]

[1]Department of Mechanical Engineering, Faculty of Engineering, University of Jember, Jalan Kalimantan No. 37, Tegalboto Campus, P.O. Box 159, Jember 68121, East Java, Indonesia
[2]Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, 26600 Pekan, Pahang, Malaysia
[3]Department of Mechanical Systems Engineering, College of Engineering, Ibaraki University, 4-12-1 Nakanarusawa, Hitachi, Ibaraki 316-8511, Japan

ABSTRACT

Slip and sinkage remain critical challenges for wheeled robot mobility on granular terrain. Previous studies on assistive grousers have examined the effects of geometric shapes, while the orientation of the contact angle, which affects traction force generation, has not been investigated. This study investigated C-shaped (180°; C-Front: lateral face forward; C-Rear: lateral face rearward) and Chevron (120°) assistive grousers using kinematic analysis, force measurements with calibrated sensors, and force-time analysis during a single operating cycle. To determine the performance of the grouser-ground interaction, we conducted static resistance tests (100% slip) and dynamic traction tests (21-23% slip) on Toyoura sand using the Single-Wheel Test Bed (SWTB) under a 14 kg normal load at a rotational speed of 1 RPM. The C-Front configuration produced the highest forces (138.7 N resistance / 38.4 N traction), followed by C-Rear (57.1 N / 27.1 N), while Chevron yielded the lowest (44.0 N / 9.0 N). Despite having identical geometry, C-Front and C-Rear configurations differ significantly due to the orientation of the lateral side, which affects the geometric soil confinement mechanism. One-way ANOVA confirmed significant differences among all configurations ($p < 0.001$, $\eta^2 > 0.99$). Force-time analysis shows three phases: penetration, traction, and retraction, with different durations. C-Front exhibits the longest traction phase (38.95% of the cycle) compared with the Chevron configuration (28.36%). These results suggest that a smaller contact angle is more suitable for lightweight robots experiencing low

ARTICLE INFO

Article history:
Received: 26 June 2026
Accepted: 21 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.04

edoward.teknik@unej.ac.id (Mochamad Edoward Ramadhan)
anajmuddin@umpsa.edu.my (Ahmad Najmuddin Ibrahim)
yasuhiro.fukuoka.phd@vc.ibaraki.ac.jp (Yasuhiro Fukuoka)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

resistance, while a forward lateral orientation maximises traction for heavier platforms in sandy terrain.

Keywords: Assistive grouser, bulldozing resistance, local traction, planetary rover, single wheel test bed

INTRODUCTION

In recent years, the mobility and traction performance of mobile platforms over unstructured or non-cohesive terrain have attracted significant research interest worldwide, particularly within European and global engineering frameworks. Reliable rover locomotion fundamentally depends on the contact mechanics between the propulsion system and the terrain, as well as the ability to avoid collisions with other objects in the surrounding environment; these are key requirements that wheeled robots must meet (Rubio et al., 2019).

Contact mechanics, the experimental basis for accurately measuring propulsion system performance related to tire-ground interface friction, remains a major challenge, requiring specialised instrumentation to accurately measure the friction coefficient. For example, the sliding friction parameter, which is strongly influenced by material properties such as contact surface roughness, humidity, lubrication, temperature, and relative velocity, remains a key parameter if environmental conditions can be determined (Kelemenová et al., 2020).

The ability to avoid collisions with other objects is essential for wheeled robots. Hruboš et al. (2016) used a 2-D scanner and utilised GPS and INS data fusion to create a final 3-D point cloud model. After generating the surface area, they applied textures automatically derived from a series of camera recordings. The final model was saved in *.obj format and consisted of individual environmental fragments that were then analysed for potential collisions with moving objects, which were then transferred through the measured environment. The term "moving object" could represent a train on a track or a vehicle on a road carrying a large load; environmental fragments searched for collisions could include trees, building parts, bridges, or power lines. The robotic manipulator developed by Blatnický et al. (2020), with a maximum load capacity of 25 kg for gripping automotive products, could also be used for evacuation to clear objects obstructing the rover's movement. It could also be used to lift the rover's body if its wheels become stuck in sandy terrain.

When combined with the robot’s inertial navigation system, which is equipped with an accelerometer and a gyroscope (Qazizada & Pivarčiová, 2016), the rover’s mobility system will be able to determine its position in space consistently and accurately. Considering the fundamental global progress and the research gaps still found in the development of wheel-mounted locomotion devices, this study focuses on wheels equipped with terrain-adaptive locomotion, which are the basis of rover mobility for exploration in all types of terrain.

In the field of Terramechanics, the efficient traversal of non-cohesive terrain is critical for the range and power autonomy of wheeled robots used in industrial, agricultural, and

aerospace exploration. The primary global challenges for these robots are slips and sinkage, which warrant detailed investigation (Tagliavini et al., 2022; Zhang et al., 2022). Soil displacement at the wheel-terrain interface occurs along longitudinal and lateral directions, inducing a bulldozing resistance that is a fundamental driver of slippage (Bekker, 1956; Wong, 1997). Watanabe et al. (2023) evaluated rigid-wheel dynamics under various conditions and identified critical trafficability thresholds. Furthermore, Sakayori and Ishigami (2024) confirmed that steady-state slip alters acting force as the terrain's load-bearing capacity fluctuates, compromising effective travel distance.

The development of sufficient tractive effort and soil shear strength is essential to mitigate these motion resistances. Classical frameworks by Bekker (1956) demonstrate that including grousers or spuds on rigid wheels facilitates soil densification, generating the thrust needed for translation while maintaining minimal slip ratios. To enhance predictions, Schäfer et al. (2010) modelled intricate wheel–soil dynamics, while Multi-Body System (MBS) frameworks, integrated with Soil Contact Models (SCM), enable sophisticated evaluations of sinkage and resistive forces. Notably, discrete grouser engagement typically induces sinusoidal oscillations in traction profiles. While conventional grousers can create voids that reduce traction (Yokoyama et al., 2020), trapezoidal shapes enhance performance by achieving superior soil compaction compared to rectangular designs (Suzuki et al., 2020). According to Nakamura et al. (2020), grouser integration fundamentally alters shear distributions, with the slip ratio increasing with increased vertical loading and diminished bearing capacity.

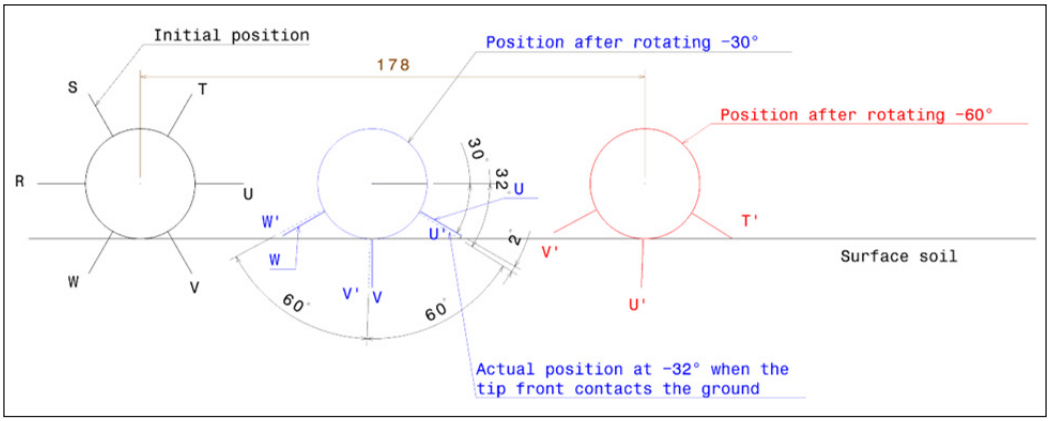
Research indicates that a 20° attack angle is optimal for maximising tractive effort on both level and sloped (0°-10°) granular terrain (Aziz et al., 2023). Experimental evaluations reveal that performance is primarily dictated by the number of grousers in active engagement with the substrate—which can cause fluctuating force profiles in multi-grouser setups due to mechanical interference—rather than the total count on the wheel (Aziz et al., 2023; Liu et al., 2024). To overcome these limitations, Ibrahim et al. (2016) introduced assistive grouser mechanisms with decoupled angular control to facilitate a consistent 0° angle of attack, defined throughout this study as the angle between the grouser blade and the vertical at the moment of soil entry, so that 0° denotes perpendicular penetration of the sand surface, resulting in improved efficiency and lower power consumption (Ibrahim & Fukuoka, 2018; Sabarudin et al., 2020a; Sabarudin et al., 2020b).

While Basri et al. (2024) established that expanding grouser width is ineffective in non-cohesive sand, grouser geometry and contact orientation remain critical for governing bulldozing resistance and soil translocation. This study specifically investigates the origins of negative traction forces (force reversals) identified in previous experimental data by comparing Chevron and C-shape designs within the assistive mechanism. Our methodology integrates kinematic analysis with the evaluation of calibrated load-cell signals to quantify how varying contact angles affect resistance and traction. These findings aim to provide

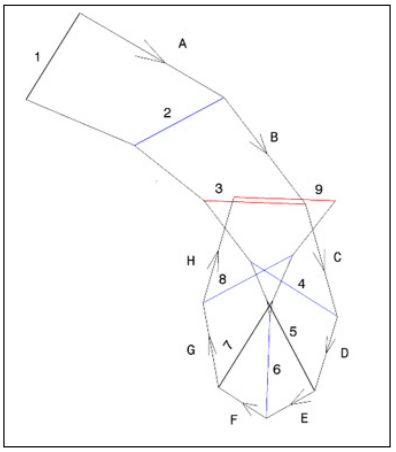
a comprehensive contribution to the development of high-traction assistive grousers for challenging terrains.

Kinematic Investigation and Trajectory Formulations

This kinematic model is the basis for the experimental design. Movement observations were modelled in two dimensions to interpret position changes relative to the wheel's movement (Figure 1). In the baseline conventional configuration, the U, V, and W grousers rotate by 30°. While grouser V aligns with the vertical axis at 270°—a position yielding maximum traction but inherent instability (Sutoh et al., 2012)—grousers U and W remain above the surface (Figure 1a). Unlike previous research focusing primarily on visualisation (Guo et al., 2021; Higa et al., 2016; Higa et al., 2019; Nakamura et al., 2020), this investigation analyses the full penetration-to-retraction cycle.

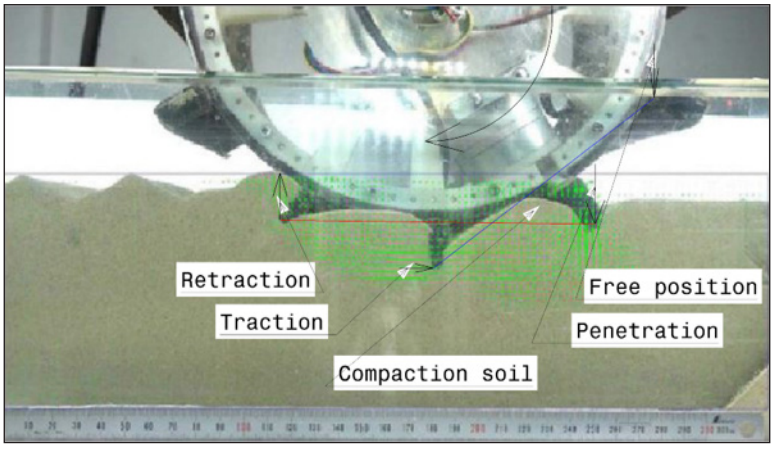


(a)

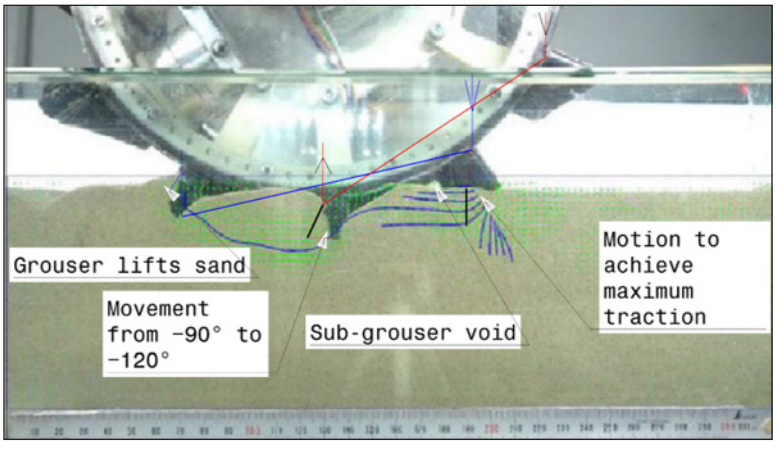


(b)

Figure 1. Motion of a conventional grouser (a) movement of the wheel with conventional grousers; (b) Trajectory of the grouser



(a)



(b)

Figure 2. Motion of Conventional Grouser PIV (a) Grouser position; (b) Grouser motion (Nakamura et al., 2020)

Using the Particle Image Velocimetry (PIV) method (Nakamura et al., 2020), the unit progresses through six phases: surface movement (Phase C), initial penetration (Phase D), and maximum penetration (Phase E, 0°), where peak traction is generated. Steps following this peak are considered passive as the system converts torque into shear forces rather than effective displacement (F to H). Figure 2 illustrates the kinematic cycle of a conventional grouser, which has been revalidated through a kinematic investigation of a single grouser movement cycle (Figure 1b) involving eight analytical positions (A to H), where the 2 initial phases, A and B, are phases where the grouser does not interact with the ground. These trajectory positions can be obtained with smoother graphics using the following formulations: the translation of movement, Equations 1 and 2; wheel trajectory, Equations 3 and 4; and tip grouser formulations, Equations 5 and 6.

Translation of movement:

$$X_c(t_i) = X_0 - (R \times (\theta_i + \omega t_i) \times \frac{\pi}{180}) \quad [1]$$

$$Y_c = R \quad [2]$$

Trajectory of wheel formulation:

$$X_{wheel}(t_i) = X_c(t_i) + R\cos(\theta_i + \omega t_i + (k_{n-1}\frac{360}{\Delta\theta})) \times \frac{\pi}{180} \quad [3]$$

$$Y_{wheel}(t_i) = Y_c + R\sin(\theta_i + \omega t_i + (k_{n-1}\frac{360}{\Delta\theta})) \times \frac{\pi}{180} \quad [4]$$

Trajectory of tip grouser formulation:

$$X_{tip\ cg}(t_i) = X_c(t_i) + (R + L)\cos(\theta_i + \omega t_i + (k_{n-1}\frac{360}{\Delta\theta})) \times \frac{\pi}{180} \quad [5]$$

$$Y_{tip\ cg}(t_i) = Y_c + (R + L)\sin(\theta_i + \omega t_i + (k_{n-1}\frac{360}{\Delta\theta})) \times \frac{\pi}{180} \quad [6]$$

Where X_c is the x-coordinate of the axle position, X_o is the initial x-coordinate of the axle, θ_i is the initial contact angle of the wheel, ω is the angular velocity of the wheel, t_i is the time at the i-th step, R is the radius wheel in mm, Y_c is the y-coordinat of the axle position, X_{wheel} is the x-coordinate of the wheel, k_n is the n-th grouser, Y_{wheel} is the y-coordinate of the wheel, $\Delta\theta$ is the angle between grouser, $X_{tip\ cg}$ is the x-coordinate of the tip conventional grouser, $Y_{tip\ cg}$ is the y-coordinate of the tip conventional grouser, and L is length of grouser in mm.

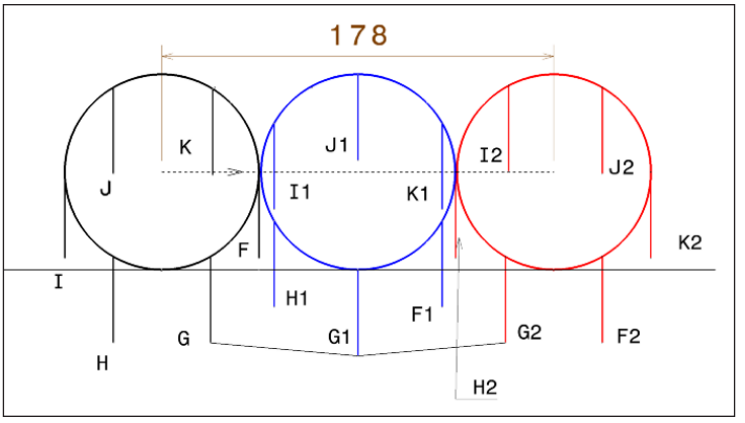
In contrast, the assistive grouser mechanism (Ibrahim et al., 2016) utilises six grousers mounted parallel to one another and offset by −90° relative to the wheel radius, which maintains the 0° angle of attack defined above throughout soil engagement (Figure 3). This configuration ensures two grousers remain embedded, providing superior stability compared to perpendicular mountings (Demishkevich et al., 2025). The assistive grouser tip trajectory is calculated using equations 7 and 8. As shown in Figure 4, the mechanism enables simultaneous penetration, traction, and retraction, optimising work distribution and reducing energy consumption compared to conventional designs (Suzuki et al., 2020).

$$X_{tip\ ag}(t_i) = X_{wheel}(t_i) + R\cos(\theta_i + \omega t_i + (k_{n-1}\frac{360}{\Delta\theta})) \times \frac{\pi}{180} \quad [7]$$

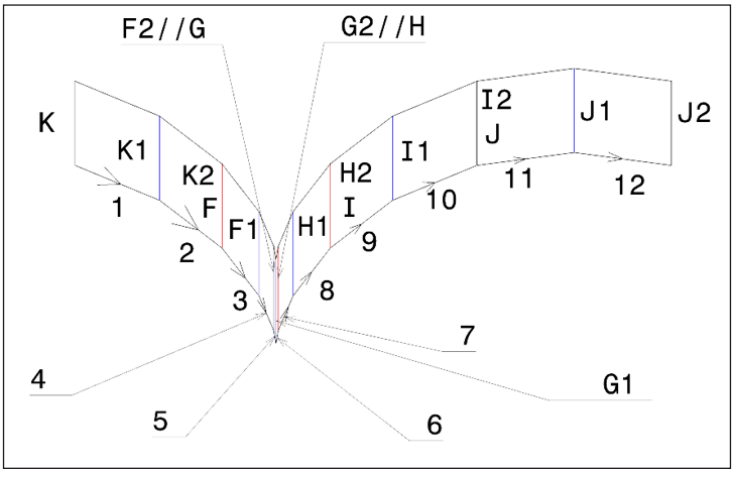
$$Y_{tip\ ag}(t_i) = Y_{wheel} + (R\sin(\theta_i + \omega t_i + (k_{n-1}\frac{360}{\Delta\theta})) \times \frac{\pi}{180}) - H \quad [8]$$

Where $X_{tip\ ag}$ is the x-coordinate of the tip assistive grouser, $Y_{tip\ ag}$ is the y-coordinate of the tip assistive grouser, and H is the height of the assistive grouser in mm.

Both grouser mechanisms share the fundamental objective of generating a traction force and achieving a longitudinal displacement equivalent to the spacing between grousers. This functional goal, supported by Elsheikh (2023), is kinematically represented by the trajectories shown in Figure 4, derived from the formulations in Equations 1 through 8.



(a)



(b)

Figure 3. Motion of Assistive Grouser (a) Movement of the wheel with assistive grouser; (b) Trajectory of the assistive grouser

As illustrated in Fig. 6b, soil aggregation occurs predominantly along the front face of the C-Front grouser during translation. This aggregation phenomenon is governed by three primary physical drivers: first, the 180º concave lateral profile restricts lateral particle escape, forcing granular sand to compress inward. Second, continuous translation forces Toyoura sand along a bounded shear failure plane ahead of the grouser. Third, high localised shear displacement compresses soil particles into a dense state.

While internal particle velocity fields were not imaged dynamically via Particle Image Velocimetry (PIV) in this setup, the presence of geometric soil confinement is strongly supported by the post-test surface aggregation patterns (Figure 6b) and the sustained elevated force-time impulse profiles.

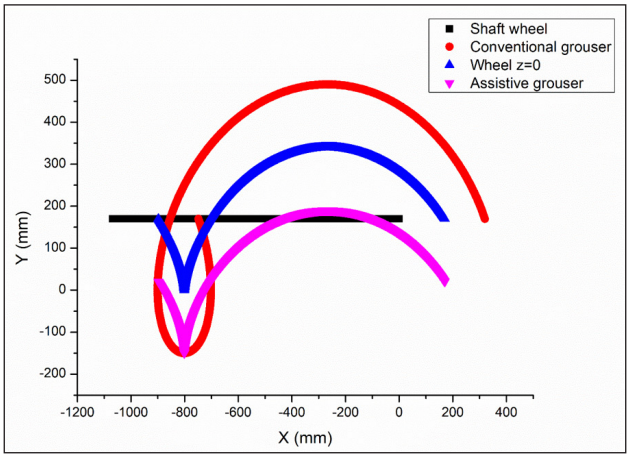
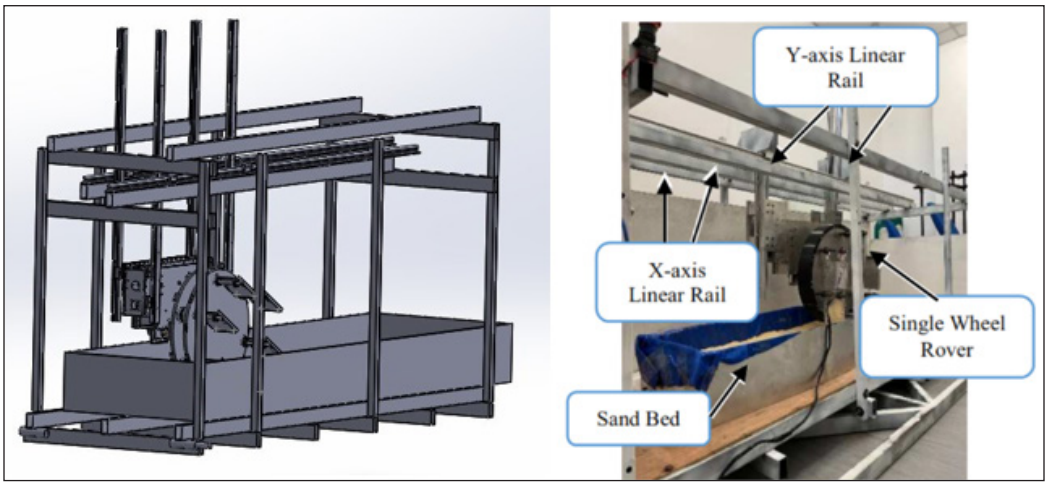


Figure 4. Comparison of trajectory

METHODOLOGY

Experimental Equipment and Substrate

To evaluate wheel-terrain interaction and tractive effort, this research employed a single-wheel test apparatus integrated into a laboratory test bed (Figure 5). The experimental wheel

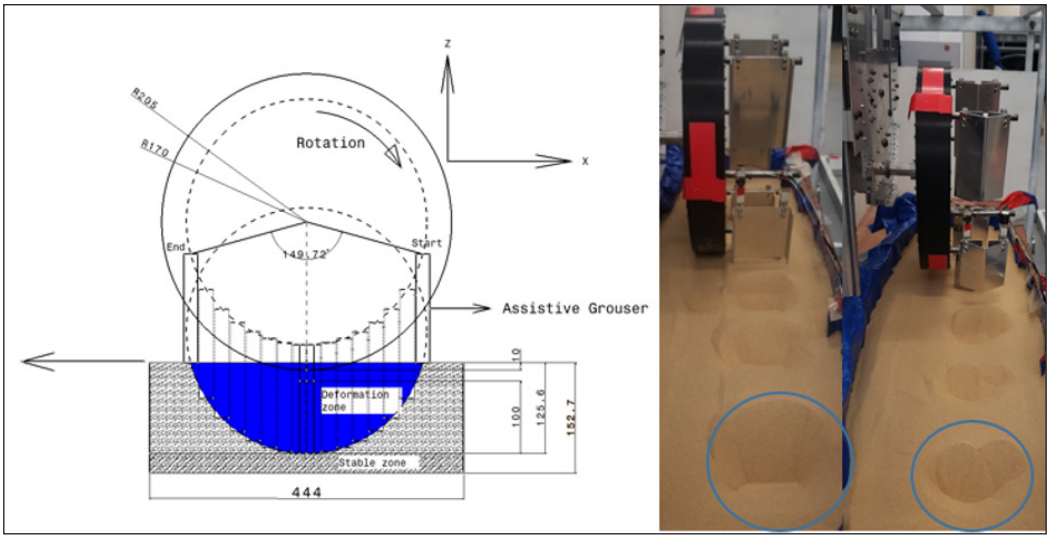


(a) (b)

Figure 5. Structural architecture (a) Design (b) Physical implementation of the single-wheel test apparatus

(radius: 205 mm, width: 60 mm) features a 6061-aluminium plate rim and a 3D-printed Poly Lactic Acid (PLA) structure. Based on a circumference division into 60º segments, the theoretical contact distance between consecutive grousers is 214.56 mm. The kinematic trajectory follows an arc length of up to 444 mm from soil entry to exit (Figure 6), with the support shaft fixed at a radius of 170 mm from the axle.

The soil bin facility (1000 × 250 × 160 mm) contained Toyoura sand maintained at a depth of 150 mm. The particle size of Toyoura sand typically ranges from 0.1 to 0.3 mm (Carow & Rackwitz, 2023; Ishikawa et al., 2014), with a bulk density between 1450 and 1600 kg/m³ (Onur et al., 2014), an internal friction angle between 32° and 46° (Safdar et al., 2021), and the cohesion is 0 kPa. To ensure environmental consistency, the laboratory was maintained at 26 °C and 60% humidity



(a) (b)

Figure 6. Operational assessment of the assistive grouser through (a) Static loading and (b) dynamic cycle analysis

Grouser Morphology and Configurations

Drawing from Aziz et al. (2023) and Basri et al. (2024), we compared two 6061 aluminium alloy geometries (thickness: 2 mm, width: 90 mm, height: 100 mm):

- C-shaped: 180° inclusion angle (Figure 7a).
- Chevron-shaped: 120° inclusion angle (Figure 7b).

For the C-shaped profile, we evaluated two orientations: the forward-facing lateral configuration (C-Front) and the rearward-facing lateral configuration (C-Rear), as illustrated in Figure 8.

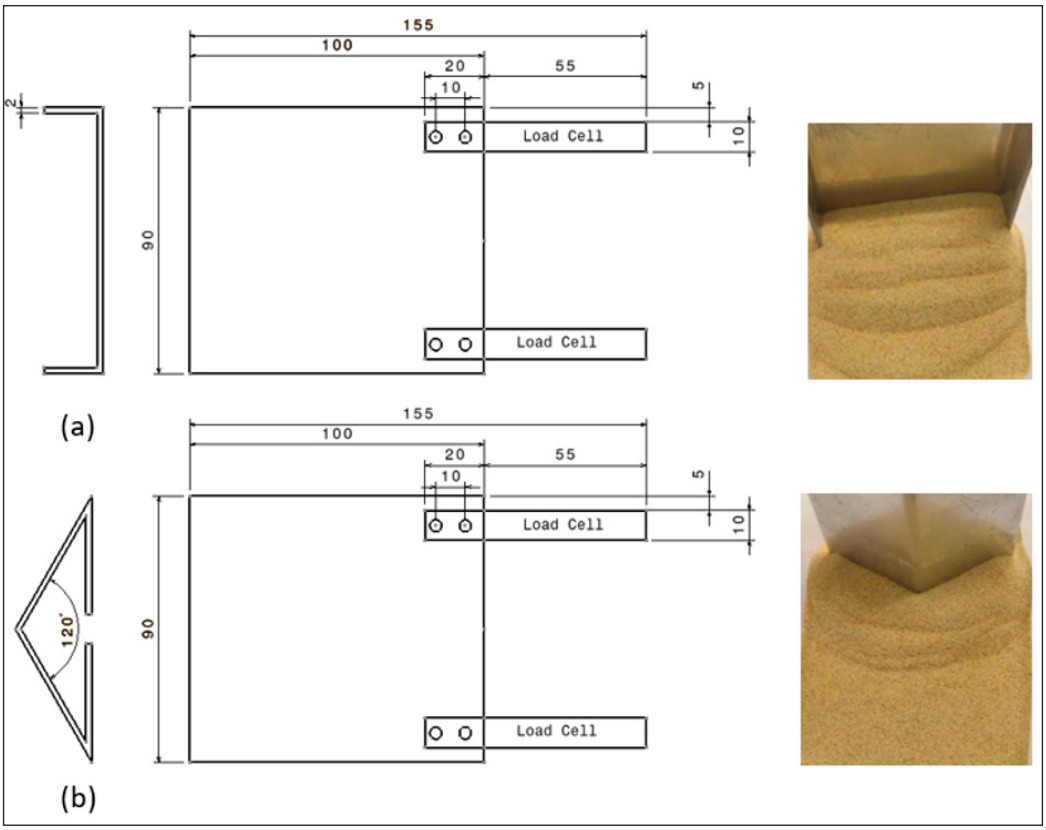


Figure 7. Integration and spatial placement of load cells within (a) C-shaped and (b) Chevron grouser assemblies

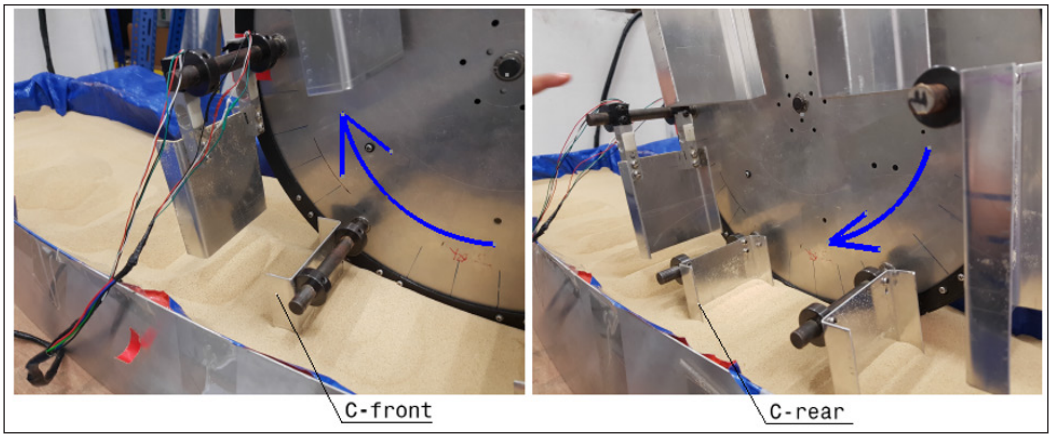


Figure 8. C-shape configuration

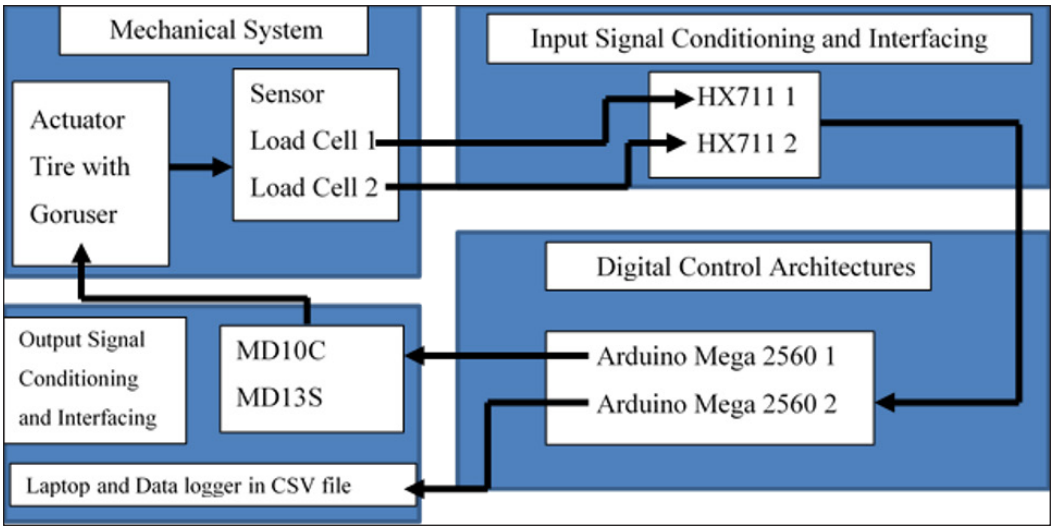


Figure 9. Integrated sensor architecture and motion control framework for force acquisition

Force Acquisition and Instrumentation

The data acquisition (DAQ) framework employs two 50-kg-capacity Load Cells (YZC-133, 0.05% FS) with HX711 converters and an Arduino Mega 2560 to manage signal flow for force measurement and motor actuation. Prior to testing, a 50 g (0.49 N) F2-class reference mass was used for zero-offset nulling and to verify signal sensitivity. Both load cells were mounted so that each responds to the full applied force, and the reported force is the mean of the two readings, (LC1 + LC2)/2. Averaging two independent measurements of the same load reduces random error and yields a single resultant estimate from two positions on the deflecting bracket. The same combination rule was applied throughout calibration, validation, and testing. Sensor accuracy was therefore determined empirically rather than assumed from the manufacturer's specification. Across the validated range of 0–10,000 g, the largest deviation of an individual cell was +40.78 g at 10,000 g, equivalent to 0.41% of reading or 0.082% of full scale, which exceeds the manufacturer's stated non-linearity tolerance of ±0.05% FS. The combined mean output, which is the quantity reported, deviated by +22.01 g at the same point, or 0.22% of reading and 0.044% of full scale, and therefore remains within that tolerance; averaging the two channels measurably improves accuracy relative to either cell alone. Measured accuracy is reported in place of the factory figure. The two channels differ by a mean of 38.2 g (SD 1.9 g) above 500 g while agreeing to within 0.12 g at 50 g, a behaviour consistent with a small seating effect in the Load Cell 2 mounting that is fully compensated by the individual calibration equations. The calibration factor disparity observed between Load Cell 1 (214.63 counts/g) and Load Cell 2 (219.03 counts/g) was due to slight strain-gauge manufacturing tolerances

and variations in mounting frame stiffness. To maintain high accuracy, each sensor was processed independently using its specific calibration equation (K_1 and K_2) rather than a single average value. Processed data are exported via USB in CSV format, while drive systems are manually operated. Theoretical tractive effort was determined by normalising axle torque against the respective radius equations 9 to 12. Peak forces during dynamic testing reached 40 N, well inside the directly validated range of 0-10,000 g (0-98 N). The peak static resistance of 138.7 N (≈14.1 kg) lies approximately 1.4 times beyond that range and rests on the fitted linearity of Equations 22 and 23 ($R^2 > 0.9999$); this extrapolation is acknowledged as a limitation of the static measurement.

$$\tau_{wheel} = \tau_{Elecric\ Motor} \times \text{Final Gear Ratio}_{Wheel} \quad [9]$$

$$F_{traction} = \tau_{wheel}(\text{Load or No-load}) \div r_{Wheel} \quad [10]$$

$$\tau_{Grouser} = \tau_{Elecric\ Motor} \times \text{Final Gear Ratio}_{Grouser} \quad [11]$$

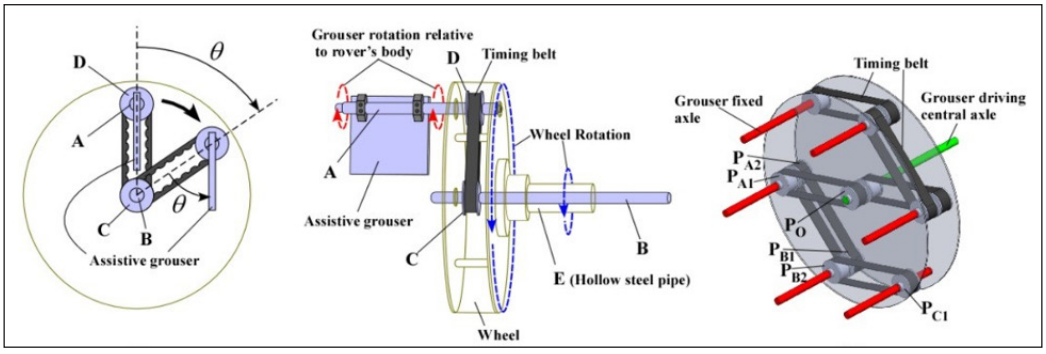
$$F_{traction} = \tau_{Grouser}(\text{Load or No-load}) \div r_{Grouser} \quad [12]$$

Actuation and Transmission Systems

The system utilises 150W DC motors (Maxon RE40 series) with harmonic drive units. As detailed in Table 1 and Figure 10, the transmission architectures differ: the wheel employs a four-stage reduction (1088:1), while the assistive grouser uses a penta-stage configuration (1440:1). This "in-wheel" design places the shaft on the lateral side, allowing two grousers to maintain ground contact during rotation. During evaluation, the wheel moves along the longitudinal x-axis while the vertical z-axis remains unconstrained to account for dynamic sinkage. This substantial torque amplification (103.25 N.m for the wheel and 136.66 N.m for the grouser) guarantees continuous output velocity without speed fluctuation or stalling under peak passive bulldozing resistance (138.7N). Assuming the force vector is applied at the geometric centre of the wheel-ground interface with a distance of the wheel radius (0.205 m) and the assistive-grouser moment arm taken to the geometric centre of the 100 mm grouser blade (0.170 + 0.050 = 0.220 m), using formulas 10 and 12 yields an average gross tractive force of 503.66 N for the wheel and 621.18 N for the assistive mechanism.

Table 1
Drive System transmission specifications

Drive System	Stages	Total Ratio	Rated Output Torque (N.m)
Wheel	4 (2 × 100 × 1.7 × 3.2)	1088:1	103.25
Grouser	5 (2.5 × 100 × 1.44 × 2 × 2)	1440:1	136.66



(a) (b) (c)

Figure 10. Wheel and assistive grouser transmission system (a) Angle between grousers; (b) Wheel shaft and grouser shaft; (c) Transmission for six grousers (Ibrahim et al., 2016)

Experimental Procedures

The sand was manually levelled before each trial to ensure a consistent terrain profile. Each run commenced with the wheel at the starting line and the instrumented grouser tip positioned 1–2 mm above the surface. The protocol followed two phases:

1. Static Calibration and Resistance: After zero-balancing the Load Cells, a 50-g F2-class weight was applied to derive scaling factors. Static resistance was then measured by rotating the grouser axle from 0° to -180 ° while the wheel was fixed.
2. Dynamic Testing: Tractive effort was characterised at 1 RPM, maintaining a 21%–23% slip ratio under a 14-kg normal load over a 170-mm displacement.

The Experimental baseline parameters- specifically, a normal vertical wheel load of 14 kg (137.34 N), a constant rotational speed of 1 RPM, and a dynamic slip ratio of 21-23% —were intentionally selected to represent the nominal operating profile of the target autonomous planetary rover. Operating at a low speed of 1 RPM enforces a quasi-static regime, eliminating inertial soil-throwing artefacts and enabling direct isolation of geometric contact forces.

RESULT AND DISCUSSIONS

This study examines the interaction between wheels equipped with an assistive grouser and Toyoura sand, focusing on quantifying resistive and resulting traction forces. Initially, the kinematic trajectories and operational dynamics of the assistive mechanism were benchmarked against traditional configurations to delineate fundamental differences between active decoupled systems and standard fixed-angle designs. Empirical evaluations show that C-shaped configurations, particularly in the C-Front orientation, generate the most significant resistive forces.

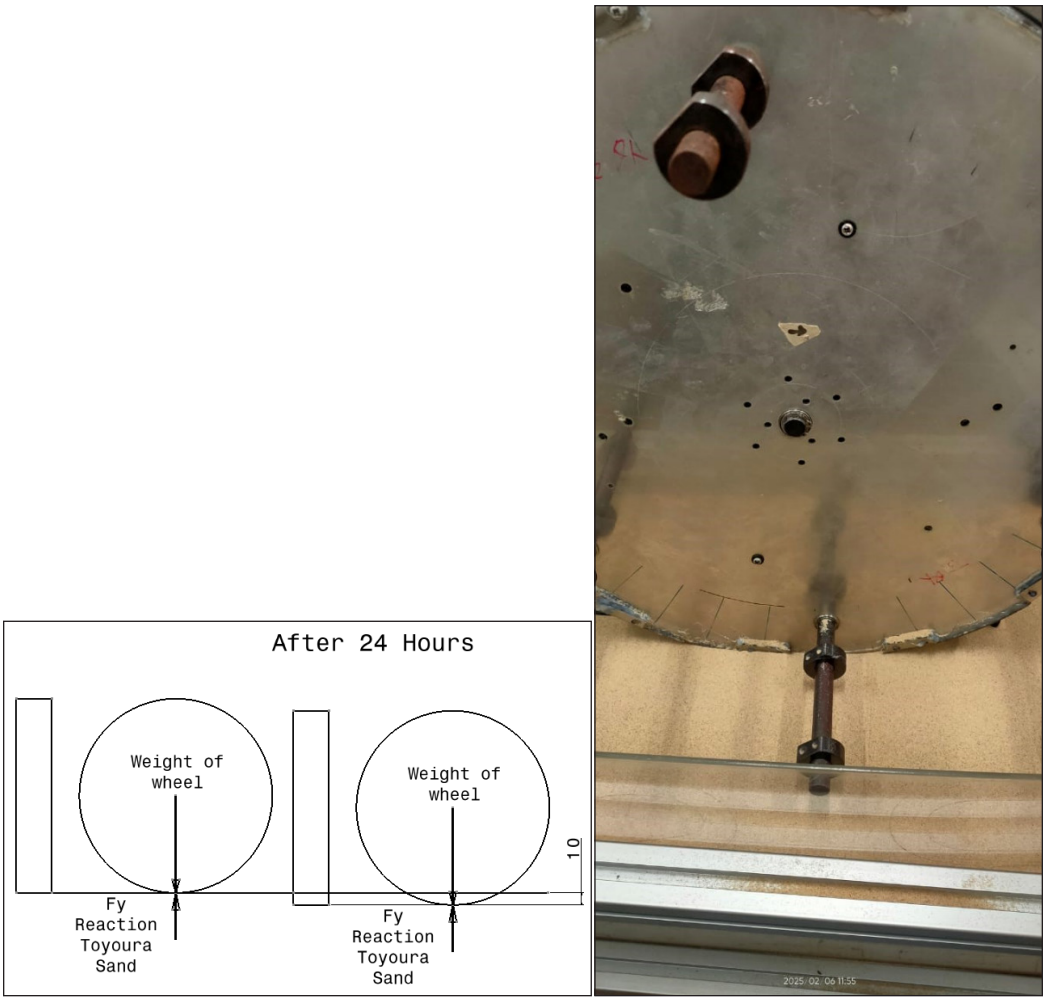


Figure 11. Static sinkage condition

Analysis of Static Sinkage and Baseline Bearing Capacity under Stationary Loading

This static test is a Bekker compression test used to determine the soil's bearing capacity under wheel loads. The difference is that the Bekker uses a single indenter and is pressed until the maximum downward movement is reached (due to applied pressure). While a stationary normal load of 14 kg (137.34 N) is applied to a natural substrate, where Newton's Third Law states that soil resistance must equal gravity in equilibrium (Boresi & Schmidt, 2001), wheel settlement results in an increasing surface area, encompassing both the underside and lateral areas. Unlike a rigid surface (W≤Soil bearing capacity), sand deforms and distributes pressure when the load exceeds its bearing capacity, causing the wheel to sink (W>Soil bearing capacity). After 24 hours of static loading, a 10 mm

settlement was observed (Figure 11). On the underside of the wheel tread and rim, the sand compacted downward from a narrow to a wide surface contact area, while in other areas, the sand deformed symmetrically toward the front, rear, and lateral surfaces of the wheel.

In conditions where the sinkage content changes from 0 to 10 mm against the wheel contact area, it can be mathematically determined using Equations 13 to 16, where Equations 13 and 14 define the symmetrical contact angle (θ0), Equation 15 gives the real curved contact area, and Equation 16 applies the wheel dimensions (R = 205 mm, B = 60 mm):

Symmetrical contact angle (θ_0)

$$\cos(\theta_0) = \frac{R-z}{R} = 1 - \frac{z}{R} \quad [13]$$

$$\theta_0 = arc\cos\ (1 - \frac{z}{R}) \quad [14]$$

Real curved contact area:

The total length of the curved contact arc from front to back is

$$A_{curved}\ = 2\,.B.R.arc\cos\ (1 - \frac{z}{R}) \quad [15]$$

Substitution of values R= 205 mm, B = 60 mm:

$$A_{curved}\ = 24600\ arc\cos\ (1 - \frac{z}{205})\ mm^2 \quad [16]$$

The area of the curve is the contact area where the soil bearing capacity is generated; this is evidence that the soil bearing capacity cannot be represented by an indenter in the form of a straight plate, as in a Bevameter.

Lateral contact area:

The lateral section consists of 2 flat, planar surfaces (the left and right sides of the wheel). The shape of the lateral contact area on each side is a circular sector with height z. Its area is given by Equation 17, and Equation 18 is obtained by substituting R = 205 mm.

$$A_{lateral}\ = 2\,.R^2.arc\cos\ (1 - \frac{z}{R}) - 2(R-z)\sqrt{2Rz - z^2} \quad [17]$$

Substitution of values R = 205 mm:

$$A_{lateral}\ = 84050\,.arc\cos\ (1 - \frac{z}{205}) - (410 - 2z)\sqrt{410z - z^2}\ (mm^2) \quad [18]$$

Note: The angle arc cos must be calculated in radians.

Total Wheel-to-Ground Surface Contact Area (A_{total}) z = 10 mm:

$$A_{total} = A_{curved} + A_{lateral} = 7715\,mm^2 + 1695\,mm^2 = 9410\,mm^2 \quad [19]$$

$$\sigma_{vertical} = \frac{W}{A_{Curved}} = \frac{137.34\,N}{7715\times10^{-6}m^2} = 17.80\,kPa \quad [20]$$

$$\sigma_{lateral} = \frac{A_{lateral}}{A_{curved}} \times \sigma_{vertical} = \frac{1695}{7715} \times 17.80 = 3.91\,kPa \quad [21]$$

At z = 10 mm, the total contact area is 9410 mm² (Equation 19), giving a vertical normal pressure of 17.80 kPa (Equation 20) and a lateral pressure of 3.91 kPa (Equation 21). Normal pressure is greatly influenced by the wheel's immersion and sinkage. The deeper the contact area, the larger it is, resulting in a smaller normal pressure. This corresponds to the soil's bearing capacity for the wheel's weight at rest. Meanwhile, when the wheel is at rest, the lateral side of the wheel also experiences uniform pressure, so the magnitude of the wheel's lateral pressure is the ratio of the lateral area to the bottom's curved area, producing normal pressure.

Evaluating the Magnitude of Soil Resistance Forces via Calibrated Instrumentation

Digital output profiles from the HX711 converters under idle and loaded states are illustrated in Figure 12. Raw ADC signals maintained a steady baseline with minimal noise (±50 counts) until the 22-second mark, where load application triggered a step-rise peaking at

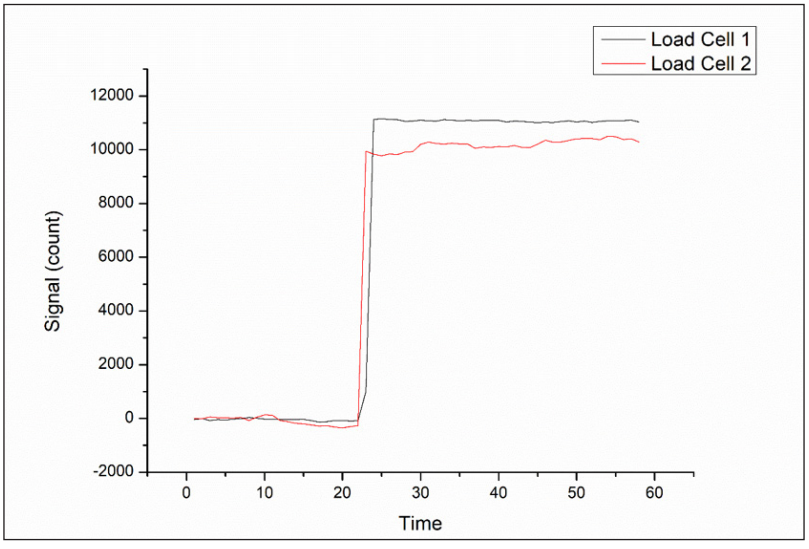


Figure 12. Raw ADC signal response of HX711 converters under no-load and applied-load

11,128 counts (Load Cell 1) and 9,953 counts (Load Cell 2). While Load Cell 2 responded approximately 1 second faster, Load Cell 1 demonstrated higher sensitivity, confirming the hardware's precision in force measurement.

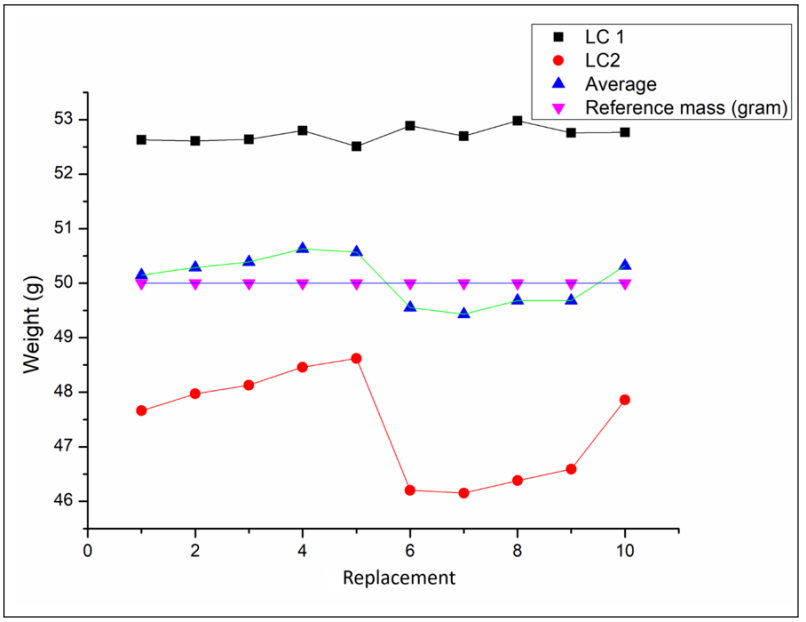


Figure 13. Systematic calibration and performance verification of the dual load cell assembly

To ensure system precision, calibration was performed using a 50 g F2-class standard weight. Across ten iterative trials, the system derived scale factors of 214.63 (LC1) and 219.03 (LC2), with integrated outputs ranging between 49.42 and 50.63 g. Statistical analysis verified the instrumentation's accuracy, yielding a mean deviation of +0.06 g (SD: 0.45 g, $p = 0.65$). As depicted in Figure 13, a detailed analysis of the calibration dataset revealed a shift in performance between the initial five trials (mean deviation: +0.40 g) and the subsequent five (mean deviation: -0.27 g). This variance was primarily caused by localised underestimation in Load Cell 2 during trials 6 through 9, which showed drops ranging from -0.32 g to -0.58 g. These deviations in Load Cell 2 readings stem from manufacturing tolerances and variations in mounting stiffness and are compensated for by applying individual Load Cell calibration factors rather than a combined average value. Despite these minor fluctuations, the absolute measurement error remained consistently below 1 g throughout the process, confirming high accuracy for subsequent dynamic testing. To ensure the accuracy of the Load Cell readings, a validation process is performed using the F2 standard over the range 0 to 10,000 grams, as shown in Table 2.

Table 2
Validation Mass balance 0-10,000 grams

Mass balance F2 (gram)	LC 1(gram)	LC 2 (gram)	Average (gram)
0	0	0	0
50	50.06	49.94	50.00
100	123.28	79.81	101.55
500	516.18	477.16	496.67
1000	1019.72	981.4	1000.56
2000	2020.81	1984.36	2002.59
3000	3019.32	2978.3	2998.81
4000	4020.56	3984.81	4002.68
5000	5021.24	4983.2	5002.22
6000	6021.56	5982.92	6002.24
7000	7026.85	6990.54	7008.69
8000	8023.45	7985.92	8004.68
9000	9028.42	8990.26	9009.34
10000	10040.78	10003.24	10022.01
Average (3975)	3995.16	3962.28	3978.72

The validation data in Table 2 were fitted by least-squares linear regression of the form Y = a + bX, where X is the F2 reference mass and Y is the load cell reading (g). The resulting calibration equations for Load Cell 1 and Load Cell 2 are given in Equations 22 and 23, respectively, and both show near-perfect linearity (R^2 = 0.99998 and 0.99999). The fitted lines against the F2 standard are plotted in Figure 14.

$$Y = a + b.X$$

Load Cell 1

$$b1 = \frac{\sum(Xi - 3975)(Y1i - 3995.16)}{\sum(Xi - 3975)^2} = 1.0022$$

$$a1 = \bar{Y} - b1.\bar{X} = 3995.16 - (1.0022 \,.\, 3975) = 11.32 \qquad [22]$$

$Y1 = 11.32 + 1.0022X$, R^2 = 0.99998

Load Cell 2

$$b2 = \frac{\sum(Xi - 3975)(Y2i - 3962.28)}{\sum(Xi - 3975)^2} = 1.0006$$

$$a2 = \bar{Y} - b2.\bar{X} = 3962.28 - (1.0006 \,.\, 3975) = -15.11 \qquad [23]$$

$Y2 = -15.11 + 1.0006\,X$, R^2 = 0.99999

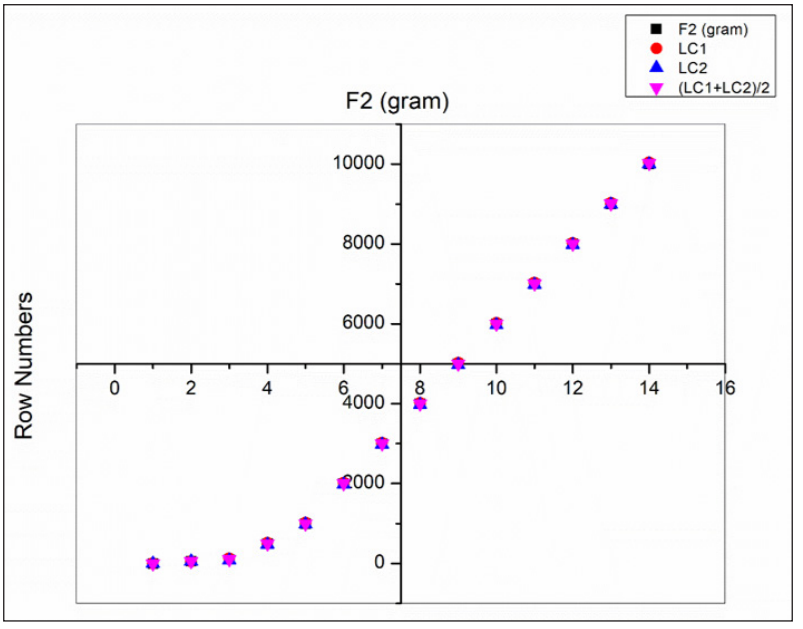


Figure 14. Validation Load Cell 1 and Load Cell 2 with F2 standard

Characterising Total Resistive Forces across Diverse Grouser Geometries and Contact Orientations

To characterise static soil resistance, the experimental procedure involves holding the wheel in a fixed state while rotating the assistive grouser axle from 0° to 180°. To ensure data reliability across 10 retrieval cycles, the displaced sand is manually flattened, and a wheel retainer is reinstalled before each trial. Bekker's original sand shear force measurements, which involved moving a rectangular plate in the longitudinal direction, could only measure the maximum force separately. Due to limitations in constructing a mechanism to approximate the cycloidal motion of the grouser, we measured the maximum shear force by rotating it half a rotation, with the goal of recording all shear forces generated by the grouser and the sand during its interaction.

As illustrated in Figure 15, the contact area angle is a decisive factor; geometric differences result in a reduction in resistive force of approximately 22.9% between C-Rear and Chevron. Statistical analysis using one-way ANOVA confirms that these differences are highly significant ($p < 0.001$, $\eta^2 = 0.999$), with the C-Front configuration yielding the highest average total resistance force at 138.7 N, followed by C-Rear (57.1 N) and Chevron (44.0 N), as detailed in Table 3. The one-way ANOVA results are summarised in Table 4, and the post-hoc Tukey HSD comparisons in Table 5 confirm that all three pairwise differences are significant ($p < 0.001$).

Table 3
Resistance force

Replication	C-Front (N)	C-Rear (N)	Chevron (N)
1	140	59	43
2	136	56	42
3	139	57	45
4	140	57	44
5	137	59	44
6	140	55	45
7	139	57	43
8	142	57	44
9	138	55	46
10	136	59	44
	Mean (N)	**SD (N)**	**SE (N)**
C-Front	138.7	1.95	0.62
C-Rear	57.1	1.52	0.48
Chevron	44.0	1.15	0.36

Table 4
One-way ANOVA of resistance force

Source of Variation	SS	df	MS	F	p-value
Between Groups	52661.0	2	26330.5	10611.0	<0.001
Within Groups	67.0	27	2.48	-	-
Total	52728.0	29	-	-	-

Table 5
Post-Hoc tukey HSD of resistance force

Comparison	Mean Difference (N)	p-value	Conclusion
C-Front vs C-Rear	+81.6	<0.001	Significant
C-Front vs Chevron	+94.7	<0.001	Significant
C-Rear vs Chevron	+13.1	<0.001	Significant

These measurements provide evidence that the total resistance force generated by each configuration with the same surface area differs in the direction of flow and in sand compaction, significantly affecting the resulting total resistance force, regardless of whether the surface area is tangential or lateral. In a 100% slip condition, the measured soil shear force equals the total resistance force, and it must be less than the net traction force. The net traction force is obtained from Equation 24 by subtracting the measured total resistance force from the gross tractive force of the assistive grouser shaft (621.18 N), giving 482.48 N for C-Front, 564.08 N for C-Rear and 577.18 N for Chevron as shown in Equations 25 to 27.

$$F_{net} = \tau_{shaft\ grouser} - Total\ Resistance\ Force \quad [24]$$

$$F_{net}\, C - Front = \frac{136.66\ N.m}{0.170\ m + 0.050\ m} - 138.7\ N = 482.48\ N \quad [25]$$

$$F_{net}\, C - Rear = \frac{136.66\ N.m}{0.170\ m + 0.050\ m} - 57.1\ N = 564.08\ N \quad [26]$$

$$F_{net}\, Chevron = \frac{136.66\ N.m}{0.170\ m + 0.050\ m} - 44.0\ N = 577.18\ N \quad [27]$$

This behaviour is theoretically grounded in the superposition principle of Granular Resistive Force Theory (GRFT), as proposed by Treers et al. (2021), which posits that the orientation of surface elements relative to the motion vector alters interaction patterns. Specifically, the forward lateral orientation of the C-Front design traps and accumulates sand, whereas the backward orientation of the C-Rear allows soil to flow toward the sides. These findings align with the classical theory of passive soil pressure, characterising clean Toyoura sand as having stable, high cumulative resistance at 100 mm penetration depths with minimal particle movement. Furthermore, the results corroborate research by Suzuki et al. (2020) regarding the functional tradeoff between resistive force magnitude and sinkage across different geometries.

Influence of Divergent Grouser Contact Angles on Tractive Effort Generation within Granular Media

Slip range of 21-23% at an angular speed of 1 RPM, it turns out that the need for soil shear force to hold the grouser to produce traction force on the assistive grouser mechanism with an angle of attack of 0 degrees, with a grouser height area of 100 mm, with 3 configurations, turns out to be smaller than static testing, assuming 100% slip occurs. The magnitude of the traction force equivalent to the soil shear force is reported as follows: according to the results visualised in Figure 16, the C-Front layout generates the highest maximum soil shear force at 38.4 N. In comparison, the C-Rear arrangement provides the next performance tier with a peak traction of 27.1 N. Statistically, this performance gap is highly significant ($p<0.001$), with the C-Front producing approximately 42% more traction than the C-Rear and 4.3 times as much as the Chevron design. Even the C-Rear configuration demonstrated a superior response, yielding nearly three times the force recorded for the Chevron system (9.0 N), with a detailed breakdown presented in Table 6. The one-way ANOVA results for tractive effort are summarised in Table 7 ($\eta^2 = 0.997$), and the post-hoc Tukey HSD comparisons in Table 8 confirm that all pairwise differences are significant ($p < 0.001$).

Table 6
Tractive effort

Replicate	C-Front (N)	C-Rear (N)	Chevron (N)
1	38	27	9
2	39	26	10
3	37	27	9
4	39	27	10
5	39	28	9
6	38	26	8
7	39	27	10
8	39	28	8
9	38	27	9
10	38	28	8
	Mean (N)	**SD (N)**	**SE (N)**
C-Front	38.4	0.7	0.22
C-Rear	27.1	0.74	0.23
Chevron	9.0	0.82	0.26

Table 7
One-way ANOVA of tractive effort

Source of Variation	SS	df	MS	F	p-value
Between Groups	4399.2	2	2199.6	3880.0	<0.001
Within Groups	15.3	27	0.57	-	-
Total	4414.5	29	-	-	-

Table 8
Post-Hoc Tukey HSD of tractive effort

Comparison	Mean Difference (N)	p-value	Conclusion
C-Front vs C-Rear	+11.3	<0.001	Significant
C-Front vs Chevron	+29.4	<0.001	Significant
C-Rear vs Chevron	+18.1	<0.001	Significant

These significant disparities are primarily attributed to the orientation of the lateral faces. Specifically, the forward-facing edges of the C-Front design facilitate a robust soil-trapping and compaction mechanism within the Toyoura sand, enabling the mechanism to collect and move soil particles more effectively than the more streamlined Chevron geometry.

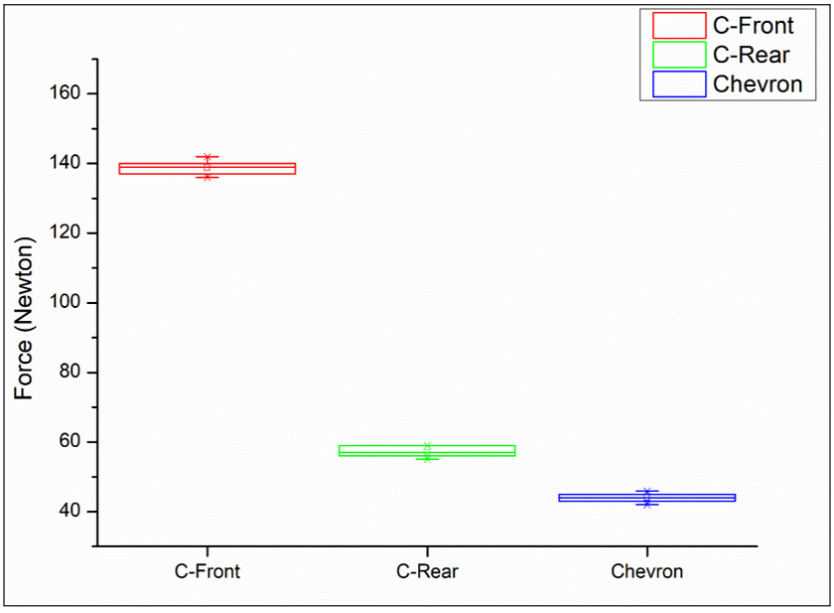


Figure 15. Peak resistive force response in Toyoura sand during static loading trials

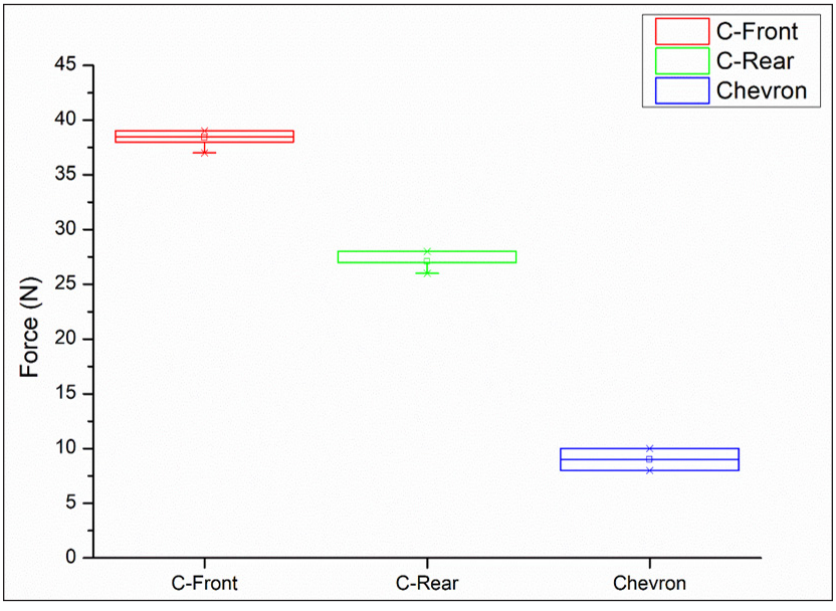


Figure 16. Magnitude of localised traction force development in Toyoura Sand

This research corroborates the findings of Aziz et al. (2023) on the role of contact angles while introducing critical methodological improvements. Unlike previous studies that lacked environmental controls, this investigation maintained a temperature of 26±1°C

and humidity of 60±2% to prevent moisture-induced pseudo-cohesion from affecting soil shear strength (Tapia et al., 2016). By implementing these rigorous controls and calibrated measurements, this study provides a more accurate and reliable data set than earlier reports based on raw values.

From both static tests with 100% slip and natural dynamic tests with 21-23% slip, it can be concluded that the traction force is generated by the grouser through shear-force compaction (tangential) and soil-bearing capacity (normal), without exceeding the total resistance force. If it exceeds that, the sand flows from beneath the wheel surface and collects behind the wheel.

Comprehensive Analysis of Force-Time Profiles Throughout the Assistive Grouser Operational Cycle

The advantage of using two Load Cells placed on one grouser is its ability to report the grouser deflection movement due to soil shear forces during the clockwise test. While previously assumed to be driven only from behind the wheel (in front of the grouser in its direction of travel), it turns out that the sand in front also contributes to generating traction force. The values derived from the force-time analysis represent instantaneous peak values from a single representative cycle, rather than trial averages.

The force-time history across all geometries reveals a consistent tri-phase cycle—penetration, traction generation, and retraction—within a 19-second operational window as reported in Figure 17. The penetration phase (6.0–6.4 s) is temporally uniform, though peak forces vary significantly: C-Front (+38 N), C-Rear (+28 N), and Chevron (+8 N), with all peaks occurring between the 3 and 4-second marks. These findings expand upon Aziz et al. (2023) by quantifying the interplay between contact angles and tractive effort through calibrated instrumentation.

During the traction phase, tractive effort increases progressively, peaking at −35 N for the C-Front, compared to −20 N (C-Rear) and −10 N (Chevron). This performance is driven by a superior soil-trapping mechanism—characterised by forward-facing lateral edges—that facilitates more effective soil collection and compaction. This mechanism not only increases force magnitude but also extends traction persistence; the phase duration expanded from 28.36% (Chevron) and 32.26% (C-Rear) to a maximum of 38.95% for the C-Front. This increased persistence enables the C-Front to dominate the net impulse, ensuring robust mobility on granular substrates such as Toyoura sand.

Experimental results registered positive penetration impulses of +26.5 N.s (Chevron), +83.2 N.s (C-Rear), and +112.4 N.s (C-Front). The peak negative impulse—representing the tractive effect as the grouser engages the substrate—was most pronounced in the C-Front design (−153.1 N.s), followed by the C-Rear (−50.0 N.s) and Chevron (−24.3 N.s). The net impulse generated by the C-Front was −40.7 N.s, followed by the Chevron

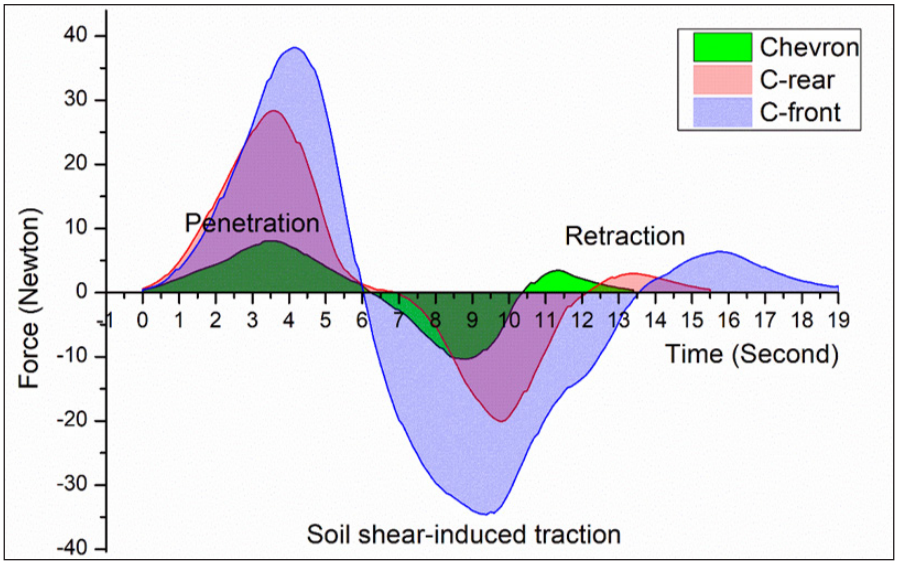


Figure 17. Sequential three-phase operational cycle of the assistive grouser mechanism

and C-Rear configurations of +2.2 N.s and +33.2 N.s, respectively, facilitating smoother forward motion.

The maximum change in position, with a value of + to -, is the magnitude of the generated traction force, which is the average value of the total amplitude formed by one cycle because of the interaction of the sheer force of the sand in front of and behind the grouser.

These disparities highlight unique force-response patterns within a single wheel–ground interaction cycle. The C-Front configuration excels at initial penetration and generates sustained traction, but it also triggers significant densification and shear failure in the granular material. This behaviour aligns with Terramechanics principles, in which increased effective contact area reinforces soil slice formation while elevating compaction resistance and energy dissipation. Conversely, the Chevron and C-Rear designs exhibit lighter ground interactions; while their peak forces are lower, they facilitate a faster release of granular material, optimising efficiency.

Ultimately, these findings underscore a fundamental trade-off between maximum penetration capacity and force-generation efficiency. An optimal assistive grouser must balance aggressive soil engagement with the ability to efficiently release compacted materials, a critical factor for wheeled robot mobility in granular terrain like Toyoura sand.

CONCLUSION

This study demonstrates that assistive grouser geometry, specifically contact angle and orientation, significantly dictates traction performance on Toyoura sand. The C-Front

configuration (180° angle, forward-facing) proved most effective, generating a peak traction force of 38.4 N by forming a robust passive soil wedge. In contrast, the C-Rear and Chevron geometries produced lower forces (27.1 N and 9.0 N, respectively), a disparity confirmed as highly significant ($p<0.001$). Force-time analysis identified a consistent tri-phase interaction:

- Penetration: Lasting 6.0–6.4 s, with the C-Front reaching a peak force of +38 N.
- Traction: The C-Front maintained the highest persistence, occupying 38.95% of the cycle compared to only 28.36% for the Chevron.
- Retraction: All geometries exhibited force restoration as soil was displaced toward the surface.

These results highlight a fundamental mechanical trade-off: aggressive designs like the C-Front maximise soil engagement but increase compaction resistance, resulting in a negative local net impulse (−40.7 N.s). However, the assistive mechanism's independent angle control ensures that the overall system's mobility remains superior to that of conventional designs. Optimising the balance between aggressive engagement and efficient material release is essential for traversing granular terrain. These findings establish design guidelines for high-traction wheels. Future work will extend the evaluation across multiple normal loads and rotational speeds and across the full slip curve from 0% to 100%, rather than the single operating point examined here, and will apply Particle Image Velocimetry particle tracking to image the internal soil velocity field directly, so that the geometric soil confinement mechanism inferred in this study can be confirmed by observation of soil flow rather than by force response alone.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Higher Education Malaysia for financial support under the Fundamental Research Scheme (FRGS) No. FRGS/1/2023/TK10/UMP/02/37 (University reference RDU230135), Ministry of Education Malaysia, Ministry of Education Indonesia, University of Jember, and Universiti Malaysia Pahang Al-Sultan Abdullah for the facilities and support provided for this research.

LIST OF ABBREVIATION

SWTB	Single-Wheel Test Bed
2-D	Two-Dimensional
GPS	Global Positioning System
INS	Inertial Navigation System

3-D	Three-Dimensional
ANOVA	Analysis of Variance
CCD	Continuous Collision Detection
MBS	Multi-Body System
SCM	Soil Contact Models
PIV	Particle Image Velocimetry
PLA	Poly Lactic Acid
DAQ	Data Acquisition
USB	Universal Serial Bus
CSV	Comma-Separated Values
ADC	Analog-to-Digital Converter
LC	Load Cell
GRFT	Granular Resistive Force Theory
RPM	Revolutions Per Minute
FRGS	Fundamental Research Scheme

REFERENCES

Aziz, I. N. A. C., Ibrahim, A. N., Basri, I., & Fukuoka, Y. (2023). Rover wheel assistive grouser angle of attack effects on traction force in soft terrain. In *2023 19th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 236-241). IEEE. https://doi.org/10.1109/cspa57446.2023.10087721

Basri, I., Aziz, I. N. A. C., Ibrahim, A. N., & Fukuoka, Y. (2024). Investigation of wheel robot grouser width parameter effect on robot mobility in soft sand terrain using sand test bed. In W. H. Mohd. Isa, I. M. Khairuddin, M. A. Mohd. Razman, S. Saruchi, S. H. Teh, & P. Liu (Eds.), *Intelligent manufacturing and mechatronics* (Vol. 850, pp. 199-207). Springer. https://doi.org/10.1007/978-981-99-8819-8_16

Bekker, M. G. (1956). *Theory of land locomotion*. University of Michigan Press.

Blatnický, M., Dižo, J., Gerlici, J., Sága, M., Lack, T., & Kuba, E. (2020). Design of a robotic manipulator for handling products of automotive industry. *International Journal of Advanced Robotic Systems, 17*(1), Article 1729881420906290. https://doi.org/10.1177/1729881420906290

Boresi, A. P., & Schmidt, R. J. (2001). *Engineering mechanics: Statics*. Brooks/Cole.

Carow, C., & Rackwitz, F. (2023). Cyclic compression of sands at constant stress ratio—experimental investigations and constitutive modelling. *Acta Geotechnica, 18*(3), 1305-1330. https://doi.org/10.1007/s11440-022-01628-z

Demishkevich, A., Thai, B. T., & Skonieczny, K. (2025). Advantages of slanted grousers for skid-steer planetary rovers with rigid wheels. *Mechatronics, 112*, Article 103408. https://doi.org/10.1016/j.mechatronics.2025.103408

Elsheikh, M. A. (2023). Design of a special rigid wheel for traversing loose soil. *Scientific Reports, 13*(1), Article 171. https://doi.org/10.1038/s41598-022-27312-6

Guo, J., Li, W., Ding, L., Gao, H., Guo, T., Huang, B., & Deng, Z. (2021). Linear expressions of drawbar pull and driving torque for grouser-wheeled planetary rovers without terrain mechanical parameters. *IEEE Robotics and Automation Letters, 6*(4), 8197-8204. https://doi.org/10.1109/lra.2021.3103641

Higa, S., Nagaoka, K., & Yoshida, K. (2019). Stress distributions of a grouser wheel on loose soil. *Journal of Terramechanics, 85*, 15-26. https://doi.org/10.1016/j.jterra.2019.07.001

Higa, S., Sawada, K., Nagaoka, K., Nagatani, K., & Yoshida, K. (2016). Measurement of stress distributions of a wheel with grousers traveling on loose soil. In *2016 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2828-2833). IEEE. https://doi.org/10.1109/icra.2016.7487445

Hruboš, M., Svetlík, J., Nikitin, Y., Pirník, R., Nemec, D., Šimák, V., Janota, A., Hrbček, J., & Gregor, M. (2016). Searching for collisions between mobile robot and environment. *International Journal of Advanced Robotic Systems, 13*(5), Article 1729881416667500. https://doi.org/10.1177/1729881416667500

Ibrahim, A. N., Aoshima, S., Shiroma, N., & Fukuoka, Y. (2016). The effect of assistive anchor-like grousers on wheeled rover performance over unconsolidated sandy dune inclines. *Sensors, 16*(9), Article 1507. https://doi.org/10.3390/s16091507

Ibrahim, A. N., & Fukuoka, Y. (2018). Effect of assistive grouser mechanism on lightweight rover power consumption pattern on steep soft sand inclines. In M. H. A. Hassan (Ed.), *Intelligent manufacturing & mechatronics* (pp. 343-351). Springer. https://doi.org/10.1007/978-981-10-8788-2_30

Ishikawa, T., Zhang, Y., Tokoro, T., & Miura, S. (2014). Medium-size triaxial apparatus for unsaturated granular subbase course materials. *Soils and Foundations, 54*(1), 67-80. https://doi.org/10.1016/j.sandf.2013.12.007

Kelemenová, T., Dovica, M., Božek, P., Koláriková, I., Benedik, O., Virgala, I., Prada, E., Miková, Ľ., Kot, T., & Kelemen, M. (2020). Specific problems in measurement of coefficient of friction using variable incidence tribometer. *Symmetry, 12*(8), Article 1235. https://doi.org/10.3390/sym12081235

Liu, J., Liu, X., Zhang, X., Chen, X., Li, H., Luan, L., Ding, C., & Gao, X. (2024). Experimental study on reinforcement of bionic grouser of deep-sea mining vehicle. *Biogeotechnics, 2*(3), Article 100088. https://doi.org/10.1016/j.bgtech.2024.100088

Nakamura, H., Nagaoka, K., & Yoshida, K. (2020). Soil flow analysis for grouser wheels based on a particle image velocimetry method. *Journal of Terramechanics, 91*, 233-241. https://doi.org/10.1016/j.jterra.2020.07.001

Onur, M. İ., Tuncan, M., & Tuncan, A. (2014). An experimental study for determining the shear modulus of Toyoura sand. In *Proceedings of the Second European Conference on Earthquake Engineering and Seismology* (Paper 829). https://eaee.org/Media/Default/2ECCES/2ecces_eaee/829.pdf

Qazizada, M. E., & Pivarčiová, E. (2016). Mobile robot controlling possibilities of inertial navigation system. *Procedia Engineering, 149*, 404-413. https://doi.org/10.1016/j.proeng.2016.06.685

Rubio, F., Valero, F., & Llopis-Albert, C. (2019). A review of mobile robots: Concepts, methods, theoretical framework, and applications. *International Journal of Advanced Robotic Systems, 16*(2), Article 1729881419839596. https://doi.org/10.1177/1729881419839596

Sabarudin, S. S., Ibrahim, A. N., & Fukuoka, Y. (2020a). Effect of grouser angle of attack on performance of adjustable robot wheel assistive grouser. In M. Osman Zahid, R. Abd. Aziz, A. Yusoff, N. Mat Yahya, F. Abdul Aziz, & M. Yazid Abu (Eds.), *iMEC-APCOMS 2019* (pp. 605-610). Springer. https://doi.org/10.1007/978-981-15-0950-6_92

Sabarudin, S. S., Ibrahim, A. N., & Fukuoka, Y. (2020b). Investigation of effect of assistive grouser attack angle when moving in unconsolidated soft sand inclination slope using discrete element method (DEM) simulation. In *2020 IEEE Student Conference on Research and Development (SCOReD)* (pp. 149-153). IEEE. https://doi.org/10.1109/scored50371.2020.9250997

Safdar, M., Newson, T., & Shah, F. (2021). Consolidated drained (CID) behavior of fibre reinforced cemented Toyoura sand in triaxial loading conditions. *International Journal of Geo-Engineering, 12*, Article 27. https://doi.org/10.1186/s40703-021-00165-0

Sakayori, G., & Ishigami, G. (2024). Modeling of slip rate-dependent traversability for path planning of wheeled mobile robot in sandy terrain. *Frontiers in Robotics and AI, 11*, Article 1320261. https://doi.org/10.3389/frobt.2024.1320261

Schäfer, B., Gibbesch, A., Krenn, R., & Rebele, B. (2010). Planetary rover mobility simulation on soft and uneven terrain. *Vehicle System Dynamics, 48*(1), 149-169. https://doi.org/10.1080/00423110903243224

Sutoh, M., Yusa, J., Ito, T., Nagatani, K., & Yoshida, K. (2012). Traveling performance evaluation of planetary rovers on loose soil. *Journal of Field Robotics, 29*(4), 648-662. https://doi.org/10.1002/rob.21405

Suzuki, H., Watanabe, Y., Kobayashi, T., Iagnemma, K., & Ozaki, S. (2020). Terramechanics-based investigation of grouser shape for rigid wheels: Comparison between rectangular and trapezoidal grousers. *Cogent Engineering, 7*(1), Article 1846254. https://doi.org/10.1080/23311916.2020.1846254

Tagliavini, L., Colucci, G., Botta, A., Cavallone, P., Baglieri, L., & Quaglia, G. (2022). Wheeled mobile robots: State of the art overview and kinematic comparison among three omnidirectional locomotion strategies. *Journal of Intelligent & Robotic Systems, 106*(3), Article 57. https://doi.org/10.1007/s10846-022-01745-7

Tapia, F., Santucci, S., & Géminard, J. C. (2016). Fracture reveals clustering in cohesive granular matter. *EPL (Europhysics Letters), 115*(6), Article 64001. https://doi.org/10.1209/0295-5075/115/64001

Treers, L. K., Cao, C., & Stuart, H. S. (2021). Granular resistive force theory implementation for three-dimensional trajectories. *IEEE Robotics and Automation Letters, 6*(2), 1887-1894. https://doi.org/10.1109/lra.2021.3057052

Watanabe, Y., Nakano, S., Suzuki, H., & Ozaki, S. (2023). Numerical analysis of wheel locomotion on soft soil using the extended terramechanics model based on cellular automata. *Journal of Terramechanics, 109*, 9-20. https://doi.org/10.1016/j.jterra.2023.05.005

Wong, J. Y. (1997). Dynamics of tracked vehicles. *Vehicle System Dynamics, 28*(2-3), 197-219. https://doi.org/10.1080/00423119708969354

Yokoyama, A., Nakashima, H., Shimizu, H., Miyasaka, J., & Ohdoi, K. (2020). Effect of open spaces between grousers on the gross traction of a track shoe for lightweight vehicles analyzed using 2D DEM. *Journal of Terramechanics, 90*, 31-40. https://doi.org/10.1016/j.jterra.2019.10.004

Zhang, W., Lyu, S., Xue, F., Yao, C., Zhu, Z., & Jia, Z. (2022). Predict the rover mobility over soft terrain using articulated wheeled bevameter. *IEEE Robotics and Automation Letters, 7*(4), 12062-12069. https://doi.org/10.1109/lra.2022.3211153

Pertanika J. Sci. & Technol. 34 (5): 2649 - 2673 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Formulation of Regulatory Sandbox for Nanotechnology Innovation

Nor Akhmal Hasmin[1]*, Juan Matmin[2], and Ainul Hafiza Zainudin[3]

[1]Faculty of Law, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[2]Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[3]Faculty of Law, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

ABSTRACT

Regulatory sandboxes represent a modern approach to regulatory governance for testing products, services, or programs derived from emerging technologies. Testing enables the identification of risks before releasing innovations to consumers. Innovators are allowed to test their innovation within a controlled environment under the supervision of regulators, without being fully subjected to standard regulations. Currently, there are about 200 regulatory sandboxes worldwide across 92 countries in finance, energy, telecommunications, and health. This widespread adoption suggests that regulatory sandboxes are regarded as a mechanism for promoting innovation and enabling regulators to manage potential risks. While the use of regulatory sandboxes for technological innovation continues to expand, there is currently no equivalent framework tailored specifically to nanotechnology innovations. Nanotechnology is associated with regulatory gaps and scientific uncertainties regarding safety and health risks. This article conceptualises the general framework for designing a regulatory sandbox for nanotechnology innovations. The formulation of this sandbox is based on the conceptual analysis of the essential components and key features adopted by international institutions such as the European Commission (EC) and the Organisation for Economic Co-operation and Development (OECD), as well as those proposed by scholars in designing regulatory sandboxes. The essential components of a regulatory sandbox for nanotechnology innovation identified in this article include the establishment of competent authorities, the evaluation of legal impediments, and the identification of sandbox prerequisites, namely eligibility criteria, the sandbox's scope, duration, monitoring and reporting mechanisms,

ARTICLE INFO

Article history:
Received: 02 July 2026
Accepted: 26 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.05

E-mail addresses:
norakhmal5559@uitm.edu.my (Nor Akhmal Hasmin)
juanmatmin@utm.my (Juan Matmin)
ainulhafiza989@uitm.edu.my (Ainul Hafiza Zainudin)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

and an exit strategy. Fundamentally, this article contributes to the ongoing discourse on innovation-friendly laws and sustainability for nanotechnology development.

Keywords: Innovation, nanotechnology, regulatory sandbox, sustainable nanotechnology development

INTRODUCTION

In the 21st century, nanotechnology has emerged as the most transformative technology offering innovative solutions across various industries. The ability to manipulate materials and devices at the nanoscale unlocks distinctive physical, chemical, and biological characteristics. This enables innovators to develop products that demonstrate enhanced efficiency, improved design, optimised processes, and superior performance (Kumar et al., 2025; Malik et al., 2023). These advantages are applicable across a broad spectrum of industries, highlighting the transformative potential of nanotechnology innovation.

In 2024, the global nanotechnology market was valued at USD 84.7 billion, and within ten years, the value is projected to increase by USD 26.5 billion (Emergen Research, 2026). The growth, inter alia, is attributable to the rapid pace of nanotechnology innovation in countries such as China, South Korea, and the United States. Likewise, Malaysia's nanotechnology market size is estimated to reach USD 8.5 billion in 2033 (NanoMalaysia, 2024). The characteristics of nanotechnology, supported by a large market size, have stimulated firms to intensify R&D and consequently increase the volume of nanotechnology innovations.

From the innovation law perspective, nanotechnology innovations represent a typical tension of innovation and safety. Despite its significant economic benefits, nanotechnology is considered a controversial technology owing to its scientific novelty (Zabeo et al., 2019). The safety and health risks remain highly uncertain and potentially severe. Scientific studies associate the utilisation of engineered nanomaterials (ENMs) with harmful interactions in biological systems and the environment, raising concerns about toxicity (Borm et al., 2006; Chávez-Hernández et al., 2024; Ma et al., 2024). Such concerns require robust regulatory oversight to safeguard consumers' well-being and the environment from unintended and potentially harmful risks. Consequently, the comprehensive framework will uphold the ethical and sustainable development of nanotechnology.

In certain jurisdictions, such as Malaysia, there is a notable absence of legislation regulating the proliferation of nanotechnology across various industries. For instance, engineered nanomaterials (ENMs) in food products are regulated by the Food Act 1983 [*Act 281*], which contains no express provisions relating to nanotechnology or nanomaterials. As a result, nanofood falls within the regulatory framework enacted before the emergence of nanotechnology in the food processing industry. The physicochemical properties of ENMs differ substantially from those of bulk materials, making nanofood unsuitable for

regulation under legislation designed for conventional bulk-scale materials. As a result, nanofoods are being introduced in the market without adequate regulatory oversight. This regulatory gap allows these products to become available to consumers before there is a comprehensive understanding of their long-term safety for human consumption (Hasmin et al., 2024).

Given the regulatory gaps and uncertainties surrounding safety and health risks, this article argues for the development of a regulatory sandbox for nanotechnology innovations. The growing market size for nanotechnology innovations exemplifies the need for a regulatory sandbox. The market expansion is closely linked to heightened innovation activity driven by the proliferation of novel products derived from ENMs or nanodevices. Originally introduced in the FinTech sector, the regulatory sandbox is now gaining traction among innovators in other industries (Johnson, 2023; Qiu et al., 2025). They allow innovators to test new products in a controlled environment, identify opportunities and risks, and facilitate real-world deployment without being fully subject to standard regulations. At the same time, regulators can use insights from these experiments to refine legal frameworks and grant approvals (Federal Ministry for Economic Affairs and Energy, 2025).

The European Commission (EC) (2023) recognised the regulatory sandbox as a tool for identifying potential regulatory barriers or new laws that affect innovation. A regulatory sandbox for nanotechnology can facilitate the testing and oversight of new products within a controlled environment, thereby addressing the regulatory gaps and ensuring the safety and health of consumers and the environment (Federal Ministry for Economic Affairs and Energy, 2025). The central objective of this article is to conceptualise the design of the regulatory framework for nanotechnology innovation.

Currently, there is no equivalent framework tailored specifically to nanotechnology innovations. A systematic search was conducted in Scopus and Web of Science using the advanced search function for article titles, abstracts, and keywords. The search used the terms “nanotechnology”, “regulatory sandbox”, and “regulatory sandboxes”, and identified three documents. However, these articles do not focus on the formulation of a regulatory sandbox for nanotechnology innovation. The article by Johnson (2023) offers a general framework for characterising regulatory sandboxes across different sectors. In contrast, Safarzadeh and Di Maria (2026) and Maric et al. (2026) focus on the importance of regulatory sandboxes in encouraging nanotechnology innovation.

This article advances a general framework applicable to nanotechnology innovation, without focusing on any product category. The proposed sandbox aims to address several key objectives: reduce regulatory gaps, foster nanotechnology innovation, safeguard consumer protection, and promote ethical and sustainable nanotechnology development. This article is divided into three parts. The first part presents the research methodology. The second part offers an overview of nanotechnology innovations. The discussion in

this section aims to understand the fundamental nature and dynamics of innovation in nanotechnology, thereby informing the design of the sandbox. The third part is the characteristics of a regulatory sandbox, including the design of the regulatory sandbox, the structure of the regulatory authority responsible for the sandbox, and the sandbox's design. The final part is the conclusion.

METHODOLOGY

In formulating the regulatory sandbox for nanotechnology innovation, this article adopts a qualitative doctrinal research approach by analysing relevant literature from reports, research journals, and books available in online databases and on institutional websites. A content analysis method was used to critically examine the design of regulatory sandboxes. The analysis was primarily guided by the Organisation for Economic Co-operation and Development (OECD) Regulatory Sandbox Toolkit, which provides comprehensive guidance for regulators on establishing and managing regulatory sandboxes. In addition, the study examined the European Union's regulatory sandbox for artificial intelligence, as the framework is expressly enacted into legislation, namely the EU Artificial Intelligence Act. The analysis was further supplemented by academic literature proposing frameworks and key elements for establishing regulatory sandboxes. This method focused on identifying the key components and features required for designing a regulatory sandbox. Even though reference is made to existing frameworks such as AI and FinTech, the relevant features are critically adapted and refined to reflect the distinctive nature of nanotechnology innovation.

NATURE OF NANOTECHNOLOGY INNOVATION

Nanotechnology has existed since at least the fourth century. The modern era of nanotechnology began in the 1980s and has continued to evolve to date (Pisano & Durlo, 2025). The term nanotechnology refers to the size scale of objects or materials, expressed in nanometres (nm) or the nanoscale. The nanoscale length range is approximately 1 nm to 100 nm (International Organization for Standardization, 2015). The nanometre is about 100,000 times smaller than the diameter of a human hair or 1,000 times smaller than a red blood cell (Science Policy Council, 2007). Hence, it is impossible to see nano-sized particles using the naked eye or with a conventional microscope.

The core components of nanotechnology innovation rested in the development and application of nanomaterials and nanodevices, as illustrated in Figure 1. The Organisation for Economic Co-operation and Development (2024) defines nanomaterials as chemicals that are either in the form of nano-objects or nano-structured. A nano-object is defined as a material confined in one, two, or three dimensions at the nanoscale, and nano-structured means having an internal or surface structure at the nanoscale. Conventional bulk-sized materials can be engineered into a nanoscale and referred to as engineered

nanomaterials (ENMs). The tiny size offers novel ways to manipulate, control, and formulate materials, structures, devices, and systems that differ greatly from their bulk counterparts. Such a distinction is mainly attributable to novel physicochemical properties, including enhanced reactivity, quantum effects, increased surface area, and electrical conductivity (Asha & Narain, 2020). Examples of ENMs include nanoparticles, nanofibers, and nanoplatelets such as carbon nanotubes, graphene, quantum dots, gold nanoparticles, and metal oxides, which are widely used in various industries.

Other than nanomaterials, the utilisation of nanotechnology also consists of nanodevices. These refer to nanoscale technologies that integrate mechanical, electronic, magnetic, and biological systems. Nanodevices may include nanosensors, nanoelectromechanical systems (NEMS), chip technologies, and implantable biomedical devices, which are commonly used in energy harvesting, medicine, the semiconductor industry (electronic devices across various sectors), environmental control, and countless other applications (Bhattacharjee & Maiti, 2021). As of January 2026, the Nanotechnology Products Database classified nanotechnology products into 15 categories. These categories encompass diverse sectors, including electronics, medicine, construction, cosmetics, textiles, automotive, environmental, renewable energy, food, home appliances, petroleum, agriculture, printing, sports and fitness, and others.

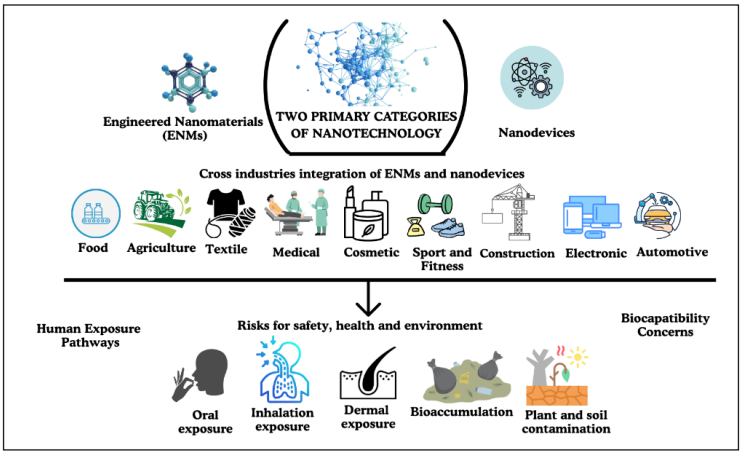


Figure 1. The applications of nanotechnology and potential risks to safety, health, and the environment

As shown in Figure 1, the applications of ENMs in certain product categories, such as foods, cosmetics, personal care, medicines, and textiles, allow these materials to come into intimate contact with the human body via oral, dermal, or inhalation exposure routes. Exposure of the human body to ENMs is associated with various safety and health concerns, as demonstrated by laboratory studies. The tiny particle has a greater surface area-to-volume ratio, and a larger surface area leads to higher chemical reactivity, thereby

increasing its ability to penetrate human tissue (Pathak et al., 2025; Xie et al., 2024). Likewise, nanomaterials are more adhesive to human tissue surfaces than larger particles. They cannot be easily removed from the lungs, gastrointestinal tract or other organs compared with conventional bulk counterparts, resulting in potential accumulation within the body (Fiordaliso et al., 2022; Oberdörster et al., 2005).

Scientific studies using in vitro, in vivo, and in silico experiments demonstrated diverse toxicology effects of ENMs on human safety and health, such as reducing the capacity of the intestinal tract (Carbone et al., 2016), damaging the intestinal glands (Shahare et al., 2013), harmful effects to liver and kidney (Baki et al., 2014), membrane damage (Cid-Samamed et al., 2023), oxidative stress (Zhang et al., 2021), respiratory inflammation (Halappanavar et al., 2021), immune dysregulation (Tang et al., 2021), and metabolic disturbances (Wojciechowska et al., 2023). Collectively, these studies provide important scientific evidence of the potential toxicity of ENMs and their possible adverse effects on human health and safety.

The utilisation of nanodevices does not involve direct contact with internal organs or tissues, as it does not cross human biological barriers. The potential adverse impacts are predominantly linked to environmental risks arising from the use of non-biocompatible devices. ENMs may enter the environment either through direct or indirect pathways. Direct release of ENMs into the environment can occur through agricultural activities, such as incorporating ENMs into agricultural products such as fertilisers and pesticides (Servin et al., 2015). On the contrary, indirect releases arise throughout the life cycle of ENMs, including manufacturing, consumer use, and end-of-life disposal, resulting in their discharge into air, water, and soil (Giese et al., 2018). These release pathways make the accumulation of ENMs in soil and groundwater effectively unavoidable, leading to soil and water contamination (Kokilavani et al., 2023). Therefore, the introduction of ENMs into products must be rigorously monitored to mitigate the potential risks to safety, health, and the environment.

Based on the above analysis, the design for the nanotechnology regulatory sandbox must take into consideration the nature and specific modes of nanotechnology utilisation. This includes whether the innovation involves the utilisation of ENMs or nanodevices. Furthermore, it is essential to evaluate the potential impacts of these technologies on human safety, health, and the environment. The design of the regulatory sandbox should be tailored to address these unique characteristics of nanotechnology.

General Description of Regulatory Sandboxes

A regulatory sandbox is a legal framework that allows innovators to test and experiment with innovative products, services, programs, or businesses with actual consumers under the supervision of regulatory authorities for limited periods (Madiega & Van De Pol, 2022).

It operates like a real laboratory for testing new products and brings together stakeholders to ensure responsible, ethical technological innovation by protecting consumers from unintended risks and losses. The regulatory sandbox has been established as a responsive regulatory framework to address the potential risks posed by fast-moving technological innovations (Allen, 2019; Attrey et al., 2020; Ringe & Ruof, 2020). A landmark adoption of a regulatory sandbox was to nurture FinTech innovations. The United Kingdom's Financial Conduct Authority (FCA) is the world's pioneer for the regulatory sandbox, launched in 2016. This initiative allows FinTech companies to test new solutions with real consumers under controlled conditions. The success of the FCA regulatory sandbox has been replicated in nearly 60 jurisdictions worldwide (Fahy, 2022; Jeník & Duff, 2020). Later, lessons from regulatory sandboxes for FinTech have been applied in other industries, such as AI (Artificial Intelligence Act, 2024), energy sectors (Aydın & Yardımcı, 2024), blockchain, and the Internet of Things (IoT) (Durham, 2023).

According to Parenti (2020), the regulatory sandbox benefited not only innovators but also regulators and consumers. The innovator-driven regulatory sandbox allows them to test their innovations in a controlled environment, i.e., under the watchful eyes of the regulators, to help mitigate the unintended consequences and reduce legal risks. Likewise, the worldwide presence of regulatory sandboxes has positively promoted technological innovation (Gumbo & Chude-Okonkwo, 2025), influenced venture investment in innovation (Goo & Heo, 2020), and reduced time to market (Alaassar et al., 2020). Meanwhile, a regulator-driven regulatory sandbox helps regulatory authorities assess whether the current regulatory framework is effective in regulating innovation arising from emerging technologies.

The sandbox is seen as pivotal to modern regulatory governance, especially for new technologies. The formulation of a regulatory sandbox for nanotechnology innovation reflects the adoption of anticipatory governance (Morgan, 2023). In an era of rapid technological advancement, anticipatory governance serves as an important decision-making approach that prepares society for future challenges and risks. It does so by actively exploring, experimenting with, and continuously learning about emerging technological developments (Guston, 2014). A regulatory sandbox reflects anticipatory governance by encouraging innovators, policymakers, regulators, and other stakeholders to reflect on their roles and responsibilities in shaping nanotechnology development.

The OECD also emphasises the importance of incorporating a robust consumer protection mechanism in a regulatory sandbox (Organisation for Economic Co-operation and Development, 2025). From consumers' perspective, the sandbox enables them to access the latest innovations in a supervised environment, without being exposed to the typical uncertainties associated with unregulated technological experimentation. The mechanisms include a requirement that the measure under test have a beneficial effect on consumers and the disclosure of potential risks (Poncibò & Zoboli, 2020). The potential

risks can be properly monitored, preventing further harm to consumers. Furthermore, the transparency of sandboxes is also high. The participating firms are required to disclose comprehensive information regarding product performance, limitations, and potential risks. Such transparency enhances consumer trust and addresses information asymmetry, a persistent challenge in markets driven by emerging technologies.

As explained in the previous subtopics, nanomaterials research in various industries is highly associated with scientific uncertainties. There are major gaps in knowledge regarding the potential safety and health risks associated with ENMs (Heidmann & Milde, 2013). Currently, the safety and risks associated with ENMs across various industries are only assessed through in vitro, in vivo, and in silico studies conducted in laboratories. There are concerns that these laboratory studies are not validated, are contradictory or inconsistent, or are not reproducible. Furthermore, the use of animal data from in vivo studies is debatable because test results often cannot be transferred to humans due to potential biological differences between animals and humans (Jahnel, 2015). Therefore, the regulatory sandbox for nanotechnology offers a practical way to identify or measure safety and health risks associated with ENMs beyond lab-based studies. Additionally, the data generated can be used to collaterally produce information relevant to the authorisation process for products that require post-market approval.

Furthermore, there is a lack of systematic, mandatory reporting to identify ENM-related accidents. Although there are accident cases involving ENMs, it was often difficult to find detailed information, leading to uncertainties on the actual number of accident cases, the nature or physicochemical properties of ENMs that contribute to the accidents, and the nature of safety and risks (Organisation for Economic Co-operation and Development, 2022). This gap hampers effective risk governance of ENMs due to the refusal to enact legislation to govern potential safety and health risks. For instance, in Malaysia, nanotechnology risks are managed through soft law, with no specific legislation or amendments addressing nanomaterials, as safety concerns have only been observed in laboratory settings. The country relies on international and local guidance documents. Although the United States enacted the 21st Century Nanotechnology Research and Development Act of 2003, the aim is to coordinate nanotechnology programs and R&D under the National Nanotechnology Initiative (NNI).

On the other hand, the European Union has incorporated specific provisions on nanotechnology, nanomaterials, and nanoparticles in regulations for cosmetics and foods. These provisions aim to regulate nanotechnology activities and protect consumers against the potential risks (Angelopoulou et al., 2022). The incorporation of nanotechnology-specific provisions is supported by the European Commission's adoption of the precautionary principle in regulating nanotechnology activities (Heselhaus, 2010). The precautionary principle is a legal principle that allows regulators to take the earliest possible measures when there is a reasonable indication of risks to the environment, safety, and health, and

in which lack of scientific certainty is not a reason for inaction (Todt & Luján, 2014). For the European Union, the regulatory sandbox for nanotechnology innovation aligns with the precautionary principle that has been expressly adopted in regulating nanotechnology activities across member countries.

The article argues that regulatory sandboxes are justified for nanotechnology innovation because of scientific and legal uncertainties. These sandboxes can help address those challenges. With the systematically recorded and reported data from the sandbox, innovators could identify the potential safety and health risks associated with their innovations. The data collected from the sandbox could overcome the limitations of in vivo, in vitro, and in silico studies and strengthen consumer protection by providing evidence on how the innovations perform under a controlled environment.

Additionally, cooperation in the mutual exchange of expertise between innovators and regulators in the sandbox enables assessment of the adequacy of the regulatory framework, ensuring that existing rules become more resilient and that unregulated activities enter a regulated perimeter. The regulatory sandbox for nanotechnology innovation also serves as a mechanism aligned with the precautionary principle for governing the scientific uncertainties associated with nanotechnology products. The sandbox will accelerate nanotechnology innovation, increase nanotechnology investment, and promote responsible nanotechnology development. The following discussion is on the design of a regulatory sandbox for nanotechnology innovations.

The Design of a Regulatory Sandbox for Nanotechnology Innovation

According to Jeník and Duff (2020), regulatory sandboxes vary in design, as no two sandboxes are the same. The most important consideration in designing a regulatory sandbox is that it should serve its purpose and be customised to the nature of the technological products. The discussion in this section outlines the elements to consider when implementing a sandbox, including the organisational and governance of the sandbox and the operational prerequisites.

Organisational and Governance of Regulatory Sandbox for Nanotechnology Innovation

The design of a regulatory sandbox begins with establishing competent authorities to supervise and govern the sandbox. These authorities play a central role in ensuring effective oversight throughout the sandbox's operation and management, facilitating coordinated governance and compliance with applicable regulatory requirements. The European Union Artificial Intelligence Act (AI Act) may serve as a useful reference in developing the framework for the competent authorities responsible for the nanotechnology innovations sandbox.

Article 70 designates national competent authorities (NCAs) for the AI regulatory sandbox. The NCA for the AI sandbox shall consist of at least a notifying authority and a market surveillance authority that exercises its power independently. Article 3(19) states that the notifying authority is responsible for establishing and carrying out the necessary procedures for the assessment, designation, and notification of conformity assessment bodies, as well as for their monitoring. Additionally, Article 70(3) contains express provisions mandating that the NCA be equipped with adequate technical, financial, and human resources, as well as comprehensive infrastructure, to effectively fulfil its responsibilities.

This article suggests that the framework for competent authorities overseeing the nanotechnology innovation sandbox should also consist of (i) the notifying authority and (ii) the market surveillance authority. The illustration in this subtopic depicts the nanotechnology ecosystem in Malaysia, as shown in Figure 2. The notifying authority should comprise national agencies responsible for coordinating, developing, and governing the regulatory sandbox and nanotechnology. For instance, in Malaysia, the National Technology and Innovation Sandbox (NTIS) and the National Nanotechnology Centre (NNC), under the Ministry of Science, Technology, and Innovation (MOSTI), must be among the notifying authorities. The NTIS is responsible for creating a sandbox for innovators across sectors, including capacity building, funding facilitation, regulatory pathway facilitation, and commercialisation facilities (National Technology & Innovation Sandbox, n.d.). The NNC focuses on coordinating research, technology development activities, and standards and regulatory issues related to nanotechnology in Malaysia (National Nanotechnology Centre, n.d.).

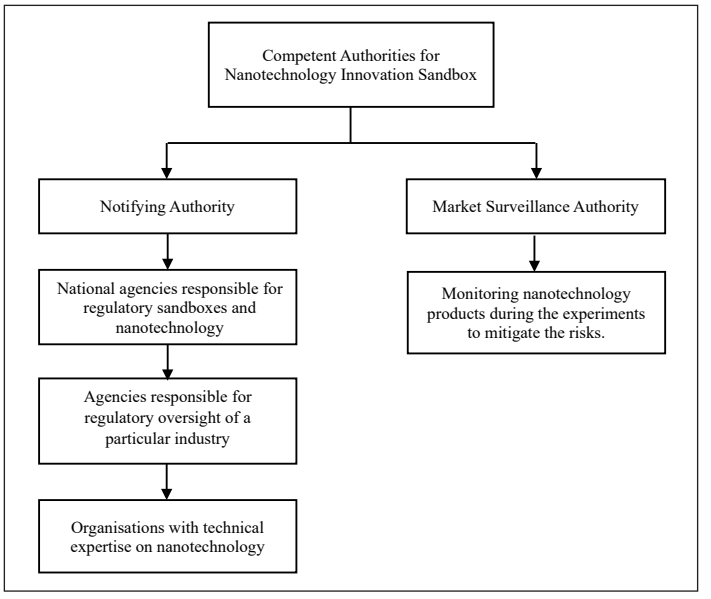


Figure 2. Competent authority for the nanotechnology sandbox

The notifying authority should also consist of the regulatory body and the designated standard organisation. The identification of the relevant regulatory body should follow the sectoral approach, i.e., the regulatory authority with jurisdiction over a particular industry serves as the lead regulator for legal and safety compliance. For instance, if the nanotechnology innovation is a nanofood that utilises new chemical compounds at the nanoscale, the Food Safety and Quality Division (FSQD) would assume the role of the regulatory authority. In this capacity, the FSQD is responsible for overseeing compliance with the food safety standards and legislative requirements.

Standard organisations provide essential technical expertise for nanotechnology across all industries. These bodies provide a horizontal layer of support for characterisation, measurement, and safety, which shall include the National Institute of Occupational Safety and Health (NIOSH), the Department of Chemistry Malaysia, SIRIM QAS International, the National Metrology Institute of Malaysia (NMIM), and NanoVerify. The market surveillance authority is responsible for conducting appropriate checks on product characteristics, including physical and laboratory checks based on adequate samples, prioritising its resources and actions to ensure effective market surveillance, including the possible risks.

Additionally, the competent authorities must be accorded the necessary power to act effectively, independently, impartially and without bias. This is to safeguard the experiment's principles of objectivity. According to the Organisation for Economic Co-operation and Development (2025), the competent authorities should be empowered in three different ways, namely, a binding decision for the grant of approval, simple discretion giving room to the authority to decide within its scope and intended discretion that should generally be exercised in the manner specified.

The Operational Requirements of Regulatory Sandboxes

The planning and designing process is pivotal for the success of regulatory sandboxes. According to Chen and Taeihagh (2026), the design of regulatory sandboxes shall start with a decision flow process for introducing a sandbox and identifying operational prerequisites consisting of (i) eligibility criteria, (ii) scope, (iii) duration of testing, (iv) monitoring and reporting, (v) exit rules, and (vi) a transition plan, as shown in Figure 3.

The initial stage in designing a regulatory sandbox involves establishing a decision flow process. This process begins by assessing the appropriate basis for experimentation, including an evaluation of applicable legal requirements and any prohibitions on testing the proposed innovation. If a prohibition exists, the regulator shall exercise their power to grant legal derogations or issue participants a no-enforcement letter exempting them from a specified set of regulations for a certain period. This is to ensure that the existing regulation never constitutes a significant barrier to tested innovation (Buocz et al., 2023).

The power to grant legal derogation and issue a no-enforcement letter is also connected with the enactment of experimental legislation for the sandbox.

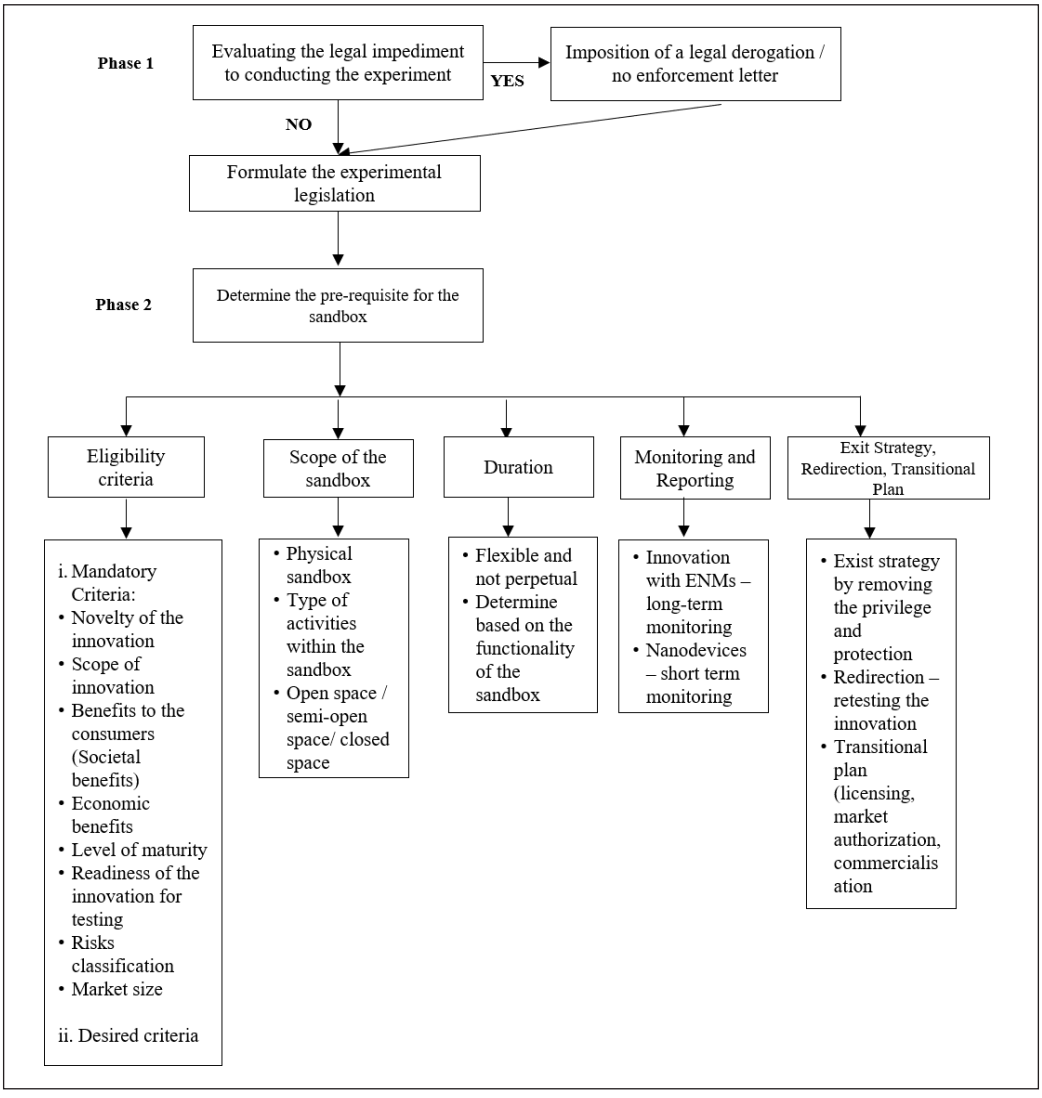


Figure 3. The regulatory sandbox framework for nanotechnology innovation

The experimental legislation shall contain express authorisation and conditions for the participant to deviate from the legal requirements. Thereby offering innovators the freedom to experiment with their innovations without incurring legal liabilities. The experimental legislation must safeguard not only the innovator's interests but also the public interest, particularly that of consumers. Accordingly, it shall contain clauses to protect a group of citizens or a territory which is subject to the experiment (Eichengreen, 2023).

To compensate for the derogations, the regulatory authorities should impose compensatory obligations and conditions on innovators for any adverse effects arising during the experimental period (Raudla et al., 2025). Innovators must bear responsibility for compensating for any injury or loss affecting safety, health, and the environment that arises during the experimental period. The compensation also operates as a risk-balancing mechanism to promote innovation and discourage negligent practices throughout the testing phase.

As previously discussed, in certain jurisdictions, nano-specific legislation is not part of the regulatory framework. The absence of nano-specific legislation removes formal regulatory impediments to establishing a nanotechnology sandbox, thereby allowing experimental activities to proceed within a legally unencumbered environment. However, other legislation must also be examined, as nanotechnology activities continue to fall under the regulatory framework enacted before the proliferation of such technology. The regulatory authority must carefully analyse whether legal derogation and experimental legislation are appropriate. The regulatory requirements for clinical trials and pre-market approval, such as substance or dossier authorisation, must be strictly observed and shall not be bypassed to protect public health and safety.

For example, nanotechnology innovation in medicine and medical devices may not be appropriate for experimentation within a regulatory sandbox. These innovations are subject to strict clinical trials and stringent compliance with regulatory requirements, protocols and ethical standards. Establishing a regulatory sandbox for nanotechnology-based medicines and medical devices would therefore necessitate targeted derogations from existing pharmaceutical and medical device laws. Allowing the innovation to enter the market prior to approval will expose the public to unassessed risks. For instance, the use of ENMs as new food additives must comply with the Codex General Standard for Food Additives (GSFA, Codex STAN 192-1995) (Liang, 2022) and obtain authorisation in accordance with the conditions specified, including prior written approval of the director, as stipulated under the Food Act 1983. Under such circumstances, granting legal derogation may contravene medical ethics principles and public health protection. Therefore, the primary distinction between a regulatory sandbox and a clinical trial is that a clinical trial is subject to strict regulatory requirements, whereas a regulatory sandbox may allow legal derogations that relax certain legal requirements.

Experimental legislation may also introduce new regulatory mandates that are currently absent from the existing regulatory frameworks. The introduction of new regulatory requirements is to embed essential protections for consumers participating in the experiments. For instance, in the context of food nanotechnology, international food labelling standards do not mandate disclosure of the presence of ENMs, resulting in some countries, such as the United States and Malaysia, adopting voluntary nano-labelling. The sandbox could insert a clause mandating the innovators to label their innovations as part

of the experimental legislation. Consumers must be informed about the presence of ENMs to enable them to make an informed decision on using the products and participating in the experiment.

The next stage is defining the operational prerequisites for the regulatory sandboxes. First, the eligibility criteria for the participant. It is the responsibility of the regulatory authority, i.e., the notifying authority, to determine the eligibility criteria for the nanotechnology regulatory sandbox. A wealth of literature has discussed the eligibility criteria for a regulatory sandbox. Seferi (2025) highlights three primary criteria to consider: the degree of innovativeness, public and societal benefit, and level of maturity. Meanwhile, the UK's FCA imposed five criteria: the scope of innovation, the novelty of the innovation, consumer benefits, the readiness of the innovation for testing, and the need for support to test the innovation in the sandbox (Financial Conduct Authority, 2026). The other eligibility requirements would be firm-specific criteria, benefits to the economy and society, and risk classification (Nabil, 2024).

This article is of the view that the other factor that should be considered in deciding the eligibility criteria for the regulatory sandbox for nanotechnology innovation is the market size of nanotechnology products. In countries with small markets and limited nanotechnology innovation, the eligibility criteria should be sufficiently flexible to avoid deterring potential participants from engaging in experimental activities. For instance, in Malaysia, the utilisation of nanotechnology across various industries remains relatively nascent compared with global nanotech hubs. As of January 2026, the Nanotechnology Products Database only listed 163 nanotechnology products from Malaysia across various industries (Nanotechnology Products Database, n.d.-b). Hence, the eligibility criteria should ensure equal access for both large corporations and small- to medium-sized enterprises, fostering inclusive innovation ecosystems. Such an approach is essential to accelerate nanotechnology development, expand market availability of nanotech products, and enhance regulatory learning through diverse participation.

Based on the above discussion, this article suggests the following eligibility criteria as mandatory, namely the novelty of the innovation, scope of innovation, benefits to the consumers (societal benefits), economic benefits (commercial value of the innovation), level of maturity, readiness of the innovation for testing, risks classification (including life-cycle risk assessment), and market size. The notifying authority may impose other eligibility criteria it deems appropriate, based on the nature of the nanotechnology innovation.

The second operational prerequisite is the scope of a regulatory sandbox. According to Chen and Taeihagh (2026), the scope should be delineated across three factors, namely the (i) parameter or scale of experiments, (ii) boundaries placed on its physical and virtual spaces, and (iii) types of activities allowed within a sandbox. The parameters and scales of experiments are fundamentally influenced by the characteristics and nature of

nanotechnology innovations. ENMs are widely used across diverse industries, including the manufacturing of consumer products such as fertilisers, foods, cosmetics, personal care products, textiles, medicines, sports equipment, and electronic products.

The sandbox for nanotechnology innovation should be designed as a physical rather than a virtual one, given the tangible nature of nanotechnology products. Unlike digital innovation, which can be experimented with online to assess performance and risks, nanotechnology innovation involves products containing ENMs or nanosized devices. Consumers ingest, apply, or physically operate these products in daily life. The assessment of safety and health risks associated with ENMs must be based on robust empirical data (Justo-Hanani, 2025), post-market monitoring (Mitra et al., 2024), and direct supervision. Therefore, online or virtual testing environments are inappropriate for evaluating safety and health risks, environmental impacts, and performance.

Given the nature of risks and benefits of nanotechnology, the activities conducted within the regulatory sandbox should extend beyond mere product development and improvement. The physical experiments must also assess the potential adverse effects of ENM use on human health and safety. The geographical scope pertains to the regulatory sandbox's spaces. Molitorisová and Purnhagen (2025) conceptualise the spaces of sandboxes into three categories: open, semi-open, and closed. These spaces will determine the extent of control that regulators and innovators have over the sandbox's object.

A closed regulatory sandbox is highly rigid, as the object cannot leave the sandbox during testing or be accessed from the outside. Only participants, innovators, and regulators are allowed in the sandbox, without interference from third parties. The rigidity will lower risks, as there is a high level of traceability, which serves as a risk-mitigation measure for human safety and health. On the contrary, an open regulatory sandbox entails a higher level of risk due to its openness. It allows the public to enter the experimental space, which is not separated from other facilities of ordinary commercial activity. The open space also reduces traceability because the product under test is neither labelled nor documented. A semi-open regulatory sandbox is not open to the public; only selected participants are allowed to engage with the labelled and recorded objects. The objects are also allowed to leave the sandbox but remain under regulatory oversight.

This article asserts that the sandbox for nanotechnology innovation should be designed to reflect the distinct characteristics of various nanotechnology innovations. As previously discussed, the integration of nanotechnology into innovation activities involves ENMs or nanodevices. The closed regulatory sandbox is appropriate for innovation that utilises novel ENMs with uncertainties regarding potential safety and health risks, in which experiments are conducted under regulatory supervision with predefined parameters, monitoring protocols, and containment measures. On the other hand, open or semi-open sandboxes are

appropriate for innovation that utilises nanodevices, as they do not involve direct exposure to human safety and health risks. When the sandbox for nanotechnology innovation operates in a closed environment, its geographical boundaries must be clearly defined and confined to domestic territory to enable continuous, close-range oversight by both innovators and regulatory authorities. In contrast, open and semi-open regulatory sandboxes may extend beyond national boundaries, allowing experimental nanotechnology products to circulate within international markets and across jurisdictions, considering the global nature of nanotechnology trade. Hence, it is also necessary for the notifying authorities to evaluate the potential interaction between the regulatory sandbox and international trade laws. It is also important to note that closed, open, and semi-open sandboxes should all be physical in nature, given the tangible nature of nanotechnology innovations, as explained in the previous paragraph.

The third operational prerequisite is the duration of the regulatory sandbox. According to Seferi (2025), the duration of the experimentation phase is typically pre-defined and may range from 3 to 36 months. Meanwhile, Novelli et al. (2026) argued that the sandbox duration must remain flexible within a reasonable timeframe. The sandbox may end early if the experiment achieves its objectives. Likewise, Ranchordás (2021) suggested that no specific duration be set for the regulatory sandbox, as it needs to be operated in line with the specificities of the innovations. Chen and Taeihagh (2026) contended that the duration of the regulatory sandbox depends on its functionality. They classified the regulatory sandbox's functionality as exploratory, confirmatory, and on demand. The exploratory regulatory sandbox aims to explore and test innovations. Therefore, the duration of the regulatory sandbox should be longer than that of a confirmatory regulatory sandbox, which is primarily intended to validate and substantiate previously explored technological applications and can be completed within a shorter time frame. Meanwhile, the on-demand regulatory sandbox can be long- or short-term, as it is used when the intention is to make the innovation available to the public upon exiting the sandbox.

Scholars have suggested different approaches to determine the appropriate duration of a regulatory sandbox. The general rule is that the regulatory sandbox shall not be perpetual. Instead, it must be within a reasonable timeframe given the objectives of the experiments. This article argues that the duration of the sandbox for nanotechnology innovation should be based on functionality, as suggested by Chen and Taeihagh (2026). Given this dynamic landscape, exploratory, confirmatory, and on-demand regulatory sandboxes are all appropriate for nanotechnology. For instance, if an innovation utilises ENMs commonly used across industries and regarded as safe, a confirmatory regulatory sandbox with a shorter time frame is more appropriate. If an innovation utilises ENMs with unknown toxicity, it should be placed in an exploratory sandbox with a longer experimental duration. Regulatory authorities must assess the nature of these innovations and the objectives of innovators, particularly regarding their intent to commercialise the technology.

The next operational requisite is monitoring and reporting. The purpose is to determine whether the nanotechnology products continue to meet relevant standards, especially in terms of safety and health. The activities conducted within the regulatory sandbox must be subjected to formal ongoing monitoring. This article suggests that the frequency and detail of monitoring should depend on the complexity of the products, their life cycles, and the potential risks they pose. High-risk products will require more detailed and frequent monitoring, while lower-risk products could benefit from lighter monitoring. The information obtained from monitoring activities must be reported, shared among participants in the regulatory sandbox, and evaluated (Organisation for Economic Co-operation and Development, 2025).

This article suggests that nanotechnology-based products, such as foods and feeds, cosmetics, personal care items, and fertilisers, should undergo long-term monitoring due to their high-risk profile stemming from direct consumer exposure to ENMs. This contention is based on various laboratory studies demonstrating potential adverse effects of ENMs on human health and safety. Data obtained from the regulatory sandbox provides critical empirical evidence that may validate or challenge laboratory study findings, thereby strengthening the scientific basis for risk assessment of nanotechnology products.

On the other hand, nanotechnology products such as devices, semiconductors, and fabrics with no added ENMs are suitable for short-term monitoring, as there is no direct exposure of ENMs to human safety and health. Conversely, nanotechnology products without added ENMs, such as devices, semiconductors, and fabrics, are appropriate for short-term monitoring, as they do not involve direct exposure to ENMs that could potentially compromise human health and safety. Most importantly, the risk classification should consider the unique characteristics of each ENM category, including physicochemical properties such as size distribution, surface properties, solubility, persistence, routes of exposure, and exposure duration.

The last operational requisite is the exit strategy. There are two ways for the participant to exit the regulatory sandbox. First, the regulatory authority may revoke the sandbox when the participant fails to comply with the sandbox conditions or unacceptable risks arise. There must be a matrix outlining acceptable and unacceptable risks that can serve as a guideline for determining when the regulatory sandbox should be terminated. It is the responsibility of the regulatory authority and the participant to formulate the exit strategy from the experiment (Ringe & Ruof, 2020). For the regulatory authority, the exit strategy shall include issuing at least one month's written notice of its intention to terminate the sandbox and removing privileges or protections granted to participants under legal derogation.

For the participant, when unacceptable risks arise, the exit strategy must specify the participant’s responsibility to provide compensation or redress mechanisms to any consumers who have suffered financial losses or safety and health issues arising from

the tested nanotechnology innovation (Bank Negara Malaysia, 2024). The participant must also clearly outline the procedure for recalling the nanotechnology products from consumers' possession. The exit strategy may also involve redirection, in which the participant will restart the experiment in an improved environment to reassess the risks (Stazi & Jovine, 2025). The second way to exit the sandbox is after the successful completion of the experiment within the stipulated timeframe. When the experiment is successfully completed, the participant may proceed with a transitional plan, which may include licensing, market authorisation, or commercialisation.

As previously analysed, the function of the regulatory sandbox for nanotechnology innovation, inter alia, includes a systematic reporting scheme on the utilisation of ENMs and nanodevices, particularly regarding potential safety and health risks. In certain jurisdictions, it is a requirement that a formal exit report be prepared by the competent authority upon completion of the regulatory sandbox experiment (Stazi & Jovine, 2025). Therefore, this study contended that the regulatory authority and the participant should prepare a formal exit report summarising the experimental outcomes, the compliance trajectory, the regulatory implications, concerns regarding safety, health, and the environment, and lessons learned from the experiment. Such reports play a critical role in ensuring transparency and accountability in the regulatory sandbox framework. The monitoring and final reports should be made publicly available to uphold transparency in the regulatory sandbox for industries and consumers. However, such disclosure should be subject to certain limitations, particularly where the innovation is protected under intellectual property rights or confidential data. It must balance the need for transparency with the innovator's proprietary interests.

CONCLUSION

Nanotechnology is a highly versatile and transformative technology that has stimulated innovation across various industries. However, there are two major concerns regarding the adoption of nanotechnology for innovation across various industries: the risks to safety, health, and the environment remain uncertain, and there is a lack of nanotechnology-specific legislation to govern these risks. There is a need to formulate a framework to protect consumers and the environment against the risks posed by nanotechnology innovation. A regulatory sandbox is an appropriate legal framework for regulating risks by testing innovation in a controlled environment. While this framework has been adopted by other technologies, there is a lack of academic exploration of the regulatory sandbox for nanotechnology.

Figure 3 illustrates the basic framework for a nanotechnology innovation sandbox that could be adopted in governing the innovation and risks associated with nanotechnology. The formulation of the sandbox for nanotechnology innovation must commence with

establishing a competent authority comprising the notifying authority and the market surveillance authority. The design of the sandbox must take into consideration the current legal position governing the specific nanotechnology innovation and the nature of nanotechnology itself, whether ENMs or nanodevices. While testing, the innovator must systematically monitor, assess, and report the impacts of their innovation on safety, health, and the environment before fully releasing the products to market. The sandbox will be terminated if risks are identified and persist. Hence, the sandbox serves as a precautionary measure in the absence of specific nanotechnology legislation while promoting sustainable nanotechnology development. This article also highlights the need for further research into the relationship between the regulatory framework for nanotechnology innovation and international trade law, particularly the World Trade Organisation (WTO) instruments. Further research is also essential to develop a risk-based monitoring framework tailored for each category of nanotechnology innovations across various industries.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of the Faculty of Law, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia and the Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, Skudai, Johor, for providing the facilities and financial support for this research.

REFERENCES

Alaassar, A., Mention, A.-L., & Aas, T. H. (2020). Exploring how social interactions influence regulators and innovators: The case of regulatory sandboxes. *Technological Forecasting and Social Change, 160*, Article 120257. https://doi.org/10.1016/j.techfore.2020.120257

Allen, H. J. (2019). Regulatory sandboxes. *George Washington Law Review, 87*(3), 579-645. https://www.gwlr.org/regulatory-sandboxes/

Angelopoulou, P., Giaouris, E., & Gardikis, K. (2022). Applications and prospects of nanotechnology in food and cosmetics preservation. *Nanomaterials, 12*(7), Article 1196. https://doi.org/10.3390/nano12071196

Artificial Intelligence Act, Regulation (EU) 2024/1689. (2024). *Official Journal of the European Union, L*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Asha, A. B., & Narain, R. (2020). Nanomaterials properties. In R. Narain (Ed.), *Polymer science and nanotechnology* (pp. 343-359). Elsevier. https://doi.org/10.1016/b978-0-12-816806-6.00015-7

Attrey, A., Lesher, M., & Lomax, C. (2020). *The role of sandboxes in promoting flexibility and innovation in the digital age* (OECD Going Digital Toolkit Notes No. 2). OECD Publishing. https://doi.org/10.1787/cdf5ed45-en

Aydın, Z., & Yardımcı, O. (2024). Regulatory sandboxes and pilot projects: Trials, regulations, and insights in energy transition. *Engineering Science and Technology, An International Journal, 56*, Article 101792. https://doi.org/10.1016/j.jestch.2024.101792

Baki, M. E., Miresmaili, S. M., Pourentezari, M., Amraii, E., Yousefi, V., Spenani, H. R., Talebi, A. R., Anvari, M., Fazilati, M., Fallah, A. A., & Mangoli, E. (2014). Effects of silver nano-particles on sperm parameters, number of Leydig cells and sex hormones in rats. *Iranian Journal of Reproductive Medicine, 12*(2), 139-144. https://pmc.ncbi.nlm.nih.gov/articles/PMC4009568/

Bank Negara Malaysia. (2024). *Financial technology regulatory sandbox framework*. https://www.bnm.gov.my/-/sandbox-pd24

Bhattacharjee, I., & Maiti, D. K. (2021). Nanotools and nanodevices in geothermal energy. In S. Devasahayam & C. M. Hussain (Eds.), *Nano tools and devices for enhanced renewable energy* (pp. 507-518). Elsevier. https://doi.org/10.1016/b978-0-12-821709-2.00022-0

Borm, P. J. A., Robbins, D., Haubold, S., Kuhlbusch, T., Fissan, H., Donaldson, K., Schins, R., Stone, V., Kreyling, W., Lademann, J., Krutmann, J., Warheit, D., & Oberdorster, E. (2006). The potential risks of nanomaterials: A review carried out for ECETOC. *Particle and Fibre Toxicology, 3*(1), Article 11. https://doi.org/10.1186/1743-8977-3-11

Buocz, T., Pfotenhauer, S., & Eisenberger, I. (2023). Regulatory sandboxes in the AI Act: Reconciling innovation and safety? *Law, Innovation and Technology, 15*(2), 357-389. https://doi.org/10.1080/17579961.2023.2245678

Carbone, M., Donia, D. T., Sabbatella, G., & Antiochia, R. (2016). Silver nanoparticles in polymeric matrices for fresh food packaging. *Journal of King Saud University - Science, 28*(4), 273-279. https://doi.org/10.1016/j.jksus.2016.05.004

Chávez-Hernández, J. A., Velarde-Salcedo, A. J., Navarro-Tovar, G., & Gonzalez, C. (2024). Safe nanomaterials: From their use, application, and disposal to regulations. *Nanoscale Advances, 6*(6), 1583-1610. https://doi.org/10.1039/d3na01097j

Chen, G. K. H., & Taeihagh, A. (2026). Designing regulatory sandboxes: A comprehensive framework for aligning functionalities and objectives. *Policy Design and Practice, 9*(1), 1-15. https://doi.org/10.1080/25741292.2025.2570954

Cid-Samamed, A., Correa-Duarte, M. Á., Mariño-López, A., & Diniz, M. S. (2023). Exposure to oxidized multi-walled CNTs can lead to oxidative stress in the Asian freshwater clam Corbicula fluminea (Müller, 1774). *International Journal of Molecular Sciences, 24*(22), Article 16122. https://doi.org/10.3390/ijms242216122

Durham, J. (2023). Regulatory sandboxes enable pragmatic blockchain regulation. *Washington Journal of Law, Technology & Arts, 18*(1), 27. https://digitalcommons.law.uw.edu/wjlta/vol18/iss1/3/

Eichengreen, B. (2023). Financial regulation in the age of the platform economy. *Journal of Banking Regulation, 24*(1), 40-50. https://doi.org/10.1057/s41261-021-00187-9

Emergen Research. (2026, August 30). *Global nanotechnology market*. https://www.emergenresearch.com/industry-report/nanotechnology-market

European Commission. (2023). *Better regulation toolbox*. https://commission.europa.eu/system/files/2023-09/BR%20toolbox%20-%20Jul%202023%20-%20FINAL.pdf

Fahy, L. A. (2022). Fostering regulator–innovator collaboration at the frontline: A case study of the UK's regulatory sandbox for FinTech. *Law & Policy, 44*(2), 162-184. https://doi.org/10.1111/lapo.12184

Federal Ministry for Economic Affairs and Energy. (2025). *Regulatory sandboxes: Testing environments for innovation and regulation*. https://www.bundeswirtschaftsministerium.de/Redaktion/EN/Dossier/regulatory-sandboxes.html

Financial Conduct Authority. (2026, July 24). *Regulatory sandbox eligibility criteria*. https://www.fca.org.uk/firms/innovation/regulatory-sandbox/eligibility-criteria

Fiordaliso, F., Bigini, P., Salmona, M., & Diomede, L. (2022). Toxicological impact of titanium dioxide nanoparticles and food-grade titanium dioxide (E171) on human and environmental health. *Environmental Science: Nano, 9*(4), 1199-1211. https://doi.org/10.1039/d1en00833a

Food Act 1983 (Act 281) (Malaysia).

Giese, B., Klaessig, F., Park, B., Kaegi, R., Steinfeldt, M., Wigger, H., von Gleich, A., & Gottschalk, F. (2018). Risks, release and concentrations of engineered nanomaterial in the environment. *Scientific Reports, 8*(1), Article 1565. https://doi.org/10.1038/s41598-018-19275-4

Goo, J. J., & Heo, J.-Y. (2020). The impact of the regulatory sandbox on the FinTech industry, with a discussion on the relation between regulatory sandboxes and open innovation. *Journal of Open Innovation: Technology, Market, and Complexity, 6*(2), Article 43. https://doi.org/10.3390/joitmc6020043

Gumbo, L., & Chude-Okonkwo, U. A. K. (2025). Regulatory sandbox as a frontier for innovation and sustainability: A systematic review. *Cogent Business & Management, 12*(1), Article 2510555. https://doi.org/10.1080/23311975.2025.2510555

Guston, D. H. (2014). Understanding 'anticipatory governance'. *Social Studies of Science, 44*(2), 218-242. https://doi.org/10.1177/0306312713508669

Halappanavar, S., Ede, J. D., Mahapatra, I., Krug, H. F., Kuempel, E. D., Lynch, I., Vandebriel, R. J., & Shatkin, J. A. (2021). A methodology for developing key events to advance nanomaterial-relevant adverse outcome pathways to inform risk assessment. *Nanotoxicology, 15*(3), 289-310. https://doi.org/10.1080/17435390.2020.1851419

Hasmin, N. A., Zainol, Z. A., Mahmood, A., & Matmin, J. (2024). Adoption of the precautionary principle on the safety and health risks of nanofood in Malaysia. *UUM Journal of Legal Studies, 15*(1), 271-302. https://doi.org/10.32890/uumjls2024.15.1.12

Heidmann, I., & Milde, J. (2013). Communication about scientific uncertainty: How scientists and science journalists deal with uncertainties in nanoparticle research. *Environmental Sciences Europe, 25*(1), Article 25. https://doi.org/10.1186/2190-4715-25-25

Heselhaus, S. (2010). Nanomaterials and the precautionary principle in the EU. *Journal of Consumer Policy, 33*(1), 91-108. https://doi.org/10.1007/s10603-009-9123-8

International Organization for Standardization. (2015). *Nanotechnologies—Vocabulary—Part 1: Core terms* (ISO/TS Standard No. 80004-1:2015). https://www.iso.org/standard/68058.html

Jahnel, J. (2015). Conceptual questions and challenges associated with the traditional risk assessment paradigm for nanomaterials. *NanoEthics, 9*(3), 261-276. https://doi.org/10.1007/s11569-015-0235-0

Jeník, I., & Duff, S. (2020). *How to build a regulatory sandbox: A practical guide for policy makers*. CGAP. https://www.cgap.org/research/publication/how-to-build-regulatory-sandbox-practical-guide-for-policy-makers

Johnson, W. G. (2023). Caught in quicksand? Compliance and legitimacy challenges in using regulatory sandboxes to manage emerging technologies. *Regulation & Governance, 17*(3), 709-725. https://doi.org/10.1111/rego.12487

Justo-Hanani, R. (2025). The sources of EU's international influence on nanotechnology risk regulation and governance. *European Journal of Risk Regulation, 16*(4), 1492-1503. https://doi.org/10.1017/err.2025.10042

Kokilavani, S., Janani, B., Balasurya, S., & Khan, S. S. (2023). Accumulation of engineered nanomaterials in soil, water, and air. In F. Fernández-Luqueño & J. K. Patra (Eds.), *Agricultural and environmental nanotechnology* (pp. 551-582). Springer. https://doi.org/10.1007/978-981-19-5454-2_21

Kumar, R., Ratnesh, R. K., Singh, J., & Chandra, R. (2025). An overview of Industry 4.0 applications in the area of nanotechnology. In R. Kumar, R. K. Ratnesh, & J. Singh (Eds.), *Nanotechnology applications for Industry 4.0* (pp. 1-27). CRC Press. https://doi.org/10.1201/9781003634553-1

Liang, J. (2022). Food additives. In L. Zhang (Ed.), *Nutritional toxicology* (pp. 167-180). Springer. https://doi.org/10.1007/978-981-19-0872-9_6

Ma, X., Tian, Y., Yang, R., Wang, H., Allahou, L. W., Chang, J., Williams, G., Knowles, J. C., & Poma, A. (2024). Nanotechnology in healthcare, and its safety and environmental risks. *Journal of Nanobiotechnology, 22*(1), Article 715. https://doi.org/10.1186/s12951-024-02901-x

Madiega, T., & Van De Pol, A. L. (2022). *Artificial intelligence act and regulatory sandboxes*. European Parliamentary Research Service. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)733544

Malik, S., Muhammad, K., & Waheed, Y. (2023). Nanotechnology: A revolution in modern industry. *Molecules, 28*(2), Article 661. https://doi.org/10.3390/molecules28020661

Maric, T., Tanaka, Y., Li, Z., Wu, J., Li, J., Guan, J., & Boisen, A. (2026). Micro/nanomotors for biofilm remediation: Nanoarchitectonic breakthroughs with a focus on regulatory and clinical translation challenges. *Biomaterials, 326*, Article 123715. https://doi.org/10.1016/j.biomaterials.2025.123715

Mitra, S., Singh, P. K., Mohapatra, R. K., Mohapatra, N. P., Sarkar, B., & Mishra, S. (2024). A perspective on the global market of micro- and nano- smart materials in pharmaceutical industries. In A. Behera, A. K. Nayak, R. K. Mohapatra, & A. A. Rabaan (Eds.), *Smart micro- and nanomaterials for pharmaceutical applications* (pp. 17-38). CRC Press. https://doi.org/10.1201/9781003468431-2

Molitorisová, A., & Purnhagen, K. (2025). Regulatory sandboxes for novel foods. *European Journal of Risk Regulation, 16*(3), 1063-1081. https://doi.org/10.1017/err.2025.10

Morgan, D. (2023). Anticipatory regulatory instruments for AI systems. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 980-981). Association for Computing Machinery. https://doi.org/10.1145/3600211.3604732

Nabil, R. (2024). Artificial intelligence regulatory sandboxes. *Journal of Law, Economics & Policy, 19*(2), 295-348. https://www.jlep.net/s/Nabil-Final-for-PDF.pdf

NanoMalaysia. (2024). *Strategic report 2024*. https://nanomalaysia.com.my/wp-content/uploads/2025/12/FINAL-NanoMalaysia-Strategic-Report-2024.pdf

Nanotechnology Products Database. (n.d.-a). *Introduction*. Retrieved October 20, 2026, from https://product.statnano.com/

Nanotechnology Products Database. (n.d.-b). *List of products*. Retrieved October 20, 2026, from https://product.statnano.com/search?keyword=Malaysia

National Nanotechnology Centre. (n.d.). *Pusat Nanoteknologi Kebangsaan*. Ministry of Science, Technology and Innovation. https://www.mosti.gov.my/en/pusat-nanoteknologi-kebangsaan/

National Technology and Innovation Sandbox. (n.d.). *About us*. https://sandbox.gov.my/

Novelli, C., Hacker, P., McDougall, S., Morley, J., Rotolo, A., & Floridi, L. (2026). Getting regulatory sandboxes right: Design and governance under the AI Act. *European Journal of Risk Regulation, 17*(3), 926-952. https://doi.org/10.1017/err.2026.10096

Oberdörster, G., Maynard, A., Donaldson, K., Castranova, V., Fitzpatrick, J., Ausman, K., Carter, J., Karn, B., Kreyling, W., Lai, D., Olin, S., Monteiro-Riviere, N., Warheit, D., & Yang, H. (2005). Principles for characterizing the potential human health effects from exposure to nanomaterials: Elements of a screening strategy. *Particle and Fibre Toxicology, 2*(1), Article 8. https://doi.org/10.1186/1743-8977-2-8

Organisation for Economic Co-operation and Development. (2022). *Chemical accidents involving nanomaterials: Potential risks and review of prevention, preparedness and response measures—Project report*. OECD Publishing. https://doi.org/10.1787/6395f7ce-en

Organisation for Economic Co-operation and Development. (2024). *Developments in delegations on the safety of manufactured nanomaterials and advanced materials* (ENV/CBC/MONO(2024)1). https://one.oecd.org/document/ENV/CBC/MONO(2024)1/en/pdf

Organisation for Economic Co-operation and Development. (2025). *Regulatory sandbox toolkit: A comprehensive guide for regulators to establish and manage regulatory sandboxes effectively*. OECD Publishing. https://doi.org/10.1787/de36fa62-en

Parenti, R. (2020). *Regulatory sandboxes and innovation hubs for FinTech* (PE 652.752). European Parliament. https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2020)652752

Pathak, G., Mangla, S., Gupta, G. K., Bhan, V., & Kapoor, R. K. (2025). Toxicological assessment and risk management of nanoparticles mediated composite materials-critical review: State of the art. *Discover Polymers, 2*(1), Article 12. https://doi.org/10.1007/s44347-025-00023-7

Pisano, R., & Durlo, A. (2025). Historical foundations of nanotechnology: Emergences, inventions and discoveries. In R. Pisano (Ed.), *Nanoscience & nanotechnologies: Critical problems, science in society, historical perspectives* (pp. 581-629). Springer. https://doi.org/10.1007/978-3-031-85122-3_15

Poncibò, C., & Zoboli, L. (2020). Sandboxes and consumer protection: The European perspective. *International Journal on Consumer Law and Practice, 8*, 1-22. https://repository.nls.ac.in/ijclp/vol8/iss1/1/

Qiu, Y., Yao, H., Ren, P., Tian, X., & You, M. (2025). Regulatory sandbox expansion: Exploring the leap from FinTech to medical artificial intelligence. *Intelligent Oncology, 1*(2), 120–127. https://doi.org/10.1016/j.intonc.2025.03.001

Ranchordás, S. (2021). Experimental regulations and regulatory sandboxes – law without order? *Law and Method*. https://doi.org/10.5553/REM/.000064

Raudla, R., Juuse, E., Kuokštis, V., Cepilovs, A., Cipinys, V., & Ylönen, M. (2025). To sandbox or not to sandbox? Diverging strategies of regulatory responses to FinTech. *Regulation & Governance, 19*(3), 917-932. https://doi.org/10.1111/rego.12630

Ringe, W.-G., & Ruof, C. (2020). Regulating FinTech in the EU: The case for a guided sandbox. *European Journal of Risk Regulation, 11*(3), 604-629. https://doi.org/10.1017/err.2020.8

Safarzadeh, H., & Di Maria, F. (2026). Sustainable energy storage solutions using nanomaterials: Bridging the gap between policy and technology in Italy. *Environmental Sustainability, 9*(2), 535-545. https://doi.org/10.1007/s42398-026-00444-5

Science Policy Council. (2007). *Nanotechnology white paper* (EPA 100/B-07/001). U.S. Environmental Protection Agency. https://archive.epa.gov/osa/pdfs/web/pdf/epa-nanotechnology-whitepaper-0207.pdf

Seferi, F. (2025). A comparative analysis of regulatory sandboxes from selected use cases: Insights from recurring operational practices. In F. Bagni & F. Seferi (Eds.), *Regulatory sandboxes for AI and cybersecurity: Questions and answers for stakeholders* (pp. 145-176). CINI Cybersecurity National Lab. https://cybersecnatlab.it/wp-content/uploads/2025/02/CybersecNatLab-White-Paper-Regulary-Sandboxes.pdf

Servin, A., Elmer, W., Mukherjee, A., De la Torre-Roche, R., Hamdi, H., White, J. C., Bindraban, P., & Dimkpa, C. (2015). A review of the use of engineered nanomaterials to suppress plant disease and enhance crop yield. *Journal of Nanoparticle Research, 17*(2), Article 92. https://doi.org/10.1007/s11051-015-2907-7

Shahare, B., Yashpal, M., & Gajendra. (2013). Toxic effects of repeated oral exposure of silver nanoparticles on small intestine mucosa of mice. *Toxicology Mechanisms and Methods, 23*(3), 161-167. https://doi.org/10.3109/15376516.2013.764950

Stazi, A., & Jovine, R. (2025). A comparative analysis of regulatory sandboxes: Models, evolution and strategic implications in EU, USA and China. *Comparative Law Review, 16*(2), 22-66. https://comparativelawreview.org/wp-content/uploads/2025/11/CoLRe_vol16_2_winter_issue_finale.pdf

Tang, M., Li, S., Wei, L., Hou, Z., Qu, J., & Li, L. (2021). Do engineered nanomaterials affect immune responses by interacting with gut microbiota? *Frontiers in Immunology, 12*, Article 684605. https://doi.org/10.3389/fimmu.2021.684605

Todt, O., & Luján, J. L. (2014). Analyzing precautionary regulation: Do precaution, science, and innovation go together? *Risk Analysis, 34*(12), 2163-2173. https://doi.org/10.1111/risa.12246

Wojciechowska, O., Costabile, A., & Kujawska, M. (2023). The gut microbiome meets nanomaterials: Exposure and interplay with graphene nanoparticles. *Nanoscale Advances, 5*(23), 6349-6364. https://doi.org/10.1039/d3na00696d

Xie, W., Xiong, Z., Wang, H., Liu, X., Cui, H., Huang, Q., & Tang, Y. (2024). The nanosafety assessment of ENMs under a dermal exposure scenario: From key molecular events to in silico modeling tools. *Environmental Science: Nano, 11*(3), 708-738. https://doi.org/10.1039/d3en00585b

Zabeo, A., Keisler, J. M., Hristozov, D., Marcomini, A., & Linkov, I. (2019). Value of information analysis for assessing risks and benefits of nanotechnology innovation. *Environmental Sciences Europe, 31*(1), Article 11. https://doi.org/10.1186/s12302-019-0194-0

Zhang, C., Chen, X., & Ho, S.-H. (2021). Wastewater treatment nexus: Carbon nanomaterials towards potential aquatic ecotoxicity. *Journal of Hazardous Materials, 417*, Article 125959. https://doi.org/10.1016/j.jhazmat.2021.125959

Pertanika J. Sci. & Technol. 34 (5): 2675 - 2690 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

The Use of FSR-Based Plantar Pressure Patterns for Deception Detection

Gerelbat Batgerel and Chunki Kwon*

Department of Medical IT Engineering, Soonchunhyang University, 22 Soonchunhyang-ro, Sinchang-myeon, Asan-si, Chungcheongnam-do 31538, South Korea.

ABSTRACT

This study investigates plantar pressure patterns as a behavioural indicator of deception by analysing postural balance responses using force-sensitive resistor (FSR) sensors. A mock interrogation task was conducted in which subjects drew an apple on one of two sheets and responded to a question about whether they had drawn the apple, while plantar pressure was recorded. A brief perturbation was applied to induce a balance disturbance. Trials were classified as truthful or non-truthful based on the consistency between the subject's verbal responses and actual behaviour. Results showed that truthful trials exhibited stable balance and rapid recovery, whereas non-truthful trials demonstrated greater instability and delayed recovery, reflected in prolonged plantar pressure shifts. Importantly, consistent differences in recovery dynamics were observed from three additional participants, suggesting a potentially reproducible behavioural pattern. These findings indicate that deception-related cognitive and emotional processes may influence postural control. While this study represents a proof-of-concept feasibility investigation, it suggests that plantar pressure analysis may provide a non-invasive approach for detecting deception-related behavioural responses.

Keywords: 2D Matrix FSR sensors, deception detector, foot pressure pattern, involuntary postural reactions

INTRODUCTION

Deception detection plays a critical role in forensic investigations, legal proceedings, and security screening. Conventional approaches, particularly polygraph testing, infer deception indirectly by measuring physiological responses such as heart rate, respiration, and electrodermal activity, based on the assumption that deceptive responses are associated with increased psychological

ARTICLE INFO

Article history:
Received: 17 June 2026
Accepted: 06 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.06

Gerelbat47@gmail.com (Gerelbat Batgerel)
chunkikwon@sch.ac.kr (Chunki Kwon)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

stress and autonomic nervous system activation (Grubin & Madsen, 2005; Krapohl & Shaw, 2015). However, these physiological signals require direct physical contact via electrodes and sensors, which can cause discomfort and increase psychological stress for examinees. They may be influenced by intentional countermeasures.

To overcome these problems in polygraph testing, recent studies have explored non-contact approaches, including vocal characteristics (Gallardo-Antolín & Montero, 2021; Kim, 2020) and facial expressions (Nam et al., 2021; Owayjan et al., 2012) as behavioural and cognitive indicators of deception. While promising, these approaches often exhibit variability across individuals due to differences in personality, cultural background, communication style, and expressive behaviour. Some individuals naturally exhibit minimal facial expressions or stable vocal patterns regardless of internal cognitive or emotional states, which can reduce the robustness of such approaches.

This study explores an alternative behavioural perspective by focusing on postural balance responses. The motivation for this work originates from a traditional observational test in which individuals appear to respond differently to unexpected perturbations depending on whether they are truthful or deceptive. Based on this observation, we hypothesise that deception may disrupt postural control and produce measurable changes in plantar pressure distribution (Batgerel & Kwon, 2026).

To test this hypothesis, plantar pressure patterns were measured using force-sensitive resistor (FSR) sensors embedded in insole devices. These sensors enable high-resolution monitoring of foot pressure distribution across the foot, providing quantitative information on postural balance and weight shifting.

The observation of plantar pressure patterns revealed a distinct difference between truthful and non-truthful trials when a brief external perturbation was applied. A truthful response was associated with minimal displacement and rapid balance recovery, whereas a deceptive response was associated with greater postural instability and a prolonged recovery time.

This balance disturbance was reflected in the plantar pressure distribution, with non-truthful trials showing larger and more sustained shifts in foot pressure than truthful trials. These observations suggest that postural balance may serve as a meaningful behavioural indicator of deception.

This balance disturbance was reflected in the plantar pressure distribution, with non-truthful trials exhibiting larger and more sustained shifts in foot pressure compared to truthful trials. These observations suggest that postural balance may serve as a meaningful behavioural indicator associated with deception. Furthermore, plantar pressure analysis using FSR sensors provides a non-invasive and quantitative method for detecting balance-related responses associated with deceptive behaviour.

Existing deception detection methods can generally be categorised into physiological signal-based approaches, such as polygraph testing, and non-contact approaches based

on facial expressions or speech analysis. While these methods have shown promising performances, they primarily rely on autonomic physiological responses or expressive behaviours. In contrast, the proposed method investigates deception from a different perspective by analysing perturbation-induced postural recovery through FSR-based plantar pressure measurements. By evaluating involuntary balance responses rather than physiological or expressive signals, this study introduces plantar pressure-based postural recovery dynamics as a novel behavioural indicator associated with deception. To the best of our knowledge, this is the first proof-of-concept study exploring perturbation-induced plantar pressure patterns for deception-related behavioural assessment.

This paper has been rewritten to improve the level of completion of the conference paper presented at the International Conference on Next-generation Convergence Technology Winter Conference 2026 (ICNCT2026) in Macau, China (Batgerel & Kwon, 2026).

LITERATURE REVIEW

Lie Detection and the Polygraph: A Historical Review

A historical overview of lie detection techniques is provided, with particular focus on the development, scientific foundation, and controversies surrounding polygraph testing. It aims to examine how polygraph technology has evolved and to evaluate its reliability and validity as a deception-detection tool (Grubin & Madsen, 2005).

Polygraph Practice as a Deception Detector

Fundamentals of Polygraph Practice explains lie detection by describing modern, evidence-based polygraph techniques that measure physiological responses associated with deception. It outlines validated methods for detecting lies through psychophysiological testing, data analysis, ethical considerations, and legal standards (Krapohl & Shaw, 2015).

Facial Expression as an Indicator of Deception Detector

Facial expressions are investigated as a non-contact deception-detection approach. In addition, the limitations of traditional contact-based lie detection methods are addressed. The results demonstrated the feasibility of detecting deception through facial expression dynamics. (Nam et al., 2021).

Real facial motion is analysed to derive the spatial and temporal patterns exhibited by the human face when attempting to lie. Facial micro-expressions are used to design and develop a real-time lie detection system (Owayjan et al., 2012).

Speech as an Indicator of Deception Detector

Speech is investigated as a non-contact method for deception detection. In addition, the limitations of traditional polygraph-based lie detection methods are addressed. Furthermore,

an important factor for a speech-based deception detection approach is introduced (Kim, 2020).

The development of an automatic deception detection system based on gaze and speech features. This research states that the dynamics of gaze and speech feature sequences can be an important indicator for LSTM network-based deception detection(Gallardo-Antolín & Montero, 2021).

METHODOLOGY

Experimental Setup

Plantar pressure patterns were measured using the Kitronyx MP2513 (2024) Insole Sensor Kit, which records pressure distribution at 30 Hz. This kit consists of a pair of insoles embedded with force-sensitive resistor (FSR) sensors. Each insole contains a 16x10 grid of FSR sensors (Figure 1a). For data collection using the MP2513 Insole sensor kit, the left and right insoles were positioned beneath the subject's feet (Figure 1b). Each insole independently captures a 16x10 plantar pressure matrix. When combined, the two insoles form a unified 16 × 20 sensing surface representing the plantar pressure distribution across both feet (Figure 1c). This configuration enabled real-time capture of plantar pressure distribution across the entire foot contact area. The device records plantar pressure pattern data at 30 frames per second (30Hz), providing dynamic, frame-by-frame measurements of the participant's balance adjustments throughout each 6-second trial.

The sampling rate used in this study is sufficient to capture overall balance displacement and postural recovery patterns following a perturbation. Considering that the observed recovery times ranged from approximately 0.4s to 1.7s, corresponding to 12-51 frames at 30 Hz, this temporal resolution was adequate for quantifying recovery dynamics. A rate of 30 Hz is also considered relatively high for real-time acquisition of a 16 × 10 FSR sensor

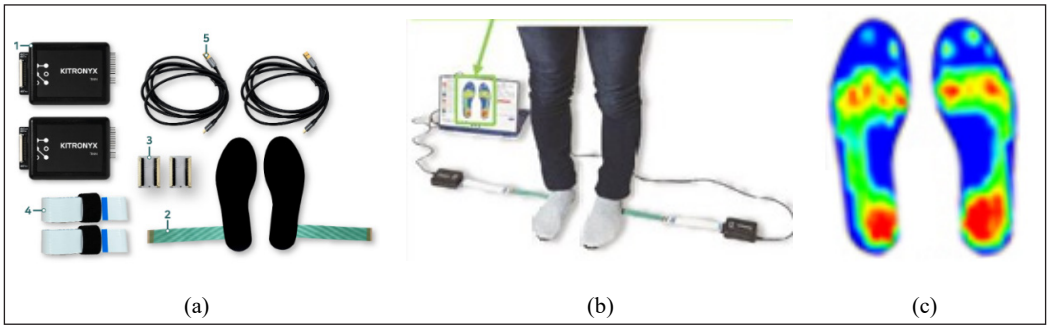


Figure 1. Kitronyx MP2513 Insole Sensor Kit: (a) Kit parts; (b) Kit usage; (c) real-time single frame of 16x20 foot pressure data captured
Source: Adapted from the website of Kitronyx.co.kr

array. Due to inherent hardware limitations, including multiplexed sensor scanning and signal stabilisation requirements, achieving sampling rates in the order of several hundred hertz is not practically feasible without significantly increasing system cost and complexity.

Prior to each experiment, the 16 × 10 FSR sensor array was calibrated under unloaded conditions. With no external force applied to the 16 × 10 FSR sensor array, all sensing elements were expected to produce an identical baseline output. Sensors exhibiting slight deviations due to manufacturing tolerances or electronic offsets were compensated via software-based offset calibration, ensuring all sensors shared the same baseline value under zero-load conditions.

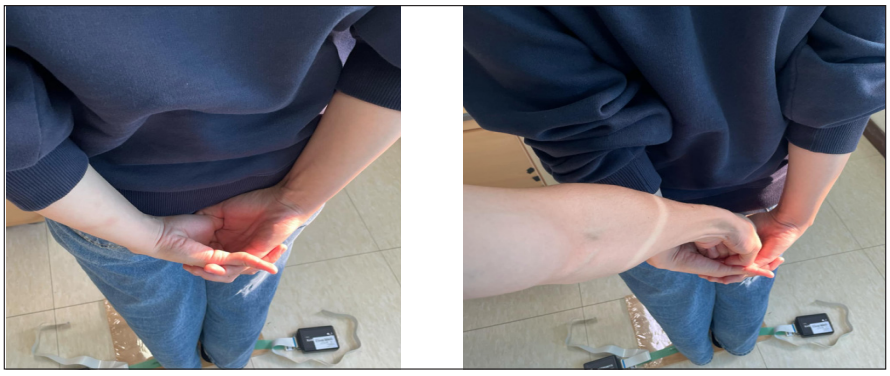
It is also noted that plantar pressure distribution may exhibit inherent asymmetry due to individual foot morphology and habitual stance patterns. To address this, baseline pressure distribution was recorded prior to perturbation and used as a subject-specific reference condition. All analyses were performed using relative deviations from this baseline, thereby reducing the influence of inter- and intra-individual asymmetry.

Perturbation Control

The downward push was applied manually to simulate a natural balance disturbance. To minimise potential bias, the experimenter applying the perturbation was blinded to the ground-truth condition (truthful vs non-truthful trial). To reduce variability in the magnitude of the applied force, the experimenter conducted repeated practice trials prior to data collection to ensure a consistent level of force across trials. While this approach cannot eliminate variability entirely, it was intended to maintain the perturbation within a consistent operational range. This approach is conceptually aligned with established experimental practices in polygraph testing, where efforts are made to standardise stimulus delivery (e.g. questioning tone and procedure) to reduce unintended variability.

Deception Task Protocol

To collect experimental data, a protocol was designed to mimic a traditional playful test using a mock crime paradigm with one physically healthy adult participant. The experimental procedure was conducted as follows. The subject is given two A4-sized sheets of paper and instructed to draw an apple on one sheet while leaving the other blank. Since the subject personally drew the apple on one of two sheets, the subject is aware of which sheet contains the apple drawing. However, the experimenter conducting the test is blinded to the sheet containing the apple. This is to prevent potential experimental bias. If the experimenter knew which sheet contained the apple, there would be a risk that they might unintentionally vary the magnitude of the downward push, which could influence the outcome of the experiment.



(a) (b)

Figure 2. Subject's Posture: (a) Standing upright with the subject's feet close together while clasping one's hands behind the back; (b) A short, forceful downward push applied to the clasped hands behind the back

Following this, the subject is instructed to stand upright with one's feet close together on the Kitronyx MP2513 (2024) Insole sensors while clasping one's hands behind one's back, as shown in Figure 2a. Plantar pressure data collection then commenced. This posture is selected to ensure that variations in balance primarily result from internal responses rather than external movement artefacts. A narrow base of support (feet close together) reduces lateral stability and increases reliance on fine postural control, thereby amplifying perturbation-induced balance responses. In addition, restricting arm movement minimises compensatory upper-body adjustments, allowing perturbation effects to be more directly reflected in plantar pressure distribution. Future work will compare shoulder-width stance and free-arm posture to evaluate the sensitivity-ecological validity trade-off.

Approximately 1s after the start of data recording, the experimenter asks the subject a single question: *"Have you truly draw an apple on this sheet?"* while pointing at one of two sheets. The subject is instructed to respond "*Yes, I did* " regardless of whether the sheet actually contains the apple. Between 3.0s and 3.8s, a short, forceful downward push is applied to the subject's clasped hands. This perturbation served as a controlled balance disturbance, simulating an unexpected postural shift and eliciting a natural balance response, as shown in Figure 2b.

The perturbation was applied approximately 3 seconds after trial initiation to ensure a stable baseline posture and allow the cognitive state (truthful vs non-truthful) to be active. Future studies will vary the timing of perturbations to analyse the temporal dynamics of cognitive influence.

For data interpretation, each trial is labelled based on whether or not the subject's verbal response matches their actual behaviour. Labels were binary: truthful or non-truthful. A trial is classified as truthful when the subject's verbal response is consistent with their actual behaviour. This included cases in which the subject responded "*Yes, I did* " when the experimenter pointed to the sheet containing the apple.

Conversely, a trial is classified as non-truthful when the subject's verbal response is inconsistent with the subject's actual behaviour. This included cases in which the subject responded "*Yes, I did*" when the experimenter pointed to the blank sheet on which the subject did not draw the apple.

In the present study, factors such as deliberate countermeasures, intentional balance control, and cognitive distraction were not explicitly addressed. These conditions pose known challenges not only for the proposed approach but also for conventional polygraph-based methods, in which intentional modulation of physiological responses remains an active area of research.

The experimental design employed a predetermined response paradigm, in which the subject was instructed to consistently answer "*Yes, I did*" regardless of the actual condition. This approach was adopted to ensure consistent verbal output across trials and to minimise variability associated with speech-related factors. Furthermore, the use of a fixed response was intended to simulate a simplified interrogation scenario under controlled conditions. While this design may introduce additional cognitive demand, it allows controlled comparison between conditions by reducing extraneous variability. Future work will incorporate more refined experimental paradigms, including variable response conditions and additional behavioural controls, to better approximate realistic settings and more precisely isolate deception-specific effects.

It is important to note that the mock crime task used in this study is intentionally simplified to maintain experimental control. While such a design does not fully replicate real-world investigative scenarios, which often involve stronger emotional and motivational factors, it enables isolation of the fundamental relationship between deception-related processes and postural responses. The present study should therefore be interpreted as a proof-of-concept feasibility study aimed at exploring whether plantar pressure patterns can serve as behavioural indicators of deception. These labels are listed in Table 1.

Table 1
The label of truthful or non-truthful trial

Label of a trial	Subject's Response	Subject's behaviour
Truthful	"Yes, I did"	Drew an apple on the sheet
Non-truthful	"Yes, I did"	Did not draw an apple on the sheet (left it blank)

Data Structuring and Visualisation

Data collection in each experimental trial lasts 6 seconds, during which 180 frames of foot pressure data are recorded. Each frame consists of a 16 × 20 matrix representing instantaneous pressure distribution across both feet, as shown in Figure 3(a).

To analyse these time series 180-frame sequences (noted as F_1, F_2, ⋯, F_{180}) visually, we transformed them into a structured grid of 60 rows (noted as R_1, R_2, ⋯, R_{60}) and 3 columns

(noted as C_1, C_2, C_3), which forms a 60x3 image matrix. This layout allowed the data to be viewed chronologically from bottom to top and left to right, preserving temporal continuity while offering a spatially compact representation. Each frame, F_i, is 1/30 s long, each row (R_i) having three columns (C_1, C_2, C_3) is 0.1s long, and 180 frames are 6s long. With that knowledge, frame F_1 at (R_1, C_1) is taken at $1/30$ s, frame F_2 at (R_1, C_2) at $2/30$ s, frame F_3 at (R_1, C_3) at $3/30$ s, frame F_4 at (R_2, C_1) at $4/30$ s, frame F_5 at (R_2, C_2) at $5/30$ s, frame F_6 at (R_2, C_3) at $6/30$ s, and so on. The final frame, F180, at (R60, C3) is captured at 6s. These frame indexing and timeline schemes for a 60 × 3 image matrix are illustrated in Figure 3(b). The mapping of frames to their indexed positions and time points facilitated quantitative and visual comparisons of postural balance responses between truthful and non-truthful trials.

Figure 3 illustrates the actual plantar pressure pattern of a subject maintaining an upright posture with feet close together without any external perturbation on clasped hands behind the back. When the body is in a balanced state, plantar pressure is evenly distributed across the foot, with comparable pressure levels in both the heel and toe regions. Also, greater foot pressure is observed on the right foot than on the left.

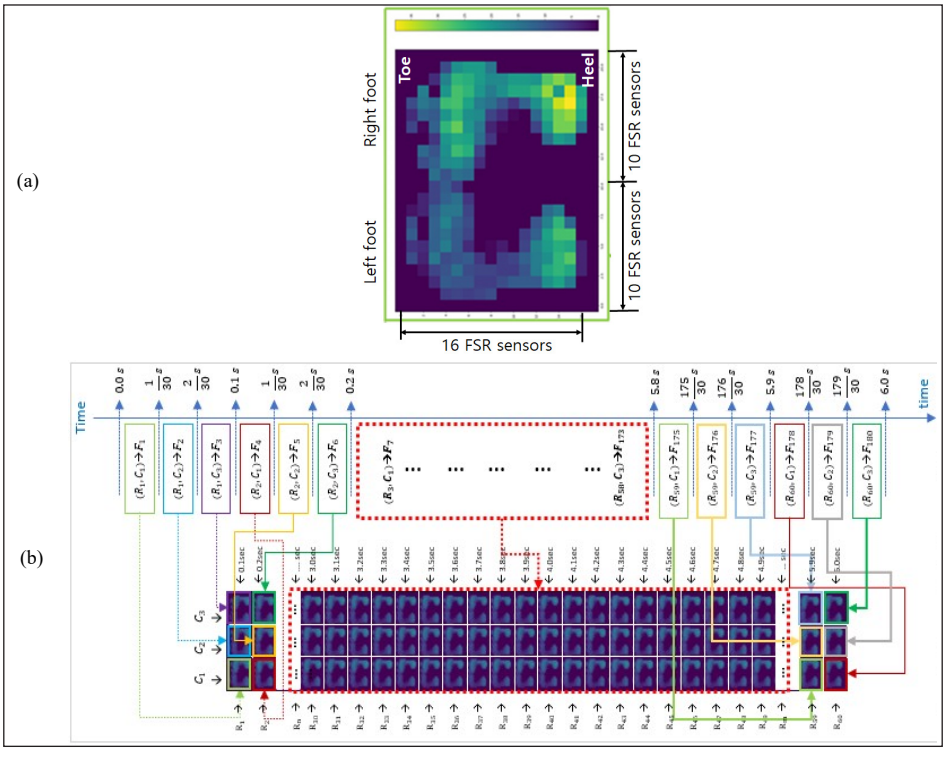


Figure 3. Frame indexing and timeline scheme for a 6s-long trial: (a) One frame of 16 × 20 FSR sensors (1/30s); (b) 180 sequential pressure frames (16 × 20 FSR sensor data) transformed into a 60 × 3 image matrix, which is organised by row(R1~R60) and column (C1~C3), progressing left to right and bottom to top in chronological order

Although the present study primarily relied on visual inspection as part of an exploratory analysis, the proposed 60 × 3 temporal representation naturally supports the extraction of several quantitative metrics that can enable objective statistical comparison, such as Centre of Pressure displacement/trajectory, Heel-toe pressure ratio, Recovery latency, Pressure variance, etc. In future work, these quantitative features will be systematically extracted and analysed using statistical methods and machine learning approaches to objectively distinguish between truthful and non-truthful conditions

RESULT AND DISCUSSIONS

Analysis of the structured 60 × 3 pressure matrices revealed distinct postural response characteristics between truthful and non-truthful trials, as shown in Figure 4 and Figure 5.

In a Truthful Trial

In a truthful trial shown in Figure 4(b), the subject is given two A4-sized sheets of paper and instructed to draw an apple on one sheet while leaving the other blank. And then stood upright with one's feet close together on the Kitronyx MP2513 (2024) Insole sensors while clasping one's hands behind one's back. Plantar pressure data collection then commenced. Approximately 1s after data recording began, the experimenter then pointed to the sheet containing the apple and asked the subject for the question: "*Have you truly draw an apple on this sheet?*". The subject responded, "*Yes, I did* ".

Around 3.1s, corresponding to frame F_{94} at (R_{32}, C_1), a short, forceful downward push was applied to the subject's clasped hands. Around $3.4\frac{1}{30}$ s, corresponding to frame F_{103} at (R_{35}, C_1), pressure in the toe region began to decrease. At approximately 3.5s, corresponding to frames $F_{106\text{-}108}$ at (R_{36}, C_1-C_3), pressure in the toe region nearly disappeared, while pressure was concentrated primarily in the heel region. This indicates a short-term backward shift of body weight.

However, postural balance was rapidly restored at approximately $3.7\frac{2}{30}$ s, corresponding to frame F_{110} at (R_{37}, C_2). By approximately 3.9s, corresponding to frame F_{115} at (R_{39}, C_1), plantar pressure returned to a more evenly distributed pattern across the feet, indicating complete recovery of postural stability.

The trial returned to baseline posture within approximately $0.4\frac{2}{30}$. The following perturbation reflects stable postural control and rapid balance recovery. In a truthful trial, the plantar pressure patterns exhibited quick stabilisation with minimal changes in postural pressure.

In a Non-Truthful Trial

In a non-truthful trial shown in Figure 4(a), the subject is given two A4-sized sheets of paper and instructed to draw an apple on one sheet while leaving the other blank. And

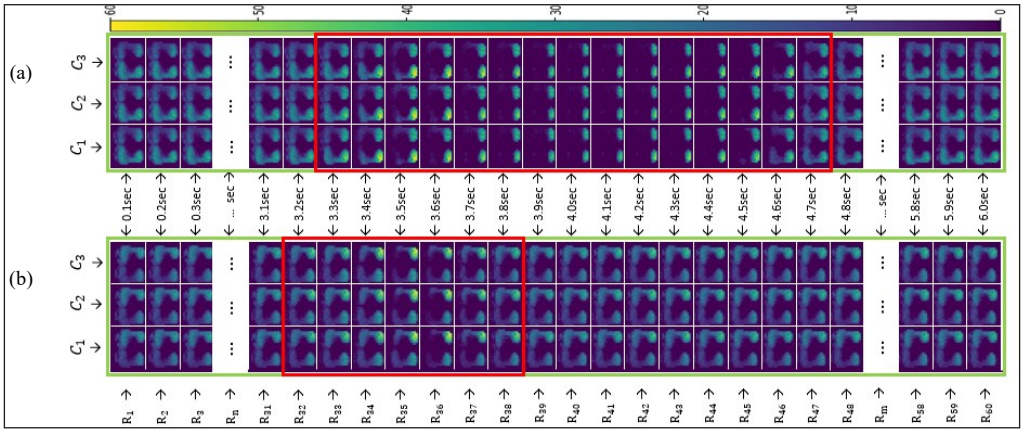


Figure 4. Comparisons of frame sequences from truthful and non-truthful trials: (a) The plantar pressure responses for the truthful trial; (b) The plantar pressure response for the non-truthful trial. Red-coloured regions highlight the plantar pressure responses following a short, forceful downward push applied to the clasped hands

then stood upright with one's feet close together on the Kitronyx MP2513 (2024) Insole sensors while clasping one's hands behind one's back. Plantar pressure data collection then commenced. Approximately 1s after data recording began, the experimenter then pointed to the blank sheet on which the subject had not drawn an apple and asked the same question: "*Have you truly drawn an apple on this sheet?*" The subject responded "*Yes, I did* ", thereby falsely claiming that an apple had been drawn on that sheet.

Around 3.2s, corresponding to frame F_{97} at (R_{33}, C_1), a short, forceful downward push was applied to the subject's clasped hands. At approximately $3.3\frac{1}{30}$ s, corresponding to frame F_{100} at (R_{34}, C_1), pressure in the toe region began to decrease. By approximately 3.4s, corresponding to frame F_{102} at (R_{34}, C_3), pressure in the toe region had further weakened.

From approximately $3.4\frac{1}{30}$ s to 4.1s, corresponding to frames $F_{103\text{-}121}$ from (R_{35}, C_1) to (R_{41}, C_3), pressure in the toe region nearly disappeared, while pressure became concentrated primarily in the heel region and began to decrease as the body moved further backwards. This indicates a pronounced backward shift of body weight. Around 4.1s, pressure in the heel region appears to be lowest, marking the initiation of forward recovery. Subsequently, from approximately 4.1s to 4.6s, corresponding to frame $F_{121\text{-}138}$ from (R_{41}, C_3) to (R_{46}, C_3), pressure in the heel region gradually increased again as the body attempted to regain balance.

At approximately $4.6\frac{1}{30}$s, corresponding to frame F_{139} at (R_{47}, C_1), pressure in the toe region reappeared and progressively increased in intensity. By approximately $4.7\frac{1}{30}$ s, corresponding to frame F_{142} at (R_{48}, C_1), plantar pressure distribution had returned to a baseline state, indicating recovery of postural stability.

Overall, the body began learning backwards at approximately $3.3\frac{1}{30}$. Following perturbation and initiated forward recovery at around 4.1s, with full stabilisation achieved near $4.7\frac{1}{30}$s. Compared to the truthful trial, the non-truthful trial exhibited a more pronounced backward displacement and delayed recovery.

The trial returned to baseline posture within approximately 1.4s following perturbation, reflecting a substantial loss of balance and slower stabilisation. In the non-truthful trial, the plantar pressure patterns showed greater deviation from baseline and larger fluctuations in postural loading than in the truthful trial.

These experimental results are listed in Table 2.

Table 2
Comparison of the time taken during foot pressure changes for one subject.

Label	Foot pressure change		
	Initiated	**Recovered**	**Duration**
Truthful	$3.4\frac{1}{30}$s	3.9s	$0.4\frac{2}{30}$ s
Non-truthful	$3.3\frac{1}{30}$s	$4.7\frac{1}{30}$ s	1.4s

Truthful .vs. Non-Truthful Trials with Three Additional Participants

The same experimental protocol was conducted with three additional participants. Although these additional trials were not included in the main analysis due to space limitations, consistent qualitative patterns were observed. Specifically, non-truthful responses exhibited larger backward displacement and slower postural recovery following perturbation compared with truthful responses. The experimental results from three additional participants are listed in Table 3 and depicted in Figure 5.

Despite inter-individual differences in baseline characteristics, a consistent trend was observed: recovery time under non-truthful trials was approximately 2-3 times longer than under truthful trials. A consistent increase in recovery time under non-truthful trials across

Table 3
The comparison in recovery time after the perturbation was applied to four subjects

Subject	Recovery Time		Ratio (Non-Truthful / Truthful)
	Truthful Trial	**Non-Truthful Trial**	
Subject A	0.5s	1.7s	3.4
Subject B	0.5s	1.3s	2.6
Subject C	0.4s	1.3s	3.25
Primary Subject	0.5s	1.4s	2.29

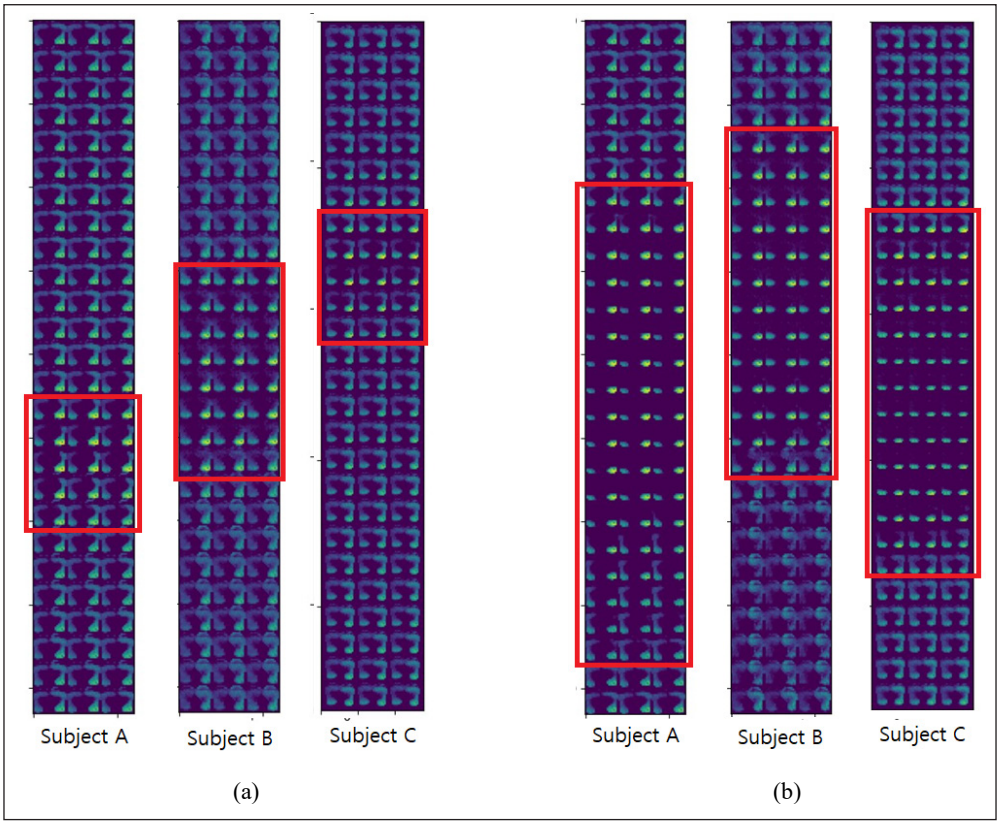


Figure 5. Comparisons of frame sequences from truthful and non-truthful trials for three additional subjects: (a) The plantar pressure responses for the truthful trial; (b) The plantar pressure response for the non-truthful trial. Red-coloured regions highlight the plantar pressure responses following a short, forceful downward push applied to the clasped hands

multiple participants suggests a potentially reproducible behavioural pattern rather than an individual-specific response. These observations suggest that the phenomenon reported in the primary experiment may not be limited to a single individual. The observed recovery time differences were consistently larger than intra-subject variability, suggesting that the effect may indicate a reproducible behavioural pattern.

Although the present analysis primarily focuses on anterior–posterior heel–toe pressure shifts, these dynamics are closely related to the trajectory of the centre of pressure (COP) along the anterior–posterior axis. Variations in heel–toe pressure distribution also provide indirect information on pressure variance across the foot during balance disturbance and recovery. While additional spatial characteristics such as mediolateral COP displacement and lateral sway may provide complementary information, the applied perturbation predominantly induced forward–backward postural adjustments. As a result, mediolateral

responses were relatively smaller and less sensitive for distinguishing between truthful and non-truthful conditions in the current experimental setup.

DISCUSSIONS

From the Viewpoint of the Theoretical Background

Deception is widely associated with increased cognitive load, as individuals must suppress truthful information while constructing and maintaining a false response. These processes rely on executive control functions, primarily mediated by prefrontal cortical networks, including working memory, inhibitory control, and conflict monitoring.

Postural control, although often considered an automatic motor function, depends on continuous sensorimotor integration involving the central nervous system. This includes integrating visual, vestibular, and proprioceptive inputs with motor outputs to maintain balance. Importantly, postural stability is not entirely independent of higher-level cognitive processes; prior studies have shown that increased cognitive load can interfere with motor coordination and balance regulation**,** particularly under conditions requiring rapid adaptation, such as perturbations.

From a neurophysiological perspective, cognitive load and motor control share limited neural and attentional resources. When cognitive demand increases—such as during deception—these shared resources may be partially reallocated from sensorimotor processes to cognitive control processes. This dual-task interference can reduce efficiency in postural regulation, resulting in delayed or less stable balance recovery following an external perturbation.

In the context of the present study, the key observation is not the initial loss of balance induced by the perturbation, which is expected in all conditions, but rather the difference in recovery dynamics**.** Specifically, deceptive responses may be associated with prolonged postural recovery, reflecting a temporary disruption in sensorimotor control due to increased cognitive demand.

This conceptual framework is consistent with principles underlying conventional polygraph-based approaches, which do not directly measure deception itself but instead infer it from relative changes in physiological responses associated with autonomic nervous system activation (e.g., heart rate, respiration, and skin conductance). Similarly, the present study does not assume a direct "deception signal," but rather examines behavioural changes in postural regulation that may arise from cognitive processes associated with deception.

While the current findings provide preliminary support for this interpretation, further studies involving larger participant samples and controlled cognitive conditions will be necessary to more precisely characterise the relationship between deception-related cognitive load and motor balance regulation.

From the Viewpoint of Experimental Results

The experimental results demonstrated consistent differences in postural recovery dynamics between truthful and non-truthful conditions. Truthful responses were associated with faster stabilisation and relatively symmetric pressure distribution, whereas non-truthful responses exhibited delayed recovery and greater deviation from baseline.

From a mechanistic perspective, these differences may arise from the combined influence of cognitive load and emotional arousal. Deceptive responses require suppressing truthful information and maintaining a false response, thereby increasing cognitive demand. At the same time, emotional factors such as anxiety, anticipation, or fear of being judged may further influence postural control. Therefore, the observed delayed recovery in non-truthful trials may reflect an interaction between cognitive and emotional processes rather than a single isolated factor.

Importantly, the present study adopts a within-subject comparison framework based on deviations from baseline, which helps reduce the influence of individual differences in emotional state or posture. However, the current design does not fully disentangle cognitive and emotional contributions. Future studies will incorporate controlled experimental conditions, including cognitive load-only tasks and emotion-inducing tasks without deception, to more precisely isolate these effects.

The current study focuses on two behavioural states (truthful vs non-truthful). However, additional conditions such as deliberate countermeasures, intentional balance control, and cognitive distraction may influence postural responses. These challenges are also present in conventional polygraph-based approaches, where intentional modulation of physiological responses remains an open issue. Future studies will therefore evaluate the robustness of plantar-pressure-based indicators under such conditions.

The mock crime task used in this study was intentionally simplified to maintain experimental control. While this design does not fully replicate real investigative scenarios, which involve stronger emotional and motivational factors, it allows isolation of the relationship between deception and postural response. Notably, even under these low-stakes conditions, consistent differences were observed. Future work will extend the paradigm to include higher-stakes and more ecologically valid scenarios.

It is important to emphasise that the present approach does not directly detect deception itself. Instead, it identifies behavioural correlates associated with cognitive and emotional processes underlying deceptive responses. This perspective is conceptually aligned with polygraph-based approaches, which infer deception from relative physiological changes rather than direct measurement.

To establish this method as a reliable deception detection approach, a comprehensive validation framework is required. This includes large-scale participant studies, standardised experimental protocols, quantitative feature extraction, statistical analysis, and machine

learning-based classification. In particular, sensitivity, specificity, and predictive reliability must be evaluated using cross-validation and independent datasets. Robustness under countermeasures and variability in experimental conditions must also be systematically assessed.

CONCLUSION

This study proposes a novel approach for deception detection based on plantar pressure pattern analysis using FSR sensors. The results demonstrate that non-truthful conditions are associated with delayed postural recovery and increased instability following perturbation. While preliminary, the results indicate a consistent behavioural pattern that may reflect underlying cognitive and emotional processes associated with deception.

Unlike conventional polygraph systems, the proposed method does not require electrodes or direct physiological monitoring, making it non-invasive, comfortable, and practical for real-world applications.

The present study should be interpreted as a proof-of-concept feasibility study, and substantial additional work is necessary to establish clinical- or forensic-level validity, including large-scale, diverse-participant studies and standardised experimental protocols. Future research will focus on large-scale validation, quantitative feature extraction, and machine learning-based classification to establish sensitivity, specificity, and predictive reliability. Additional studies will also examine the effects of cognitive load, emotional factors, and countermeasures to evaluate robustness under realistic conditions.

The present study demonstrates the feasibility of using plantar pressure–based postural dynamics as a behavioral indicator of deception.

ACKNOWLEDGEMENTS

This research was supported in part by the Soonchunhyang University Research Fund and in part by the National Research Foundation of Korea (NRF), funded by the Korean government (MSIT) (No.2021R111A3043994)

REFERENCES

Batgerel, G., & Kwon, C. (2026, February 7-9). *The use of FSR-based foot pressure patterns for deception detection* [Conference presentation]. International Conference on Next-generation Convergence Technology (ICNCT2026), Macau, China.

Gallardo-Antolín, A., & Montero, J. M. (2021). Detecting deception from gaze and speech using a multimodal attention LSTM-based framework. *Applied Sciences, 11*(14), Article 6393. https://doi.org/10.3390/app11146393

Grubin, D., & Madsen, L. (2005). Lie detection and the polygraph: A historical review. *The Journal of Forensic Psychiatry & Psychology, 16*(2), 357-369. https://doi.org/10.1080/14789940412331337353

Kim, Y. (2020). *LSTM based development of speech deception detection algorithm* [Master's thesis, Hanyang University].

Kitronyx Inc. (2024, June). *Quick start guide – MP2513 Insole Sensor Kit*. https://www.kitronyx.com/manual/?bmode=view&id=2850543

Krapohl, D., & Shaw, P. (2015). *Fundamentals of polygraph practice*. Academic Press.

Nam, B. R., Kim, J. Y., Kim, M. S., Choi, J., Cho, B. H., & Kim, I. Y. (2021). Technical note: Development of non-contact deception detection using facial expression and analysis of significant time period. *Journal of Science Criminal Investigation, 15*(3), 238-243. https://doi.org/10.20297/jsci.2021.15.3.238

Owayjan, M., Kashour, A., Al Haddad, N., Fadel, M., & Al Souki, G. (2012). The design and development of a lie detection system using facial micro-expressions. In *2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)* (pp. 33-38). IEEE. https://doi.org/10.1109/ICTEA.2012.6462897

Pertanika J. Sci. & Technol. 34 (5): 2691 - 2713 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Urban Flooding as a Hydrological Connectivity Problem: An Entropy-Based GIS Assessment in a Low-Relief Tropical City

Fachruddin Abdul Chaidir[1], Masha Nur Salsabiela Menhat[1]*, Mohd Azhafiz Abdullah[1], Muhammad Azraie Abdul Kadir[2], Effi Helmy Ariffin[3], Khairul Nizam Abdul Maulud[4], Ahmad Shakir Mohd Saudi[5], and Hasdinar Umar[6]

[1]Faculty of Maritime Studies, Universiti Malaysia Terengganu, 20130 Kuala Nerus, Terengganu, Malaysia
[2]Neo Environment Technology, Centre for Diploma Studies, Universiti Tun Hussein Onn Malaysia, 84600 Panchor, Johor, Malaysia
[3]Institute of Oceanography and Environment, Universiti Malaysia Terengganu, 20130 Kuala Nerus, Terengganu, Malaysia
[4]Department of Civil Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[5]Center for Water Engineering Technology, Malaysia France Institute, Universiti Kuala Lumpur, 43650 Bandar Baru Bangi, Selangor, Malaysia
[6]Department of Ocean Engineering, Faculty of Engineering, Hasanuddin University, 92171 Gowa, South Sulawesi, Indonesia

ABSTRACT

Rapid urban development can intensify drainage limitations and urban flooding in low-lying areas, where hydrological connectivity may be more informative than terrain gradient alone. This study applied an entropy-weighted Geographic Information System (GIS) framework to Taman Sri Muda, Shah Alam, Malaysia, integrating a 0.5 m UAV-derived spatial dataset, seven flood-conditioning factors, and community-informed spatial validation. The factors were elevation, slope, distance to discharge channel, runoff volume, impervious surface ratio, street-drainage intersection density, and Index of Downstream Potentiality Ratio (IDPR). IDPR (0.424) and street-drainage intersection density (0.277) together accounted for 70.1% of the entropy-derived weight. Spatial validation yielded a hit rate of 73.08% (19 of 26 observed flood locations within high- and very-high-hazard zones; 95% binomial confidence interval: 52.4%-88.4%). A targeted sensitivity analysis of the dominant IDPR weight showed only small changes in the combined high- and very-high-hazard areas under perturbations of up

ARTICLE INFO

Article history:
Received: 02 July 2026
Accepted: 02 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.07

E-mail addresses:
didirfachrudin24@gmail.com (Fachruddin Abdul Chaidir)
mashanur@umt.edu.my (Masha Nur Salsabiela Menhat)
azhafiz@umt.edu.my (Mohd Azhafiz Abdullah)
mazraie@uthm.edu.my (Muhammad Azraie Abdul Kadir)
effihelmy@umt.edu.my (Effi Helmy Ariffin)
knam@ukm.edu.my (Khairul Nizam Abdul Maulud)
ahmadshakir@unikl.edu.my (Ahmad Shakir Mohd Saudi)
hasdinar.umar@gmail.com (Hasdinar Umar)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

to ±75%. The framework is therefore presented as a relative, drainage-oriented spatial screening tool for flood-hazard prioritisation in low-relief, data-scarce urban settings, rather than as a predictive model of flood depth or extent.

Keywords: Drainage system performance, entropy-based weighting, GIS-based framework, hydrological connectivity, low-relief urban watershed, UAV-derived mapping, urban flood hazard mapping

INTRODUCTION

Due to the urban heat island effect and rising average rainfall, urban flooding is currently the most common natural disaster and has the potential to affect 379 million people (Manandhar et al., 2023). Urbanisation drives flooding, due to an expansion of impermeable surfaces and development outpacing drainage (Bibi et al., 2023; Du et al., 2019). In low-relief areas, minimal local topographic variation and gentle gradients can reduce drainage efficiency and favour water accumulation. Low-relief landscapes are characterised by limited local relief, so even small elevation differences may strongly influence surface-water pathways (Leopold et al., 1964). This situation is even more pronounced when drainage is obstructed, the floodplain is used inappropriately, and permeability is reduced in small catchment areas (Agonafir et al., 2023; Hassan et al., 2022).

Mapping areas prone to urban flooding is one of the most effective methods for reducing flooding risk across communities by increasing awareness and preparedness for flooding (Fernández & Lutz, 2010). Many methods for flood management have been developed, including remote sensing with machine learning, hydrological process modelling, and participatory methods (Abdessamed & Abderrazak, 2019; Guo & Guan, 2022; Khan et al., 2022). Hydrological and hydraulic models such as HEC-RAS and SWMM offer high physical fidelity but require extensive calibration data and significant computational resources, limiting applicability in data-scarce settings (Abdessamed & Abderrazak, 2019; Feng et al., 2020; Guo & Guan, 2022). Machine learning approaches, including random forest and convolutional neural networks, demonstrate strong predictive performance when large, labelled inventories are available, but are prone to overfitting in small-sample contexts (Guo & Guan, 2022; Khan et al., 2022). Multi-Criteria Decision Analysis (MCDA) integrated with GIS offers a transparent, data-efficient alternative suited for low-data environments, and entropy-based weighting further replaces subjective expert assignment with an objective measure derived from the spatial variability of each conditioning factor (Al-Hinai & Abdalla, 2021; Fernández & Lutz, 2010; Sharma et al., 2024). Most existing methods, however, require large quantities of data, are extremely computationally intensive, and difficult to implement in data-poor regions (Feng et al., 2020; Hassan et al., 2022; Manandhar et al., 2023). There is a clear demand for efficient urban flood-mapping methods that can adapt to the various conditions in which they will be used.

Recent studies in similarly ranked journals have advanced flood susceptibility mapping in several methodological directions. Sharma et al. (2024) applied Shannon entropy, combined with frequency ratio, within a GIS framework to assess flood susceptibility in Chandrapur District, India, demonstrating the utility of entropy-based weighting in data-scarce flat terrain, though without UAV-derived inputs or drainage connectivity factors. Systematic reviews of urban flood susceptibility mapping have identified the entropy weighting method as an objective, data-driven alternative to expert-opinion approaches such as AHP, while also noting that drainage infrastructure parameters remain consistently underrepresented across the literature (Islam et al., 2025). Goumghar et al. (2025) demonstrated the applicability of AHP-MCDM in data-scarce, semi-arid environments (Upper Draa Basin, Morocco; AUC = 0.882), though the study relied on conventional 30 m ASTER DEM inputs rather than high-resolution topographic data. The present work advances on these precedents by: (1) UAV-derived DSM at 0.5 m resolution; (2) explicit operationalisation of hydrological connectivity through IDPR and street-drain intersection density collectively accounting for 70.1% of entropy-derived weight; and (3) multi-source community-informed validation in a Malaysian low-relief tropical urban township.

Hydrological connectivity, the water-mediated transfer within hydrological systems (Dwivedi et al., 2025), encompasses structural and functional dimensions (Bracken et al., 2013). Urbanisation fundamentally alters these dynamics: impervious surfaces reduce infiltration and shorten flow travel times, while engineered infrastructure introduces artificial flow pathways (Du et al., 2015; Singh et al., 2018). In low-relief urban settings, flood susceptibility is governed more by drainage network geometry than terrain gradient (Dell'Aira & Meier, 2025), as terrain-based indices such as the Topographic Wetness Index (TWI) lose discriminatory power in flat areas (Moore et al., 1991). Operationalising hydrological connectivity is therefore essential for characterising flood mechanisms in low-relief tropical cities.

A widely used framework for integrating spatial parameters in flood assessment is Geographic Information Systems (GIS). Although they offer great utility, most GIS methodologies rely on ad hoc, arbitrary weighting, resulting in a high degree of uncertainty. The entropy method is a more sophisticated approach and determines weighting schemes based on the spatial data distribution, improving the consistency, clarity, and replicability of the approach (Al-Hinai & Abdalla, 2021; Liu et al., 2017; Malekinezhad et al., 2021; Sepehri et al., 2019; Sharma et al., 2024).

In addition to methodological improvements, the effectiveness of flood hazard mapping is critically dependent on map validation. Trust and resilience within the community are further enhanced through participation (Abid et al., 2024; Halim et al., 2024), and a comprehensive range of participants increases the utility of the maps for evacuation (Fernández & Lutz, 2010).

Unfortunately, validation methods primarily depend on historical flood inventories. These inventories may be incomplete or updated only intermittently. There is therefore a need for validation approaches that incorporate available on-site evidence, while recognising the limitations of sparse flood inventories for evaluating model performance.

This study addresses an operational gap in drainage-oriented flood-hazard screening for data-scarce, low-relief tropical urban areas. Building on established entropy-MCDA approaches (Malekinezhad et al., 2021; Sepehri et al., 2019), the study contributes: (1) application of a 0.5 m UAV-derived spatial dataset for fine-scale urban analysis; (2) explicit representation of hydrological connectivity through IDPR and street-drainage intersection density, which together account for 70.1% of the entropy-derived weight; (3) multi-source flood-location validation supplemented by community evidence from the December 2021 event; and (4) a targeted sensitivity analysis testing the dominant IDPR weight over ±25%, ±50%, and ±75% perturbations. The objective was not to propose entropy weighting as a new method, but to evaluate a reproducible drainage-oriented screening configuration for prioritising flood-prone locations when calibrated hydraulic model data are unavailable.

METHODOLOGY

Study Area

Taman Sri Muda (Figure 1) is a highly populated, low-relief, flood-prone area along the Klang River (Section 25, Shah Alam, Selangor, Malaysia; 3.0471° N, 101.5275° E). It has



Figure 1. Area of study

a population of 38,812 people in an area where the average population density for Shah Alam is 3,068 people/km². This flood-prone area has poor natural drainage, excessive impervious surfaces, and significant elevation reduction (Dewan Negeri Selangor, 2022; Muhamad Azizi & Teriman, 2023). The location and infrastructure of Taman Sri Muda make it an ideal setting for studying drainage-based flood-reduction methods in densely populated, low-relief urban areas. December 2021 was the most extreme flooding event to date and was caused by 280 mm of rainfall (Nooryati, 2022).

Flood Conditioning Factors: Selection Criteria and Data Preparation

Numerous interconnected physical and anthropogenic factors create complex, nonlinear direct drivers of flood hazard. The initial step was determining the conditioning factors. Since environmental and hydrological conditions vary greatly across geographic locations, determining conditioning factors is essential. In accordance with past studies, flood conditioning factors can be described in the following three ways. First, discrete criteria, which are primarily static physical characteristics of the landscape, for example, slope and elevation; second, integrated criteria, which are processes of hydrology that are sustained and intensified, for example, the volume of runoff; and third, the criteria of connectivity, which describe the spatial connection and interaction of flow between the elements of drainage combined. These three categories of conditioning factors are summarised in Table 1.

Table 1
Factors Influencing Flood Hazard Mapping

Category	Criterion	Resolution	Source	Former Studies
Topography	Elevation	0.5 m x 0.5 m	UAV (2023)	Fernández & Lutz, 2010; Malekinezhad et al., 2021; Tang et al., 2018
	Slope	0.5 m x 0.5 m	UAV (2023)	Fernández & Lutz, 2010; Malekinezhad et al., 2021; Tang et al., 2018
Hydrology	Distance to discharge channel	0.5 m x 0.5 m	UAV drainage (2023)	Fernández & Lutz, 2010; Malekinezhad et al., 2021
	IDPR	0.5 m x 0.5 m	UAV elevation/ drainage (2023)	Fernández & Lutz, 2010; Malekinezhad et al., 2021
	Runoff Volume	0.5 m x 0.5 m	UAV land cover FAO soil 2021 rainfall	Malekinezhad et al., 2021
	ISR	0.5 m x 0.5 m	UAV orthomosaic (2023)	Du et al., 2015; Feng et al., 2020
Human activity	Street-drainage intersection	0.5 m x 0.5 m	UAV drainage + OSM roads	Malekinezhad et al., 2021

The proposed analytical framework combines the findings of research by Sepehri et al. (2019), Malekinezhad et al. (2021), and Fernández and Lutz (2010) using a structured Multi-Criteria Decision Analysis (MCDA) framework to integrate evidence from raster and vector datasets and the field (Figure 2). The following seven primary factors constitute the flood-hazard assessment: elevation, slope, distance to discharge channel, runoff volume, street-drainage network intersection density, Index of Downstream Potentiality Ratio (IDPR), and Impervious Surface Ratio (ISR). The UAV-derived dataset at 0.5 m resolution provides a breakdown of fine-scale hydrological connectivity and topographical variance. The resulting flood-hazard index is a relative spatial screening index and should not be interpreted as a direct estimate of flood depth or flood extent.

Elevation

The study used an existing UAV-derived elevation dataset from the 2023 project survey, resampled to a 0.5 m raster for the present GIS analysis. Because the UAV acquisition and primary photogrammetric processing were undertaken as part of the source project rather than specifically for this study, only processing steps documented in the project materials are reported here. The project workflow used Agisoft Metashape for image alignment, pixel-based dense stereo reconstruction, mesh and point-cloud generation, and export of orthomosaic and Digital Surface Model (DSM) products in TIFF format. The elevation products were referenced to the project coordinate framework in UTM Zone 47N and related to the JUPEM mean sea level datum. Detailed flight metadata, formal ground-control-point documentation, and an independent vertical RMSE assessment were not available to the present study; this limitation was therefore acknowledged rather than inferred. The UAV-derived elevation surface was used for elevation classification and as the terrain input for

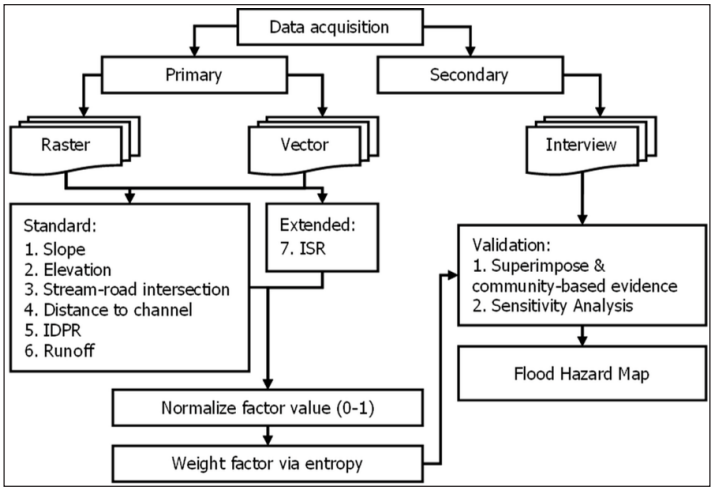


Figure 2. Methodology framework

slope and surface-flow analyses in ArcGIS Pro 3.4. The elevation surface was segmented into classes based on its value distribution and surface characteristics (Figure 3a). Low-elevation areas were assigned higher flood-hazard relevance because they represent preferential zones for surface-water accumulation (Malekinezhad et al., 2021).

Slope

Slope was derived in degrees from the UAV-derived elevation surface using ArcGIS Pro 3.4 (Figure 3b). Slope affects flow direction and velocity and is a well-documented flood-conditioning factor (Fernández & Lutz, 2010). In low-relief terrain, lower slopes reduce surface-flow velocity and favour water accumulation and waterlogging; accordingly, lower slope values were interpreted as representing greater flood-hazard potential (Malekinezhad et al., 2021; Tang et al., 2018).

Street-Drainage Network Intersection Density

The drainage-channel network was derived from the UAV-based terrain surface using flow-direction and flow-accumulation analyses in ArcGIS Pro 3.4, whereas the road network was obtained from OpenStreetMap (OSM) as road centreline data. Network-intersection points were identified by spatially overlaying the derived drainage network and the OSM road network. Kernel Density Estimation was then applied to the intersection points using a 100 m search radius and a 0.5 m output cell size, consistent with the analysis raster resolution. The resulting density surface represents the spatial concentration of street-drainage intersections per unit area and was normalised to a 0-1 scale before entropy weighting (Figure 3c). Roads can rapidly convey surface runoff toward drainage channels, while drainage-road intersections may act as local convergence nodes during intense rainfall; therefore, higher intersection density was interpreted as greater flood-hazard potential (Pimenta et al., 2025; Singh et al., 2018).

Runoff

Runoff volume was estimated using the Soil Conservation Service-Curve Number (SCS-CN) method, widely applied in urban flood studies for its ability to link rainfall with catchment retention characteristics (Malekinezhad et al., 2021). The method quantifies runoff as:

$$Q = \frac{(P-0.2S)^2}{(P+0.8S)}, for P \geq 0.2S \quad [1]$$

$$S = \frac{25400}{CN} - 254 \quad [2]$$

Equation 1 was used to calculate runoff depth (Q), while Equation 2 was applied to determine the potential maximum retention (S) based on the Curve Number (CN). Here, Q represents runoff depth (mm), P represents rainfall (mm), S represents potential maximum retention (mm), and CN represents land-surface infiltration characteristics. The Gleysol classification used by the source project was taken from FAO global soil-mapping resources accessed through the FAO Soils Portal. Rainfall represented the December 2021 extreme event (approximately 280 mm). Antecedent Moisture Condition (AMC) II was assumed to represent average antecedent soil moisture. CN values were assigned from UAV orthomosaic-based land-cover classes: fully connected impervious surfaces (roads and rooftops), CN = 98; partially connected impervious surfaces, CN = 90; and vegetated or open areas, CN = 61, following the adopted event-based SCS-CN setup and the approach of Malekinezhad et al. (2021). The runoff layer, therefore, represented a static, event-based index rather than a calibrated multi-event runoff model, and its applicability beyond the 280 mm event should be interpreted with caution.

Index of Downstream Potentiality Ratio (IDPR)

IDPR was implemented from the UAV-derived elevation surface through a surface-flow workflow in ArcGIS Pro 3.4. The raw IDPR was expressed as the ratio of the flow length from a cell to its downstream sink relative to the maximum flow length in the study area. Subsurface drainage pipes were not included in this surface-routing representation. Because the study area is low-relief, the D8 representation and the omission of below-ground drainage may introduce uncertainty in local flow paths; these limitations were explicitly considered in interpreting the resulting hazard index.

$$IDPR_i = \frac{L_i}{L_{max}} \quad [3]$$

Equation 3 was used to calculate the IDPR for each pixel using the downstream flow length and its maximum value within the study area. Where IDPR is the hydrological-connectivity index of a pixel (dimensionless), L is the flow length from the pixel to the nearest theoretical downstream sink, and Lmax is the maximum flow length used for normalisation (Malekinezhad et al., 2021). A smaller raw IDPR corresponds to a shorter downstream flow path and was interpreted as greater flood-hazard potential. For hazard standardisation, IDPR was therefore treated as an inverse (cost-type) criterion, so that smaller raw values receive higher normalised hazard scores. This treatment preserves the intended physical interpretation of short downstream flow paths in the weighted overlay. IDPR was preferred to the Topographic Wetness Index (TWI) because flow-path length can retain spatial discrimination in low-relief terrain where slope-dependent indices may become unstable as local slope approaches zero (Lane et al., 2009; Moore et al., 1991).

Impervious Surface Ratio (ISR)

The Impervious Surface Ratio (ISR) is a metric used to gauge how changes in land surfaces caused by human activities affect local hydrological systems. In cities, surging runoff and flood risks are a direct result of rapid land development via building construction, road and pavement installation, and other similar activities (Dell'Aira & Meier, 2025; Du et al., 2015; Ogden et al., 2011). ISR data for each land unit came from UAV land-cover classifications (Figure 3f). ISR data were used to indicate potential runoff, with higher ISR values indicating a greater probability of runoff.

Distance to Discharge Channel

Distance to discharge channels indicates proximity to the surface drainage network (Figure 3g). Euclidean Distance was calculated in ArcGIS Pro 3.4, with shorter distances interpreted as greater flood-hazard potential because locations close to drainage pathways may be more directly exposed to local surcharge, overflow, or convergent surface flow during intense rainfall. For weighted integration, the distance layer was inversely standardised so that smaller raw distances receive larger normalised hazard scores. Euclidean distance was selected as a transparent and computationally efficient approximation of drainage accessibility in the densely built study area. The method does not account for barriers such as buildings and elevated road embankments, or for the hydraulic capacity of drains; these simplifications are acknowledged as limitations (Fernández & Lutz, 2010; Malekinezhad et al., 2021).

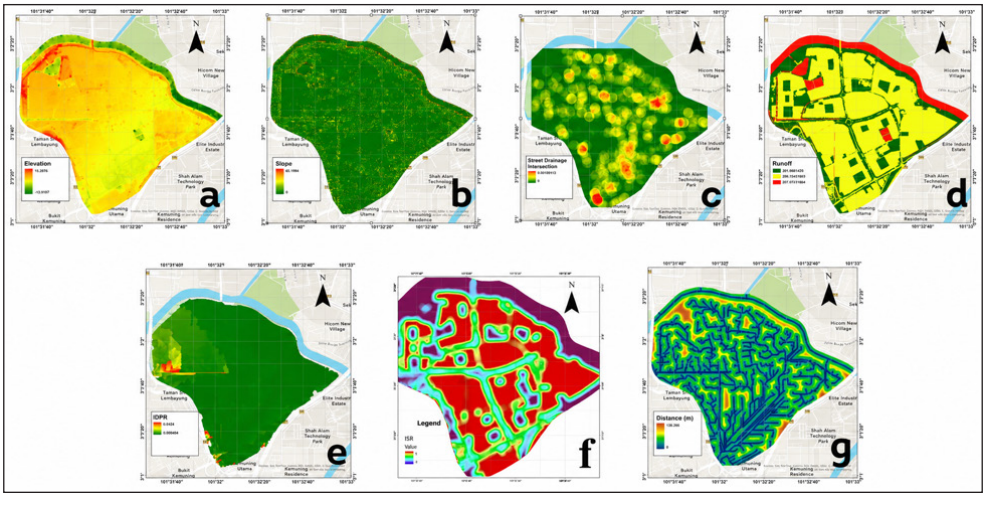


Figure 3. Flood conditioning factors used in this study: (a) Elevation, (b) Slope, (c) Street drainage network intersection density, (d) Runoff, (e) IDPR, (f) ISR, (g) Distance to discharge channel

Processing: Determination of Weights for Flood Conditioning Factors and Mapping

The entropy method reduces reliance on subjective MCDA weighting by quantifying the information contribution of each factor from its spatial variability (Malekinezhad et al., 2021; Sharma et al., 2024). Before entropy calculation, factor values were normalised to a common 0-1 hazard scale. Runoff volume, ISR, and street-drainage intersection density were treated as benefit-type criteria, for which larger raw values represent greater hazard. Elevation, slope, IDPR, and distance to discharge channel were treated as cost-type criteria and inversely standardised so that smaller raw values represent greater hazard. This directionality ensured that all normalised layers have a consistent interpretation: larger normalised values indicate greater flood-hazard potential.

$$P(x) = \frac{\mathrm{x} - \min(x)}{\max(x) - \min(x)} \quad [4]$$

$$P(x) = \frac{\max(\mathrm{x}) - \mathrm{x}}{\max(x) - \min(x)} \quad [5]$$

Equation 4 was used to normalise benefit-type criteria, while Equation 5 was applied to inversely normalise cost-type criteria, ensuring a consistent flood-hazard scale across all conditioning factors.

The equation for measuring the entropy for factor j is expressed as follows:

$$H_j = -k \sum_{i=1}^{n} p_{ij} \ln(p_{ij}) \quad [6]$$

Equation 6 was used to calculate the entropy value for each flood-conditioning factor based on the distribution of its normalised values across the study area.

where p_{ij} defines the quantity for the factor j at the spatial unit i., n is the number of spatial units, and $k = 1/(\ln n)$ is the normalisation constant. Smaller entropy values result in greater spatial variability and a greater contribution to information. The determination of the entropy-based weight for each factor was as follows:

$$w_j = \frac{1 - H_j}{\sum_{j=1}^{m} (1 - H_j)} \quad [7]$$

Equation 7 was applied to determine the entropy-based weight of each flood-conditioning factor, reflecting its relative contribution to the overall flood-hazard assessment. Here, m represents the total number of flood-conditioning factors.

Normalisation was performed across all raster cells within the complete study area boundary rather than only at validation points, so the entropy weights reflect the full spatial distribution of each factor. Zero values in the normalised raster were replaced with a small constant (p = 0.0001) before entropy calculation to avoid undefined logarithmic terms. All seven factors were directionally standardised to a common hazard orientation prior to entropy weighting: higher normalised values consistently indicate higher hazard. Runoff volume, ISR, and street-drainage intersection density were standardised as benefit-type criteria, whereas elevation, slope, IDPR, and distance to channel were inversely standardised as cost-type criteria. The continuous FHI was classified into five hazard classes (very low, low, moderate, high, and very high) using Jenks Natural Breaks.

$$\mathrm{FHI} = \sum_{j=1}^{m} w_j \cdot P_j \qquad [8]$$

Equation 8 was used to calculate the final Flood Hazard Index (FHI) by integrating the normalised flood-conditioning factors with their corresponding entropy-derived weights, where P_j represents the normalised raster layer of factor j, and w_j represents its entropy-derived weight. The continuous FHI values were then divided into five hazard classes (very low, low, moderate, high, and very high) using the Natural Breaks (Jenks) classification, which maximises differences between classes and minimises variance within a class.

Validation and Robustness Assessment of Flood Hazard Mapping

The flood hazard map was evaluated using spatial overlay validation, community-based confirmation, and a targeted robustness analysis given the limited official flood inventory. The 26 observed flood-occurrence locations were compiled from on-site documentation, official post-flood reports, media records, and interviews with residents who directly experienced the December 2021 event; source-specific counts were not retained separately and were therefore not reconstructed retrospectively. The model agreement was summarised as the proportion of observed flood locations falling within the high and very high hazard classes. Robustness was evaluated as a one-factor-at-a-time sensitivity test of the dominant IDPR weight only, using perturbations of ±25%, ±50%, and ±75% relative to its baseline value. The analysis, therefore, tests sensitivity to uncertainty in the highest-weighted factor and should not be interpreted as a full multivariate uncertainty analysis.

RESULTS AND DISCUSSIONS

Results of Weight Determination for Flood Conditioning Factors

Entropy-based weighting estimates each factor's relative contribution from spatial variability. Factors with lower entropy values exhibit greater spatial heterogeneity and

contribute more information for flood hazard differentiation (Al-Hinai & Abdalla, 2021; Haghizadeh et al., 2017; Malekinezhad et al., 2021).

IDPR showed the lowest entropy (0.935) and the largest entropy-derived weight (0.424), indicating the greatest contribution to spatial information among the seven factors. Street-drainage intersection density ranked second (entropy = 0.958; weight = 0.277). Runoff volume (entropy = 0.979; weight = 0.140) and ISR (entropy = 0.981; weight = 0.124) had intermediate contributions. Distance to drainage, slope, and elevation each had entropy values of 0.998, with corresponding weights of 0.015, 0.010, and 0.010, respectively. The near-uniform elevation and slope surfaces therefore contributed comparatively little spatial discrimination in this low-relief setting (Table 2).

In Taman Sri Muda, the low-relief terrain limits the spatial contrast provided by elevation and slope, whereas the organisation of drainage pathways and their interaction with the built environment provide stronger spatial differentiation. The large weights assigned to IDPR and street-drainage intersection density were therefore interpreted as evidence that drainage connectivity is a dominant discriminator within this relative hazard model. Runoff volume and ISR remain important secondary factors, representing runoff generation and imperviousness, but their effects were interpreted together with the capacity and organisation of the drainage system (Ogden et al., 2011).

The model interpretation is consistent with post-flood accounts that described drainage-system overload during the December 2021 event. This agreement does not establish hydraulic causality because drain capacity, pipe geometry, pumping, and sluice operations were not explicitly simulated. Rather, the entropy-weighted spatial pattern indicates that connectivity-related factors provide stronger discrimination than topographic variation within the study area. The result supports the use of the framework for first-order drainage-oriented screening and prioritisation, while calibrated hydraulic modelling remains necessary for quantitative assessment of flood depth, conveyance capacity, and infrastructure performance (Liu et al., 2017; Singh, 1997).

Table 2
Results of Entropy-based weighting showing the relative contribution of flood-conditioning factors to urban flood hazard mapping

Rank	Flood-conditioning factor	Entropy (Hj)	Divergence (dj)	Weight (wj)
1	IDPR	0.935	0.065	0.424
2	Street drainage intersection density	0.958	0.042	0.277
3	Runoff volume	0.979	0.021	0.140
4	Impervious surface ratio (ISR)	0.981	0.019	0.124
5	Distance to drainage	0.998	0.002	0.015
6	Slope	0.998	0.002	0.010
7	Elevation	0.998	0.002	0.010

Flood Hazard Map Based on Entropy Weighting

Flood hazard mapping is shown in Figure 4. The map was constructed from the entropy-derived weights through weighted overlay in ArcGIS Pro 3.4 and classified into five Natural Breaks (Jenks) classes. The resulting pattern shows localised concentrations of high and very high relative hazard associated with the drainage framework and the dominant connectivity factors.

The hazard map (Figure 4) exhibits an uneven spatial distribution. High and very high zones are concentrated mainly within densely built residential areas, particularly in the central and eastern portions of the study area, where impervious surfaces and drainage-road interactions are prominent. Based on the GIS areal output, very low, low, moderate, high, and very high classes occupy 0.483, 0.424, 1.067, 1.087, and 0.338 km², respectively. These correspond to approximately 14.22%, 12.46%, 31.40%, 31.99%, and 9.94% of the study area. Thus, moderate-to-very-high classes account for 73.32%, whereas low and very low classes together account for 26.68%. These classes represent relative spatial screening priorities rather than deterministic inundation depth or extent.

These results match previously documented findings. IDPR and intersection density received the highest-ranking weights, and the analysis concludes that the pattern of flood hazards is influenced primarily by the drainage system's connectivity and performance. In most flat, built-up, urban areas, drainage and road networks concentrate flows. Blockage of drainage and road networks leads to rapid water accumulation at intersections (Zhou et al., 2022). These results are consistent with other studies (DeSousa et al., 2024; Liu & Zhang, 2025), which show that the characteristics of urban drainage systems and the

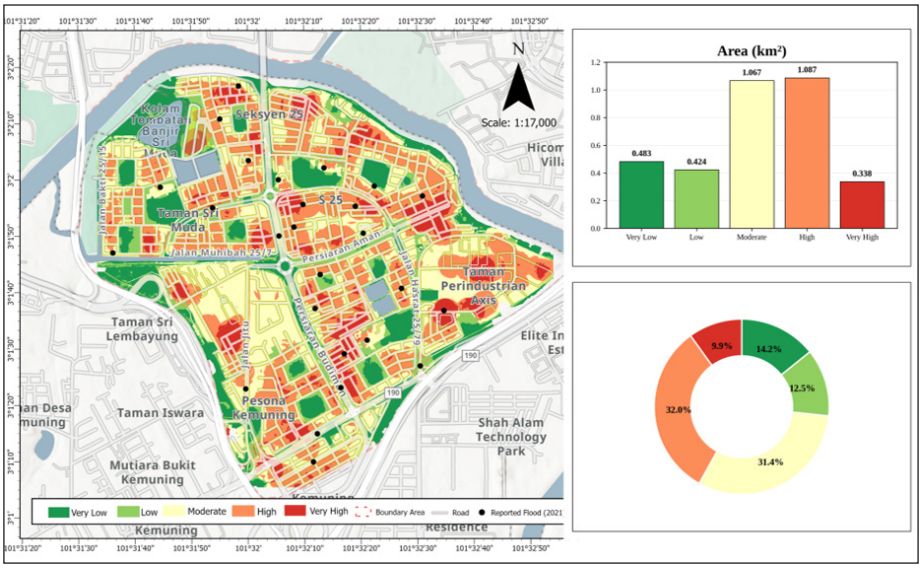


Figure 4. Spatial distribution and areal proportion of urban flood hazard classes in Taman Sri Muda

degree of connectivity dominate the location and the intensity of flooding, more so than the topography of the area. The moderate zones are areas that experience substantial runoff but have drainage systems only partially overloaded. The intermediate weights assigned to runoff and ISR indicate that runoff formation depends on the drainage system's capacity to receive flow (Ahmad et al., 2024; Hassan et al., 2022).

Low- to very-low-hazard zones are characterised by an increasing proportion of green open space, resulting in enhanced permeability and decreased flow velocities (Bai et al., 2018; De Vleeschauwer et al., 2014; Liu & Zhang, 2025). However, green zones adjacent to rivers or floodplains are susceptible to backwater and overflow (Yusuf et al., 2023). The 2021 flood fully inundated almost the entire Taman Sri Muda, including the green zones. This shows that the hazard map expresses the level of risk in relative terms, not that the zones are risk-free.

While some flood-susceptibility studies emphasise topographic controls (Gebremichael et al., 2025; Nizar et al., 2024), the present results suggest that topography provides less spatial discrimination in this flat, highly urbanised setting, where built surfaces and drainage connectivity exert stronger local contrasts (Komori et al., 2022). Post-flood documentation also describes drainage-system overload during the 2021 event. The spatial correspondence between the relative hazard pattern and available flood-positive locations supports prioritising drainage assessment, connectivity bottlenecks, and obstruction management in high and very high zones. Measures addressing imperviousness and urban green space should be considered complementary rather than interpreted as substitutes for hydraulic-capacity assessment.

Results of Spatial Validation and Robustness Analysis

Flood points were integrated into the model to evaluate spatial agreement between observed flood impacts and the mapped hazard pattern in a context where a complete formal flood inventory was unavailable. The 26 geocoded flood-occurrence locations were compiled from four source types: on-site documentation, official post-flood reports (Dewan Negeri Selangor, 2022; Nooryati, 2022), media records, and community interviews. Source-specific counts were not retained separately. Community interviews involved five (n = 5) flood-affected residents of Taman Sri Muda who directly experienced the December 2021 event. Verbal consent was obtained before face-to-face interviews covering flood experience, water depth and duration, rescue response, drainage conditions, and post-flood recovery. Interview-derived locations were referenced to respondents' residential/locality information within the study area rather than interpreted as instrument-surveyed flood coordinates. Respondents' identities were not reported, and personally identifiable information was not retained in the analysis. Interview evidence was used as contextual corroboration and triangulated with documentary sources rather than treated as a statistically representative

sampling. Potential recall bias, uncertainty from locality-based geocoding, media selection bias, and the absence of probabilistic sampling were therefore recognised.

The use of community-reported and digitally mediated flood-location data also requires explicit attention to privacy and data governance. In this study, interview-derived information was used only for spatial validation; identities were not reported, and personally identifiable information was not retained. This conservative approach is consistent with broader privacy-preserving principles in connected data environments, including secure auditing of cloud-based IoT data and decentralised/federated approaches that reduce unnecessary centralisation of user information (Khan et al., 2024; Ullah et al., 2025). These studies are cited to frame data-governance practice rather than to claim that their technical architectures were implemented in the present flood study.

As shown in Table 3, 73.08% of the observed flood-occurrence locations (19 of 26) fell within the high and very high hazard classes, with a 95% binomial confidence interval of 52.4%-88.4%. A further 26.92% (7 of 26) fell within the moderate class, and none of the available flood points fell within the low or very low classes. This result is reported as a spatial hit rate (success ratio), not as a spatial correlation coefficient. Because the inventory contains only flood-positive locations and no systematically verified non-flood controls, the result supports spatial concordance with observed flood locations but cannot assess specificity or false-positive performance.

The December 2021 flood is the most extensively documented event used in this study. Community accounts indicated recurrent local flooding and described the 2021 event as particularly severe, with rapid water-level rise and prolonged inundation. These observations provide contextual support for interpreting the mapped hazard pattern but do not constitute independent calibration data. More complete predictive validation would require a structured inventory containing both flood-affected and verified non-flood locations, allowing measures such as AUC-ROC, sensitivity, specificity, precision-recall, and false-positive rates. Such non-flood reference data were not available for the same

Table 3
Distribution of observed flood-occurrence points across flood-hazard classes used for spatial validation

Flood hazard class	Number of points (n)	Percentage (%)
Very low	0	0.00
Low	0	0.00
Moderate	7	26.92
High	13	50.00
Very high	6	23.08
High + Very high	19	73.08
Total	26	100.00

Note. Source Adapted from: (flood case 2021)

event; consequently, these metrics were not calculated, and the 73.08% value was retained only as a flood-point hit rate.

Robustness Analysis

A targeted robustness analysis tested the sensitivity of the mapped hazard pattern to the dominant IDPR weight. The baseline IDPR weight was perturbed by ±25%, ±50%, and ±75% on a per-factor basis, while the remaining factor structure was retained for comparison. Results are summarised in Table 4.

Across the tested scenarios, the area classified as high- or very-high hazard changed only modestly (approximately 1%-3%) even under a ±75% perturbation of the IDPR weight. Some weight-reduction scenarios produced slight increases rather than monotonic decreases in hazard area because the relative contribution of the remaining factors, particularly street-drainage intersection density and runoff, becomes more influential in the weighted overlay. These results indicate that the mapped high-hazard pattern is only weakly sensitive to large changes in the dominant IDPR weight. However, the analysis was restricted to one-factor-at-a-time perturbation of IDPR and does not demonstrate robustness to simultaneous uncertainty in all seven factors. Leave-one-factor-out analysis, simultaneous weight perturbation, and Monte Carlo propagation of terrain uncertainty were not conducted. The result should therefore be interpreted as a targeted sensitivity check supporting the stability of first-order screening priorities, not as comprehensive uncertainty quantification.

Improvement in Drainage-Based Urban Flood Management Framework

In low-relief urban areas, drainage capacity and connectivity can strongly influence where flooding concentrates in response to rainfall forcing. In this case study, the dominance of IDPR (weight 0.424), street-drainage intersection density (0.277), and the spatial

Table 4
Results of the robustness analysis showing changes in the areas of high and very high flood hazard zones under systematic perturbations to the IDPR weight.

Scenario	Description	High & Very High Areas (km²)	Change from Original (%)
Original	Entropy-derived weights	1.425	–
Scenario +25%	IDPR weight increased 25%	1.448	2%
Scenario -25%	IDPR weight decreased 25%	1.439	1%
Scenario +50%	IDPR weight increased 50%	1.455	2%
Scenario -50%	IDPR weight decreased 50%	1.464	3%
Scenario +75%	IDPR weight increased 75%	1.479	3%
Scenario -75%	IDPR weight decreased 75%	1.472	3%

correspondence between modelled hazard and the 2021 flood impact in Taman Sri Muda indicate that drainage-related factors are major spatial controls on flood concentration. The results are consistent with the findings of Agonafir et al. (2023) and Bibi et al. (2023) in that inadequate drainage systems are the primary factor in flood hazards in rapidly urbanising areas.

Flood risk management, especially in the urban context, requires integrated solutions. Stand-alone interventions, such as channel enlargement, are inadequate in the long term but can be more effective when used in conjunction with restorative nature-based interventions (Zhou, 2014). In Malaysia, flood risk management is the result of many different agencies operating at different administrative levels, often leading to funding limitations that leave drainage systems incomplete, with the most critical system bottlenecks unresolved (Rosmadi et al., 2023).

High and very high hazard zones represent the first prioritisation level for drainage inspection, capacity assessment, identification of hydraulic bottlenecks, and obstruction management, whereas moderate zones warrant preventive maintenance and runoff management. An adaptive flood-management framework can combine spatial diagnosis, periodic monitoring, cross-sectoral collaboration, and stepwise implementation. Citizen observatories and digital reporting systems can improve local flood information when public feedback is integrated into formal planning processes (Wehn et al., 2015).

Although developed for precision-agriculture applications rather than flood modelling, context-aware IoT studies demonstrate a transferable principle: distributed local sensing can support data-driven environmental decision-making. IoT-assisted soil-fertility mapping links real-time field conditions to context-specific recommendations (Khan et al., 2022), while data-driven evapotranspiration studies show how limited environmental observations can support water-management decisions (Bashir et al., 2023; Khan et al., 2022). These studies are cited only as methodological analogues for future adaptive monitoring and are not used as evidence for the validity of the present flood-hazard model.

The framework illustrated in Figure 5 presents an adaptive, multi-level, drainage-based urban flood-management pathway that links local diagnosis and mitigation with longer-term, coordinated government action. At the local level, GIS-based flood vulnerability mapping and the drainage-connectivity indicators identified in this study provide the basis for diagnosing likely drainage failure mechanisms and defining management priorities, including drainage-oriented mitigation, flood-hotspot zoning, emergency-access planning, and rescue assembly points.

The framework then links ad hoc drainage mitigation and community self-mitigation measures to monitoring of flood recurrence and drainage performance. If local conditions remain unstable, the process escalates to cross-agency discussion, budgeted drainage upgrading, and coordinated implementation, followed by renewed monitoring. Figure 5

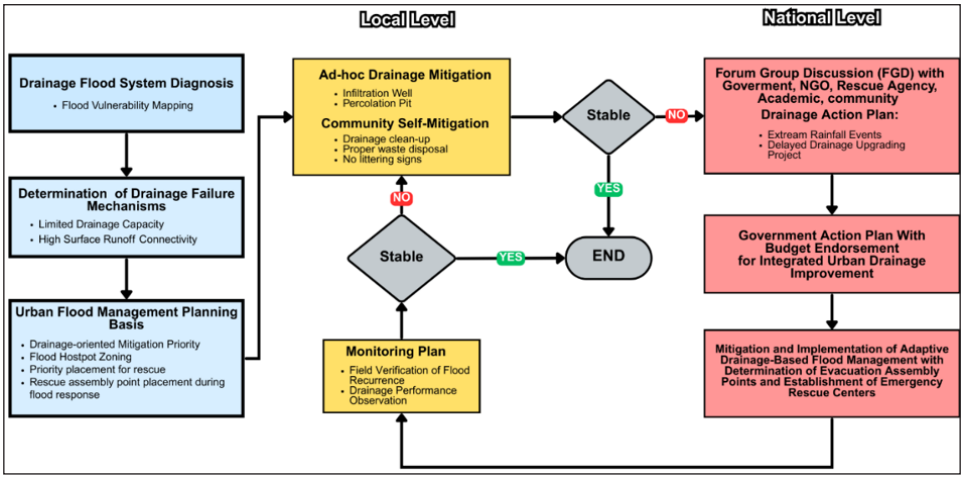


Figure 5. Framework for adaptive drainage-based urban flood management

is therefore intended as an adaptive management pathway informed by the study's spatial screening outputs, rather than as a calibrated hydraulic or engineering design model.

LIMITATIONS

Several limitations should be considered when interpreting the results. First, the flood hazard map is a relative spatial screening index and should not be interpreted as a predictive model of flood depth or extent. Second, the validation event occurred in December 2021, whereas the UAV dataset was acquired in 2023. Infrastructure and land-cover changes between these dates could not be independently verified; this temporal mismatch, therefore, introduces uncertainty that cannot be quantified from the available records. Third, the SCS-CN runoff factor was derived from a single approximately 280 mm rainfall event and was not calibrated across multiple storms. Fourth, rainfall was treated as spatially uniform, so local sub-catchment rainfall variability was not represented. Fifth, subsurface drainage pipes, pipe dimensions, invert levels, pumping, and sluice-gate operations were not explicitly modelled in the surface-routing analysis. Sixth, Kernel Density-based street-drainage intersection density and Euclidean distance to channel are simplified spatial proxies rather than hydraulic-capacity measures. Seventh, potential conceptual overlap exists among ISR, runoff volume, and street-drainage intersection density because all partially reflect urban imperviousness and drainage organisation. Eighth, the community component is based on five residents and is vulnerable to recall, locality-geocoding, and selection biases; the 26-point flood inventory also lacks verified non-flood controls. Ninth, the robustness analysis perturbed only the IDPR weight and therefore does not constitute a comprehensive multivariate uncertainty analysis. Finally, detailed UAV flight metadata, formal GCP documentation, and an independent vertical RMSE were unavailable to this

study. These limitations reinforce the interpretation of the FHI as a first-order, drainage-oriented prioritisation tool rather than a substitute for calibrated hydraulic simulation.

CONCLUSION

This study shows that within the low-relief, densely urbanised setting of Taman Sri Muda, spatial differentiation of relative flood hazard is dominated by drainage connectivity indicators rather than by topographic variation alone. IDPR and street-drainage intersection density received the two largest entropy-derived weights (0.424 and 0.277, respectively), together accounting for 70.1% of the total weight. The spatial hit rate of 73.08% (19/26; 95% binomial CI: 52.4%-88.4%) indicates agreement between the high/very high hazard classes and the available flood-positive locations, while the targeted IDPR sensitivity analysis indicates that the mapped high-hazard area changes only modestly under substantial perturbation of the dominant weight. These findings support the use of the framework as a relative drainage-oriented screening approach, not as a deterministic flood-depth model.

The principal contribution of the study is the operational integration of fine-resolution UAV-derived spatial data, entropy weighting, and explicit drainage-connectivity factors in a data-scarce low-relief urban setting. This combination reduces dependence on subjective expert weighting and provides a transparent basis for prioritising locations for further drainage inspection. The framework should nevertheless be applied with caution because the validation inventory is flood-positive only, the UAV and validation dates are temporally mismatched, subsurface hydraulic infrastructure is not represented, and the available sensitivity analysis addresses only the dominant IDPR weight.

This study highlights a critical research gap in coupling connectivity-based GIS frameworks with dynamic hydrological-hydraulic modelling to capture the temporal evolution of urban flooding. The absence of standardised and continuously updated urban flood inventories further constrains robust validation and model transferability. Addressing these gaps through integrated modelling and systematic data collection is essential for advancing evidence-based, drainage-oriented urban flood risk management.

ACKNOWLEDGEMENTS

This research was supported by the Ministry of Higher Education (MOHE) through the Fundamental Research Grant Scheme (FRGS), Grant No. FRGS/1/2023/TK02/UMT/02/1. The authors would like to thank the MOHE for the financial support. We also thank our colleagues and institutions for their contributions to this research.

REFERENCES

Abdessamed, D., & Abderrazak, B. (2019). Coupling HEC-RAS and HEC-HMS in rainfall–runoff modeling and evaluating floodplain inundation maps in arid environments: Case study of Ain Sefra city, Ksour

Mountain. SW of Algeria. *Environmental Earth Sciences, 78*(19), Article 586. https://doi.org/10.1007/s12665-019-8604-6

Abid, S. K., Sulaiman, N., Al-Wathinani, A. M., & Goniewicz, K. (2024). Community-based flood mitigation in Malaysia: Enhancing public participation and policy effectiveness for sustainable resilience. *Journal of Global Health, 14*, Article 04290. https://doi.org/10.7189/jogh.14.04290

Agonafir, C., Lakhankar, T., Khanbilvardi, R., Krakauer, N., Radell, D., & Devineni, N. (2023). A review of recent advances in urban flood research. *Water Security, 19*, Article 100141. https://doi.org/10.1016/j.wasec.2023.100141

Ahmad, S., Jia, H., Ashraf, A., Yin, D., Chen, Z., Ahmed, R., & Israr, M. (2024). A novel GIS-SWMM-ABM approach for flood risk assessment in data-scarce urban drainage systems. *Water, 16*(11), Article 1464. https://doi.org/10.3390/w16111464

Al-Hinai, H., & Abdalla, R. (2021). Mapping coastal flood susceptible areas using Shannon's entropy model: The case of Muscat governorate, Oman. *ISPRS International Journal of Geo-Information, 10*(4), Article 252. https://doi.org/10.3390/ijgi10040252

Bai, T., Mayer, A. L., Shuster, W. D., & Tian, G. (2018). The hydrologic role of urban green space in mitigating flooding (Luohe, China). *Sustainability, 10*(10), Article 3584. https://doi.org/10.3390/su10103584

Bashir, R. N., Khan, F. A., Khan, A. A., Tausif, M., Abbas, M. Z., Shahid, M. M. A., & Khan, N. (2023). Intelligent optimization of reference evapotranspiration (ETo) for precision irrigation. *Journal of Computational Science, 69*, Article 102025. https://doi.org/10.1016/j.jocs.2023.102025

Bibi, T. S., Reddythta, D., & Kebebew, A. S. (2023). Assessment of the drainage systems performance in response to future scenarios and flood mitigation measures using stormwater management model. *City and Environment Interactions, 19*, Article 100111. https://doi.org/10.1016/j.cacint.2023.100111

Bracken, L. J., Wainwright, J., Ali, G. A., Tetzlaff, D., Smith, M. W., Reaney, S. M., & Roy, A. G. (2013). Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews, 119*, 17-34. https://doi.org/10.1016/j.earscirev.2013.02.001

De Vleeschauwer, K., Weustenraad, J., Nolf, C., Wolfs, V., De Meulder, B., Shannon, K., & Willems, P. (2014). Green–blue water in the city: Quantification of impact of source control versus end-of-pipe solutions on sewer and river floods. *Water Science and Technology, 70*(11), 1825-1837. https://doi.org/10.2166/wst.2014.306

Dell'Aira, F., & Meier, C. I. (2025). Beyond total impervious area: A new lumped descriptor of basin-wide hydrologic connectivity for characterizing urban watersheds. *Hydrology and Earth System Sciences, 29*(4), 1001-1032. https://doi.org/10.5194/hess-29-1001-2025

DeSousa, S., Bhaskar, A. S., Kelleher, C., & Livneh, B. (2024). Understanding spatiotemporal patterns and drivers of urban flooding using municipal reports. *Hydrological Processes, 38*(12), Article e70028. https://doi.org/10.1002/hyp.70028

Dewan Negeri Selangor. (2022). *Penyata Jawatankuasa Pilihan Mengenai Pihak Berkuasa Tempatan (JP-PBT) bagi Dewan Negeri Selangor berhubung bencana banjir pada bulan Disember 2021 di Negeri Selangor* (Kertas Bil. 7 Tahun 2022). https://dewan.selangor.gov.my/statement/kertas-bil-7-tahun-2022-jp-pbt-bencana-banjir-pada-bulan-disember-2021-di-negeri-selangor/

Du, J., Cheng, L., Zhang, Q., Yang, Y., & Xu, W. (2019). Different flooding behaviors due to varied urbanization levels within river basin: A case study from the Xiang River basin, China. *International Journal of Disaster Risk Science, 10*(1), 89-102. https://doi.org/10.1007/s13753-018-0195-4

Du, S., Shi, P., Van Rompaey, A., & Wen, J. (2015). Quantifying the impact of impervious surface location on flood peak discharge in urban areas. *Natural Hazards, 76*(3), 1457-1471. https://doi.org/10.1007/s11069-014-1463-2

Dwivedi, D., Poeppl, R. E., & Wohl, E. (2025). Hydrological connectivity: A review and emerging strategies for integrating measurement, modeling, and management. *Frontiers in Water, 7*, Article 1496199. https://doi.org/10.3389/frwa.2025.1496199

Feng, B., Wang, J., Zhang, Y., Hall, B., & Zeng, C. (2020). Urban flood hazard mapping using a hydraulic–GIS combined model. *Natural Hazards, 100*(3), 1089-1104. https://doi.org/10.1007/s11069-019-03850-7

Fernández, D. S., & Lutz, M. A. (2010). Urban flood hazard zoning in Tucumán province, Argentina, using GIS and multicriteria decision analysis. *Engineering Geology, 111*(1-4), 90-98. https://doi.org/10.1016/j.enggeo.2009.12.006

Gebremichael, A., Gebremariam, E., & Desta, H. (2025). GIS-based mapping of flood hazard areas and soil erosion using analytic hierarchy process (AHP) and the universal soil loss equation (USLE) in the Awash River basin, Ethiopia. *Geoscience Letters, 12*(1), Article 12. https://doi.org/10.1186/s40562-025-00382-w

Goumghar, L., Fri, R., Hajaj, S., Taia, S., & El Mansouri, B. (2025). Integrating geospatial data and analytic hierarchy process for flood-prone zones mapping in the Upper Draa basin, Morocco. *Ecological Engineering & Environmental Technology, 26*(4), 251-268. https://doi.org/10.12912/27197050/201160

Guo, K., & Guan, M. (2022). Physical complexity to model rainfall-driven urban surface water flooding. In *Proceedings of the 39th IAHR World Congress* (pp. 4667-4671). International Association for Hydro-Environment Engineering and Research. https://doi.org/10.3850/IAHR-39WC252171192022246

Haghizadeh, A., Siahkamari, S., Haghiabi, A. H., & Rahmati, O. (2017). Forecasting flood-prone areas using Shannon's entropy model. *Journal of Earth System Science, 126*(3), Article 39. https://doi.org/10.1007/s12040-017-0819-x

Halim, B. M. K., Shamsuddin, N. S. A., Abdullah, R., & Kamarudin, K. H. (2024). Factors contributing to participation in community-led DRR programmes in Malaysia. *Disaster Advances, 17*(8), 41-48. https://doi.org/10.25303/178da041048

Hassan, B. T., Yassine, M., & Amin, D. (2022). Comparison of urbanization, climate change, and drainage design impacts on urban flashfloods in an arid region: Case study, New Cairo, Egypt. *Water, 14*(15), Article 2430. https://doi.org/10.3390/w14152430

Islam, T., Zeleke, E. B., Afroz, M., & Melesse, A. M. (2025). A systematic review of urban flood susceptibility mapping: Remote sensing, machine learning, and other modeling approaches. *Remote Sensing, 17*(3), Article 524. https://doi.org/10.3390/rs17030524

Khan, A. A., Driss, M., Boulila, W., Sampedro, G. A., Abbas, S., & Wechtaisong, C. (2024). Privacy preserved and decentralized smartphone recommendation system. *IEEE Transactions on Consumer Electronics, 70*(1), 4617-4624. https://doi.org/10.1109/TCE.2023.3323406

Khan, A. A., Faheem, M., Bashir, R. N., Wechtaisong, C., & Abbas, M. Z. (2022). Internet of things (IoT) assisted context aware fertilizer recommendation. *IEEE Access, 10*, 129505-129519. https://doi.org/10.1109/ACCESS.2022.3228160

Khan, A. A., Nauman, M. A., Bashir, R. N., Jahangir, R., Alroobaea, R., Binmahfoudh, A., Alsafyani, M., & Wechtaisong, C. (2022). Context aware evapotranspiration (ETs) for saline soils reclamation. *IEEE Access, 10*, 110050-110063. https://doi.org/10.1109/ACCESS.2022.3206009

Khan, Q., Kalbus, E., Zaki, N., & Mohamed, M. M. (2022). Utilization of social media in floods assessment using data mining techniques. *PLOS ONE, 17*(4), Article e0267079. https://doi.org/10.1371/journal.pone.0267079

Komori, D., Nakaguchi, K., Inomata, R., Oyatsu, Y., Tachikawa, R., & Kazama, S. (2022). Topographical characteristics of frequent urban pluvial flooding areas in Osaka and Nagoya cities, Japan. *Water, 14*(18), Article 2795. https://doi.org/10.3390/w14182795

Lane, S. N., Reaney, S. M., & Heathwaite, A. L. (2009). Representation of landscape hydrological connectivity using a topographically driven surface flow index. *Water Resources Research, 45*(8), Article W08423. https://doi.org/10.1029/2008WR007336

Leopold, L. B., Wolman, M. G., & Miller, J. P. (1964). *Fluvial processes in geomorphology*. W. H. Freeman.

Liu, N., & Zhang, F. (2025). Urban green spaces and flood disaster management: Toward sustainable urban design. *Frontiers in Public Health, 13*, Article 1583978. https://doi.org/10.3389/fpubh.2025.1583978

Liu, R., Chen, Y., Wu, J., Gao, L., Barrett, D., Xu, T., Li, X., Li, L., Huang, C., & Yu, J. (2017). Integrating entropy-based naïve Bayes and GIS for spatial evaluation of flood hazard. *Risk Analysis, 37*(4), 756-773. https://doi.org/10.1111/risa.12698

Malekinezhad, H., Sepehri, M., Pham, Q. B., Hosseini, S. Z., Meshram, S. G., Vojtek, M., & Vojteková, J. (2021). Application of entropy weighting method for urban flood hazard mapping. *Acta Geophysica, 69*(3), 841-854. https://doi.org/10.1007/s11600-021-00586-6

Manandhar, B., Cui, S., Wang, L., & Shrestha, S. (2023). Urban flood hazard assessment and management practices in South Asia: A review. *Land, 12*(3), Article 627. https://doi.org/10.3390/land12030627

Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes, 5*(1), 3-30. https://doi.org/10.1002/hyp.3360050103

Muhamad Azizi, M. A. A., & Teriman, S. (2023). The consequences and mitigation of floods affecting Taman Sri Muda, Shah Alam. In *e-Proceeding of 6th Undergraduate Seminar on Built Environment and Technology (USBET) 2023* (pp. 1931-1940). Universiti Teknologi MARA. https://ir.uitm.edu.my/id/eprint/95286/

Nizar, N., Hariyani, S., Rahmasari, P., Rafani, M., Hidayati, E., & Arianti, I. (2024). Integration of digital elevation model (DEM) for accurate slope and contour maps as fundamental data in spatial modeling. *International Journal of Research Publication and Reviews, 5*(3), 7011-7015. https://doi.org/10.55248/gengpi.5.0324.0890

Nooryati. (2022, October 18). *Bencana banjir kilat di Selangor, faktor perubahan iklim?* Institut Pengajian Sains Sosial, Universiti Putra Malaysia. https://ipsas.upm.edu.my/artikel/bencana_banjir_kilat_di_selangor_faktor_perubahan_iklim-69663

Ogden, F. L., Pradhan, N. R., Downer, C. W., & Zahner, J. A. (2011). Relative importance of impervious area, drainage density, width function, and subsurface storm drainage on flood runoff from an urbanized catchment. *Water Resources Research, 47*(12), Article W12503. https://doi.org/10.1029/2011WR010550

Pimenta, L., Duarte, L., Teodoro, A. C., Beltrão, N., Gomes, D., & Oliveira, R. (2025). GIS-based flood susceptibility mapping using AHP in the urban Amazon: A case study of Ananindeua, Brazil. *Land, 14*(8), Article 1543. https://doi.org/10.3390/land14081543

Rosmadi, H. S., Ahmed, M. F., Mokhtar, M. B., & Lim, C. K. (2023). Reviewing challenges of flood risk management in Malaysia. *Water, 15*(13), Article 2390. https://doi.org/10.3390/w15132390

Sepehri, M., Malekinezhad, H., Hosseini, S. Z., & Ildoromi, A. R. (2019). Assessment of flood hazard mapping in urban areas using entropy weighting method: A case study in Hamadan city, Iran. *Acta Geophysica, 67*(5), 1435-1449. https://doi.org/10.1007/s11600-019-00342-x

Sharma, A., Poonia, M., Rai, A., Biniwale, R. B., Tügel, F., Holzbecher, E., & Hinkelmann, R. (2024). Flood susceptibility mapping using GIS-based frequency ratio and Shannon's entropy index bivariate statistical models: A case study of Chandrapur district, India. *ISPRS International Journal of Geo-Information, 13*(8), Article 297. https://doi.org/10.3390/ijgi13080297

Singh, P., Sinha, V. S. P., Vijhani, A., & Pahuja, N. (2018). Vulnerability assessment of urban road network from urban flood. *International Journal of Disaster Risk Reduction, 28*, 237-250. https://doi.org/10.1016/j.ijdrr.2018.03.017

Singh, V. P. (1997). The use of entropy in hydrology and water resources. *Hydrological Processes, 11*(6), 587–626. https://doi.org/10.1002/(SICI)1099-1085(199705)11:6%3C587::AID-HYP479%3E3.0.CO;2-P

Tang, Z., Zhang, H., Yi, S., & Xiao, Y. (2018). Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis. *Journal of Hydrology, 558*, 144-158. https://doi.org/10.1016/j.jhydrol.2018.01.033

Ullah, F., Pun, C.-M., Mohmand, M. I., Mahendran, R. K., Khan, A. A., Alhammad, S. M., Rodrigues, J. J. P. C., & Farouk, A. (2025). Privacy-aware secure data auditing for cloud-based intelligence of things environment. *IEEE Internet of Things Journal, 12*(11), 15288-15303. https://doi.org/10.1109/JIOT.2025.3528117

Wehn, U., Rusca, M., Evers, J., & Lanfranchi, V. (2015). Participation in flood risk management and the potential of citizen observatories: A governance analysis. *Environmental Science & Policy, 48*, 225-236. https://doi.org/10.1016/j.envsci.2014.12.017

Yusuf, A., Ojo, R., Idrees, M. O., Balogun, A.-L., Salami, I. B., & Sani, O. S. (2023). Modelling flood hazards impacted by ungauged river in urbanised area using HEC-RAS and GIS. *Nigerian Journal of Technological Development, 20*(2), 83-92. https://doi.org/10.4314/njtd.v20i2.1405

Zhou, Q. (2014). A review of sustainable urban drainage systems considering the climate change and urbanization impacts. *Water, 6*(4), 976-992. https://doi.org/10.3390/w6040976

Zhou, R., Zheng, H., Liu, Y., Xie, G., & Wan, W. (2022). Flood impacts on urban road connectivity in southern China. *Scientific Reports, 12*(1), Article 16866. https://doi.org/10.1038/s41598-022-20882-5

Pertanika J. Sci. & Technol. 34 (5): 2715 - 2738 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Capacity-Limited Cooling Explains Simultaneous Energy Waste and Thermal Discomfort in Two University Buildings in the Southern Philippines

Fejaycris O. Pillodar[1], Dioniela Mae C. Ellos[1], Jiza Gay P. Jarabe[1], John Russel A. Dalogdog[1], Monaliza Joy Z. Magsayo[1,2], Charlene Mae Z. Gaab[1], Manuel Anthony P. Jondonero[1,3], Melgie A. Alas[4], Rovick Pelos P. Tarife[5], Noel R. Estoperez[5], Mark June L. Aporador[4], Napoleon A. Enteria[6], Kim Emissary C. Magarin[7], Ryan Lloyd R. Carpentero[7], Hernando P. Bacosa[1]*, and Peter D. Suson[1,7]

[1]Department of Environmental Science, School of Interdisciplinary Studies, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Tibanga, Iligan City, Lanao del Norte 9200, Philippines
[2]Environmental Science Department, College of Agriculture and Environmental Studies, Northwestern Mindanao State College of Science and Technology, Labuyo, Tangub City, Misamis Occidental 7214, Philippines
[3]Natural Sciences Department, College of Arts and Sciences, Bukidnon State University, Fortich Street, Malaybalay City, Bukidnon 8700, Philippines
[4]Office of Monitoring and Evaluation, Office of the Vice Chancellor for Strategic Initiatives, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Tibanga, Iligan City, Lanao del Norte 9200, Philippines
[5]Department of Electrical Engineering, College of Engineering, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Tibanga, Iligan City, Lanao del Norte 9200, Philippines
[6]Department of Mechanical and Engineering Science, College of Engineering, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Tibanga, Iligan City, Lanao del Norte 9200, Philippines
[7]MSU-IIT Centre for Resiliency, Office of the Vice Chancellor for Strategic Initiatives, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Tibanga, Iligan City, Lanao del Norte 9200, Philippines

ARTICLE INFO

Article history:
Received: 04 September 2025
Accepted: 03 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.08

E-mail addresses:
fejaycris.pillodar@g.msuiit.edu.ph (Fejaycris O. Pillodar)
dionielamae.ellos@g.msuiit.edu.ph (Dionela Mae C. Ellos)
jizagay.jarabe@g.msuiit.edu.ph (Jiza Gay P. Jarabe)
johnrussel.dalogdog@g.msuiit.edu.ph (John Russel A. Dalogdog)
monalizajoy.zaragoza@g.msuiit.edu.ph (Monaliza Joy Z. Magsayo)
charleenmae.gaab@g.msuiit.edu.ph (Charlene Mae Z. Gaab)
manuelanthony.jondonero@g.msuiit.edu.ph (Manuel Anthony P. Jondonero)
melgie.alas@g.msuiit.edu.ph (Melgie A. Alas)
rovick.tarife@g.msuiit.edu.ph (Rovick Pelos P. Tarife)
noel.estoperez@g.msuiit.edu.ph (Noel R. Estoperez)
markjune.aporador@g.msuiit.edu.ph (Mark June L. Aporador)
napoleon.enteria@g.msuiit.edu.ph (Napoleon A. Enteria)
kimemissary.magarin@g.msuiit.edu.ph (Kim Emissary C. Magarin)
ryanlloyd.carpentero@g.msuiit.edu.ph (Ryan Lloyd R. Carpentero)
hernando.bacosa@g.msuiit.edu.ph (Hernando P. Bacosa)
peter.suson@g.msuiit.edu.ph (Peter D. Suson)
* Corresponding author

ABSTRACT

Institutional buildings in the humid tropics are commonly assumed to waste energy because occupants set thermostats to low setpoints. This study tested that assumption in the two largest academic buildings at the Mindanao State University-Iligan Institute of Technology (MSU-IIT), combining 2,491 equipment records, 642 temperature and humidity observations, illuminance in 212 rooms, and a survey of 175 occupants. Room-level records were analysed using Welch t-tests, Mann-Whitney U tests, Hedges' g, chi-square tests, and ordinary least squares regression at a significance level of

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

0.05. Combined annual use was 1,151 MWh, costing PHP 13.81 million and emitting 639.5 t CO_2, with air conditioning taking 47.2% and 56.3% of demand. Set-points averaged 20.11 °C and fell below the 24 °C national guideline in 70.9% of observations, yet were attained in only 2.0%, and delivered temperature was almost insensitive to set-point (slope 0.097, 95% CI 0.053 to 0.140). Air conditioning held rooms only 1.75 °C below free-running rooms, indicating capacity-limited rather than control-limited operation, and self-reported set-points corroborated the measured values to within 0.5 °C. Separately, 87.3% of rooms fell below 300 lx, despite an average lighting power density of only 2.9 W m^{-2}, a provision deficit. These findings invert the conventional retrofit hierarchy, because thermostat policy cannot save energy while units remain capacity-limited, so cooling capacity and the building envelope must be restored first. Sequenced this way, five measures deliver 21.6% of baseline consumption at a 3.3-year payback, twice the reduction required of state universities. Testing set-point attainment before prescribing behavioural measures is a transferable diagnostic for tropical institutional estates.

Keywords: Building energy audit, canopy cover, energy conservation measures, greenhouse gas emissions, illuminance, indoor environmental quality, lighting power density, set-point attainment

INTRODUCTION

Energy underpins socio-economic development, yet its production and consumption carry a large environmental footprint through greenhouse gas (GHG) emissions (Masoso & Grobler, 2010). In 2019, the Philippines emitted 237 million t of CO_2 equivalent, a 3.2% increase over 2018, ranking 38th globally and accounting for 0.49% of global emissions (United Nations Development Programme, 2026). Power generation is the largest domestic source of GHG emissions, accounting for 56.2% of national emissions (Department of Energy Philippines, 2024), and the Philippine Energy Plan 2023-2050 identifies the commercial and institutional sectors as among the largest consumers.

Depletion of the Malampaya gas field increases reliance on imported liquefied natural gas (Department of Energy Philippines, 2022), exposing the country to global price volatility (International Energy Agency, 2020) and to geopolitical disruption of alternative domestic supply (Abinales & Amoroso, 2017), so demand-side efficiency is not merely an environmental preference but a measure of energy resilience. Republic Act 11285, the Energy Efficiency and Conservation Act of 2019, places state universities and colleges under the Government Energy Management Program (Republic of the Philippines, 2019), which targets a reduction of at least 10% in electricity and fuel consumption in government facilities and requires each agency to maintain an energy management program, define conservation measures, target savings, and designate an energy conservation officer.

University energy audits in the tropics constitute a substantial body of literature, but most of it stops at accounting. They report end-use disaggregation, benchmark energy use

intensity (EUI), and conclude with generic recommendations that are rarely tested against the building's measured performance (Chan et al., 2020; Ferrer-Balas et al., 2010; Leal Filho et al., 2019; Lopez et al., 2017). Three consequences follow. Energy audits and indoor environmental quality (IEQ) surveys are usually conducted separately, so it cannot be asked whether the energy consumed is for purchasing comfort. Where high cooling energy is observed, it is attributed either to inefficient equipment or to low thermostat setpoints, without a measurement that can distinguish between them. Recommended measures are seldom quantified, so institutions receive priorities that are not ranked by savings, cost, or abatement.

The distinction between control-limited and capacity-limited operation has not been tested in a Philippine institutional setting. In a control-limited system, the plant has sufficient capacity, cycles under thermostat control, and the set-point governs both delivered temperature and energy consumed, so the set-point policy is the correct intervention. In a capacity-limited system, the plant runs continuously without reaching its set-point because capacity, refrigerant charge, or envelope and infiltration losses constrain it; thus, the set-point governs neither comfort nor consumption, and behavioural policy will fail. The regimes are separable because the elasticity of delivered temperature with respect to set-point approaches unity in the first case and zero in the second, a test that requires recording set-point and room temperature together, which routine audits rarely do.

This study integrates an equipment-level energy audit with concurrent thermal, hygric and luminous measurements in the same rooms of two large academic buildings and uses the combined dataset to diagnose mechanisms before prescribing measures. The objectives were to quantify energy use, cost and GHG emissions by end use per unit floor area, to decompose the difference in cooling energy into installed capacity, technology mix and operating schedule, to test whether the cooling systems are control-limited or capacity-limited by estimating set-point attainment and the set-point elasticity of delivered temperature, to assess thermal, hygric and luminous compliance against ASHRAE and national criteria using psychrometric analysis that separates real moisture removal from temperature-driven changes in relative humidity (RH), to determine whether under-lighting reflects inefficient luminaires or insufficient installed lighting power, to examine how site geometry and afternoon solar exposure relate to the diurnal structure of the indoor difference, and to appraise a prioritized package of conservation measures by annual saving, avoided emissions, capital cost and simple payback. The study introduces a diagnostic test that distinguishes capacity from control limitation using data that an ordinary audit can collect, shows that the intervention hierarchy conventionally recommended for tropical institutional buildings is inverted in these buildings, and delivers a quantified retrofit package for a state university subject to a statutory savings target.

MATERIALS AND METHODS

Study Buildings and Audit Period

The audit covered the two largest academic buildings of Mindanao State University-Iligan Institute of Technology (MSU-IIT) in Iligan City, Philippines, at 8.24° N, 124.24° E, in a tropical rainforest climate under the Köppen classification (Af). The College of Engineering (COE) has a gross floor area of 4,987.27 m² across five stories and presents its principal glazed façade to the west, toward Iligan Bay, with no intervening obstruction. The College of Science and Mathematics (CSM) has 5,507.31 m² across the main and annex blocks, three storeys high, and lies immediately east of COE, bounded on its western side by tall, mature trees. Both operate on a five-day academic week. Data were collected from September to November 2024 during the first semester under normal occupancy, as shown in Figure 1.

The buildings were selected because institutional electricity is metered at the feeder level rather than at the building level, and prior estimates ranked these two colleges highest in connected load and gross floor area. Establishing building-level baselines was therefore itself an objective, since no energy audit had previously been conducted at MSU-IIT.

Equipment Inventory and Energy Estimation

Every electrical device in every accessible room was enumerated, recording the device type, brand, quantity, nameplate power, functional status, and daily operating hours, the

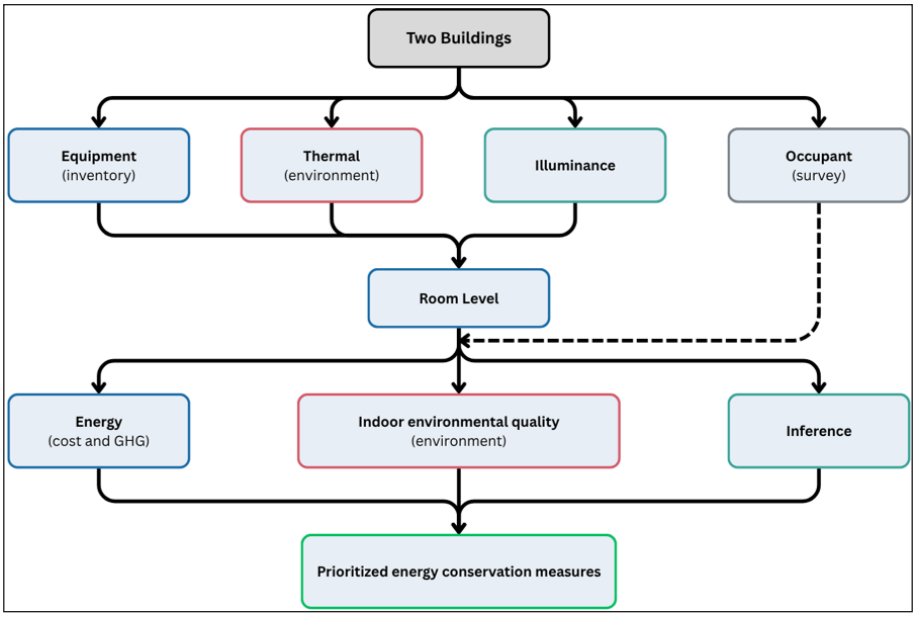


Figure 1. Design of the study, showing the five data streams, their sampling design and sample sizes, the room-level harmonisation step and the analysis branches leading to the prioritised conservation measures

latter of which was established through interviews with the room custodian and direct observation. The inventory yielded 2,491 records, 1,309 at COE and 1,182 at CSM, in five end-use categories (Table 1).

The daily energy consumption for each record was estimated using Equation 1, where P is the nameplate power in kW, N is the quantity, and h is the daily operating hours. Annual consumption was calculated using Equation 2, where D is the effective operating days, taken as 168 to match the academic calendar, excluding weekends, semestral breaks, and holidays, or 365 for continuously operating equipment. Energy cost was computed using Equation 3 at the prevailing Iligan Light and Power, Inc. tariff of PHP 12 per kWh, and emissions using Equation 4 with the Mindanao grid-weighted factor of 0.5555 kg CO_2 per kWh.

$$E_{\text{daily}} = P \times N \times h \quad [1]$$

$$E_{\text{annual}} = E_{\text{daily}} \times D \quad [2]$$

$$C = E \times 12 \quad [3]$$

$$G = E \times 0.5555 \quad [4]$$

All quantities were normalised by gross floor area to give energy use intensity in kWh m^{-2} y^{-1}, cost intensity in PHP m^{-2} y^{-1}, and emission intensity in kg CO_2 m^{-2} y^{-1}. This approach estimates consumption from rated power and reported schedules rather than measurement. Nameplate power is a rated maximum, so variable-load devices draw less on average, and the method tends to overestimate; reported hours are also subject to recall bias. Part-load efficiency, compressor cycling, and simultaneity are not represented, and standby power is not captured, which biases in the opposite direction. Non-functional equipment, 178 records at COE and 92 at CSM, was excluded from consumption but retained in the asset inventory. The estimates are therefore a consistent relative baseline rather than a substitute for metered consumption, and because both buildings were audited by the same team using the same protocol in the same period, systematic bias should be largely common to both.

Thermal Environment Measurement

Air temperature and RH were measured using a handheld thermo-hygrometer with an accuracy of ±1 °C and ±3% RH, respectively. Five stations were used in each room: the four corners and the centre, at approximately 1.1 m above floor level, corresponding to the seated occupant head height specified in ASHRAE Standard 55 (ASHRAE, 2020). The

instrument was allowed to stabilise at each station, and the five readings were averaged to give the room value.

Each room was surveyed at approximately 08:00, 12:00 and 17:00, spanning the occupied period. Outdoor temperature and RH were recorded concurrently immediately outside the envelope, and, for air-conditioned rooms, the thermostat setpoint displayed on the unit was logged. Rooms were measured under normal occupancy and normal air-conditioning operation as found. The campaign yielded 642 room-period observations, 276 at COE and 366 at CSM, across 206 physical rooms, of which 443 were air-conditioned and 199 free-running.

Two derived quantities were used. The set-point gap is the difference between the measured room temperature and the thermostat set-point, so a system meeting its set-point yields a gap of zero or less. The humidity ratio W, in g kg^{-1} of dry air, was computed from the dry-bulb temperature and RH at standard atmospheric pressure using the Magnus formulation and is reported because RH alone conflates moisture content with temperature. Compliance thresholds follow ASHRAE Standard 55 (ASHRAE, 2020) and the Guidelines on Energy Conserving Design of Buildings (Department of Energy Philippines, 2021), namely a comfort band of 23 to 27 °C, an RH ceiling of 60% and a recommended set-point of 24 °C.

Illuminance and Lighting Power Measurement

Task-plane illuminance was measured with a digital lux meter with an accuracy of ±2%. Readings were taken on a grid ranging from 8 to 40 points per room, depending on room dimensions, at a desk height of approximately 0.75 m, following the grid-averaging procedure in the IESNA lighting handbook (Rea, 2000). Measurements were taken during daytime occupied hours with the installed luminaires switched on and blinds or curtains in their normal in-use position, so the recorded values represent the combined artificial and daylight conditions experienced by occupants. Room dimensions were measured to derive floor area, and total connected luminaire wattage was recorded, giving the installed lighting power density (LPD) in W m^{-2} and installed system efficacy in lx m^2 W^{-1}. Rooms were classified against the 300-lx minimum for general classroom and office tasks as under-lit below 300 lx, acceptable from 300 to 500 lx and over-lit above 500 lx. Complete records were obtained for 212 rooms, 93 at COE and 119 at CSM.

Site Geometry and Canopy Cover Estimation

Site geometry was characterised from a georeferenced satellite image of the campus, which recorded the orientation of each building and the disposition of surrounding obstructions, and an indicative estimate of canopy cover in plan was derived from the same image. Vegetation pixels were classified using the excess green index, defined as

twice the normalised green channel minus the normalised red and blue channels, with a threshold of 0.02 and the additional requirement that the green channel exceed both the blue channel and 95% of the red channel. Canopy fraction was computed within circular buffers of 30, 50, 75, and 100 m around each building centroid. Because the centroids are 59 m apart, only the 30 m buffers are effectively non-overlapping, and this radius is the primary comparison. Sensitivity was tested by displacing each centroid by up to 20 pixels, about 7 m. The estimate is a plan-view screening measure that does not resolve canopy height or the azimuth of obstruction.

Occupant Survey

A structured questionnaire was administered during the audit period to provide an independent check on the operating conditions recorded at the units, covering thermal adjustment method, the conditions under which air conditioning is operated, the temperature at which units are set, daily operating duration, and agreement with candidate conservation policies scored on five-point scales. It returned 175 valid responses, 114 from COE and 61 from CSM, all from undergraduate students, since the staff and faculty branch returned none. Survey findings are therefore treated as the perspective of student occupants and used only for corroboration, not as primary evidence. Responses were not room-linked or entered into the room-level models, and self-reported set points outside the plausible range of 15 to 30 °C were excluded, removing 5 responses. Participation was voluntary and anonymous, and respondents were informed of the purpose of the study beforehand.

Data Harmonization

The three technical datasets were harmonised on a common room key derived from room numbers or labels, with combined spaces retained as merged units. They covered 207 rooms in the audit, 212 in the illuminance survey, and 206 in the thermal survey, and 138 rooms carried complete records across all three, forming the matched dataset used for correlation and regression. Building-level aggregates use the full datasets, and raw source files were not modified.

Statistical Analysis

Analyses were performed in Python 3.12 using NumPy 2.4, SciPy 1.17, and pandas 3.0, with a significance level of 0.05 throughout, and figures were produced with Matplotlib 3.10. Continuous variables are summarised as mean ± standard deviation with 95% confidence intervals (CI), median, and interquartile range. Normality was assessed with the Shapiro-Wilk test and homogeneity of variance with the median-centred Levene test. Because most distributions departed from normality at $p < 0.05$, even with large sample sizes, two-group

comparisons are reported using the Welch unequal-variance t-test as the primary test and the Mann-Whitney U test as a distribution-free confirmation; the two agreed in direction and significance for every comparison. Effect sizes are reported as Hedges' g with 95% CI and mean differences with 95% CI, so that both practical and statistical significance can be assessed. Comparisons across the three times of day used one-way analysis of variance with the Kruskal-Wallis test as a check, and categorical compliance outcomes were compared with the Pearson chi-square test.

Associations among room-level variables were assessed using Pearson correlation. Two ordinary least squares (OLS) models were estimated on the matched dataset of air-conditioned rooms, for which $n = 92$. The first regressed air-conditioning energy intensity on installed input power density, set-point, daily operating hours, window-type capacity share, and building; the second regressed the set-point gap on the same predictors, excluding operating hours. The set-point elasticity of delivered temperature was estimated by simple linear regression of measured room temperature on set-point, and internal consistency of the survey constructs was assessed with Cronbach's alpha.

Appraisal of Energy Conservation Measures

Five energy conservation measures (ECMs) were appraised, covering the replacement of window-type units with inverter split systems, set-point control at 24 °C, LED relighting to 300 lx, plug-load and standby control, and envelope sealing with solar-control glazing. Annual savings were estimated by applying measure-specific savings fractions to the relevant audited end-use baseline, and the fractions, their justification and the unit bases of the capital costs are stated in the note to Table 5. Set-point governance was credited only after the replacement measure because set-point changes cannot save energy while units remain capacity-limited. Simple payback is the capital cost divided by the annual cost savings, assuming PHP 12 per kWh, and excludes discounting, escalation, and maintenance. The package is benchmarked against the target under Republic Act 11285, with at least a 10% reduction in consumption.

RESULT AND DISCUSSIONS

Inventory of Electrical Equipment

The inventory assigned 2,491 records to five categories (Table 1). Cooling and ventilation equipment dominated in both buildings. COE carried more operations equipment, consistent with its computer laboratories, while CSM carried more electrical appliances. Laboratory equipment was numerous in both, but it contributed modestly to energy use because it operates intermittently. A total of 178 records at COE and 92 at CSM were non-functional yet remained installed, an asset-management finding.

Table 1
Equipment categories, representative equipment and inventory counts

Category	Examples	COE records	CSM records	COE units	CSM units
HVAC	Air conditioner, ceiling fan, wall fan, stand fan, exhaust fan, humidifier	133	178	156	207
Electrical appliances	Refrigerator, rice cooker, microwave, oven	70	84	70	93
Laboratory equipment	Microscope, oscilloscope, digital analytical balance, spectrophotometer	132	155	295	252
Lighting	Fluorescent, bulb, pin light	199	178	1,490	1,405
Operations	Monitor, keyboard, mouse, CPU, speaker, projectors	775	587	1,241	940
Total		1,309	1,182		

Energy Consumption, Cost and Greenhouse Gas Emissions

Combined annual electricity consumption was 1,151 MWh, costing PHP 13.81 million and emitting 639.5 t of CO_2 (Table 2, Figure 2). Building EUI values were closely matched at 107.2 kWh m^{-2} y^{-1} for COE and 112.0 kWh m^{-2} y^{-1} for CSM, placing both at the lower end of the range reported for tropical university buildings, which is unsurprising given single-split cooling, the absence of central plant and daytime-only occupancy.

The end-use composition differed materially. Cooling and ventilation accounted for 47.2% of COE and 56.3% of CSM electricity, and were dominant in both, consistent with the literature on cooling-dominated institutional buildings (Pérez-Lombard et al., 2008; Ryu & Yoo, 2025; Santamouris, 2016). The second-largest end use diverged sharply, with operations equipment taking 29.3% at COE against 16.2% at CSM, reflecting COE's computer laboratories, whereas CSM's appliance load at 11.8% exceeded COE's at 7.8%.

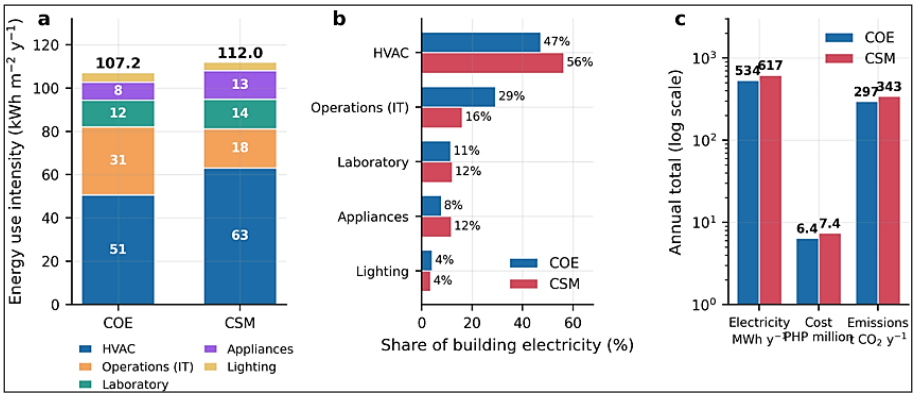


Figure 2. Energy, cost and emissions. (a) End-use energy intensity by building. (b) End-use composition as a share of building electricity. (c) Building-level annual totals for electricity, cost and emissions on a logarithmic scale

Table 2
Annual energy, cost and emission intensity by end use

Type of equipment	COE kWh m^{-2}	CSM kWh m^{-2}	COE PHP m^{-2}	CSM PHP m^{-2}	COE kg CO_2 m^{-2}	CSM kg CO_2 m^{-2}
HVAC	50.56	63.09	606.7	757.1	28.08	35.05
Electrical appliances	8.36	13.25	100.3	159.0	4.64	7.36
Laboratory equipment	12.30	13.57	147.6	162.9	6.83	7.54
Lighting	8.36	13.25	100.3	159.0	4.64	7.36
Operations	31.45	18.10	377.4	217.2	17.47	10.06
Total	107.16	111.98	1,285.8	1,343.8	59.51	62.21

Lighting was the smallest end use in both at 4.2% and 3.5% because the illuminance deficit cannot be remedied by cutting lighting energy and there is ample headroom to increase it.

Determinants of the Difference in Cooling Energy

CSM used 12.5 kWh m^{-2} y^{-1} more cooling energy than COE. Such differences are often attributed to equipment age, which is not a measurable quantity in an audit of this kind, so the difference was instead decomposed into its physical components (Figure 3), as in multi-dimensional analyses of air-conditioning energy use in university teaching buildings (Li et al., 2020).

The two buildings differ fundamentally in technology. Of the 147 units at CSM, 120 (81.6%) are window-type or portable, whereas of the 104 units at COE, 93 (89.4%) are split systems. This matters thermodynamically because window units house the compressor, condenser, and evaporator in a single casing, so condenser heat rejection is compromised by the small coil face area and by short-circuiting of the discharge air. Their seasonal efficiency is typically well below that of split systems, which is why labelling schemes

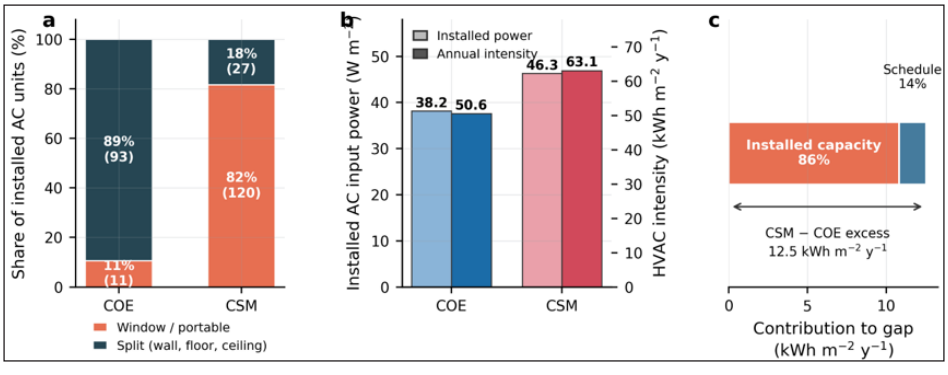


Figure 3. Origin of the difference in cooling energy. (a) Composition of the installed air-conditioning fleet by technology, with unit counts in parentheses. (b) Installed input power density on the left axis against annual cooling energy intensity on the right axis. (c) Logarithmic decomposition of the difference between CSM and COE into installed capacity and operating schedule

such as the Department of Energy's Cooling Seasonal Performance Factor rating distinguish between the classes.

The consequence appears in installed capacity. Air-conditioning input power density was 46.3 W m^{-2} at CSM against 38.2 W m^{-2} at COE, so CSM installs 21% more electrical input to condition a comparable area, while capacity-weighted operating hours were nearly identical at 8.05 and 7.80 h per day. A logarithmic decomposition attributes 86% of the gap to installed capacity and 14% to schedule, and the model in Table 4 confirms the mechanism, since cooling energy intensity is explained almost entirely by installed input power density and operating hours, with no independent building effect once these are controlled ($p = 0.743$). CSM therefore has more and less efficient installed cooling power for the same task, a more actionable diagnosis than equipment age, because it identifies the 120 window units as the intervention target.

Set-Point Attainment and the Capacity Limitation of Cooling

Set points were low, averaging 20.11 ± 3.72 °C, and 70.9% of observations fell below the 24 °C national guideline, the pattern conventionally read as wasteful occupant behaviour (Figure 4). Measured room temperatures in the same rooms averaged 26.79 ± 1.76 °C. The set-point was attained in only 9 of 443 observations, or 2.0%, and the mean gap was 6.68 °C with a 95% CI of 6.33 to 7.03, being 6.02 ± 3.36 °C at COE and 7.17 ± 3.99 °C at CSM (Welch $t = -3.29$, $p = 0.001$, $g = -0.31$).

Regressing delivered temperature on set-point yields a slope of 0.097 with a 95% CI of 0.053 to 0.140 and an R^2 of 0.042, whereas a control-limited system would approach 1.0. A slope near zero means that lowering the thermostat by 10 °C buys roughly 1 °C

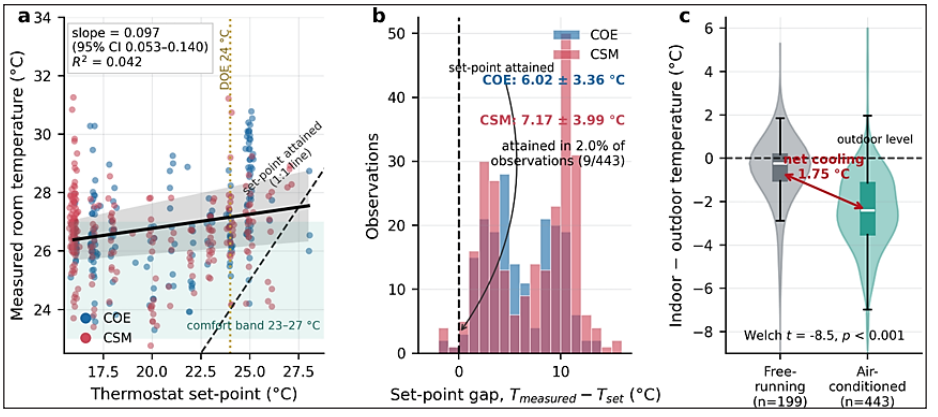


Figure 4. Set-point attainment. (a) Measured room temperature against thermostat set-point for all air-conditioned observations, with the 1:1 attainment line shown dashed, the fitted regression shown solid with its 95% confidence band, the comfort band of 23 to 27 °C and the national guideline of 24 °C. (b) Distribution of the set-point gap by building. (c) Indoor-outdoor temperature difference for free-running and air-conditioned rooms, where the difference between group means represents, the net cooling delivered

of delivered cooling, so the rooms are not over-cooled but barely cooled at all. Installed capacity density did not predict the gap ($r = 0.027$, $p = 0.802$), and in the second model of Table 4, the set-point coefficient on the gap was −1.007 with a 95% CI of −1.075 to −0.939, indistinguishable from −1.0, the signature of a delivered temperature invariant to the set-point. An independent measure confirms the magnitude: air-conditioned rooms sat 2.40 °C below outdoor air, while free-running rooms sat 0.65 °C below, so the net cooling attributable to air conditioning is only 1.75 °C (Welch $t = -8.52$, $p < 0.001$). For roughly half of the electricity these buildings consume, the thermal return is very small.

The occupant survey corroborates this account independently. Respondents reported setting or preferring a mean air-conditioning temperature of 19.65 ± 3.02 °C against the 20.11 ± 3.72 °C recorded at the units, and reported operating durations consistent with the audited schedule, so two measurements sharing no common source of error agree to within 0.5 °C. The low set-points are real and settled, rather than artefacts of logging, while occupant practice is otherwise unremarkable. Several mechanisms are consistent with the shortfall and are not mutually exclusive. Envelope and infiltration losses are likely dominant, because rooms are entered and left continuously between classes, many have jalousie or awning windows and undercut doors, and no room has a vestibule. Degraded capacity is also implicated, since window units predominate at CSM, the refrigerant charge is not on a maintenance schedule, and the condenser coils show visible fouling, while under-sizing relative to internal gains is plausible in the computer laboratories at COE. The data nonetheless exclude one candidate because the set point demonstrably does not govern the outcome. The behavioural corollary is a feedback loop: a room that fails to cool invites the occupant to lower the thermostat further, which cannot improve comfort in a saturated system but does keep the compressor at full load, a pattern consistent with evidence that occupant interaction with air-conditioner controls is shaped by the response of the equipment itself (Winkler et al., 2018; Zhang et al., 2018).

Thermal Environment and Comfort Compliance

Given the preceding result, the two buildings should differ little in delivered thermal conditions, and they do not (Figure 5, Table 3). COE averaged 27.79 ± 2.08 °C and CSM 27.36 ± 1.98 °C, a difference of 0.43 °C with a 95% CI of 0.11 to 0.75 (Welch $t = 2.64$, $p = 0.009$; Mann-Whitney $p = 0.024$). The effect size is small ($g = 0.21$), so the magnitude of the difference alone is not evidence of superior thermal performance at CSM; its structure across the day is the more informative feature.

Compliance is the more meaningful comparison and is poor in both buildings, since 59.1% of COE and 50.3% of CSM observations exceeded 27 °C, with only 40.9% and 48.9% inside the 23 to 27 °C band, a range in which classroom temperature has been linked to measurable differences in student thermal comfort and cognitive performance (Jiang et

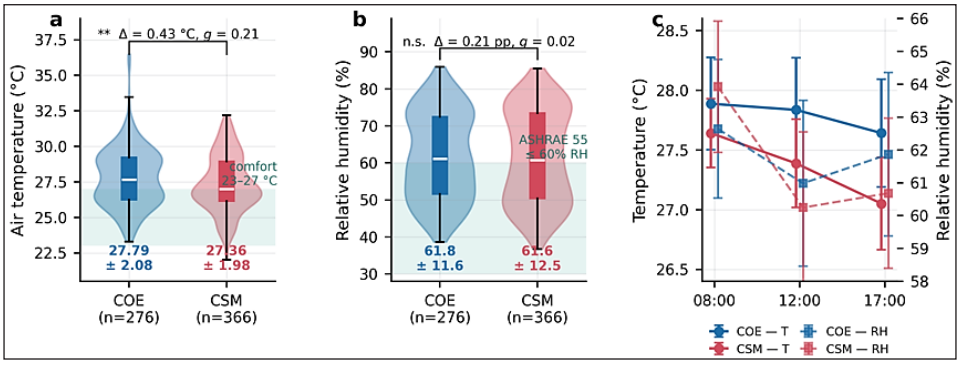


Figure 5. Indoor thermal and hygric environment. (a) Air temperature and (b) relative humidity by building, shown as violin plots with embedded box plots, where shaded bands mark the comfort criteria. (c) Diurnal pattern of temperature shown solid and relative humidity shown dashed, as mean with 95% confidence interval

Table 3
Descriptive statistics and between-group comparisons for indoor environmental variables

Variables	Groups	Group 1 (n, M ± SD)	Group 2 (n, M ± SD)	Difference (95% CI)	t (df)	p (W/ MW)	g (95% CI)
Temperature (°C)	COE, CSM	276, 27.79 ± 2.08	366, 27.36 ± 1.98	0.43 (0.11 to 0.75)	2.64 (578)	0.009 / 0.024	0.21 (0.06 to 0.37)
Relative humidity (%)	COE, CSM	276, 61.83 ± 11.63	366, 61.62 ± 12.47	0.21 (−1.66 to 2.09)	0.22 (612)	0.823 / 0.845	0.02 (−0.14 to 0.17)
Temperature (°C)	AC, free-run	443, 26.79 ± 1.76	199, 29.23 ± 1.52	−2.45 (−2.71 to −2.18)	−17.95 (439)	<0.001 / <0.001	−1.45 (−1.63 to −1.26)
Relative humidity (%)	AC, free-run	443, 56.42 ± 9.35	199, 73.49 ± 8.78	−17.07 (−18.58 to −15.57)	−22.33 (404)	<0.001 / <0.001	−1.86 (−2.05 to −1.66)
Humidity ratio (g kg^{-1})	AC, free-run	443, 12.58 ± 3.03	199, 19.00 ± 3.10	−6.42 (−6.94 to −5.90)	−24.45 (374)	<0.001 / <0.001	−2.10 (−2.31 to −1.90)
Set-point gap (°C)	COE, CSM	188, 6.02 ± 3.36	255, 7.17 ± 3.99	−1.15 (−1.84 to −0.46)	−3.29 (433)	0.001 / 0.001	−0.31 (−0.50 to −0.12)
Illuminance (lx)	COE, CSM	93, 166.5 ± 105.0	119, 158.0 ± 115.1	8.5 (−21.4 to 38.4)	0.56 (205)	0.576 / 0.355	0.08 (−0.19 to 0.35)
RH, laboratories (%)	COE, CSM labs	123, 59.16 ± 11.41	168, 62.62 ± 12.66	−3.47 (−6.26 to −0.68)	−2.45 (277)	0.015 / 0.015	−0.29 (−0.52 to −0.05)
Outdoor temperature (°C)	COE, CSM site	276, 29.36 ± 2.30	366, 29.44 ± 1.25	−0.08 (−0.38 to 0.22)	−0.52 (396)	0.607 / 0.274	−0.04 (−0.20 to 0.11)

al., 2025; Li et al., 2025). Upper storeys were consistently warmer, the third floor of COE averaging 29.18 ± 2.64 °C against 27.28 ± 1.42 °C on the first, and the third floor of the CSM annex 28.35 ± 1.95 °C against 26.73 ± 2.03 °C on the second floor of the main block, a documented consequence of roof solar gain and stack effect in the tropics (Rajapaksha,

2019). Both buildings peaked in the morning and declined slightly throughout the day ($F = 2.59$, $p = 0.076$; Kruskal-Wallis $p = 0.013$), a diurnal pattern in which morning temperature and humidity covary (Zhmakin, 2023), consistent with thermal mass charging overnight and with cooling systems that, once started, run continuously without recovering to set point.

Site Geometry, Shading and the Diurnal Divergence

The two buildings sit 59 m apart in the same dense low-rise urban fabric, and their outdoor air temperatures were indistinguishable at every time of day, differing by 0.09 °C in the morning, −0.15 °C at noon, and −0.18 °C in the afternoon, nonsignificant. Ambient conditions, therefore, cannot account for the indoor difference, and attention turns to solar exposure.

The exposures differ in geometry rather than in the quantity of surrounding vegetation. COE presents its principal glazed façade to the west, where the ground falls away toward Iligan Bay, so the façade receives direct low-angle radiation through the afternoon without obstruction, whereas CSM lies east of COE and is screened in the afternoon both by tall, mature trees along its western boundary and by the mass of COE itself. Plan-view canopy cover does not capture this asymmetry, and estimated cover within the non-overlapping 30 m buffer was, in fact, higher around COE, at 11.8% versus 8.6%, and remained higher at 50, 75, and 100 m (Figure 6). A planar fraction is insensitive to canopy height and to the azimuth of obstruction, and those are precisely the properties that determine afternoon shading, so canopy fraction describes the site rather than testing the hypothesis.

The diurnal structure of the indoor data is consistent with the geometry. Expressed relative to outdoor air, so that ambient variation is removed, COE sat 1.14 °C below outdoor

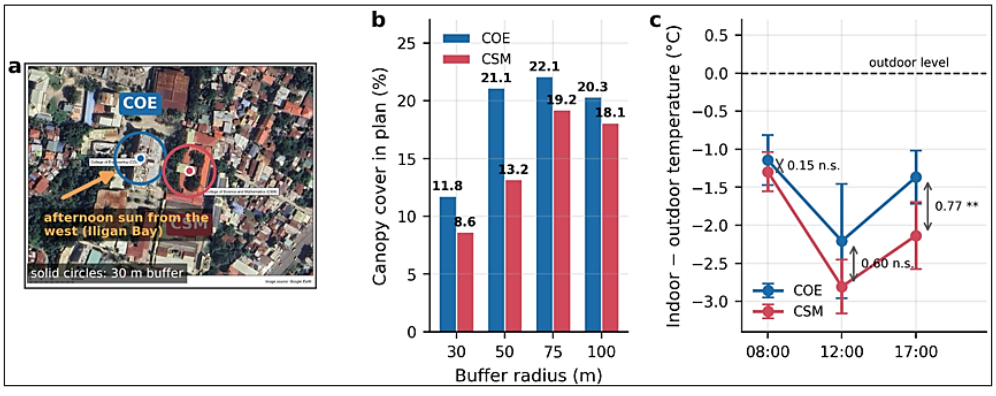


Figure 6. Site geometry, canopy cover and the diurnal divergence. (a) Satellite view of the two buildings with 30 m buffers showing COE's unobstructed western exposure toward Iligan Bay. (b) Estimated canopy cover in plan within buffers of 30 to 100 m, a measure insensitive to canopy height and to the azimuth of obstruction. (c) Indoor minus outdoor temperature by time of day, as mean with 95% confidence interval, with the difference between buildings annotated

in the morning, against 1.30 °C at CSM, a difference of 0.15 °C, widening to 0.60 °C at noon and 0.77 °C in the afternoon ($p = 0.008$). The same monotonic widening appears in air-conditioned rooms, from 0.28 °C in the morning to 0.62 °C at noon ($p = 0.043$) and 0.80 °C in the afternoon ($p = 0.012$). Between morning and afternoon, indoor temperature at CSM fell by 0.59 °C while its outdoor temperature rose by 0.25 °C, whereas at COE, indoor temperature fell by only 0.25 °C against a flat outdoor trend. CSM decouples from ambient conditions as the afternoon progresses, and COE does not; the signature expected of a building shielded from afternoon solar gain, as opposed to one that is not.

This is an inference from indoor thermal behaviour and site geometry rather than a measurement of the mechanism, since façade irradiance, external surface temperatures, and canopy geometry were not measured, and the divergence is absent in the small subset of free-running rooms, which are predominantly corridors of uncertain orientation. The data support the conclusion that the thermal difference is associated with afternoon solar exposure rather than with the amount of vegetation per se, even though green space can moderate humid heat where canopy cover is extensive (Yang et al., 2024), so the appropriate intervention at COE is solar control on the western façade, which is included in the appraisal as part of ECM-5.

Relative Humidity and Psychrometric Performance

Between buildings, RH was effectively identical at 61.83 ± 11.63% for COE and 61.62 ± 12.47% for CSM, a difference of 0.21 percentage points with a 95% CI of −1.66 to 2.09 ($p = 0.823$, $g = 0.02$), so no difference in relative humidity between the buildings can be claimed.

The substantive hygric finding concerns what air conditioning does, and psychrometric analysis is required to see it (Figure 7). Air-conditioned rooms averaged 56.42 ± 9.35% RH against 73.49 ± 8.78% in free-running rooms ($g = −1.86$). Because RH falls when air is merely warmed, this alone would not establish dehumidification, but converting to humidity ratio removes the ambiguity. Air-conditioned rooms held 12.58 ± 3.03 g kg^{-1} against 19.00 ± 3.10 g kg^{-1}, a reduction of 6.42 g kg^{-1} with a 95% CI of 5.90 to 6.94 and $g = −2.10$, the largest effect in the study. The coils are performing genuine latent-load removal, indicating that sensible capacity, rather than the refrigeration cycle, is the binding constraint. Only 35.4% of observations fell inside the combined comfort zone of 23 to 27 °C and 30 to 60% RH, and free-running rooms clustered almost entirely outside it at high humidity ratios, consistent with the role of humidity as a governing comfort parameter in warm and humid climates (Amaripadath et al., 2023; Jing et al., 2013), confirming that natural ventilation alone cannot deliver comfort in this climate.

Laboratories carry equipment-protection requirements and merit separate comment. Those at CSM averaged 62.62 ± 12.66% RH, compared with 59.16 ± 11.41% at COE (p =

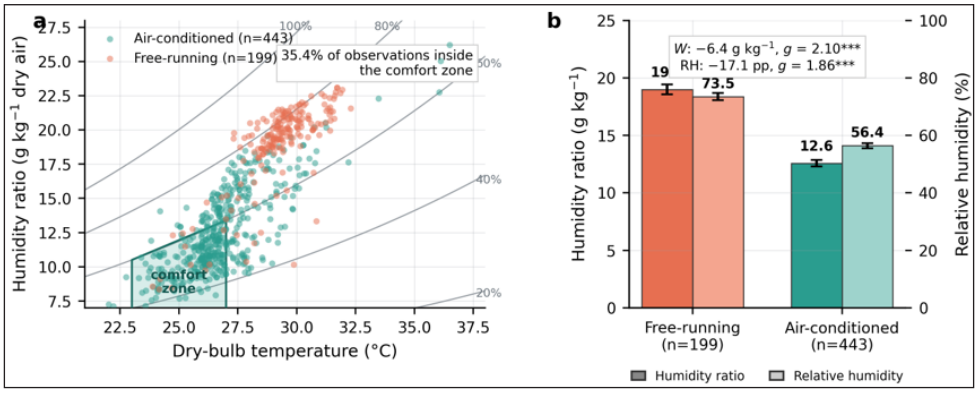


Figure 7. Psychrometric performance. (a) All observations plotted on temperature and humidity-ratio axes with constant-relative-humidity curves and the comfort zone of 23 to 27 °C and 30 to 60% relative humidity. (b) Humidity ratio and relative humidity for free-running and air-conditioned rooms, as mean with 95% confidence interval, where the humidity-ratio difference demonstrates real moisture removal rather than a temperature artefact

0.015, g = −0.29). Within CSM, they were more humid than non-laboratories ($p = 0.025$), whereas within COE, they were indistinguishable ($p = 0.974$). A plausible explanation, consistent with the inventory, is the different character of the wet work, since the chemistry and biology laboratories at CSM operate water baths, autoclaves, distillation apparatus, and sinks that release latent load directly into the space, whereas those at COE are predominantly dry. This is an inference from the inventory rather than a measured moisture balance. It matters because ASHRAE guidance recommends an RH band of 35 to 60% for laboratories, and the CSM mean is above it, so dedicated dehumidification, rather than lower setpoints, is the appropriate response (Lim et al., 2018; Shehadi, 2018).

Illuminance and Lighting Power Density

Illuminance was inadequate throughout both buildings, and the two do not differ (Figure 8). Mean task-plane illuminance was 166.5 ± 105.0 lx at COE, and 158.0 ± 115.1 lx at CSM ($p = 0.576$, $g = 0.08$), and under the 300-lx criterion, 86.0% of COE rooms and 88.2% of CSM rooms were under-lit, with indistinguishable compliance distributions (chi-square = 1.25, $p = 0.535$). Under-lighting is therefore a condition of the estate rather than of one college.

The diagnostic question is whether luminaires are inefficient or if too few are installed. Installed LPD averaged 3.29 ± 3.49 W m^{-2} at COE and 2.58 ± 2.45 W m^{-2} at CSM ($p = 0.099$), against a typical code allowance for educational occupancies of roughly 8 to 12 W m^{-2}. A total of 95.6% of rooms fall below that allowance, and 86.1% are simultaneously below the allowance and below 300 lx, while LPD correlates positively with delivered illuminance ($r = 0.52$, $p < 0.001$), as it must if the deficit is one of provision. Installed system efficacy was similar in both buildings, with medians of 52.2 and 47.8 lx m² W^{-1},

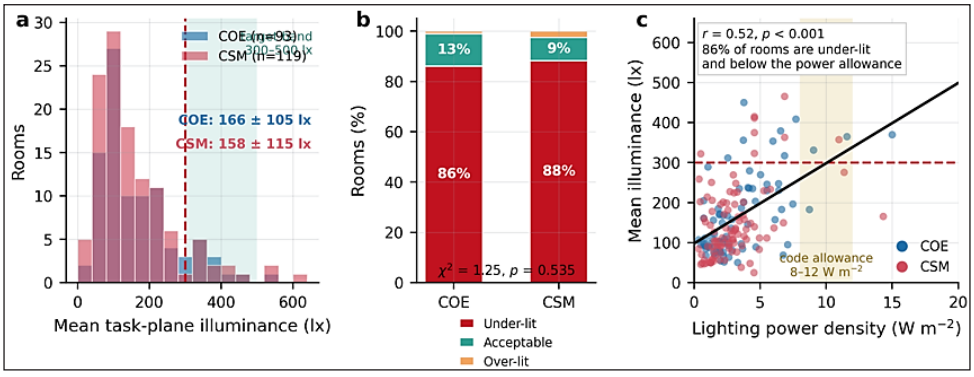


Figure 8. Illuminance and lighting power. (a) Distribution of mean task-plane illuminance by building against the 300-lx minimum and the target band of 300 to 500 lx. (b) Compliance classification by building. (c) Mean illuminance against installed lighting power density, where the shaded band marks a typical code allowance for educational occupancies

so the luminaires are not unusually poor; these buildings are under-lit because too little lighting power is installed.

This inverts the usual retrofit recommendation. A like-for-like LED replacement would halve lighting energy but leave 87% of rooms non-compliant, improving an energy metric while entrenching a health, safety and academic performance problem, since adequate illuminance supports visual comfort and task accuracy (Boyce, 2014; Oliveira et al., 2024; Rea, 2000). Relighting to 300 lx with LED luminaires instead requires a median installed density of only 1.9 W m^{-2} at COE and 2.1 W m^{-2} at CSM, and 99.4% of rooms would remain within a 10.5 W m^{-2} allowance. The measure is close to energy-neutral overall and net-negative at COE, saving 8.0 MWh y^{-1} despite roughly tripling delivered light, so adequate lighting and lower lighting energy are here the same intervention.

Room-Level Determinants of Energy Intensity

Across the 138 matched rooms, annual energy correlated with equipment count (r = 0.61) and equipment diversity (r = 0.69), and negatively but weakly with temperature (r = −0.26), RH (r = −0.25), and illuminance (r = −0.12), consistent with a causal direction running from conditioning to environment. Table 4 reports the two OLS models on air-conditioned rooms. The energy model explains 98.6% of variance in cooling energy intensity, essentially entirely through installed input power density and operating hours, with no residual building effect and no contribution from set-point (p = 0.520), while the gap model explains 91.2% of variance in the set-point gap, driven almost wholly by the set-point itself. Set point, therefore, predicts the shortfall almost perfectly and predicts consumption not at all, so lowering the thermostat changes what is asked for but neither what is delivered nor what is spent.

Table 4
Ordinary least squares models for air-conditioned rooms (n = 92)

Model	Predictor	B	SE	t	p	95% CI
a	Intercept	−66.45	9.997	−6.65	<0.001	−86.33 to −46.58
	Building (CSM)	1.18	3.584	0.33	0.743	−5.95 to 8.30
	Input power density (W m^{-2})	1.333	0.018	75.35	<0.001	1.297 to 1.368
	Set-point (°C)	−0.244	0.378	−0.65	0.520	−0.996 to 0.507
	Daily operating hours	8.979	0.682	13.17	<0.001	7.623 to 10.334
	Window-type capacity share	−3.064	3.571	−0.86	0.393	−10.16 to 4.04
b	Intercept	27.105	0.752	36.03	<0.001	25.61 to 28.60
	Building (CSM)	−0.174	0.324	−0.54	0.593	−0.818 to 0.470
	Input power density (W m^{-2})	−0.0024	0.0016	−1.48	0.144	−0.0056 to 0.0008
	Set point (°C)	−1.007	0.034	−29.37	<0.001	−1.075 to −0.939
	Window-type capacity share	0.045	0.323	0.14	0.889	−0.597 to 0.687

Appraisal of Energy Conservation Measures

Table 5 and Figure 9 appraise five measures, and the sequence matters as much as the content. ECM-1, replacement of window-type units with inverter split systems (Ebrahimi et al., 2023; Olatunde et al., 2024), is the largest single measure at 107.1 MWh y^{-1}, worth PHP 1.28 million and 59.5 t CO_2 y^{-1}, for an indicative PHP 5.90 million and a 4.6-year payback, concentrated at CSM where 120 window units are installed. Within the sensitivity range of 30% to 55% input reduction, payback ranges from 3.7 to 6.7 years. It is a prerequisite because it restores the capacity without which no control measure can act.

ECM-2, set-point governance at 24 °C, is credited at 58.0 MWh y^{-1} with a 0.4-year payback, but only after ECM-1, which is the practical expression of the central finding. Applied to the buildings as they stand, a thermostat policy would deliver no savings, because units that never reach 19 °C will likewise never reach 24 °C and will run continuously in either direction, or it would remove the only lever available to occupants without restoring any comfort. Sequenced after capacity restoration, it becomes one of the cheapest measures available, and it is not an unpopular imposition, since survey respondents endorsed switching air conditioning off during defined periods at 4.31 out of 5 and scheduled use of classrooms and laboratories at 4.34 out of 5. Audits recommending a set-point policy without first testing attainment risk prescribe an intervention that cannot work.

ECM-5, envelope sealing and solar-control glazing, delivers 48.0 MWh y^{-1} at a 4.0-year payback, addressing both the loss mechanism consistent with the observed net cooling of 1.75 °C and the unobstructed western exposure of COE, where vertical greenery

Table 5
Appraisal of energy conservation measures for both buildings combined

Measure	MWh y^{-1}	PHP y^{-1}	t CO_2 y^{-1}	CapEx (PHP)	Payback (y)	% base
ECM-1 Inverter split replacement of window-type units	107.1	1,284,786	59.5	5,895,000	4.6	9.3
ECM-2 Set-point governance at 24 °C, after ECM-1	58.0	695,999	32.2	300,000	0.4	5.0
ECM-5 Envelope sealing and solar control	48.0	575,620	26.6	2,308,808	4.0	4.2
ECM-4 Plug-load and standby governance	25.7	307,833	14.3	160,000	0.5	2.2
ECM-3 LED relighting to 300 lx	9.8	117,102	5.4	1,302,750	11.1	0.8
Package total	248.4	2,981,340	138.0	9,966,558	3.3	21.6

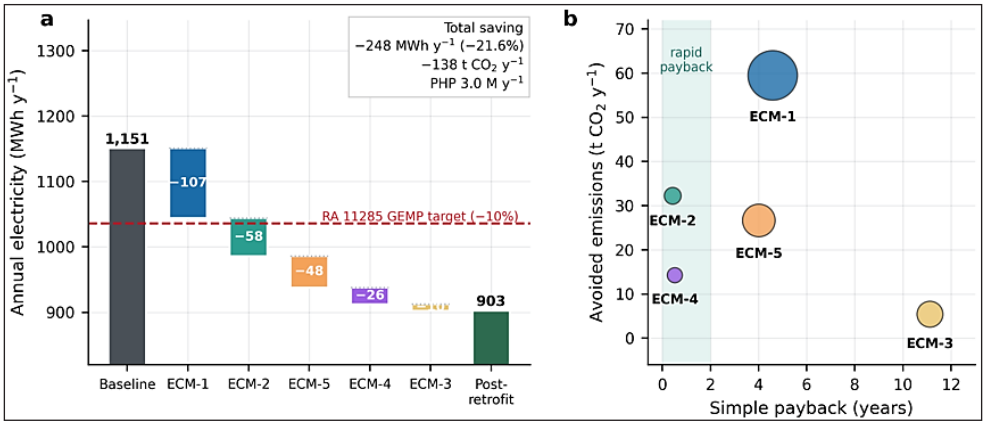


Figure 9. Appraisal of energy conservation measures. (a) Cumulative effect of the five measures on annual consumption, referenced to the 10% reduction target of the Government Energy Management Program under Republic Act 11285. (b) Avoided emissions against simple payback, where the bubble area is proportional to the indicative capital cost

and window gardens offer additional passive options in tropical and subtropical settings (Mohammad Shuhaimi et al., 2022; Ren et al., 2022). ECM-4, plug-load and standby governance, offers 25.7 MWh y^{-1} at a 0.5-year payback and is attractive at COE, where operations equipment is 29.3% of load. ECM-3, LED relighting to 300 lx, yields 9.8 MWh y^{-1} and an 11.1-year payback on energy grounds alone, and is recommended as a compliance measure that happens to be energy-neutral to negative.

Together, the package delivers 248 MWh y^{-1}, equivalent to 21.6% of baseline, worth PHP 2.98 million and 138 t CO_2 y^{-1}, at a blended payback of 3.3 years, more than twice the reduction of at least 10% required of state universities under the Government Energy Management Program of Republic Act 11285.

Limitations of the Study

Six limitations bound the interpretation. Energy is estimated from nameplate power and reported schedules rather than from metered data, so absolute EUI values are scoping estimates, whereas comparisons between end uses and buildings are more robust. The campaign covers three months of the first semester and omits the second semester, the summer term, and the examination periods. The thermal survey uses repeated spot measurements at three times of day rather than continuous logging, so it characterises the occupied-hours envelope but not overnight recovery or cycling. The canopy estimate is a plan-view screening analysis from a single-date image without ground truth, so it describes the site but cannot quantify a shading effect, for which façade-level irradiance and surface temperature measurement is required. The occupant survey, used only as corroboration, drew responses solely from undergraduate students and is not room-linked. The appraisal of conservation measures uses literature-derived savings fractions and indicative capital costs, so measurement and verification should accompany implementation.

CONCLUSION

This audit quantified 1,151 MWh y^{-1} of electricity, PHP 13.81 million in annual cost, and 639.5 t CO_2 y^{-1} across two large academic buildings at MSU-IIT, with cooling accounting for 47.2% of demand at COE and 56.3% at CSM, at energy use intensities of 107.2 and 112.0 kWh m^{-2} y^{-1}. Set points averaged 20.1 °C and fell below the 24 °C guideline in 71% of observations, a pattern routinely read as occupant profligacy. Measurements in the same rooms show the set-point attained in 2.0% of observations, with delivered temperature almost insensitive to it (slope of 0.097), and air-conditioning holding rooms only 1.75 °C below free-running conditions, so these buildings are capacity- and envelope-limited rather than control-limited. An occupant survey corroborates the low set-points to within 0.5 °C, placing the constraint in the plant rather than in behaviour. The difference in cooling energy between buildings is explained by installed capacity, which accounts for 86% of the logarithmic difference and reflects a technology mix at CSM dominated by window-type units. Cooling nonetheless dehumidifies 6.42 g kg^{-1}, isolating sensible capacity and envelope losses as the binding constraint. The thermal difference between the buildings is associated with afternoon solar exposure rather than with the quantity of vegetation, since outdoor temperatures were indistinguishable at every time of day while the indoor difference relative to ambient widened from 0.15 °C in the morning to 0.77 °C in the afternoon, matching a geometry in which the western façade of COE faces Iligan Bay unobstructed while CSM is screened by tall, mature trees and by COE itself. Separately, 87.3% of rooms fall below 300 lx, although the lighting power density averages 2.9 W m^{-2} against an allowance of 8 to 12 W m^{-2}, indicating a provision deficit in under-lighting.

The practical implication is an ordered program. Window-type units should be replaced with correctly sized inverter split systems and refrigerant charge and condenser condition restored, the envelope sealed and solar control applied to the western façade of COE in particular, and only then a 24 °C set-point policy instituted, which becomes among the cheapest measures once units can reach it and which occupants already endorse. Plug loads should be governed, and rooms relit to 300 lx with LED luminaires, raising illuminance roughly threefold while reducing installed lighting power. Together, these deliver 248 MWh y^{-1}, or 21.6% of the baseline, with a 3.3-year blended payback period, more than twice the statutory 10% target under Republic Act 11285.

Set-point and room temperature should be recorded together as a routine audit variable, because their relationship distinguishes interventions that can work from those that cannot at negligible cost, and energy and indoor environmental quality should be assessed in the same rooms, because the central question for a cooling-dominated building is not how much energy it uses but how much comfort that energy buys. Future work should extend the campaign across a full academic year, install sub-metering, log thermal conditions continuously alongside façade irradiance, surface temperature and resolved canopy height, and extend the survey to staff.

ACKNOWLEDGEMENT

The authors thank Mindanao State University-Iligan Institute of Technology (MSU-IIT) for providing resources and facilities essential to this study, and the DOST-ASTHRDP. Special thanks go to the MSU-IIT Centre for Resiliency (MCR) for its support with facilities, manpower, and equipment. This work was funded by MSU-IIT's internal research program through the Office of the Vice Chancellor for Research and Enterprise under Special Order No. 00786 series of 2024. No grant number was assigned to this funding.

REFERENCES

Abinales, P. N., & Amoroso, D. J. (2017). *State and society in the Philippines* (2nd ed.). Rowman & Littlefield.

Amaripadath, D., Rahif, R., Velickovic, M., & Attia, S. (2023). A systematic review on the role of humidity as an indoor thermal comfort parameter in humid climates. *Journal of Building Engineering, 68*, Article 106039. https://doi.org/10.1016/j.jobe.2023.106039

ASHRAE. (2020). *ANSI/ASHRAE Standard 55-2020: Thermal environmental conditions for human occupancy*.

Boyce, P. R. (2014). *Human factors in lighting* (3rd ed.). CRC Press. https://doi.org/10.1201/b16707

Chan, M., Caram, F. B., Contreras, M. A., San Miguel, C. A., & Tamayao, M.-A. (2020). Developing an energy audit for baseline and scenario analysis of a university library. *Philippine Engineering Journal, 41*(2). https://hdl.handle.net/11343/275168

Department of Energy Philippines. (2021). *Guidelines on energy conserving design for buildings (2020 ed.)*. https://legacy.doe.gov.ph/energy-efficiency/2020-guidelines-energy-conserving-design-buildings?q=energy-efficiency/downloads-energy-conserving-design-for-buildings

Department of Energy Philippines. (2022). *Natural gas development plan (NGDP)*. https://doe.gov.ph/articles/group/plans-2?category=Downstream%20Oil%20and%20Natural%20Gas&display_type=Card

Department of Energy Philippines. (2024). *Philippine energy plan 2023–2050 (Vol*. 1). https://doe.gov.ph/site/eppb/articles/group/plans?category=Philippine%20Energy%20Plan&display_type=Card

Ebrahimi, P., Ridwana, I., & Nassif, N. (2023). Solutions to achieve high-efficient and clean building HVAC systems. *Buildings, 13*(5), Article 1211. https://doi.org/10.3390/buildings13051211

Ferrer-Balas, D., Lozano, R., Huisingh, D., Buckland, H., Ysern, P., & Zilahy, G. (2010). Going beyond the rhetoric: System-wide changes in universities for sustainable societies. *Journal of Cleaner Production, 18*(7), 607–610. https://doi.org/10.1016/j.jclepro.2009.12.009

International Energy Agency. (2020). *Energy efficiency 2020*. https://www.iea.org/reports/energy-efficiency-2020

Jiang, J., Liu, F., Jiang, Z., Zhang, J., & Ma, C. (2025). Experimental evaluation of short-term thermal experiences on students' thermal comfort and cognitive performance. *Building and Environment, 285*, Article 113613. https://doi.org/10.1016/j.buildenv.2025.113613

Jing, S., Li, B., Tan, M., & Liu, H. (2013). Impact of relative humidity on thermal comfort in a warm environment. *Indoor and Built Environment, 22*(4), 598–607. https://doi.org/10.1177/1420326X12447614

Leal Filho, W., Salvia, A. L., do Paço, A., Anholon, R., Gonçalves Quelhas, O. L., Rampasso, I. S., Ng, A., Balogun, A.-L., Kondev, B., & Brandli, L. L. (2019). A comparative study of approaches towards energy efficiency and renewable energy use at higher education institutions. *Journal of Cleaner Production, 237*, Article 117728. https://doi.org/10.1016/j.jclepro.2019.117728

Li, X., Chen, S., Li, H., Lou, Y., & Li, J. (2020). Multi-dimensional analysis of air-conditioning energy use for energy-saving management in university teaching buildings. *Building and Environment, 185*, Article 107246. https://doi.org/10.1016/j.buildenv.2020.107246

Li, Y., Zhang, Q., Wang, P., Li, Y., Yang, J., & Wang, S. (2025). Study on the effect of summer classroom temperature on cognitive performance and heart rate variability in elementary school students. *Building and Environment, 285*, Article 113618. https://doi.org/10.1016/j.buildenv.2025.113618

Lim, D. K., Ahn, B. H., & Jeong, J. H. (2018). Method to control an air conditioner by directly measuring the relative humidity of indoor air to improve the comfort and energy efficiency. *Applied Energy, 215*, 290-299. https://doi.org/10.1016/j.apenergy.2018.02.004

Lopez, N. S., Gonzaga, J., & Lim, L. A. G. (2017). Energy audit and analysis of the electricity consumption of an educational building in the Philippines for smart consumption. In *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)* (pp. 1-4). IEEE. https://doi.org/10.1109/HNICEM.2017.8269505

Masoso, O. T., & Grobler, L. J. (2010). The dark side of occupants' behaviour on building energy use. *Energy and Buildings, 42*(2), 173-177. https://doi.org/10.1016/j.enbuild.2009.08.009

Mohammad Shuhaimi, N. D. A., Mohamed Zaid, S., Esfandiari, M., Lou, E., & Mahyuddin, N. (2022). The impact of vertical greenery system on building thermal performance in tropical climates. *Journal of Building Engineering, 45*, Article 103429. https://doi.org/10.1016/j.jobe.2021.103429

Olatunde, T. M., Okwandu, A. C., Akande, D. O., & Sikhakhane, Z. Q. (2024). Review of energy-efficient HVAC technologies for sustainable buildings. *International Journal of Science and Technology Research Archive, 6*(2), 012-020. https://doi.org/10.53771/ijstra.2024.6.2.0039

Oliveira, A. P., Ferreira, G., & Martinez-Perez, C. (2024). Effects of stage lighting on visual comfort at summer festivals: A study in Portugal. *Healthcare, 12*(23), Article 2441. https://doi.org/10.3390/healthcare12232441

Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings, 40*(3), 394-398. https://doi.org/10.1016/j.enbuild.2007.03.007

Rajapaksha, U. (2019). Heat stress pattern in conditioned office buildings with shallow plan forms in metropolitan Colombo. *Buildings, 9*(2), 35. https://doi.org/10.3390/buildings9020035

Rea, M. S. (Ed.). (2000). *The IESNA lighting handbook: Reference and application* (9th ed.). Illuminating Engineering Society of North America.

Ren, J., Tang, M., Zheng, X., Lin, X., Xu, Y., & Zhang, T. (2022). The passive cooling effect of window gardens on buildings: A case study in the subtropical climate. *Journal of Building Engineering, 46*, Article 103597. https://doi.org/10.1016/j.jobe.2021.103597

Republic of the Philippines. (2019). *Republic Act No. 11285: An act institutionalizing energy efficiency and conservation, enhancing the efficient use of energy, and granting incentives to energy efficiency and conservation projects*. Official Gazette.

Ryu, D., & Yoo, W. (2025). Ventilation-dominated energy savings in large commercial buildings: Multi-measure assessment revealing HVAC optimization priorities for hot-humid climates. *Case Studies in Thermal Engineering, 74*, Article 107034. https://doi.org/10.1016/j.csite.2025.107034

Santamouris, M. (2016). Cooling the buildings – past, present and future. *Energy and Buildings, 128*, 617-638. https://doi.org/10.1016/j.enbuild.2016.07.034

Shehadi, M. (2018). Review of humidity control technologies in buildings. *Journal of Building Engineering, 19*, 539-551. https://doi.org/10.1016/j.jobe.2018.06.009

United Nations Development Programme. (2026, April 30). *UNDP Climate Promise | Helping countries reach their climate goals*. https://climatepromise.undp.org/

Winkler, J., Munk, J., & Woods, J. (2018). Effect of occupant behavior and air-conditioner controls on humidity in typical and high-efficiency homes. *Energy and Buildings, 165*, 364-378. https://doi.org/10.1016/j.enbuild.2018.01.032

Yang, Y., Cao, C., Bogoev, I., Deetman, C., Dietz, G., Hang, J., Howard, L., Huang, X., Kendall, N., Lai, J., Lam, H., Tam, K., Yoo, C., Zhang, K., & Lee, X. (2024). Regulation of humid heat by urban green space across a climate wetness gradient. *Nature Cities, 1*(12), 871-879. https://doi.org/10.1038/s44284-024-00157-y

Zhang, Y., Bai, X., Mills, F. P., & Pezzey, J. C. V. (2018). Rethinking the role of occupant behavior in building energy performance: A review. *Energy and Buildings, 172*, 279-294. https://doi.org/10.1016/j.enbuild.2018.05.017

Zhmakin, V. M. (2023). *Morning temperature decrease with increasing humidity*. [Preprint]. Research Square. https://doi.org/10.21203/rs.3.rs-3752685/v1

Pertanika J. Sci. & Technol. 34 (5): 2739 - 2763 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Regenerative Braking-Based Energy Recovery for Electric Motorcycles in Hilly Terrain

Mawada Ahmed[1], Siti Fauziah Toha[1]*, Khairul Affendy[1], Muhammad Abdullah[2], Salmiah Ahmad[3], Mohammad Osman Tokhi[4], and Ahmad Syahrin Idris[5]

[1]Department of Mechatronics Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Gombak, Selangor, Malaysia
[2]Department of Mechanical and Aerospace Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Gombak, Selangor, Malaysia
[3]College of Engineering and Technology, University of Doha for Science and Technology, Al Tarafa, Jelaiah Street, Duhail North, P.O. Box 24449, Doha, Qatar
[4]School of Engineering, London South Bank University, 103 Borough Road, London SE1 0AA, United Kingdom
[5]Department of Electrical and Electronic Engineering, University of Southampton Malaysia, 79100 Iskandar Puteri, Johor, Malaysia

ABSTRACT

Limited battery energy availability remains a major challenge for electric motorcycles, especially in hilly terrain, where repeated climbing and downhill braking increase variability in energy consumption. This study investigates regenerative braking as an energy-recovery strategy for an electric motorcycle using a MATLAB/Simulink-based model. The proposed model integrates longitudinal vehicle dynamics, road-gradient effects, an H-bridge inverter, an L-C filter, a DC machine, a lithium-ion battery, and feedforward/PID control strategies. The Worldwide Motorcycle Test Cycle (WMTC) was used as the reference speed profile to evaluate speed tracking, motor response, battery behaviour, recovered energy, and state of charge (SOC) variation. The results show that downhill operation creates suitable conditions for regenerative braking, allowing the electric machine to operate as a generator and return energy to the battery. Compared with feedforward control, the PID controller improved the simulated regenerative braking response by reducing the RMSE from 14.10 m/s to 11.546 m/s. The recovered energy increased from 1.00×10^5 J to 2.20×10^5 J, while the SOC variation improved from −13.0% to −10.2%. The SOC response also showed a temporary recovery

ARTICLE INFO

Article history:
Received: 04 June 2026
Accepted: 26 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.09

E-mail addresses:
Mawadaahmed184@gmail.com (Mawada Ahmed)
tsfauziah@iium.edu.my (Siti Fauziah Toha)
mohd_abdl@iium.edu.my (Muhammad Abdullah)
salmiah.ahmad@udst.edu.qa (Salmiah Ahmad)
affendy@iium.edu.my (Khairul Affendy)
tokhim@lsbu.ac.uk (Mohammad Osman Tokhi)
A.S.Idris@soton.ac.uk (Ahmad Syahrin Idris)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

during braking, confirming that part of the mechanical energy was returned to the battery. The battery current and power responses verified reversed energy flow during regeneration, while the torque response confirmed generator operation. Overall, the findings demonstrate that regenerative braking with PID control can improve energy recovery and reduce battery depletion in electric motorcycles operating in hilly terrain.

Keywords: Battery state of charge, electric motorcycle, energy recovery, hilly terrain, MATLAB/Simulink, PID control, regenerative braking system, WMTC

INTRODUCTION

Electric motorcycles (EMs) are increasingly important in the transition toward sustainable and energy-efficient transportation. Compared with conventional internal combustion engine motorcycles, EMs produce lower local emissions, generate less noise, have simpler drivetrain structures, and offer lower operating costs. The global shift toward electric vehicles (EVs) is driven by environmental concerns, fossil-fuel depletion, and the need to reduce carbon dioxide emissions from the transportation sector. Recent developments in battery technology, power electronics, electric machines, charging infrastructure, and energy management have further supported the growth of EV systems (International Energy Agency, 2022; Madichetty et al., 2021). Karyatanti et al. (2025) also emphasised that EV adoption can reduce air pollution, greenhouse gas emissions, and dependence on petroleum-based fuels.

Despite these advantages, limited driving range remains a major barrier to wider adoption of EVs and EMs. This limitation is mainly caused by limited battery capacity, long charging times, battery ageing, and varying power demands under real driving conditions. The issue is more critical for electric motorcycles because their compact design limits the space available for large battery packs. Therefore, improving onboard energy utilisation is essential, especially for lightweight vehicles operating in areas with limited charging infrastructure or challenging road profiles (Aduama et al., 2023; Kadav & Asher, 2019; Yang, 2022).

Hilly terrain represents a demanding operating condition for electric motorcycles. During uphill motion, additional energy is required to overcome gravitational resistance and maintain the desired speed. During downhill motion, some of the energy previously expended during climbing is again available as potential and kinetic energy. In conventional braking systems, this energy is dissipated as heat through friction brakes. Regenerative braking systems (RBSs) offer a practical method for recovering part of this energy by allowing the electric machine to operate as a generator during braking or downhill descent. The recovered electrical energy can then be returned to the battery, improving energy utilisation and reducing battery energy depletion. For two-wheeled vehicles, this

is especially relevant because lightweight design and efficient energy use are essential for operation on hilly roads (Nama et al., 2022; Rego & Castro, 2023).

Regenerative braking performance depends on several factors, including machine characteristics, inverter switching, braking control, road gradient, battery condition, and energy storage capability. Joseph Godfrey and Sankaranarayanan (2018) showed that braking mode selection and switching strategy affect the amount of recovered energy in BLDC motor-driven EVs. Heydari et al. (2019) demonstrated that energy recovery can be improved by dynamically adjusting the low-speed cutoff point, while Heydari et al. (2020) showed that optimal brake allocation can increase harvested braking energy. Other studies have investigated regenerative energy recovery through switched-capacitor voltage control, low-speed battery charging, and hybrid energy storage systems (Jung & Ko, 2021; Naseri et al., 2017; Podder et al., 2021; Yeo et al., 2020). These studies confirm that effective regenerative braking requires proper coordination between the motor, power electronics, control strategy, and battery system.

Regenerative braking has also been examined in terms of braking-force distribution and battery state-of-charge (SOC) performance. Ghazali et al. (2020) reported that regenerative braking is an energy-saving solution because conventional braking produces unnecessary heat losses, while RBS technology enables part of the braking energy to be recovered. Their study also highlighted that braking force distribution, SOC, motor performance, and overall vehicle performance must be considered together when designing an RBS. This is highly relevant to the present study because the proposed EM model evaluates regenerative braking not only through recovered energy but also through motor response and SOC behaviour during hilly-terrain operation.

Battery behaviour is also central to regenerative braking performance. The recovered energy is useful only if the battery can safely receive and store the regenerated current. Badran and Toha (2024) explained that lithium-ion batteries are widely used in EV applications due to their high energy density and long lifespan, but their performance depends on stability, efficiency, life cycle, SOC estimation, fault detection, and cell balancing. Similarly, studies on battery cooling, cell balancing, and energy storage management show that battery condition directly affects EV range and long-term performance (Riczu & Bauman, 2021; Xie et al., 2020). Therefore, state of charge (SOC), battery current, and battery voltage are important indicators when evaluating regenerative braking systems.

Several studies have examined regenerative braking with respect to braking-force distribution, battery SOC, and overall EV performance. Ghazali et al. (2020) reported that regenerative braking reduces energy loss from conventional friction braking and improves battery SOC by recovering braking energy. Their work also highlighted the need to consider the distribution of braking force, motor response, SOC, and vehicle performance together when designing an RBS. This supports the need for a coupled mechanical-electrical model

capable of evaluating vehicle motion, motor-generator operation, and battery response within a single simulation framework.

Electric motorcycles require dedicated modelling because their dynamics, packaging constraints, and energy capacity differ from those of four-wheeled EVs (Madichetty et al., 2021). The selection of the electric motor affects propulsion efficiency and vehicle performance. Rego and Castro (2023) showed that regenerative braking in electric motorcycle prototypes requires attention to vehicle-specific dynamics and energy recovery limits. Nama et al. (2022) further demonstrated the design, modelling, and hardware implementation of regenerative braking for electric two-wheelers on hilly roads. These studies show that motorcycle-specific regenerative braking models should include longitudinal dynamics, road-gradient effects, braking force, motor operation, and battery SOC response.

The present work also builds on previous studies related to lightweight electric mobility and electric drive systems. Mohamed et al. (2025) developed a smart IoT-based energy optimisation and localisation monitoring system for e-bike sharing, showing the importance of energy-aware monitoring in electric two-wheeler applications. Mohamed et al. (2025) investigated voltage regulation of buck converter-fed DC motors under dynamic load conditions, which is relevant to stable motor-drive operation under changing electrical and mechanical loads. These studies support the broader need for reliable energy management and controlled motor-drive behaviour in compact electric mobility systems.

Although regenerative braking has been widely studied in electric cars, hybrid vehicles, buses, rail systems, and charging infrastructure, fewer studies focus on electric motorcycles operating in hilly terrain. This gap is important because hilly routes involve repeated high-energy demands during climbs and repeated recovery opportunities during descents. A simplified and practical simulation framework is therefore needed to evaluate how regenerative braking can improve SOC behaviour and energy recovery in electric motorcycles under gradient-rich operating conditions.

Therefore, this paper investigates regenerative braking as an energy-recovery strategy for electric motorcycles operating in hilly terrain. A MATLAB/Simulink model is developed to link motorcycle longitudinal dynamics, road-gradient effects, braking force, motor-generator operation, H-bridge switching, battery current, voltage, SOC, and recovered energy. The study evaluates a practical regenerative braking configuration using feedforward and PID control strategies. The main contribution of this work is a coupled mechanical-electrical simulation framework that assesses whether regenerative braking can recover useful downhill energy, reduce battery depletion, and improve the energy performance of electric motorcycles in hilly environments.

MATERIALS AND METHODS

Research Design

This study used a MATLAB/Simulink-based modelling approach to evaluate regenerative braking as an energy-recovery strategy for an electric motorcycle operating in hilly terrain. The model integrates motorcycle longitudinal dynamics, a lithium-ion battery, an H-bridge inverter, an L-C filter, a DC machine, switching logic, and performance measurement blocks within a single simulation environment. The main purpose of the model is to analyse the transition between motoring and regenerative braking modes and to evaluate the electrical energy recovered during deceleration.

The scope of the study was limited to regenerative energy recovery, battery state-of-charge behaviour, speed-tracking response, and energy-flow direction. Direct driving-range estimation, no-regeneration baseline comparison, and hardware-level component sizing were not included in the present model.

The overall structure of the proposed regenerative braking system for electric motorcycles is shown in Figure 1. The model receives the Worldwide Motorcycle Test Cycle (WMTC) as the reference speed input. The vehicle dynamics subsystem calculates the force demand, torque, angular speed, and mechanical power. The battery supplies energy during motoring mode and receives energy during regenerative braking mode. The H-bridge inverter controls the current direction through four MOSFET switches, while the L-C filter reduces switching ripple. The performance-output block calculates speed, torque, braking force, electrical power, mechanical power, recovered energy, and efficiency.

The operating principle of the electric motorcycle is shown in Figure 2. During motoring mode, electrical energy flows from the battery to the motor, which produces driving torque.

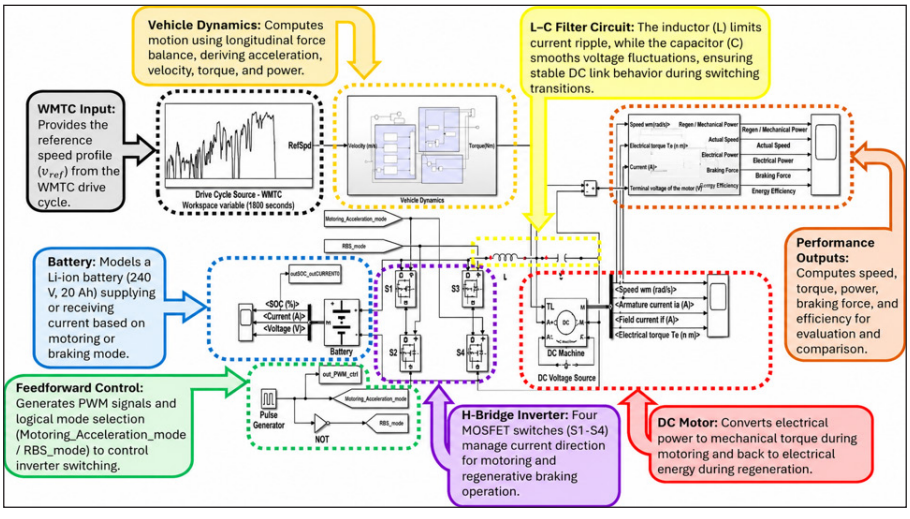


Figure 1. Overall MATLAB/Simulink architecture of the electric motorcycle regenerative braking system

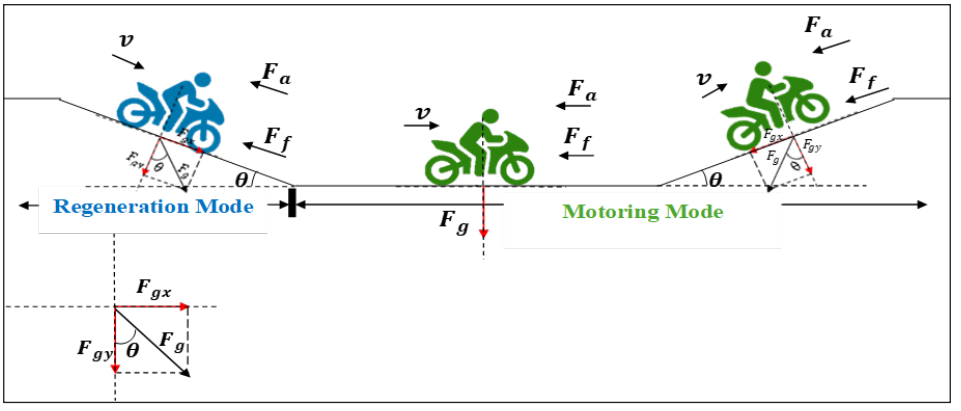


Figure 2. Operating principle of the electric motorcycle during motoring and regenerative braking modes

During regenerative braking mode, the wheel drives the machine mechanically, causing it to operate as a generator. In this condition, part of the mechanical energy is converted into electrical energy and returned to the battery.

Electric Motorcycle Model Parameters

The electric motorcycle was represented as a longitudinal single-track vehicle. The model focuses on straight-line motion, road-gradient effects, braking force, wheel torque, motor operation, and energy recovery. Lateral dynamics, steering behaviour, tyre slip, rider posture variation, and detailed mechanical brake blending were not considered, since the main purpose of this study is to analyse regenerative braking energy flow.

All vehicle-dynamics equations were defined using SI units. The vehicle speed, wheel angular speed, force, torque, and power variables were treated consistently according to the longitudinal vehicle model. A direct wheel-to-machine speed relationship was assumed through $v = r\omega$, and no separate gear ratio, transmission loss model, or detailed drivetrain efficiency map was included.

The parameters used in the longitudinal dynamics model are listed in Table 1. These parameters were used to calculate rolling resistance, aerodynamic drag, gradient resistance, inertial force, wheel torque, angular speed, and mechanical power.

Longitudinal Vehicle Dynamics

The longitudinal motion of the electric motorcycle was described using the force balance along the direction of travel. The model includes rolling resistance, aerodynamic drag, gradient resistance, and inertial force. These force components determine the total force demand at the wheel and are then used to calculate wheel torque and mechanical power.

Table 1
Motorcycle and Simulation Parameters Used in the Longitudinal Dynamics Model

Parameter	Symbol	Value	Unit	Description
Total mass, vehicle and rider	m	255	kg	Combined mass of the electric motorcycle and rider
Gravitational acceleration	g	9.81	m/s²	Standard gravitational acceleration
Wheel radius	r	0.30	m	Tyre rolling radius
Air density	ρ	1.18	kg/m³	Ambient air density
Frontal area	A	0.60	m²	Projected frontal area of the motorcycle
Aerodynamic drag coefficient	C_d	0.90	–	Coefficient representing aerodynamic resistance
Rolling resistance coefficient	μ_{rr}	0.01	–	Coefficient representing tyre-road rolling resistance
Rotational inertia coefficient	λ	1.03	–	Correction factor for wheel and drivetrain rotational inertia
Road angle	α	Variable	°	Road-gradient condition for flat, uphill, and downhill operation
Desired braking deceleration	a_{brake}	−15	m/s²	Target deceleration during braking mode
Initial angular speed	ω_0	300-400	rad/s	Initial wheel angular speed range
Simulation step size	T_s	0.1	s	Sampling time used for drive-cycle tracking
Drive cycle	–	WMTC	–	Worldwide Motorcycle Test Cycle reference speed profile
Drive-cycle duration	T	1800	s	Total duration of the WMTC profile

The rolling resistance force was calculated using Equation 1:

$$F_{rr} = mg\mu_{rr}cos(\alpha) \quad [1]$$

where F_{rr} is the rolling resistance force, m is the total motorcycle-rider mass, g is the gravitational acceleration, μ_{rr} is the rolling resistance coefficient, and α is the road angle.

The aerodynamic drag force was calculated using Equation 2:

$$F_{ad} = \frac{1}{2}\rho C_d A v^2 \quad [2]$$

where F_{ad} is the aerodynamic drag force, ρ is the air density, C_d is the aerodynamic drag coefficient, A is the frontal area, and v is the motorcycle speed.

The gradient resistance force was calculated using Equation 3:

$$F_g = mgsin(\alpha) \quad [3]$$

where F_g is the component of the gravitational force acting along the slope. During uphill motion, this force increases the required tractive effort. During downhill motion, it assists the motorcycle's motion and creates an opportunity for regenerative braking.

The inertial force was calculated using Equation 4:

$$F_i = \lambda m \frac{dv}{dt} \quad [4]$$

where F_i is the inertial force, λ is the rotational inertia coefficient, and dv/dt is the longitudinal acceleration.

The total longitudinal force was then obtained using Equation 5:

$$F_{total} = F_{rr} + F_{ad} + F_g + F_i \quad [5]$$

The required wheel torque was calculated using Equation 6:

$$T_W = F_{total}\, r \quad [6]$$

where T_W is the wheel torque and r is the wheel radius.

The relationship between motorcycle speed and wheel angular speed was calculated using Equation 7:

$$v = r\omega \quad [7]$$

where ω is the wheel angular speed.

The mechanical power at the wheel was calculated using Equation 8:

$$P_{mech} = T_w \omega \quad [8]$$

The mechanical variables calculated in the longitudinal subsystem represent the vehicle-side force, torque, angular speed, and power demand used to evaluate the simulated regenerative braking response. These variables were not intended to represent experimentally validated motor-shaft limits. The Simulink implementation of the vehicle dynamics subsystem is shown in Figure 3. The subsystem calculates rolling resistance, aerodynamic drag, gradient resistance, acceleration resistance, total force, wheel torque, and mechanical power.

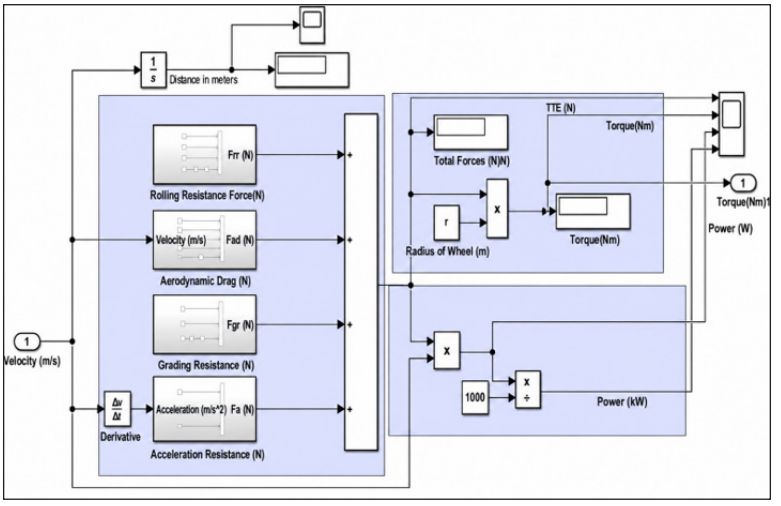


Figure 3. Vehicle dynamics subsystem for calculating force, torque, and mechanical power

Regenerative Braking System Architecture

The regenerative braking system was designed using a battery, an H-bridge inverter, an L-C filter, and a DC machine. The H-bridge inverter consists of four MOSFET switches that control the direction of current flow. During motoring mode, the battery supplies electrical energy to the DC machine. During regenerative braking mode, the DC machine operates as a generator, returning electrical energy to the battery.

The block-level structure of the regenerative braking system is shown in Figure 4. The controller selects either motoring or regenerative braking depending on the speed condition and braking demand. This selected mode controls the inverter switching state and determines whether electrical power flows from the battery to the motor or from the motor back to the battery.

The electrical and switching parameters used in the regenerative braking model are listed in Table 2.

The DC machine rating in Table 2 represents the nominal motor-generator parameter used in the model. The present work does not include detailed component-level saturation, torque-speed envelope limits, current protection limits, thermal constraints, battery charge-acceptance limits, or hardware derating. Therefore, the electrical and mechanical waveforms are interpreted as simulated response indicators for regenerative braking behaviour and controller comparison rather than final hardware operating ratings. Detailed component sizing and experimentally validated operating limits are recommended for future work. The detailed Simulink circuit of the regenerative braking model is shown in Figure 5. The circuit contains the battery, four MOSFET switches, an L-C filter, a DC voltage source, and a DC machine. The L-C filter is placed between the inverter and the machine side to reduce current ripple and smooth voltage fluctuation during switching.

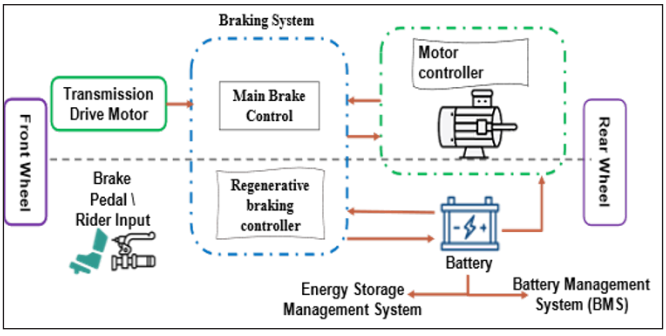


Figure 4. Block diagram of the regenerative braking system for the electric motorcycle drive

Table 2
Electrical and Switching Parameters of the Regenerative Braking Model

Component	Parameter	Symbol	Value	Unit	Description
Battery	Initial state of charge	SOC_0	50	%	Initial battery charge level
Battery	Battery type	–	Lithium-ion	–	Main energy source and storage element
DC machine	Rated power	P_m	5	HP	Motor-generator power rating
DC machine	Rated voltage	V_m	240	V	Rated machine voltage
DC machine	Rated speed	n_m	1750	rpm	Rated machine speed
DC voltage source	Source voltage	V_{dc}	300	V	DC supply voltage used in the circuit
Inductor	Inductance	L	5	mH	Limits current ripple
Capacitor	Capacitance	C	150	μF	Smooths DC-link voltage fluctuation
MOSFET bridge	Power switches	$S_1 - S_4$	4	–	Switches used for motoring and regeneration
Pulse generator	Time period	T_P	2×10^{-5}	s	Corresponds to 50 kHz switching frequency
Pulse generator	Pulse width	D	25	%	Switch-on duration within one PWM period

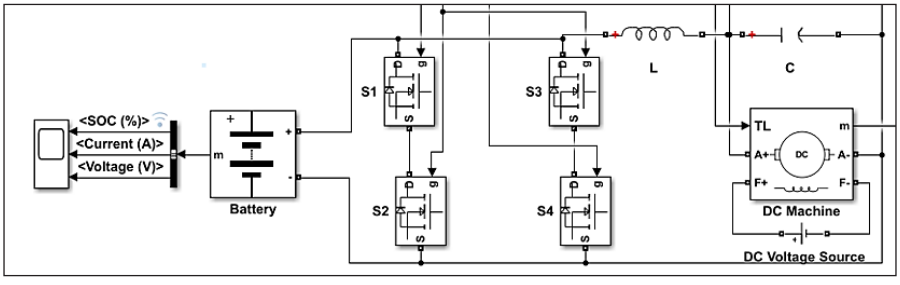


Figure 5. MATLAB/Simulink regenerative braking circuit with H-bridge inverter, L-C filter, battery, and DC machine

Motoring and Regenerative Braking Modes

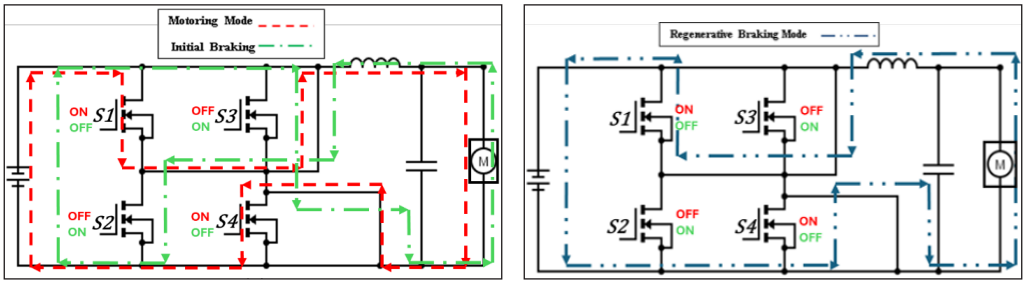
The H-bridge inverter was controlled in two modes: motoring and regenerative braking. In motoring mode, the battery discharges and supplies current to the DC machine, enabling it to produce driving torque. In regenerative braking mode, the wheel drives the DC machine, and the generated electrical energy is directed back to the battery.

The switching states used in the H-bridge inverter are shown in Table 3. In motoring mode, switches S_1 and S_4 are turned ON, while S_2 and S_3 are turned OFF. In regenerative braking mode, switches S_2 and S_3 are turned ON, while S_1 and S_4 are turned OFF.

The current path during motoring mode is shown in Figure 6(a), while the current path during regenerative braking mode is shown in Figure 6(b). These figures show how the active switch pair changes the current direction and changes the DC machine operation between motor and generator modes.

Table 3
H-Bridge Switching States During Motoring and Regenerative Braking

Operating mode	S1	S2	S3	S4	Machine operation	Battery operation
Motoring mode	ON	OFF	OFF	ON	Motor	Discharging
Regenerative braking mode	OFF	ON	ON	OFF	Generator	Charging



(a) (b)

Figure 6. H-bridge current paths: (a) motoring mode and (b) regenerative braking mode

Mode Selection Logic

The operating mode was selected using switching logic based on braking demand and the relationship between the reference speed and the actual speed. The reference speed was obtained from the WMTC drive-cycle profile, while the actual speed was calculated from the vehicle dynamics subsystem.

The speed error was calculated using Equation 9:

$$e(t) = v_{ref}(t) - v(t) \quad [9]$$

where $e(t)$ is the speed error, $v_{ref}(t)$ is the reference speed, and $v(t)$ is the actual motorcycle speed.

Regenerative braking was activated when the actual speed exceeded the reference speed or when braking input was applied. The speed-based switching condition is shown in Equation 10:

$$v(t) > v_{ref}(t) \quad [10]$$

When Equation 10 is satisfied, the inverter switches from motoring mode to regenerative braking mode. This allows the DC machine to generate opposing torque and return current to the battery.

The complete mode-selection process is shown in Figure 7. The process begins with rider or brake-pedal input. The system then checks whether the electric motorcycle is active and selects either motoring mode or regenerative braking mode by activating the corresponding switch pair.

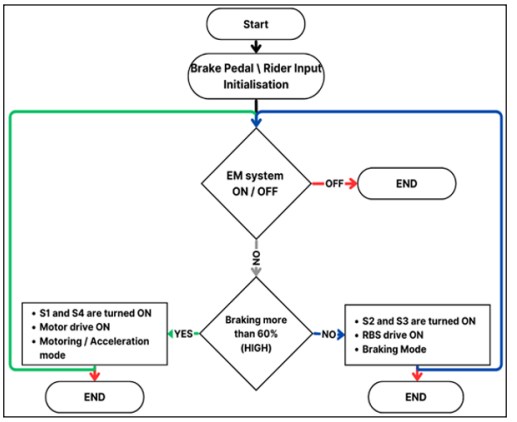


Figure 7. Flowchart of the motoring and regenerative braking mode-selection process

Feedforward and PID Control

A feedforward command was used to provide the initial switching and PWM signal based on the reference drive-cycle condition. The feedforward signal was defined as a function of the reference speed and reference acceleration, as shown in Equation 11:

$$u_{ff}(t) = f(v_{ref}(t), a_{ref}(t)) \quad [11]$$

where $u_{ff}(t)$ is the feedforward command, $v_{ref}(t)$ is the reference speed, and $a_{ref}(t)$ is the reference acceleration obtained from the WMTC profile.

A PID controller was used to improve speed-tracking response by adjusting the duty cycle based on the difference between the reference and actual speeds. The PID control signal was calculated using Equation 12:

$$u_{PID}(t) = K_p e(t) + K_i \int e(t)dt + K_d \frac{de(t)}{dt} \quad [12]$$

where $u_{PID}(t)$ is the PID output, K_p is the proportional gain, K_i is the integral gain, K_d is the derivative gain, and $e(t)$ is the speed error.

The PWM duty cycle was limited using Equation 13.

$$0 \leq D(t) \leq 1 \quad [13]$$

where $D(t)$ is the inverter duty cycle.

The PID-based Simulink implementation is shown in Figure 8. The WMTC reference speed is compared with the actual motorcycle speed, and the PID controller generates the duty-cycle correction signal. The corrected PWM signal is then applied to the H-bridge inverter to improve speed tracking and regulate regenerative braking.

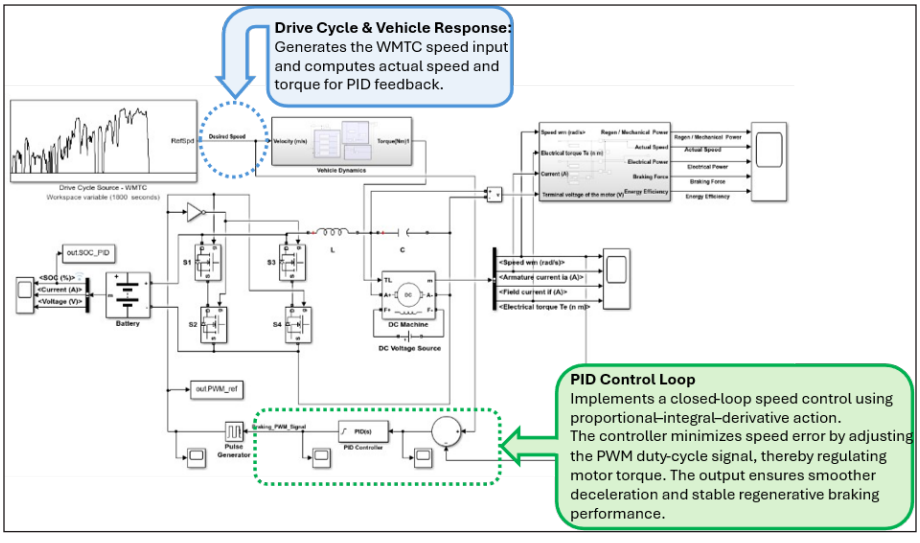


Figure 8. PID-based Simulink control structure for regenerative braking

Battery Model and Energy Flow Calculation

The lithium-ion battery was modelled as the main energy source during motoring and as the storage element during regenerative braking. During motoring, the battery current is positive because electrical energy flows from the battery to the motor. During regenerative braking, the battery current becomes negative because electrical energy flows from the generator back to the battery. The battery voltage, current, power, and SOC were recorded to evaluate the system's charging and discharging behaviour.

The instantaneous battery power was calculated using Equation 14:

$$P_{batt}(t) = V_{batt(t)}\, I_{batt(t)} \quad [14]$$

where $P_{batt(t)}$ is the battery power, $V_{batt(t)}$ is the battery terminal voltage, and $I_{batt(t)}$ is the battery current.

The recovered regenerative energy was calculated by integrating the charging portion of the battery power curve, as shown in Equation 15:

$$E_{regen} = \int_0^T max\,(0, -P_{batt}\,(t))dt \quad [15]$$

where E_{regen} is the recovered energy, and T is the simulation time.
The discharge energy was calculated using Equation 16:

$$E_{dis} = \int_0^T max\,(0, P_{batt}\,(t))dt \quad [16]$$

where E_{dis} is the energy supplied by the battery during motoring operation.
The regenerative energy ratio was calculated using Equation 17:

$$\eta_{regen} = (\frac{E_{regen}}{E_{dis}}) \times 100 \quad [17]$$

where η_{regen} is the regenerative energy ratio.
The SOC variation was calculated using Equation 18:

$$\Delta SOC = SOC_{final} - SOC_{initial} \quad [18]$$

where SOC_{final} is the final state of charge and $SOC_{initial}$ is the initial state of charge. These energy-related outputs were used to evaluate battery energy-flow direction and SOC behaviour. They were not used to calculate absolute driving range, Wh/km energy consumption, or percentage range improvement in the present study.

Braking Force, Power, and Efficiency Calculation

The regenerative braking force was calculated from the electromagnetic torque and wheel radius, as shown in Equation 19:

$$F_b = \frac{T_e}{r} \quad [19]$$

where F_b is the regenerative braking force, T_e is the electromagnetic torque, and r is the wheel radius. During regenerative braking, the electromagnetic torque opposes the wheel's rotation, producing a braking effect.

The regenerative mechanical power was calculated using Equation 20:

$$P_{regen} = T_e \omega \quad [20]$$

where P_{regen} is the mechanical regenerative power, and ω is the wheel angular speed. The regeneration efficiency was calculated using Equation 21:

$$\eta_{energy} = \frac{P_{elec}}{P_{regen}} \times 100 \quad [21]$$

where η_{energy} is the regeneration efficiency, P_{elec} is the electrical power returned to the battery, and P_{regen} is the available mechanical power during regenerative braking.

The braking force, power, and efficiency calculations were used to compare feedforward and PID control. Since detailed component-level saturation and experimental validation were outside the scope of the present work, these outputs were interpreted as model-based performance indicators. The performance-output subsystem used for these calculations is shown in Figure 9. The subsystem receives angular speed, electromagnetic torque, current, and terminal voltage as inputs. It then calculates actual speed, regenerative mechanical power, electrical power, braking force, and energy efficiency.

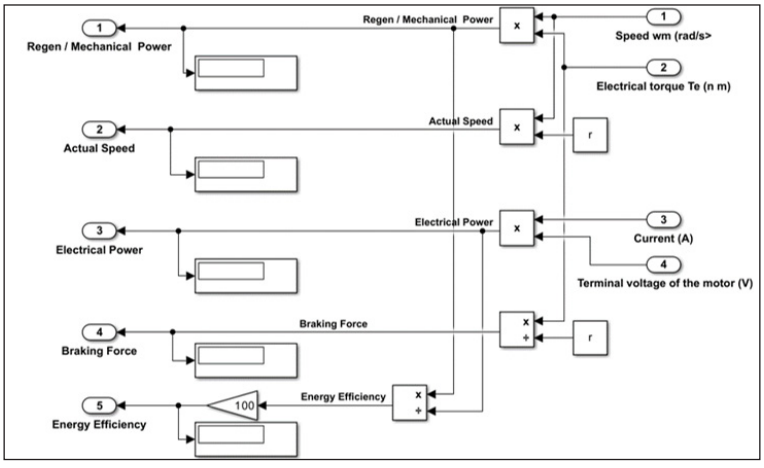


Figure 9. Performance-output subsystem for speed, torque, power, braking force, and regeneration efficiency

Model Validation Procedure

The model was validated by monitoring the mechanical and electrical responses during motoring and regenerative braking. The reference speed was compared with the actual speed to evaluate speed-tracking behaviour, while battery voltage, battery current, battery power, SOC, electromagnetic torque, angular speed, and armature current were used to verify regenerative braking operation. The validation in this study refers to internal simulation consistency and verification of regenerative braking behaviour. It does not represent experimental validation against a physical electric motorcycle prototype.

The speed-tracking error was calculated using the root mean square error (RMSE), as shown in Equation 22:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(v_{ref,i} - v_i)^2} \quad [22]$$

where N is the number of samples, $v_{ref,i}$ is the reference speed at sample i, and v_i is the actual motorcycle speed at sample i.

Successful regenerative braking was identified by negative battery current during braking, opposing electromagnetic torque during deceleration, and reduced SOC drop or temporary SOC recovery during regenerative intervals. These validation outputs were used to compare the feedforward and PID-based regenerative braking responses. Therefore, the model was used to assess regenerative braking activation, controller response, and battery energy recovery trends rather than to establish final motor sizing, range extension, or hardware operating limits.

RESULT AND DISCUSSION

Force Response During Hilly-Terrain Operation

The longitudinal force response was evaluated to verify the behaviour of the electric motorcycle under downhill, flat-road, and uphill operating conditions. Figure 10 shows the rolling resistance, air resistance, slope resistance, acceleration resistance, net motoring force, net regenerative force, and regenerative braking force acting on the motorcycle.

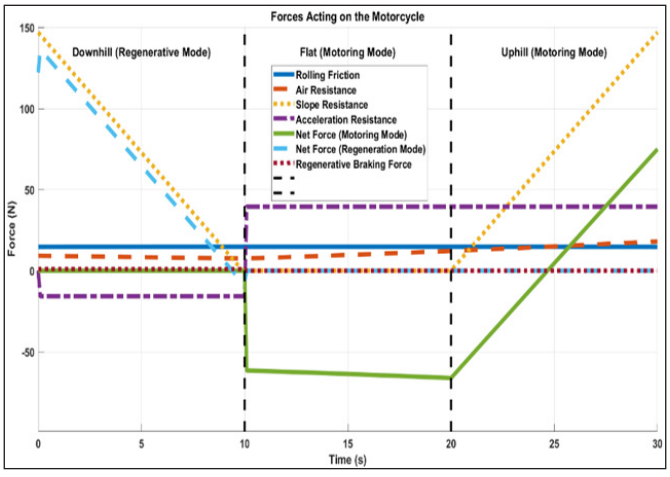


Figure 10. Forces acting on the electric motorcycle during downhill, flat-road, and uphill operation

In the downhill region from 0 s to 10 s, the slope resistance assists the motorcycle's motion. The net force in regenerative mode becomes negative, reaching approximately −60 N. This indicates that the electric machine produces an opposing braking effect as the motorcycle descends. At the same time, the regenerative braking force opposes the wheel's rotation, allowing part of the available mechanical energy to be converted into electrical energy.

In the flat-road region from 10 s to 20 s, the slope resistance is nearly absent, and the force demand becomes more stable. The motorcycle primarily needs to overcome rolling and aerodynamic resistance. In the uphill region from 20 s to 30 s, the slope resistance increases strongly, reaching approximately 75 N near the end of the profile. As a result, the required motoring force becomes positive, showing that the motor must supply additional torque to overcome the gravitational load.

This confirms that road gradient directly affects electric motorcycle energy demand. Downhill sections create an opportunity for regenerative braking, while uphill sections increase the required motoring effort. Therefore, the hilly-terrain force response supports the use of regenerative braking as an energy-recovery strategy during downhill and deceleration conditions.

Speed Tracking Under WMTC Operation

The reference and actual speed responses were compared under the WMTC-based simulation profile. Figure 11 shows the feedforward speed response, while Figure 12 shows the PID speed response.

For feedforward control, the actual speed generally follows the reference speed, but noticeable deviations occur during rapid transients. This shows that feedforward control

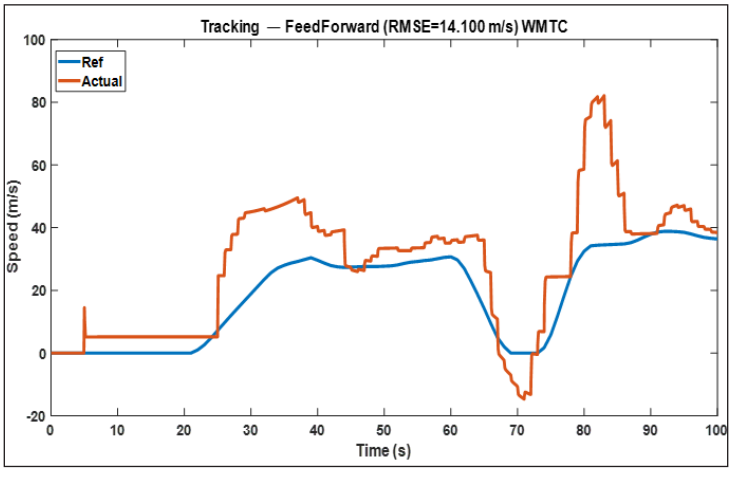


Figure 11. Reference and actual speed response using feedforward control under the WMTC profile

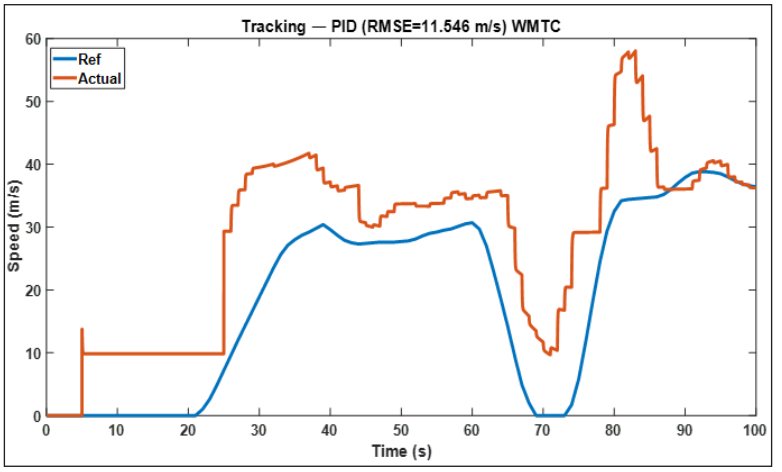


Figure 12. Reference and actual speed response using PID control under the WMTC profile

is limited under fast-transient conditions because it does not correct the speed error via feedback.

For PID control, the actual speed still deviates from the reference speed during some transient regions, but the overall tracking response is improved compared with feedforward control. The PID controller adjusts the PWM duty cycle according to the difference between the reference and actual speeds. To avoid overinterpreting transient simulated peak values as physical hardware limits, the speed response is evaluated using RMSE rather than by emphasising maximum speed values. The numerical speed-tracking comparison is shown in Table 4.

The PID controller reduced the RMSE from 14.10 m/s to 11.546 m/s. This represents an 18.1% reduction in speed-tracking error compared with feedforward control. Since RMSE measures the difference between the reference and actual speeds, a lower RMSE indicates that PID control provides better speed-tracking performance.

Table 4
WMTC Speed-Tracking Performance of Feedforward and PID Controllers

Controller	RMSE (m/s)	Tracking performance
Feedforward	14.10	Higher tracking error
PID	11.546	Lower tracking error

Battery State of Charge Response

The battery state of charge was evaluated to determine whether regenerative braking returned energy to the battery. Figure 13 shows the SOC response during motoring and regenerative braking operation.

The SOC begins at approximately 90% and decreases during motoring operation because the battery supplies electrical energy to the motor. Between approximately 60 s

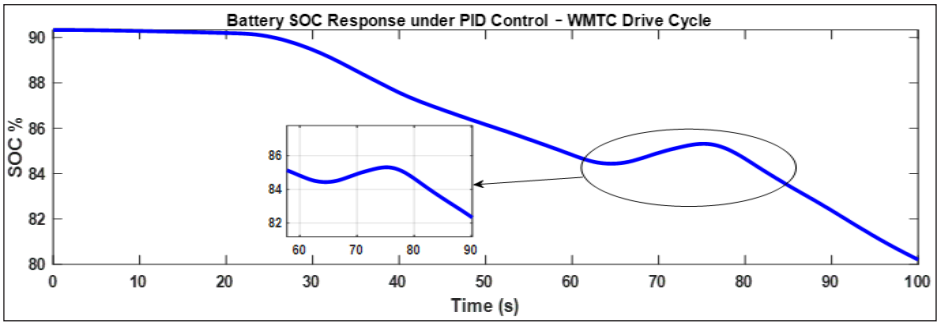


Figure 13. Battery state of charge response during motoring and regenerative braking

and 80 s, a temporary SOC recovery is observed. In this region, the SOC increases from approximately 84.5% to about 85.3%, showing that part of the braking energy is converted into electrical energy and stored back in the battery.

After the regenerative interval, the SOC decreases again as the motorcycle returns to motoring operation. By the end of the 100-second simulation window, the SOC reaches approximately 80%. This behaviour confirms that regenerative braking does not eliminate battery discharge, but it reduces the rate of depletion by recovering part of the energy during braking or downhill operation.

The SOC recovery region is important because it proves that the regenerative braking system is functioning correctly. Instead of dissipating all braking energy as heat, the system recovers a portion of the mechanical energy and returns it to the battery. Therefore, the SOC response is used as evidence of energy recovery and reduced battery depletion, not as a direct quantified driving-range improvement.

Battery Voltage, Current, and Power Response

Battery voltage, current, and power were analysed to verify the direction of energy flow during motoring and regenerative braking. Figure 14 shows the feedforward battery response, while Figure 15 shows the PID battery response.

In the feedforward case, the battery voltage, current, and power responses show alternating motoring and regenerative intervals. During motoring operation, the battery current and battery power are positive, indicating that energy flows from the battery to the motor. During regenerative braking, the battery current and battery power become negative, confirming that the direction of energy flow is reversed and that electrical energy is returned to the battery.

In the PID case, the battery current and power responses also show clear positive and negative regions. The negative current and negative power regions indicate that the DC machine operates as a generator during braking, returning energy to the battery. Compared with feedforward control, the PID response exhibits more effective regenerative charging

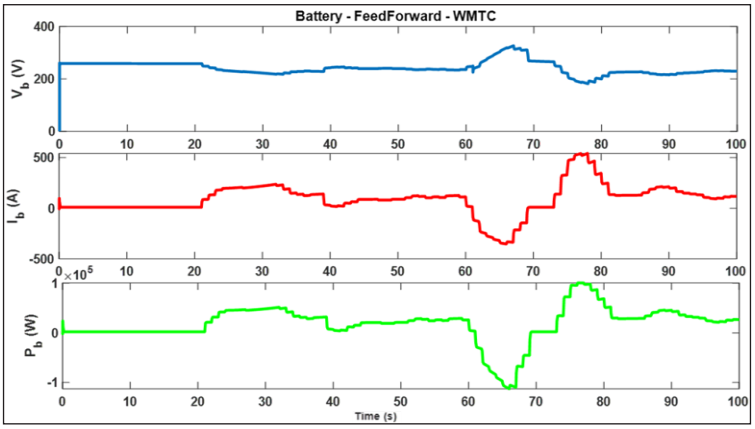


Figure 14. Battery voltage, current, and power response using feedforward control

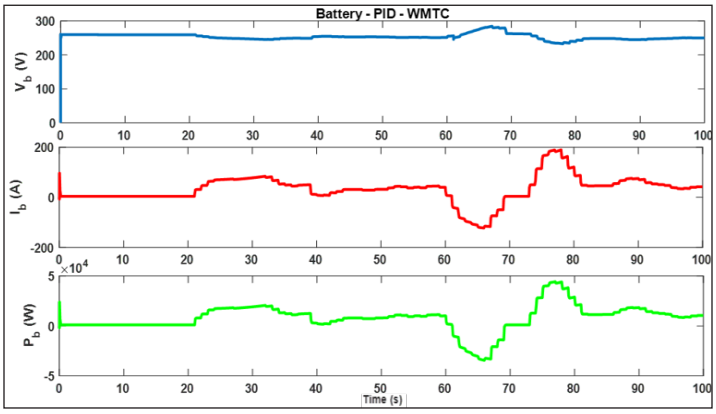


Figure 15. Battery voltage, current, and power response using PID control

behaviour, consistent with the higher recovered energy reported in Table 5. The transient battery current and power waveforms are interpreted as indicators of simulation response for energy-flow direction and controller comparison. They are not presented as experimentally validated component operating limits.

The PID controller recovered 2.20×10^5 J, while the feedforward controller recovered 1.00×10^5 J. This corresponds to a 120% increase in recovered energy. The regeneration

Table 5
WMTC Regenerative Energy and SOC Performance

Controller	Recovered energy, Eregen (J)	Regeneration fraction	SOC variation, ΔSOC (%)	Regeneration efficiency indicator
Feedforward	1.00×10^5	0.090	−13.0	0.06
PID	2.20×10^5	0.160	−10.2	0.40

fraction increased from 0.090 to 0.160, while SOC variation improved from −13.0% to −10.2%. This means the PID controller reduced SOC depletion by 2.8 percentage points compared with feedforward control.

These results show that PID control improves both battery charging behaviour and regenerative energy recovery. However, since a no-regeneration baseline was not included, the results are discussed in terms of recovered energy and SOC variation rather than absolute increase in range, Wh/km consumption, or percentage driving-range improvement.

DC Machine Torque and Current Response

The DC machine response was evaluated using angular speed, electromagnetic torque, and armature current. Figure 16 shows the feedforward response, while Figure 17 shows the PID response.

In the feedforward response, the electromagnetic torque becomes positive during motoring operation and negative during regenerative braking. The negative torque region

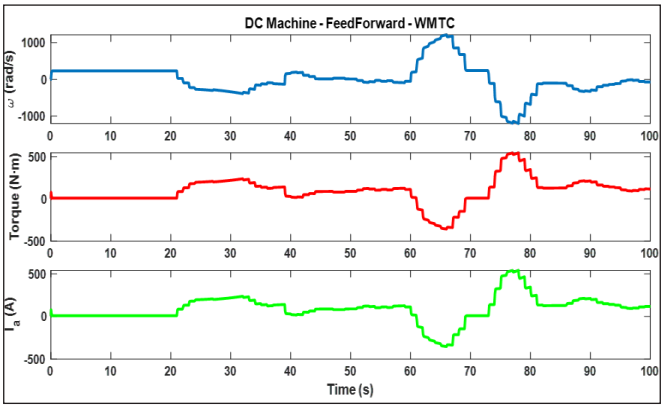


Figure 16. DC machine response using feedforward control under the WMTC profile

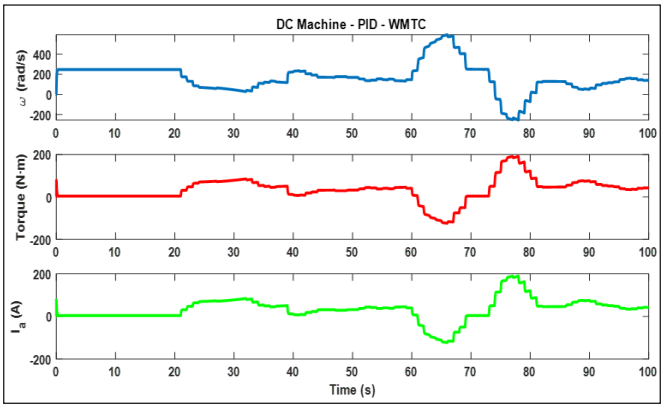


Figure 17. DC machine response using PID control under the WMTC profile

shows that the DC machine opposes the wheel motion and operates as a generator during braking. The armature current follows the torque behaviour, where the change in current direction supports the transition from motoring to regenerative operation.

In the PID response, the electromagnetic torque and armature current show a clearer transition between motoring and regenerative braking intervals. The negative torque and corresponding current reversal confirm generator operation during regenerative braking. Compared with feedforward control, PID provides a more effective regenerative response, which agrees with the higher recovered energy and lower SOC depletion reported in Table 5.

The DC machine results are used to verify the direction of torque and current flow during regenerative braking. The transient magnitudes shown in the simulation are interpreted as model-response indicators for controller comparison, while detailed motor operating limits, torque-speed constraints, current saturation, and hardware validation are outside the scope of the present study.

Integrated Performance Assessment

The results demonstrate that the proposed regenerative braking system can recover part of the mechanical energy available during downhill and deceleration conditions. The force response shows a net force of approximately −60 N during downhill operation, supporting the activation of regenerative braking. This behaviour is also reflected in the SOC response, where the battery SOC temporarily increases from about 84.5% to 85.3% between approximately 60 s and 80 s. This confirms that part of the braking energy is converted into electrical energy and returned to the battery.

The comparison between feedforward and PID control shows that PID provides better speed regulation and energy recovery. The PID controller reduced the RMSE from 14.10 m/s to 11.546 m/s, representing an 18.1% reduction in tracking error. The PID controller also increased recovered energy from 1.00×10^5 J to 2.20×10^5 J, while SOC depletion improved from −13.0% to −10.2%.

Overall, the results confirm that regenerative braking can reduce battery energy depletion during downhill and deceleration operations. The revised findings support energy recovery and SOC improvement, but they do not claim a directly quantified driving-range extension because no no-regeneration baseline, Wh/km calculation, estimated range, or percentage range improvement was included. Therefore, the present study demonstrates the energy-recovery potential of regenerative braking for electric motorcycles in hilly terrain, while direct quantification of range should be addressed in future work.

CONCLUSION

This study investigated regenerative braking as an energy-recovery strategy for electric motorcycles operating in hilly terrain. The developed model represented the transition

between motoring and regenerative braking modes by integrating longitudinal vehicle dynamics, motor-generator operation, battery response, H-bridge switching, and control operation under the WMTC driving profile.

The findings confirm that regenerative braking can return useful energy to the battery during deceleration and downhill operation. This was shown through reversed battery current, opposing electromagnetic torque, temporary SOC recovery, and reduced battery charge depletion. The comparison between feedforward and PID control also showed that controller selection directly influences both speed response and energy recovery performance. In the present simulation model, PID control outperformed feedforward control, as evidenced by lower RMSE, higher recovered energy, higher regeneration fraction, and lower SOC depletion.

The study is limited to a modelling-based evaluation, and experimental validation on a real electric motorcycle platform was not conducted. The model also focuses on longitudinal motion and does not include detailed tyre slip, rider posture variation, thermal effects, battery ageing, mechanical brake blending, hardware derating, or component-level operating limits. In addition, a no-regeneration baseline was not included; therefore, the study does not report Wh/km energy consumption, estimated driving range, or percentage range improvement.

Overall, the results show that regenerative braking can support energy recovery and reduce battery depletion in electric motorcycles operating in hilly terrain. Future work should include hardware-based validation, integration of regenerative and mechanical braking, detailed component sizing, and direct quantification of range using a no-regeneration baseline across different slope, speed, and rider-demand conditions.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Ministry of Higher Education Malaysia for supporting this research under the Fundamental Research Grant Scheme, grant number FRGS/1/2023/TK02/UIAM/01/1, entitled “Optimal Deceleration Formulation Based on Reinforcement Deep Learning for Energy Generation from Regenerative Braking Mode of an Electric Motorcycle.” The authors also acknowledge the support of the Kulliyyah of Engineering, International Islamic University Malaysia, through the KOE IIUM Engineering Merit Scholarship 2023.

REFERENCES

Aduama, P., Al-Sumaiti, A. S., & Al-Hosani, K. H. (2023). Electric vehicle charging infrastructure and energy resources: A review. *Energies, 16*(4), Article 1965. https://doi.org/10.3390/en16041965

Badran, M. A., & Toha, S. F. (2024). Employment of artificial intelligence (AI) techniques in battery management system (BMS) for electric vehicles (EV): Issues and challenges. *Pertanika Journal of Science and Technology, 32*(2), 859-881. https://doi.org/10.47836/pjst.32.2.20

Ghazali, A. K., Hassan, M. K., Radzi, M. A. M., & As'arry, A. (2020). Integrated braking force distribution for electric vehicle regenerative braking system. *Pertanika Journal of Science and Technology, 28*(Special issue 2), 173-182. https://doi.org/10.47836/pjst.28.S2.14

Heydari, S., Fajri, P., Rasheduzzaman, M., & Sabzehgar, R. (2019). Maximizing regenerative braking energy recovery of electric vehicles through dynamic low-speed cutoff point detection. *IEEE Transactions on Transportation Electrification, 5*(1), 262-270. https://doi.org/10.1109/TTE.2019.2894942

Heydari, S., Fajri, P., Sabzehgar, R., & Asrari, A. (2020). Optimal brake allocation in electric vehicles for maximizing energy harvesting during braking. *IEEE Transactions on Energy Conversion, 35*(4), 1806-1814. https://doi.org/10.1109/TEC.2020.2994520

International Energy Agency. (2022). *Global EV outlook 2022: Securing supplies for an electric future*. https://www.iea.org/reports/global-ev-outlook-2022

Joseph Godfrey, A., & Sankaranarayanan, V. (2018). A new electric braking system with energy regeneration for a BLDC motor driven electric vehicle. *Engineering Science and Technology, an International Journal, 21*(4), 704-713. https://doi.org/10.1016/j.jestch.2018.05.003

Jung, S., & Ko, J. (2021). Study on regenerative energy recovery of electric vehicle through voltage control using switched capacitor. *IEEE Transactions on Vehicular Technology, 70*(5), 4324-4339. https://doi.org/10.1109/TVT.2021.3073148

Kadav, P., & Asher, Z. D. (2019). Improving the range of electric vehicles. In *2019 Electric Vehicles International Conference (EV)* (pp. 1-5). IEEE. https://doi.org/10.1109/EV.2019.8892929

Karyatanti, I. D. P., Dewantara, B. Y., Winarno, I., Rahmatullah, D., & Nuhman. (2025). Energy management systems development in electric vehicle charging stations based on multi-source inverters with utilisation of renewable energy sources. *Pertanika Journal of Science and Technology, 33*(3), 1595-1612. https://doi.org/10.47836/pjst.33.3.22

Madichetty, S., Mishra, S., & Basu, M. (2021). New trends in electric motors and selection for electric vehicle propulsion systems. *IET Electrical Systems in Transportation, 11*(3), 186-199. https://doi.org/10.1049/els2.12018

Mohamed, M. A., Toha, S. F., Abdullah, M., Ahmad, S., Nor, K. A. M., Hassan, M. H., & Idris, A. S. (2025). Enhanced voltage regulation of buck converter-fed DC motors using fuzzy logic control under dynamic load conditions. *International Journal of Robotics and Control Systems, 5*(2), 1422-1440. https://doi.org/10.31763/ijrcs.v5i2.1836

Mohamed, M. A., Toha, S. F., Rahman, M. A., & Khairudin, M. (2025). Smart IoT energy optimisation and localisation monitoring for e-bike sharing. *IIUM Engineering Journal, 26*(2), 305-325. https://doi.org/10.31436/iiumej.v26i2.3448

Nama, T., Mondal, P., Tripathy, P., Adda, R., & Gogoi, A. K. (2022). Design, modeling and hardware implementation of regenerative braking for electric two-wheelers for hilly roads. *IEEE Access, 10*, 130602-130618. https://doi.org/10.1109/ACCESS.2022.3229597

Naseri, F., Farjah, E., & Ghanbari, T. (2017). An efficient regenerative braking system based on battery/supercapacitor for electric, hybrid, and plug-in hybrid electric vehicles with BLDC motor. *IEEE Transactions on Vehicular Technology, 66*(5), 3724-3738. https://doi.org/10.1109/TVT.2016.2611655

Podder, A. K., Chakraborty, O., Islam, S., Kumar, N. M., & Alhelou, H. H. (2021). Control strategies of different hybrid energy storage systems for electric vehicles applications. *IEEE Access, 9*, 51865-51895. https://doi.org/10.1109/ACCESS.2021.3069593

Rego, N., & Castro, R. (2023). Regenerative braking applied to a student team's electric racing motorcycle prototype: A theoretical study. *Applied Sciences, 13*(6), Article 3784. https://doi.org/10.3390/app13063784

Riczu, C., & Bauman, J. (2021). Implementation and system-level modeling of a hardware efficient cell balancing circuit for electric vehicle range extension. *IEEE Transactions on Industry Applications, 57*(3), 2883–2895. https://doi.org/10.1109/TIA.2021.3067300

Xie, Y., Wang, C., Hu, X., Lin, X., Zhang, Y., & Li, W. (2020). An MPC-based control strategy for electric vehicle battery cooling considering energy saving and battery lifespan. *IEEE Transactions on Vehicular Technology, 69*(12), 14657-14673. https://doi.org/10.1109/TVT.2020.3032989

Yang, C. (2022). Running battery electric vehicles with extended range: Coupling cost and energy analysis. *Applied Energy, 306*, Article 118116. https://doi.org/10.1016/j.apenergy.2021.118116

Yeo, W., Jung, S., Kim, S., Park, K., & Ko, J. (2020). A regenerative energy recovery system for electric vehicles charging a battery at a low speed. *Advances in Science, Technology and Engineering Systems, 5*(2), 64-73. https://doi.org/10.25046/aj050208

Pertanika J. Sci. & Technol. 34 (5): 2765 - 2790 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

An Integrated Multi-Source Framework for Explainable Android Malware Detection with SOC Integration

Pallavi M. Shimpi[1]* and Nitin Namdeo Pise[1]

Department of Computer Engineering and Technology, Faculty of Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Survey No. 124, Paud Road, Kothrud, 411038 Pune, Maharashtra, India

ABSTRACT

Neurosec Malware Scanner is an integrated malware-analysis framework that combines multi-source threat intelligence, static machine-learning classification, SHAP-based explainability, and Security Operations Centre (SOC)-oriented result presentation within a unified workflow. The system queries multiple file-hash threat-intelligence services in parallel and consolidates their outputs using a weighted aggregation mechanism. In parallel, a Python-based machine-learning pipeline extracts static malware features and evaluates four supervised classifiers: Random Forest, XGBoost, Gradient Boosting, and Logistic Regression. The framework is evaluated using five publicly available benchmark datasets: SOMLAP, TUANDROMD, Drebin, EMBER, and MH-100K. The best classification results obtained in the experiments are 99.6% on SOMLAP, 99.8% on TUANDROMD, 98.9% on Drebin, 94.3% on EMBER, and 86.7% on MH-100K. To improve interpretability, the framework incorporates SHAP-based explanations that identify the features contributing to individual model predictions. The system additionally employs cryptographic hashing, result caching, structured logging, and dashboard-based visualization to support repeatable malware-triage workflows. The proposed framework demonstrates the feasibility of integrating machine-learning detection, explainability, and threat-intelligence enrichment within an analyst-oriented malware-analysis pipeline. However, validation using recent real-world malware samples and deployment in operational SOC environments remains necessary to establish production-level effectiveness.

Keywords: Android malware detection, explainable AI, machine learning, SOC integration, static analysis, threat intelligence

ARTICLE INFO

Article history:
Received: 19 June 2026
Accepted: 02 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.10

mspallavi2710@gmail.com (Pallavi M. Shimpi)
nitin.pise@mitwpu.edu.in (Nitin Namdeo Pise)
* Corresponding author

INTRODUCTION

Android has become one of the most widely used mobile operating systems worldwide, making it an important target for malicious

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

applications and cyberattacks (Arp et al., 2014; Zhou & Jiang, 2012). The increasing availability of Android applications and the rapid evolution of malware techniques have created significant challenges for conventional security mechanisms. Malware authors increasingly employ obfuscation, code modification, repackaging, and other evasion techniques to bypass signature-based and rule-based detection mechanisms (Christodorescu & Jha, 2003; You & Yim, 2010). Moreover, the limitations of static analysis under certain malware evasion conditions have been reported in previous studies (Moser et al., 2007). Consequently, effective Android malware detection requires approaches that can identify malicious patterns beyond previously known signatures.

Machine-learning techniques offer greater adaptability to emerging malware patterns than traditional signature-based approaches, as they can learn discriminative characteristics from malware-related features (Hardy et al., 2016; Yuan et al., 2014). Machine-learning-based Android malware detection has therefore received considerable attention for analysing application permissions, Application Programming Interface (API) calls, intents, opcodes, source-code characteristics, and other static or behavioural features (Arp et al., 2014; Prasad et al., 2025; Santos et al., 2013; Yuan et al., 2014). Deep learning approaches have further demonstrated their ability to learn complex representations from high-dimensional malware data (Hardy et al., 2016; Yuan et al., 2014). However, many machine-learning and deep-learning models operate as black-box systems, which can limit analysts' understanding of the factors underlying individual predictions and may reduce confidence in automated malware detection results (Galli et al., 2024; Rjoub et al., 2023; Shimpi & Pise, 2026). A model that provides only a malicious or benign prediction may therefore be insufficient for security analysts who need to understand why an application has been classified as malicious.

This limitation motivates the use of Explainable Artificial Intelligence (XAI) techniques in malware analysis. XAI methods can provide information about the features that contribute to individual model predictions and can consequently improve the interpretability of machine-learning-based detection systems (Galli et al., 2024; Lundberg & Lee, 2017; Ribeiro et al., 2016; Rjoub et al., 2023). Lundberg and Lee (2017) introduced Shapley Additive exPlanations (SHAP), which provides feature-attribution information for explaining model predictions, while Ribeiro et al. (2016) proposed LIME as a model-agnostic approach for interpreting individual predictions. XAI has subsequently received increasing attention in cybersecurity because interpretable predictions can help analysts understand and validate automated security decisions (Rjoub et al., 2023). Recent research has also investigated explainability in AI-based malware detection systems (Galli et al., 2024). In Android malware analysis, feature-level explanations can help analysts identify permissions, API-related characteristics, source-code attributes, and other application features that contribute to a suspicious classification (Arp et al., 2014; Prasad et al., 2025;

Shimpi & Pise, 2026). However, explainability alone does not provide complete contextual information about the identified application, associated indicators, or potentially malicious infrastructure.

Machine-learning predictions can identify suspicious applications, but classification results alone may provide limited information for subsequent security investigation. Cyber Threat Intelligence (CTI) can complement automated detection by providing reputation information, indicators of compromise, and contextual information associated with previously observed malicious activities (Sun et al., 2023). Threat intelligence information can therefore provide additional evidence to help analysts assess the relevance and severity of a detected threat. However, threat-intelligence services are frequently used as independent sources of information or lookup data rather than being systematically integrated with machine-learning predictions and model explanations within a unified malware-analysis workflow. This creates an integration challenge between automated detection, interpretable decision-making, and contextual threat intelligence.

The practical importance of such integration is particularly evident in Security Operations Centre (SOC) environments, where analysts must process multiple alerts and correlate evidence from different security sources. Existing Android malware detection approaches have primarily concentrated on classification performance (Arp et al., 2014; Prasad et al., 2025; Santos et al., 2013; Yuan et al., 2014), while XAI-oriented studies have focused mainly on improving model transparency (Galli et al., 2024; Lundberg & Lee, 2017; Ribeiro et al., 2016; Rjoub et al., 2023; Shimpi & Pise, 2026). Similarly, threat-intelligence research has emphasised the collection, mining, and utilisation of security intelligence (Sun et al., 2023). These research directions provide important capabilities, but they are not necessarily integrated into a single Android malware-analysis workflow that connects a model prediction with its feature-level explanation and external threat-intelligence evidence.

Our previous research has investigated Android malware detection from complementary machine-learning perspectives. An adaptive swarm optimisation-based neural network was developed for Android malware detection (Shimpi & Pise, 2023b). This work was followed by an empirical investigation of machine-learning-based malware detection on Android mobile phones (Shimpi & Pise, 2023a). More recently, an optimised and explainable attentional convolutional LSTM model was proposed to improve Android malware detection while providing model interpretability (Shimpi & Pise, 2026). These studies establish a progression from conventional machine-learning-based detection toward optimised and explainable deep-learning models. The present study extends this research direction beyond an individual detection model by integrating classification, explainability, threat-intelligence correlation, and operational presentation into a unified SOC-oriented analytical framework.

Based on the above limitations, the key research gap this study addresses is the limited integration of automated malware classification, feature-level explainability, multi-source threat-intelligence enrichment, and SOC-oriented operational visualisation within a single Android malware-analysis workflow. The proposed work does not claim novelty from independently introducing machine learning, SHAP, or threat intelligence, since these technologies have been extensively investigated (Galli et al., 2024; Lundberg & Lee, 2017; Ribeiro et al., 2016; Rjoub et al., 2023; Sun et al., 2023). Instead, the contribution lies in integrating these established capabilities so that an analyst can move from automated malware prediction to an explanation of the prediction, and subsequently to contextual threat-intelligence evidence, within the same analytical process.

In this study, a unified SOC-oriented malware analysis framework is developed by integrating static application analysis, machine-learning-based malware classification, SHAP-based explainability, multi-source threat intelligence correlation, and operational visualisation. The framework first extracts relevant application characteristics and applies machine-learning models to classify potentially malicious applications. The prediction results are subsequently complemented by feature-level SHAP explanations and information from external threat intelligence sources. The resulting information is presented through an operational workflow intended to support analyst-oriented malware assessment and investigation.

The framework is evaluated on multiple benchmark datasets that represent heterogeneous malware feature spaces. The primary evaluation focuses on Android malware datasets, consistent with the research's main objective. The EMBER Windows Portable Executable (PE) dataset is used only as an auxiliary cross-platform evaluation environment (Anderson & Roth, 2018). The inclusion of EMBER is intended to examine whether the analytical architecture can operate across different malware representations rather than to claim that Windows malware and Android malware constitute equivalent detection tasks. Because the feature spaces and operating-system characteristics differ, Android and Windows results are analysed and reported separately. This design allows the Android experiments to remain the primary evidence for the research objective while providing additional insight into the architectural adaptability of the proposed framework.

The main contributions of this research are a unified SOC-oriented malware-analysis framework that integrates static analysis, machine-learning classification, SHAP-based explainability, external threat-intelligence correlation, and operational visualisation into a common analytical workflow.

An explainability-driven analysis mechanism is incorporated to provide feature-level information underlying malware predictions, enabling analysts to examine the factors contributing to individual classification decisions.

A multi-source threat-intelligence aggregation mechanism is integrated to supplement machine-learning predictions with external reputation information and security intelligence, thereby providing additional contextual evidence for malware investigation.

A cross-dataset evaluation strategy is employed to examine the proposed analytical pipeline across heterogeneous malware datasets, with Android datasets representing the primary evaluation environment and the EMBER Windows PE dataset providing auxiliary cross-platform validation (Anderson & Roth, 2018).

The framework is designed as an analyst-oriented decision-support pipeline rather than solely as a classification model, connecting automated detection results with interpretable evidence and contextual threat intelligence. Therefore, the novelty of the proposed work does not primarily arise from introducing a new standalone machine-learning classifier, SHAP method, or threat-intelligence service. Instead, it lies in integrating these capabilities into a unified malware-analysis workflow that connects automated detection with feature-level explanation and actionable security intelligence. This integration provides a practical approach to malware triage by enabling analysts to progress from an automated prediction to an explanation of that prediction, and subsequently to supporting threat intelligence evidence, within the same analytical process.

RELATED WORK

Signature-Based and Static/Dynamic Malware Analysis

Traditional malware detection systems have predominantly relied on signatures and predefined patterns to identify known malicious software. Signature-based techniques are effective for previously identified malware because they can provide accurate detection with relatively low computational requirements (Christodorescu & Jha, 2003). However, such approaches are less effective against malware that employs code modification, obfuscation, polymorphism, or other evasion techniques. Previous studies have demonstrated that even relatively small modifications to malicious code can reduce the effectiveness of purely signature-based detection approaches (Moser et al., 2007; You & Yim, 2010).

Malware analysis techniques can broadly be categorised into static and dynamic approaches. Static analysis examines the characteristics of an application or executable without executing the potentially malicious program. Such characteristics may include permissions, API calls, strings, structural properties, and opcode-related information. Static analysis can provide relatively efficient and scalable malware screening because the suspicious application need not be executed (Arp et al., 2014; Christodorescu & Jha, 2003). However, its effectiveness may be affected when malware deliberately modifies or hides its static characteristics (Moser et al., 2007).

Dynamic analysis, in contrast, observes malware behaviour during execution, including system calls, API interactions, network communication, and other runtime activities.

Dynamic analysis can provide behavioural information that may not be available through static inspection alone (Egele et al., 2012). However, execution-based analysis generally requires additional computational resources and controlled environments such as sandboxes. Consequently, static analysis remains attractive for large-scale malware screening, while hybrid approaches can combine static and dynamic characteristics when computational resources permit.

For Android malware detection, static analysis has been particularly important because application permissions, API calls, intents, and other application-level characteristics can be extracted without executing the application. DREBIN demonstrated the effectiveness of combining multiple lightweight static features for Android malware detection (Arp et al., 2014). Zhou and Jiang (2012) also provided an extensive characterisation of Android malware and its evolution, establishing the importance of application-level characteristics in identifying malicious Android applications.

Machine-Learning-Based Android Malware Detection

Machine-learning approaches have been increasingly adopted to overcome some limitations of signature-based malware detection. Rather than relying exclusively on predefined signatures, machine-learning models can learn discriminative patterns from extracted malware features and subsequently classify previously unseen samples. DREBIN demonstrated that supervised machine learning could effectively detect Android malware using static features derived from applications (Arp et al., 2014). Yuan et al. (2014) proposed DroidSec and demonstrated the application of deep learning to Android malware detection. Deep learning approaches can automatically learn complex feature representations and may reduce dependence on manually designed detection rules. Opcode-based representations have also been investigated for machine-learning-based malware detection. Santos et al. (2013) demonstrated that opcode sequences can provide useful features for identifying unknown malware. Such research indicates that different representations of executable or application characteristics can contribute to automated malware classification. The authors' previous research has also examined several machine-learning approaches for Android malware detection. Shimpi and Pise (2023b) proposed an adaptive swarm optimisation-based neural network for Android malware detection. An empirical investigation of machine learning techniques for Android mobile malware detection was subsequently presented in Shimpi & Pise (2023a). More recently, an optimised and explainable attentional convolutional LSTM model was proposed for Android malware detection (Shimpi & Pise, 2026). These studies demonstrate a progression from conventional machine-learning-based classification toward optimised and explainable deep-learning approaches. Prasad et al. (2025) proposed AndroMD, an Android malware detection framework based on source-code analysis and permission scanning. This work further demonstrates the value of combining multiple application-level characteristics for Android malware detection. Importantly,

AndroMD should be cited as Prasad et al. (2025), rather than You and Yim (2010). You and Yim (2010) address malware obfuscation techniques and should not be cited as the source of the AndroMD framework. Although these machine-learning approaches demonstrate strong potential for malware classification, classification performance alone does not necessarily provide sufficient information for security analysts. In operational environments, analysts may also require an explanation of why a particular application was classified as malicious, as well as additional contextual information to support the investigation.

Explainable Artificial Intelligence for Malware Detection

The increasing use of machine-learning and deep-learning models in cybersecurity has created a need for interpretable decision-making. Complex models may achieve high classification performance while offering limited information about the factors underlying individual predictions. This black-box characteristic can make it difficult for analysts to validate automated decisions and investigate suspicious applications.

SHapley Additive exPlanations (SHAP) provide a feature-attribution mechanism for explaining individual model predictions (Lundberg & Lee, 2017). SHAP assigns contribution values to input features, allowing analysts to examine whether individual features increase or decrease the predicted probability of a particular class. LIME provides another model-agnostic approach for generating local explanations of machine-learning predictions (Ribeiro et al., 2016).

XAI has increasingly been investigated in cybersecurity because interpretable model outputs can help analysts understand automated security decisions. Rjoub et al. (2023) reviewed XAI approaches for cybersecurity and highlighted the importance of interpretability for security applications. Galli et al. (2024) further investigated explainability in AI-based behavioural malware detection systems and demonstrated the relevance of feature-level explanations in malware analysis. For Android malware detection, XAI can provide additional information about the application characteristics contributing to a malicious prediction. Such explanations may include contributions from permissions, API-related features, code characteristics, or other static attributes (Arp et al., 2014; Prasad et al., 2025; Shimpi & Pise, 2026). However, an explanation of a model prediction does not, by itself, provide information about the application's reputation, associated indicators of compromise, or related malicious infrastructure. Therefore, XAI addresses the interpretability problem but does not completely address the contextual requirements of malware investigation.

Cyber Threat Intelligence for Malware Investigation

Cyber Threat Intelligence (CTI) provides contextual information that can complement automated malware detection. CTI may include indicators of compromise, reputation

information, relationships between malicious entities, previously observed attack information, and other security intelligence. Sun et al. (2023) surveyed cyber threat intelligence mining and highlighted the importance of intelligence-driven approaches for proactive cybersecurity defence.

In a malware analysis workflow, external threat intelligence can provide additional evidence once an application or file has been identified as suspicious. For example, reputation information or previously observed indicators can help analysts determine whether a detected sample is associated with known malicious activity. Such contextual information can complement the output of machine-learning classifiers. However, threat intelligence and machine-learning detection are frequently treated as separate analytical functions. A machine-learning model may produce a probability or class label, while threat-intelligence platforms may separately provide reputation or indicator information. Similarly, XAI systems may explain the model decision without incorporating external intelligence. Consequently, the outputs of these components may remain disconnected throughout the investigation. This distinction is important for the present study. The contribution is not the introduction of a new threat-intelligence service or a new XAI algorithm. Instead, the proposed framework integrates external threat intelligence with machine-learning predictions and SHAP-based explanations, allowing the three sources of evidence to be considered within a common malware-analysis workflow.

SOC-Oriented Malware Analysis and Research Gap

SOC requires analysts to process and correlate information from multiple security sources. From this operational perspective, malware detection accuracy alone is not sufficient to support a complete investigation. Analysts may need to determine whether an application is suspicious, understand the features responsible for the prediction, obtain contextual intelligence about the sample, and review the combined information through an operational interface. Existing research provides important components of this process. Android malware studies have investigated static analysis and machine-learning-based classification (Arp et al., 2014; Prasad et al., 2025; Santos et al., 2013; Shimpi & Pise, 2023a; Shimpi & Pise, 2023b; Shimpi & Pise, 2026; Yuan et al., 2014; Zhou & Jiang, 2012). XAI research has introduced methods for interpreting machine-learning decisions (Galli et al., 2024; Lundberg & Lee, 2017; Ribeiro et al., 2016; Rjoub et al., 2023). CTI research has established mechanisms for collecting and analysing external security intelligence (Sun et al., 2023). However, these capabilities generally address different stages or objectives of malware analysis. The specific research gap addressed by this study is, therefore, the limited integration of these capabilities into a unified Android malware analysis workflow. The suggested architecture consolidates five essential elements into a cohesive framework: static application analysis and feature extraction, machine-learning-driven

malware classification, feature-level elucidation of specific predictions, multi-source threat intelligence augmentation, and Security Operations Center (SOC)-focused visualisation of the aggregated evidence. Collectively, these elements facilitate malware identification, enhance the clarity of classification outcomes, augment the analysis with pertinent threat intelligence, and deliver the consolidated evidence in a format appropriate for security analysts.

The proposed framework addresses this gap by connecting these functions within a common analytical pipeline. The objective is not to claim that machine learning, SHAP, static analysis, or threat intelligence are individually novel. Instead, the proposed contribution lies in their integration and operational sequencing for analyst-oriented Android malware investigation.

Comparison with Existing Approaches

Table 1 summarises the functional differences between representative studies and the proposed framework.

Table 1
Functional comparison of existing malware-analysis approaches.

Approach / Study	Malware Analysis	Machine Learning	Explainability / XAI	Threat Intelligence	SOC-Oriented Integration	Primary Focus
Christodorescu and Jha (2003)	Static analysis	No	No	No	No	Malicious code pattern detection
Moser et al. (2007)	Static analysis	No	No	No	No	Limitations of static malware analysis
Bayer et al. (2009)	Behavioural analysis	Yes	No	No	No	Malware clustering
Zhou and Jiang (2012)	Static and behavioural analysis	Limited	No	No	No	Android malware characterisation
Arp et al. (2014)	Static analysis	Yes	Limited	No	No	Android malware detection
Yuan et al. (2014)	Android application analysis	Deep learning	No	No	No	Android malware classification
Hardy et al. (2016)	Malware analysis	Deep learning	No	No	No	Intelligent malware detection
Ribeiro et al. (2016)	General classification	Yes	Yes	No	No	Model-level explanation

Table 1 (*continued*)

Approach / Study	Malware Analysis	Machine Learning	Explainability / XAI	Threat Intelligence	SOC-Oriented Integration	Primary Focus
Lundberg and Lee (2017)	General ML analysis	Yes	Yes (SHAP)	No	No	Feature-level model explanation
Ronen et al. (2018)	Static PE analysis	Yes	No	No	No	Windows malware classification
Prasad et al. (2025)	Android static analysis	Yes	No	No	No	Source-code and permission-based Android malware detection
Galli et al. (2024)	Behavioural malware analysis	Yes	Yes	No	No	Explainable malware detection
Sun et al. (2023)	Cyber threat intelligence	No	No	Yes	Partially	Threat-intelligence mining and cybersecurity
Proposed framework	Static Android analysis	Yes	Yes (SHAP)	Yes, multi-source	Yes	Integrated SOC-oriented Android malware analysis

Previous studies focus on different features of malware detection, machine learning, explainability, and threat intelligence, as summarised in Table 1. This comparison was produced by the authors based on the relevant literature to highlight the functional and architectural differences of the existing approaches and the proposed framework. We present a framework that combines a machine-learning classification, feature-level SHAP explanation, and multi-source threat-intelligence enrichment in an analyst-centric workflow. This comparison shows the combination of various capabilities, rather than claiming superiority to any single malware detector. Benchmark datasets and standard assessment metrics are used to evaluate the quantitative classification performance.

Quantitative Positioning Against Existing Android Malware Studies

To further situate the proposed framework within the Android malware detection literature, quantitative results should be compared with studies that use comparable datasets and evaluation metrics. Table 2 should therefore be populated using performance values reported in the original publications.

Table 2
Quantitative comparison with representative Android malware detection studies.

Study	Dataset	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC (%)	XAI	SOC Integration
Arp et al. (2014) – DREBIN	DREBIN	Linear SVM	—	—	94.0	—	—	No	No
Alkahtani and Aldhyani (2022)	DREBIN	SVM	80.71	—	—	—	—	No	No
Alkahtani and Aldhyani (2022)	DREBIN	KNN	81.57	—	—	—	—	No	No
Aamir et al. (2024) – AMDDL model	DREBIN	CNN	99.92	98.61	99.16	98.88	—	No	No
Palma et al. (2024)	DREBIN	Best ML model	97.47	—	—	—	—	Yes	No
Zaidi et al. (2025)	DREBIN	CNN + XGBoost	98.76	—	—	—	—	No	No
Proposed framework	DREBIN	Best classifier: XGBoost	98.9	98.6	98.4	98.5	99.6	Yes (SHAP)	Yes
Proposed framework	TUANDROMD	Random Forest	99.8	99.7	99.7	99.8	99.8	Yes (SHAP)	Yes

Because the studies may differ in datasets, feature representations, preprocessing procedures, train-test splits, and evaluation protocols, direct numerical comparisons should be interpreted with caution. The primary purpose of Table 2 is to position the proposed framework against comparable Android malware detection approaches, rather than to claim superiority solely based on differences in reported accuracy. Where the same dataset and evaluation protocol are available, direct comparison is more meaningful. For studies using different datasets, the results should be presented as contextual evidence rather than as a strict ranking of algorithms.

Positioning of the Proposed Study

The literature reviewed above demonstrates three major research directions. The first focuses on improving malware detection using static analysis, machine learning, and deep learning (Arp et al., 2014; Hardy et al., 2016; Prasad et al., 2025; Santos et al., 2013; Shimpi & Pise, 2023a; Shimpi & Pise, 2023b; Shimpi & Pise, 2026; Yuan et al., 2014; Zhou & Jiang, 2012). The second focuses on improving model transparency through XAI techniques such as SHAP and LIME (Galli et al., 2024; Lundberg & Lee, 2017; Ribeiro et al., 2016; Rjoub et al., 2023). The third focuses on the collection and utilisation of cyber threat intelligence to provide contextual security information (Sun et al., 2023).

The proposed framework builds upon these established research directions by integrating them into a common analytical architecture. Its primary objective is Android malware detection, and therefore, the Android datasets constitute the principal evaluation environment. The EMBER Windows PE dataset is used only as an auxiliary cross-platform validation environment (Anderson & Roth, 2018). The Windows dataset is not treated as an Android malware dataset, and results from different operating systems are not combined into a single performance measure. This distinction is necessary because Android and Windows malware datasets differ in operating-system architecture, feature representation, application structure, and malware characteristics. Android datasets may contain permissions, API usage, intents, and application-level characteristics, whereas EMBER represents Windows Portable Executable characteristics. Therefore, the cross-platform experiments are interpreted as evidence of architectural adaptability rather than direct evidence that Android and Windows detection tasks are equivalent. Overall, the existing literature demonstrates that machine learning can provide automated malware classification, XAI can improve interpretability, and CTI can provide contextual security intelligence. The specific contribution of the proposed study is to integrate these capabilities into a single SOC-oriented Android malware analysis workflow. The framework, therefore, seeks to bridge the gap between automated classification and the practical evidence required for malware triage and investigation.

PROPOSED FRAMEWORK

The proposed framework is a modular system structure that contains five major layers of operational functionality, with each layer serving a specific function:

Frontend Layer	:	A dashboard for the storage of SOC datasets with operational capabilities for uploading files from external sources and visual representation of the dataset (Web-based SOC dashboard).
Backend Layer	:	Serves as the interface for file processing, coordination of APIs, and aggregated results.
Machine Learning Engine	:	Implements a Python-based machine learning classification engine pipeline.
Threat Intelligence Aggregation	:	Provides an interface to integrate results from external API sources.
Storage Database Layer	:	Provides an internal file hashing database system for caching and hashing the processing results of uploaded files.

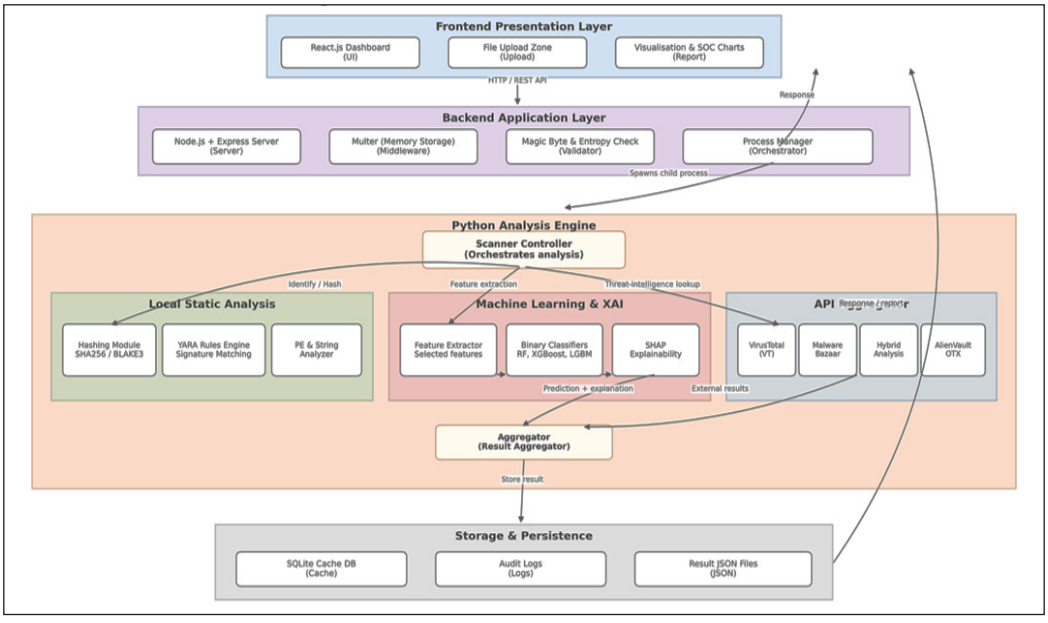


Figure 1. Proposed SOC-oriented malware analysis framework architecture

All uploaded files undergo hashing, static analysis, machine-learning inference, and threat-intelligence correlation before being aggregated into final processed results. Figure 1

depicts the operational interaction among the frontend dashboard, the backend orchestration layer, and the machine-learning Python analysis engine. Each of the three components performs either a static analysis, machine learning-based probabilistic inference, or threat intelligence correlation as part of the overall analysis pipeline before the aggregated results are output for visual representation in a SOC environment.

The novelty of the proposed framework lies primarily in its integrated operational architecture rather than in introducing a new standalone machine-learning algorithm. Existing Android malware detection studies have predominantly focused on improving classification performance, while other studies have investigated explainability or external threat intelligence as separate capabilities. In contrast, the proposed framework establishes a sequential analytical workflow that connects static malware analysis, machine-learning-based classification, feature-level SHAP explanations, multi-source threat-intelligence enrichment, and SOC-oriented visualization. This design enables an analyst to progress from automated malware prediction to an explanation of the contributing features and subsequently to contextual threat-intelligence evidence within the same workflow. Therefore, the framework is intended as an analyst-oriented decision-support architecture that complements automated classification with interpretable and contextual security information.

METHODOLOGY

System Architecture

The overall framework is designed with a modular architecture comprising a backend built in Python and a frontend built in React. The backend components include static analysis, feature preprocessing, model training, explainability, and multi-source threat intelligence enhancement. The modules are coordinated using a centralised Multiengine Scanner controller. The frontend communicates with the backend via a Node.js API (analyze.js). The API returns results only after local data processing is complete; streaming results back to the frontend will not be implemented at this time. In the early stages of integration, external API calls were embedded directly in the enrichment pipeline. If any one of these external APIs were to fail, the entire aggregation process would fail. By isolating the responses from the external APIs, we can contribute to a final threat score if at least one engine successfully processes the data. To further reduce the likelihood of receiving slow responses from any of the external APIs, server session durations have been increased, and redundant calls are minimized by introducing a caching mechanism.

Static Analysis Layer

The system is designed to perform static analysis, ensuring that files are never executed and thereby reducing operational risks during large-scale screening. The analysis includes

generating SHA-256 cryptographic hashes for file identification, applying YARA-based pattern matching to detect known suspicious patterns, and examining file structures and internal strings to identify potentially malicious characteristics. Additionally, file entropy and metadata are analysed to identify unusual patterns that may indicate malicious activity. The Shannon entropy of a given byte sequence is calculated using Equation 1, which measures the uncertainty in the byte distribution.

$$H = -\sum_{x} p\,(x) \log_2 p\,(x) \qquad [1]$$

where $p\,(x)$ represents the relative frequency of byte value \(x\) in the complete byte sequence. The calculated entropy value is then normalized to the range [0, 1] to maintain consistency in feature scaling across datasets. Additional static features are extracted to support malware analysis. These include checking for the presence of PE headers, such as the MZ header, identifying suspicious strings, and examining the byte-level structure of files. Previous studies have demonstrated that opcode sequences and other static code characteristics can be effective features for machine-learning-based malware detection (Santos et al., 2013). A binary feature is also generated to indicate the presence or absence of headers in each file. For the Android datasets, pre-extracted feature sets were used for analysis, and preprocessing was kept to a minimum. Consequently, the analysis relied primarily on the available static features, with limited scope for behavioral interpretation. Dynamic sandboxing was considered but not implemented due to hardware limitations and the need to support scalable analysis.

Feature Engineering and Preprocessing

There is a significant difference in the number of features (dimensions) between the datasets (108-2,381). To achieve consistency, numerical features were standardised using StandardScaler and fitted only to the training data, as shown in Equation 2:

$$X_{\text{scaled}} = \frac{X - \mu}{\sigma} \qquad [2]$$

Duplicate samples from each dataset were removed. Hash identifiers and metadata fields were not included to prevent data leakage into the training dataset. No dimensionality reduction was applied to the data; although PCA improved interpretability, it showed inconsistent performance across the datasets. Due to computational constraints, the data were randomly subsampled in a stratified manner for training on CPU(s) only. For example, the EMBER dataset was reduced to 50,000 records, while the MH- 100K dataset was

reduced to 10,000 records. Subsampling was done to maintain the class distribution in the data while reducing training time.

Machine Learning Models

Four distinct models were evaluated without hyperparameter optimisation to allow direct, repeatable comparison between them. A random state of 42 was used for each model.

Random Forest

The Random Forest classifier was configured with 100 decision trees, balanced class weights, and parallel processing to improve computational efficiency. These settings are represented in Equation 3.

$$N_{\text{trees}} = 100, \quad \text{class_weight} = \text{balanced}, \quad \text{n_jobs} = -1 \qquad [3]$$

Logistic Regression

The Logistic Regression classifier was configured using the LIBLINEAR solver, a regularisation parameter (C) of 1.0, a maximum of 1000 iterations, and balanced class weights to address class imbalance. The configuration settings are summarised in Table 3.

Table 3
Logistic Regression Configuration Parameters

Parameter	Value
Solver	LIBLINEAR
Regularisation parameter (\(C\))	1.0
Maximum iterations	1000
Class weights	Balanced

XGBoost

The XGBoost classifier was evaluated using the logarithmic loss (logloss) metric, while the default model parameters were retained to maintain consistency during cross-dataset modelling and avoid model-specific hyperparameter optimisation. The evaluation setting is presented in Table 4.

Table 4
XGBoost Evaluation Configuration.

Parameters	Value
Evaluation metric	Logarithmic loss (logloss)
Model parameters	Default settings

Exploratory Autoencoder

An Autoencoder was implemented using TensorFlow in a CPU-based environment as an exploratory unsupervised model for comparison with the supervised classifiers. The encoding dimension was defined as shown in Equation 4:

$$d_e = \min\left(64, \frac{d}{4}\right) \qquad [4]$$

where d denotes the input feature dimension and d_e represents the encoding dimension. The autoencoder followed a symmetric architecture, expressed as:

$$d \rightarrow 2d_e \rightarrow d_e \rightarrow 2d_e \rightarrow d$$

The hidden layers employed the Rectified Linear Unit (ReLU) activation function as shown in Equation 5:

$$f(x) = \max(0, x) \quad [5]$$

while the output layer used the sigmoid activation function as shown in Equation 6:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad [6]$$

The reconstruction error was minimised using Mean Squared Error (MSE) as shown in Equation 7:

$$\mathcal{L}_{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad [7]$$

represents the original input and $\hat{y}_i$ represents the reconstructed output. The Autoencoder was trained using the Adam optimiser with a learning rate as shown in Equation 8:

$$\eta = 0.001 \quad [8]$$

The Autoencoder was used only for exploratory analysis and comparison with the supervised classifiers and was not used as the primary malware classification model.

Model Configuration and Hyperparameter Considerations

The models were evaluated using predefined/default configurations to ensure a consistent comparison across heterogeneous datasets. Consequently, the reported results should be interpreted as comparative baseline performance rather than optimised performance for each individual classifier. No model-specific hyperparameter optimisation was performed, thereby avoiding differences in optimisation effort across classifiers and maintaining a uniform experimental protocol.

Explainability

A feature-level explanation module was developed using SHAP (Shapley Additive exPlanations) to provide interpretable explanations of model predictions. The TreeExplainer

algorithm was employed for ensemble-based models due to its computational efficiency, whereas KernelExplainer was not used for high-dimensional datasets because of the substantial computational overhead of SHAP value estimation (Lundberg & Lee, 2017).

Conduct of Experiments

All experiments were conducted using a CPU-based computing environment. The experimental pipeline was implemented in Python 3.10 on a 64-bit Linux-based environment (Ubuntu). Scikit-learn 1.3.2 was used for conventional machine-learning classifiers; XGBoost 2.0.3 was used for gradient-boosted classification; TensorFlow 2.15.0 was used for the Autoencoder implementation; and SHAP 0.44.1 was used for model explanation. The experiments were executed on an Intel(R) Xeon(R) CPU with approximately 12–13 GB of available RAM. A fixed random seed of 42 was used wherever applicable to reduce variation caused by random initialisation and data partitioning. The preprocessing pipeline, including removal of identifier-based features, scaling, feature selection where applicable, and stratified data partitioning, was applied consistently across the evaluated datasets. Training and testing data were kept separate before scaling and model evaluation to minimise the possibility of data leakage. For the EMBER dataset, the original partitions were combined and subsequently repartitioned using the same stratified evaluation procedure adopted for the other datasets. These implementation details are reported to improve the reproducibility and transparency of the experimental evaluation.

EXPERIMENTAL RESULTS

To further position the proposed framework within existing Android malware detection research, its classification performance was compared with representative studies that report results using related datasets or comparable evaluation metrics. The comparison focuses primarily on Android malware detection, as Android is the primary application domain of the proposed framework. Where studies used different datasets, preprocessing procedures, or evaluation protocols, the reported values are presented as published and are not treated as directly equivalent to the results obtained in the present study. Therefore, the comparison is intended to provide contextual positioning rather than to claim a strict one-to-one performance advantage. The classification performance of the proposed framework is compared with representative Android malware detection studies, as summarised in Table 5. The comparison considers the datasets, detection methods, reported performance metrics, explainability (XAI), and threat-intelligence or SOC integration. The proposed framework achieves an average classification accuracy of 95.9% across the evaluated datasets, while extending the detection process with SHAP-based explainability and multi-source threat-intelligence enrichment. However, the reported results should be interpreted cautiously because differences in datasets, feature representations, preprocessing, train–test splits,

and evaluation protocols may affect performance. Therefore, the comparison provides contextual positioning of the proposed framework rather than establishing a strict one-to-one performance advantage over existing methods as shown below.

Table 5
Quantitative comparison with representative Android malware detection studies

Study	Dataset	Detection Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	XAI	Threat Intelligence / SOC Integration
Arp et al. (2014) – DREBIN	DREBIN	SVM-based static analysis	—	—	94.0	—	No	No
Aamir et al. (2024) – AMDDL model	DREBIN	CNN	99.92	98.61	99.16	98.88	No	No
Alkahtani and Aldhyani (2022)	DREBIN	LSTM	99.40	99.32	99.74	99.53	No	No
Alkahtani and Aldhyani (2022)	DREBIN	CNN-LSTM	97.20	97.72	97.92	97.82	No	No
D et al. (2024) – SecuDroid	DREBIN-215	Ensemble/ ML	See source	See source	See source	See source	No	No
AndroMD (Prasad et al., 2025)	KeyCount / ZeroOne / MNF	RF + DT + Bagging ensemble	Up to 99.88	—	—	—	No	Partial*
Han et al. (2025) – LEMSOFT	DREBIN	Soft-voting ensemble	99.89	99.77	99.62	99.68	No	No
Proposed Framework	SOMLAP, TUANDROMD, DREBIN, EMBER, MH-100K	RF / XGBoost + SHAP + Threat Intelligence	95.9†	—	—	—	Yes	Yes

The comparison indicates that the proposed framework achieves competitive classification performance while extending the detection pipeline with explainability and multi-source threat-intelligence enrichment. However, direct numerical comparisons should be interpreted with caution because differences in datasets, feature representations, preprocessing procedures, train-test partitions, and evaluation protocols can substantially

affect malware detection performance. The average classification performance or accuracy across datasets is 95.9%.

Threat Intelligence Integration

Using multiple threat intelligence platforms, such as VirusTotal, AlienVault OTX, MalwareBazaar, Hybrid Analysis, ANY.RUN, MetaDefender, and AbuseIPDB provide broader behavioural and reputation-based coverage than relying on a single detection engine. Fusing the outputs of these engines into an aggregate score creates a more reliable detection process by reducing reliance on any one engine and providing real-world contextual intelligence in the final decision-making stage. Each engine's verdict, with its accompanying confidence level, is converted into normalised numeric values via a weighted aggregation method, then mapped to one of four operational threat levels (critical, high, medium, and low).

Explainable AI Integration

The SHAP explanation framework helps you understand which features were important in determining a model's outcome. The explainability layer provides an explanation of the most important factors that determine the outcome of a given event. Examples of types of explanations include (but are not limited to) entropy-based indicators, patterns of abnormality (exceeding a threshold) in the strings of characters (API endpoints), patterns of abnormality (exceeding a threshold) in the use of API types, and patterns of structural anomalies (anomalies in structure) in an executable. The design for the explainability module has been completed as a component to allow for producing an individual per-sample detailed explanation output while running the system, increasing the understanding and transparency of how a model reached its conclusion, providing more trust in the model by analysts, and aiding analysts in making decisions when evaluating cybersecurity work products.

To demonstrate the practical role of the explainability layer, SHAP analysis was used to identify the features influencing the model’s predictions. Figure 2(a) presents the confusion matrix, showing the classification outcomes for benign and malicious samples. Figure 2(b) shows the ROC curve, illustrating the model’s ability to distinguish between the two classes, while Figure 2(c) presents the Precision–Recall curve, highlighting the relationship between precision and recall. Together, these visualisations provide complementary information about the model’s classification performance. SHAP explanations further help analysts understand the contribution of individual features to a prediction. These explanations support the interpretation of model outputs but do not replace expert judgement.

Table 6 presents a summary of results demonstrating stable performance across all datasets despite varying levels of complexity.

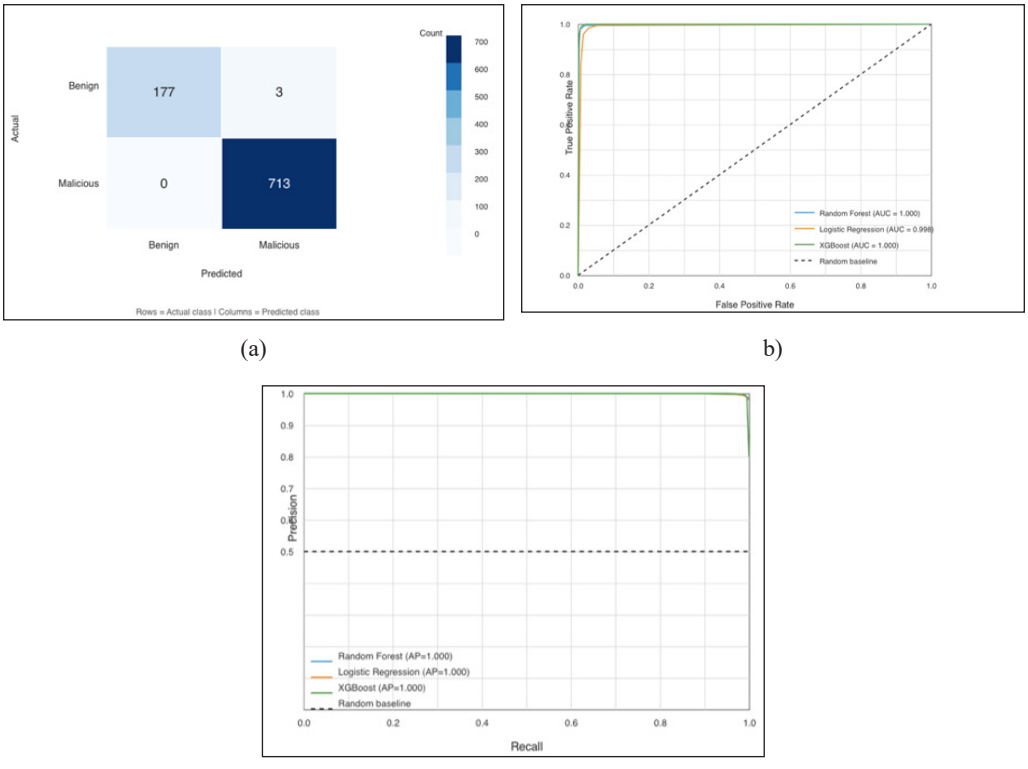


Figure 2. Representative SHAP explanation for an individual malware sample (a) Confusion Matrix, (b) ROC Curve, (c) Precision-Recall curve

Table 6
Datasets Summary

Dataset	FPR	FNR	CV F1
SOMLAP	0.4%	0.6%	0.993±0.001
TUANDROMD	0.1%	0.1%	0.994±0.004
Drebin	1.1%	1.6%	0.982±0.002
EMBER	5.7%	5.9%	0.940±0.003
MH-100K	15.1%	13.3%	0.862±0.008

DISCUSSION

The experimental findings show that the classifiers' performance varies across datasets. Instead, it largely depends on the types and characteristics of the features available in each dataset. In the present experiments, Random Forest produced the best results for SOMLAP and TUANDROMD, while XGBoost performed better on DREBIN, EMBER, and MH-100K. This variation indicates that ensemble-based models can effectively capture

nonlinear relationships among malware-related features, particularly when the underlying feature space is complex.

The results also highlight the importance of the dataset and feature representation when interpreting classification performance. The Android datasets primarily describe characteristics of mobile applications, whereas EMBER contains static features extracted from Windows Portable Executable (PE) files. These datasets, therefore, represent different malware environments and feature spaces. Consequently, the performance obtained across Android and Windows datasets should not be interpreted as a direct comparison of which operating system is easier to classify. Similar differences in malware characteristics and feature representations have also been reported in prior malware detection studies (Arp et al., 2014; Zhou & Jiang, 2012).

Another important observation is that classification performance alone may not be sufficient for an operational malware-analysis environment. A high accuracy or F1-score indicates how effectively a model performs classification, but it does not explain the reason behind an individual prediction. The use of SHAP in the proposed framework addresses this issue by identifying the features that contribute to a particular prediction (Lundberg & Lee, 2017). This information can provide analysts with additional evidence when investigating why an application has been classified as malicious or benign.

The threat-intelligence component further extends the analysis beyond the machine-learning prediction. Information obtained from external intelligence sources can provide additional context, such as malware reputation and previously observed indicators associated with an application. Cyber threat intelligence has been increasingly recognised as a useful source of contextual information for cybersecurity analysis (Sun et al., 2023). In the proposed framework, this information is used alongside the classifier output and SHAP explanation rather than being treated as an independent lookup process.

Therefore, the main contribution of the proposed framework is the integration of these complementary capabilities within a single analytical workflow. The framework enables the analysis process to move from malware classification to feature-level explanation and subsequently to contextual threat-intelligence enrichment. This design is intended to provide security analysts with more useful evidence for malware triage than a classification result alone. However, the current experimental evaluation demonstrates the framework's technical feasibility rather than its effectiveness in a production-scale SOC environment. Validation using live SOC data, real-time workloads, and analyst feedback remains an important direction for future work.

LIMITATIONS

While achieving strong classification performance, the present Application remains subject to limitations. First, the current Application is based on static analysis and therefore

would not capture the runtime behavioural patterns of Dynamic Code Loading (DCL) or Environment-Based Evasion Techniques (EBET). Second, when working with External Threat Intelligence APIs, there is a level of dependency on the External API for Access to Network Connectivity. Third, sampling for computational feasibility was performed on a variety of Large-Scale Datasets during the development of this framework, which may limit the overall scope of this framework's executability for Exhaustive Analysis. Finally, this framework was not subject to formalised benchmarking of Real-Time Performance Metrics. The Consideration and Resolution of the limitations described above will be an integral part of our immediate next steps. Although promising results have been observed, deployment in the real world could introduce various Operational Limitations. In addition, the evaluation is based primarily on publicly available benchmark datasets and may therefore not fully represent the diversity and evolving characteristics of malware encountered in operational Security Operations Centres. The current study also does not include a formal usability evaluation involving SOC analysts or deployment in a live SOC environment. Consequently, claims regarding analyst trust and operational decision support are based on the framework's functional capabilities and representative analytical outputs rather than on controlled human-subject evaluation. Future work will evaluate the framework using recent real-world malware samples, analyst-centred usability studies, large-scale threat intelligence feeds, and live or simulated SOC environments.

CONCLUSION

This study presented Neurosec Malware Scanner, an integrated framework for Android malware analysis that combines static machine-learning classification, SHAP-based explainability, multi-source threat-intelligence enrichment, and SOC-oriented result presentation. The framework was evaluated on multiple publicly available benchmark datasets, with the results demonstrating that supervised machine learning models can achieve strong classification performance across heterogeneous malware feature spaces. The principal contribution of the work is the integration of complementary analytical capabilities into a single malware-triage workflow. Machine-learning classification provides automated detection, SHAP provides feature-level interpretation, and external threat intelligence adds contextual information to assist analysts during investigations. The inclusion of caching, structured logging, hashing, and dashboard visualisation further supports repeatable analysis. However, the present study should be considered a benchmark-validated research prototype rather than a fully deployed SOC solution. The evaluation does not include a live SOC environment, a formal analyst usability study, or extensive validation using continuously evolving real-world malware. In addition, the manually specified threat-intelligence weighting scheme and fixed baseline model configurations provide opportunities for further optimisation.

Future work will therefore focus on learned threat intelligence weighting, temporal and cross-dataset validation, recent real-world Android malware samples, adversarial and obfuscation robustness, automated hyperparameter optimisation, and controlled SOC-based evaluation. These extensions will provide stronger evidence regarding the framework's generalizability, operational efficiency, and suitability for real-world cybersecurity environments.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to their institution for providing the necessary research facilities and academic support. Special thanks are extended to the research guide and colleagues for their suggestions and encouragement, which contributed to the successful completion of this work.

LIST OF ABBREVIATION

SOC	Security Operations Centre
AI	Artificial Intelligence
ML	Machine Learning
DCL	Dynamic Code Loading
EBET	Environment-Based Evasion Techniques
DCL	Dynamic Code Loading
API	Application Programming Interface (API)
CPU	Central Processing Unit (CPU)
PCA	Principal Component Analysis (PCA)
ROC	Receiver Operating Characteristic (ROC)
AUC	Area Under the Curve (AUC)
FPR	False Positive Rate (FPR)
FNR	False Negative Rate (FNR)
ReLU	Rectified Linear Unit (ReLU)
MSE	Mean Squared Error (MSE)

REFERENCES

Aamir, M., Iqbal, M. W., Nosheen, M., Ashraf, M. U., Shaf, A., Almarhabi, K. A., Alghamdi, A. M., & Bahaddad, A. A. (2024). AMDDLmodel: Android smartphones malware detection using deep learning model. *PLOS ONE*, *19*(1), Article e0296722. https://doi.org/10.1371/journal.pone.0296722

Alkahtani, H., & Aldhyani, T. H. H. (2022). Artificial intelligence algorithms for malware detection in Android-operated mobile devices. *Sensors*, *22*(6), Article 2268. https://doi.org/10.3390/s22062268

Anderson, H. S., & Roth, P. (2018). *EMBER: An open dataset for training static PE malware machine learning models* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1804.04637

Arp, D., Spreitzenbarth, M., Hübner, M., Gascon, H., & Rieck, K. (2014). DREBIN: Effective and explainable detection of Android malware in your pocket. In *Proceedings of the Network and Distributed System Security Symposium*. Internet Society. https://doi.org/10.14722/ndss.2014.23247

Bayer, U., Milani Comparetti, P., Hlauschek, C., Kruegel, C., & Kirda, E. (2009). Scalable, behavior-based malware clustering. In *Proceedings of the Network and Distributed System Security Symposium*. Internet Society. https://www.ndss-symposium.org/ndss2009/scalable-behavior-based-malware-clustering/

Christodorescu, M., & Jha, S. (2003). Static analysis of executables to detect malicious patterns. In *Proceedings of the 12th USENIX Security Symposium* (pp. 169-186). USENIX Association. https://www.usenix.org/events/sec03/tech/christodorescu.html

Egele, M., Scholte, T., Kirda, E., & Kruegel, C. (2012). A survey on automated dynamic malware-analysis techniques and tools. *ACM Computing Surveys*, *44*(2), Article 6. https://doi.org/10.1145/2089125.2089126

Galli, A., La Gatta, V., Moscato, V., Postiglione, M., & Sperlì, G. (2024). Explainability in AI-based behavioral malware detection systems. *Computers & Security*, *141*, Article 103842. https://doi.org/10.1016/j.cose.2024.103842

Han, Q., Shi, Z., Li, Y., & Zhang, T. (2025). LEMSOFT: Leveraging extraction method and soft voting for Android malware detection. *Mathematics*, *13*(21), Article 3569. https://doi.org/10.3390/math13213569

Hardy, W., Chen, L., Hou, S., Ye, Y., & Li, X. (2016). DL4MD: A deep learning framework for intelligent malware detection. In *Proceedings of the 2016 International Conference on Data Mining* (pp. 61–67). CSREA Press. https://www.worldcomp-proceedings.com/proc/proc2016/DMIN16_Final_Edition/DMIN16_Papers.pdf

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 30, pp. 4765-4774). Curran Associates. https://papers.nips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

Moser, A., Kruegel, C., & Kirda, E. (2007). Limits of static analysis for malware detection. In *23rd Annual Computer Security Applications Conference (ACSAC 2007)* (pp. 421-430). IEEE. https://doi.org/10.1109/ACSAC.2007.21

Palma, C., Ferreira, A., & Figueiredo, M. (2024). Explainable machine learning for malware detection on Android applications. *Information*, *15*(1), Article 25. https://doi.org/10.3390/info15010025

Prasad, A., Chandra, S., Alenazy, W. M., Ali, G., Shah, S., & ElAffendi, M. (2025). AndroMD: An Android malware detection framework based on source code analysis and permission scanning. *Results in Engineering*, *28*, Article 107050. https://doi.org/10.1016/j.rineng.2025.107050

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge*

Discovery and Data Mining (pp. 1135-1144). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939778

Rjoub, G., Bentahar, J., Abdel Wahab, O., Mizouni, R., Song, A., Cohen, R., Otrok, H., & Mourad, A. (2023). A survey on explainable artificial intelligence for cybersecurity. *IEEE Transactions on Network and Service Management*, *20*(4), 5115-5140. https://doi.org/10.1109/TNSM.2023.3282740

Ronen, R., Radu, M., Feuerstein, C., Yom-Tov, E., & Ahmadi, M. (2018). *Microsoft malware classification challenge* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1802.10135

Santos, I., Brezo, F., Ugarte-Pedrero, X., & Bringas, P. G. (2013). Opcode sequences as representation of executables for data-mining-based unknown malware detection. *Information Sciences*, *231*, 64-82. https://doi.org/10.1016/j.ins.2011.08.020

Shimpi, P. M., & Pise, N. N. (2023a). An empirical study of efficient malware detection analysis on Android mobile phones using machine learning. In *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 1520-1526). IEEE. https://doi.org/10.1109/ICSCDS56580.2023.10104900

Shimpi, P. M., & Pise, N. N. (2023b). Android malware detection model using the adaptive swarm optimization-based neural network. In *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1641-1648). IEEE. https://doi.org/10.1109/ICSSIT55814.2023.10061042

Shimpi, P. M., & Pise, N. N. (2026). Android malware detection using an optimized and explainable attentional convolutional LSTM model. *Journal of Intelligent Decision Making and Information Science*, *3*(2S), 212-233. https://doi.org/10.59543/jidmis.v3.639

Sun, N., Ding, M., Jiang, J., Xu, W., Mo, X., Tai, Y., & Zhang, J. (2023). Cyber threat intelligence mining for proactive cybersecurity defense: A survey and new perspectives. *IEEE Communications Surveys & Tutorials*, *25*(3), 1748-1774. https://doi.org/10.1109/COMST.2023.3273282

Vanusha, D., Singh, S., Jha, A. K., & Robinsha S, D. (2024). SecuDroid: Android malware detection using ML classifier on static features. In 2024 2nd International Conference on Networking and Communications (ICNWC) (pp. 1-9). IEEE. https://doi.org/10.1109/ICNWC60771.2024.10537417

You, I., & Yim, K. (2010). Malware obfuscation techniques: A brief survey. In *2010 International Conference on Broadband, Wireless Computing, Communication and Applications* (pp. 297-300). IEEE. https://doi.org/10.1109/BWCCA.2010.85

Yuan, Z., Lu, Y., Wang, Z., & Xue, Y. (2014). Droid-Sec: Deep learning in Android malware detection. *ACM SIGCOMM Computer Communication Review*, *44*(4), 371-372. https://doi.org/10.1145/2740070.2631434

Zaidi, A. R., Abbas, T., Daud, A., Alghushairy, O., Dawood, H., & Sarwar, N. (2025). Enhancing Android malware detection with XGBoost and convolutional neural networks. *Computers, Materials & Continua*, *84*(2), 3281-3304. https://doi.org/10.32604/cmc.2025.063646

Zhou, Y., & Jiang, X. (2012). Dissecting Android malware: Characterization and evolution. In *2012 IEEE Symposium on Security and Privacy* (pp. 95-109). IEEE. https://doi.org/10.1109/SP.2012.16

Pertanika J. Sci. & Technol. 34 (5): 2791 - 2809 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Time-Dependent Polyethene Terephthalate (PET) Biodegradation by *Aspergillus niger* and *Aspergillus oryzae* Under Solid State Fermentation

Latha Shankar[1], Norazlina Idris[1]*, Nurhafizah Ibrahim[1], Norakma Mohd Nor[1], Rodiah Mohd Hassan[1], Wan Zarina Wan Mohamed[1], Noor Fazreen Dzulkafli[2], and Nur Ain Izzati Mohd Zainudin[3]

[1]Department of Science and Biotechnology, Faculty of Engineering and Life Sciences, Universiti Selangor, Jalan Timur Tambahan, 45600 Bestari Jaya, Selangor, Malaysia
[2]Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[3]Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ABSTRACT

Polyethene terephthalate (PET) is a major plastic pollutant that contributes to environmental problems such as landfill accumulation and microplastic accumulation due to its non-biodegradable nature. This study investigates fungal-mediated PET biodegradation via solid-state fermentation (SSF) as a sustainable, eco-friendly waste management approach. The effects of fermentation time on PET weight reduction were evaluated using *Aspergillus niger* and *Aspergillus oryzae* individually over a 10-week SSF period. Xylanase activity was monitored to assess fungal enzymatic activity during PET treatment, while attenuated total reflectance Fourier-transform infrared spectroscopy (ATR-FTIR) and field emission scanning electron microscopy (FESEM) were employed to evaluate chemical and surface morphological changes. PET obtained from mineral water bottles was used as the substrate, supplemented with sucrose and yeast extract to promote fungal growth. *A. oryzae* exhibited the highest PET weight loss of 1.91% and xylanase activity of 467.8 U/mL at week 4, indicating effective early-stage PET modification. Meanwhile, *A. niger* achieved a higher peak xylanase activity of 641.9 U/mL at week 2 but showed a lower maximum PET weight loss of 0.76% at week 6. ATR-FTIR and FESEM analyses confirmed structural and surface modifications after fungal treatment,

ARTICLE INFO

Article history:
Received: 15 April 2026
Accepted: 17 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.11

E-mail addresses:
nilashankar98@gmail.com (Latha Shankar)
azlinaidris@unisel.edu.my (Norazlina Idris)
nurhafizah@unisel.edu.my (Nurhafizah Ibrahim)
norakma@unisel.edu.my (Norakma Mohd Nor)
rodiah@unisel.edu.my (Rodiah Mohd Hassan)
wzarina@unisel.edu.my (Wan Zarina Wan Mohamed)
nfazreen84@ukm.edu.my (Noor Fazreen Dzulkafli)
ainizzati@upm.edu.my (Nur Ain Izzati Zainudin)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

with *A. oryzae* showing a greater PET weight reduction than *A. niger*, which demonstrated more pronounced surface roughening and fungal colonisation. These findings highlight the potential of SSF as a low-cost, sustainable strategy to advance fungal-based PET biodegradation.

Keywords: Aspergillus niger, *Aspergillus oryzae*, Polyethene terephthalate, solid state fermentation, xylanase

INTRODUCTION

Plastics have become an essential part of daily life due to their wide applications in packaging, construction, healthcare, automotive, textiles, and electronics. They are generally composed of synthetic or semi-synthetic organic compounds that can be moulded into different shapes and forms. Their structure consists of monomers linked together to form long polymer chains (Noorjahan et al., 2022). The international market demand for plastics has surged fourfold over the past decades due to their versatility, durability, and low cost, with demand projected to double by 2050 (Jaynes, 2025). The most used plastic across many industries is polyethene terephthalate (PET). PET is a thermoplastic polymer derived from the polyester family, and it is produced by polymerising terephthalic acid and ethylene glycol or by transesterification of dimethyl terephthalate and EG (Hiraga et al., 2019). The global PET market was estimated at USD 39.12 billion in 2024 and is expected to grow at a compound annual growth rate (CAGR) of 5.6% from 2025 to 2030, reaching USD 54.47 billion (Grand View Research, 2024). This growth is driven by PET's outstanding phytochemical properties, such as chemical resistance, impermeability to liquids and gases, higher density than water, resistance to acids and bases and good thermal stability (Hiraga et al., 2019; Jovanović et al., 2025). The Asia Pacific PET market, especially China, commanded the international market and accounted for the largest share of turnover, at 38.0% in 2024, in sectors such as packaging, automotive, electronics, and healthcare (Jovanović et al., 2025).

Despite the widespread use and benefits of plastics in industries, plastic waste has become a major environmental concern. For example, single-use plastic bags are often used for only a short period before being discarded, but they can remain in the environment for many years without degrading. Each year, an estimated 14 million tonnes of plastic waste enter the oceans, causing harmful effects on marine organisms and their populations. This amount is comparable to the weight of five blue whales entering the environment every hour (Davison, 2024). PET is one of the top contributors to landfill waste, soil contamination, ocean debris, greenhouse gas emissions, and air pollution (Lionetto & Esposito Corcione, 2021). PET is widely used in food containers and beverage bottles, significantly impacting aquatic and terrestrial ecosystems. These PET wastes are gradually emerging as an environmental hazard (Jovanović et al., 2025). The density of solid PET is 1.27 g/cm^3, and it breaks down more slowly, leading to accumulation on the ocean floor,

whereas empty PET bottles and containers float on the sea surface (Hiraga et al., 2019). Surface fragmentation is a major contributor to microplastic pollution. For example, PET-based textiles that continuously release microfibers during washing will eventually reach the ocean floor and settle there (Maurya et al., 2020). Over 500 billion PET plastic bottles are used annually; however, many are not fully recycled and are instead illegally dumped into the ocean, posing a significant threat to marine species through entanglement and ingestion (Lionetto & Esposito Corcione, 2021).

Currently, PET waste is treated using various methods, including incineration, chemical and mechanical processing, burning, oxidative degradation, thermal degradation, and photodegradation (Ekanayaka et al., 2022; Zeng et al., 2023). Among these methods, biodegradation is considered a promising approach because of its pollution-free, environmentally friendly nature (Srikanth et al., 2022). Biodegradation of plastics is referred to as microbial plastic degradation, as it involves microorganisms that break them down into organic byproducts such as water and carbon dioxide (Bahl et al., 2021). One of the most effective biodegradation methods is solid-state fermentation (SSF), using filamentous fungi. Fungi can grow on insoluble substrates or solid surfaces with little to no free water (Sadh et al., 2018). SSF substrates are primarily lignocellulosic materials that promote the production of lignocellulolytic enzymes, including cellulase, xylanase, cutinase, and β-glucosidase, which are crucial for plastic degradation. The related microorganisms will release enzymes and pro-oxidant ions to facilitate the breakdown of plastic (Chandren et al., 2025; Safdar et al., 2024). The related microorganisms will release enzymes and some pro-oxidant ions, which will degrade the plastics (Srikanth et al., 2022). Biodegradation is a cost-effective and efficient process for plastic disposal compared to traditional methods like incineration or landfilling. Additionally, the secondary products generated by SSF biodegradation are often more economically valuable than those from incineration or landfilling. Biodegradation is a complex process involving several mechanisms, including depolymerisation, mineralisation, and biodeterioration. During this process, microorganisms utilise synthetic polymers as a carbon source, resulting in microbial biomass formation (Sanniyasi et al., 2021; Webb et al., 2013).

Although numerous studies have investigated bacterial PET degradation, comparatively fewer studies have examined fungal-mediated PET degradation under solid-state fermentation (SSF). Furthermore, although fungal extracellular enzyme production has been frequently reported, the temporal relationship between enzyme activity and PET surface modification during prolonged SSF remains poorly understood. Limited information is available on how changes in extracellular enzymatic activity correspond with physicochemical alterations of PET throughout the fermentation process. PET, due to its strong hydrophobicity and long carbon chain, is highly resistant to biodegradation. Fungal species such as Aspergillus, Acremonium, Fusarium, Penicillium, and Phanerochaete

have shown potential for PET degradation due to their enzymatic versatility (Zhang et al., 2020). The biodegradation of PET involves PET breakdown into mono-(2-hydroxyethyl) terephthalate (MHET), bis- (2- hydroxyethyl) terephthalate (BHET), and ethylene glycol (EG) by PETase. Then, MHETase hydrolyses the MHET into terephthalic acid (TPA) and EG, which are transported into cells and converted into protocatechuic acids (PCAs). These compounds are subsequently integrated into cellular metabolism via the tricarboxylic acid (TCA) cycle (Tegariyanto et al., 2024). *Aspergillus niger* and *Aspergillus oryzae* have the potential to degrade PET because they can grow freely in soil and at garbage sites. *Aspergillus* species produce hydrophobin protein and hydrolase enzymes that can decrease the hydrophobic properties of plastics by hydrolysing C-C and C-O bonds, which act as catalysts for the degradation of high molecular weight plastics into smaller fragments (D'Souza et al., 2021; Zhang et al., 2020). Therefore, this study investigated the time-dependent biodegradation potential of PET by *Aspergillus niger* and *Aspergillus oryzae* under solid-state fermentation. PET surface modification was evaluated over a 10-week fermentation period using weight-loss measurements, xylanase activity as an indicator of extracellular enzymatic activity, ATR-FTIR spectroscopy, and FESEM analysis to provide complementary evidence of fungal-induced changes to the PET surface.

MATERIALS AND METHODS

Preparation of PET strips

Polyethene terephthalate (PET) samples were obtained from post-consumer mineral water bottles collected within the Universiti Selangor (UNISEL) campus. The bottles underwent cleaning, such as washing with tap water to remove visible contaminants, and the labels and caps were removed. The PET bottles were cut into strips measuring approximately 2 × 2 cm. The strips were cleansed by washing with tap water, followed by distilled water, and air-dried at room temperature. For surface sterilisation, the PET strips were immersed in 70% ethanol for 30 min, rinsed with sterile distilled water, and exposed to ultraviolet (UV) light for 30 min. The sterilised strips were then dried in a hot-air oven at 40 °C for 1 h before being transferred to a desiccator containing silica gel. The initial weight of each PET strip was recorded prior to use in the biodegradation experiment (Chandren et al., 2025; Nag et al., 2021).

Preparation of Media and Fungi

Potato dextrose agar (PDA) was prepared by dissolving 7.8 g of PDA powder (TM MEDIA) in 200 mL of distilled water. The medium was sterilised by autoclaving at 121 °C for 15 min and then poured aseptically into sterile Petri dishes. After solidification, *Aspergillus niger* and *Aspergillus oryzae* were subcultured onto separate PDA plates.

The *A. niger* cultures were incubated at 30 °C for 5 days. *oryzae* cultures were incubated at 35 °C for 7 days to obtain sufficient fungal growth for inoculum preparation (Chandren et al., 2025; Ma et al., 2021; Shinkawa & Mitsuzawa, 2020). Following the incubation period, fungal spore suspensions were prepared from fully grown cultures of *Aspergillus niger* and *Aspergillus oryzae*. After sufficient sporulation was observed, the fungal spores were collected by adding 10 mL of sterile distilled water to each petri dish. The spores were gently dislodged from the culture surface with a sterile loop, and the suspension was adjusted to the required volume with sterile distilled water before use in subsequent experiments (Ahmed et al., 2025). The resulting spore suspension was filtered through a sterile filter funnel lined with sterile filter paper to remove mycelial fragments and obtain a homogeneous spore suspension. The filtered suspension was collected in a sterile conical flask and used immediately as the fungal inoculum for subsequent solid-state fermentation experiments (Latif et al., 2025).

Solid State Fermentation (SSF)

Solid-state fermentation (SSF) was performed independently with *Aspergillus niger* and *Aspergillus oryzae* as separate cultures, with each fermentation flask containing a total mixture of 10 g, including 3 g of solid components and 7 mL of liquid. The solid components consisted of 2.8 g of PET strips supplemented with 0.1 g each of sucrose and yeast extract as supplementary nutrients. Because sucrose and yeast extract were supplied in the fermentation medium, the present experimental design does not establish PET as the sole or primary carbon source for fungal growth. SSF was conducted at an initial moisture content of 70%. To achieve 70% moisture content, 6 ml of sterilised distilled water will be added to the fermentation medium, followed by inoculation of 1 mL spore suspension. The sucrose, yeast extract, and distilled water were prepared as a nutrient solution and sterilised separately by autoclaving at 121 °C for 20 min. The PET strips were not autoclaved and were instead surface treated with 70% ethanol, followed by UV irradiation, as described in Section 2.1. The use of ethanol-treated PET samples in fungal biodegradation experiments has also been reported by Safdar et al. (2024). After sterilisation and cooling of the nutrient solution, the sterile nutrient solution was aseptically combined with the prepared PET strips, followed by the addition of 1 mL of the respective fungal spore suspension. The contents were mixed aseptically to distribute the moisture and inoculum throughout the substrate.

Separate fermentation experiments were performed for *A. niger* and *A. oryzae*. The *A. niger*-inoculated flasks were incubated at 30 °C, whereas the *A. oryzae*-inoculated flasks were incubated at 35 °C. For each fungal species, 18 fermentation flasks were prepared, consisting of three independent biological replicates for each incubation period of 0, 2, 4, 6, 8, and 10 weeks (Chandren et al., 2025). PET samples were recovered at each sampling interval for the determination of PET weight loss, extracellular xylanase activity,

ATR-FTIR characterisation, and FESEM analysis. Similar approaches using fungal incubation followed by weight loss, FTIR, and SEM-based characterisation of PET degradation have been reported previously (Maheswaran et al., 2023; Safdar et al., 2024). A non-inoculated sample containing PET and the corresponding fermentation components was used as the week 0 reference to establish the initial condition. Since this sample was assessed only in week 0, it served as a baseline reference rather than an abiotic control throughout the fermentation period.

Recovery and Preparation of PET Strips After Fermentation

After each incubation period, the PET strips were carefully recovered from the fermentation flasks and separated from the fungal biomass. The recovered PET strips were first washed with 70% ethanol to remove fungal cells and lose biomass, then thoroughly rinsed with tap water to eliminate residual ethanol and other adherent materials. The washing procedure was applied consistently to all recovered PET samples (Ahmed et al., 2025; Rojas-Villacorta et al., 2025). Following washing, the PET strips were dried in a hot-air oven at 40 °C until a constant weight was achieved. The dried samples were subsequently transferred to a desiccator containing silica gel to minimise moisture absorption before weighing. The final dry weight of each recovered PET strip was recorded to determine the gravimetric PET weight loss (Zulkafli et al., 2024).

Determination of the Weight Loss of the PET Strips

The initial and final weights of PET strips from each fermentation flask were measured using an analytical balance before and after the solid-state fermentation process, respectively. The initial weight (W_i) was recorded prior to fungal inoculation, while the final weight (W_f) was obtained after recovery, washing, and drying of the PET strips following the fermentation period (D'Souza et al., 2021). The percentage of PET weight loss was calculated for each individual replicate using Equation 1:

$$Loss\ of\ weight\ (\%) = \frac{[(Wi - Wf) \times 100]}{Wi} \quad [1]$$

where W*i* represents the initial dry weight of PET strips, and W_f represents the final dry weight of PET strips after fermentation. Three independent fermentation flasks were analysed for each incubation period, and the average percentage of weight loss was calculated from the triplicate measurements. The results were expressed as mean ± standard deviation. The gravimetric method was used to quantify changes in the mass of recovered PET following fungal treatment. Because gravimetric mass loss alone cannot distinguish polymer degradation from other potential sources of mass change, such as incomplete

biomass removal, washing-related variation, or the physical loss of small PET fragments, the measured parameter was interpreted as PET weight loss rather than as definitive evidence of PET biodegradation.

Xylanase Assay

Extracellular xylanase activity during SSF was measured using the 3,5-dinitrosalicylic acid (DNS) method as described by Miller (1959), with xylan serving as the substrate. The activity was monitored in accordance with established protocols (Bailey et al., 1992) as an indicator of extracellular enzyme production by the fungi. Since the assay quantified reducing sugars released from xylan hydrolysis rather than PET breakdown, the xylanase activity was not used as a measure of PET depolymerisation and was analysed separately from PET weight loss.

After each fermentation period, 70 mL of distilled water was added directly into each fermentation flask containing the fermented substrate. The flask contents were manually shaken to facilitate the extraction of extracellular enzymes into the aqueous phase. The extract was then filtered to separate the PET strips and fungal biomass from the liquid fraction. The collected liquid extract was centrifuged at 15,000 rpm for 10 min to remove suspended particles. The resulting supernatant was further filtered using a vacuum filtration system with Whatman No. 1 qualitative filter paper (90 mm diameter; Cat. No. 1001-090) to obtain a clarified enzyme extract used for xylanase analysis. The xylanase assay was performed by mixing 0.5 mL of crude enzyme extract with 1.0 mL of 1% corn Xylan prepared in 0.05 M citrate buffer. The reaction mixture was incubated in a water bath at 50 °C for 20 min. Subsequently, 3 mL of dinitrosalicylic acid (DNS) reagent was added to terminate the enzymatic reaction, and the mixture was heated at 100 °C for 5 min to develop colour. After cooling, the absorbance was measured at 540 nm using a spectrophotometer. Xylanase activity was calculated based on the amount of reducing sugar released. One unit (U) of xylanase activity was defined as the amount of enzyme required to release 1 µmol of D-xylose per minute under the assay conditions. The enzyme activity was expressed as U/mL (Chandren et al., 2025). The enzyme activity for each sample was calculated utilising the formula given in Equation 2.

$$\text{Xylanase activity} = \frac{\mu\,\text{mol}}{\text{mL}} = \frac{\text{U}}{\text{ml}} = \left(\text{xylose concentration}\,\frac{\text{g}}{\text{mL}}\right) \times \frac{1}{0.5\,\text{mL}} \times \frac{1}{20\,\text{min}} \times \frac{1\,\text{mol}}{150.13\,\text{g}} \qquad [2]$$

Statistical Analysis

All experiments were conducted using three independent biological replicates for each fermentation period. Data on PET weight loss and extracellular xylanase activity were

analysed using IBM SPSS Statistics. Differences in PET weight loss and xylanase activity among fermentation periods were evaluated using one-way analysis of variance (ANOVA), followed by Tukey's honestly significant difference (HSD) post hoc test where appropriate. Statistical significance was considered at $p < 0.05$. Pearson correlation analysis was performed as an exploratory analysis to assess the linear association between PET weight loss and xylanase activity during the fermentation period. Because xylanase is not a PET-specific depolymerising enzyme, any observed correlation was interpreted only as a statistical association between the two measured variables and was not considered evidence of a causal relationship between xylanase activity and PET degradation (Agbangba et al., 2024).

Attenuated Total Reflectance – Fourier Transformed (ATR-FTIR) Analysis

The structural changes of PET samples before and after solid-state fermentation (SSF) were analysed using Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) spectroscopy. Following the fermentation period, PET samples were collected, washed with distilled water to remove attached fungal biomass and residual fermentation components, and dried prior to analysis. The dried PET samples were placed directly onto the ATR crystal of a PerkinElmer FTIR spectrometer for spectral analysis. The FTIR spectra were recorded over a wavenumber range of 4000–400 cm^{-1}, and the resulting spectra were presented as transmittance (%) against wavenumber (cm^{-1}). Changes in the characteristic absorption bands associated with PET functional groups were compared between untreated PET samples and PET samples subjected to SSF. The obtained spectra were used to evaluate alterations in the chemical structure of PET following fungal treatment (D'Souza et al., 2021).

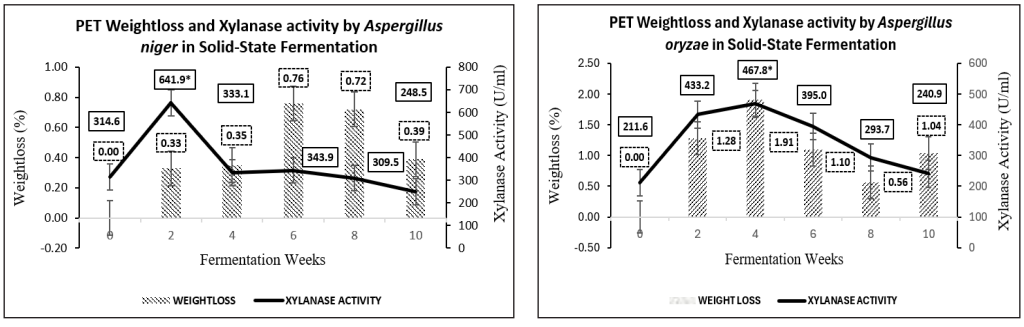
Field Emission Scanning Electron Microscopy (FESEM) Analysis

The surface morphology of untreated PET and PET samples recovered after solid-state fermentation (SSF) was examined using field emission scanning electron microscopy (FESEM). Following the fermentation period, the PET strips were carefully removed from the fermentation flasks, washed thoroughly with distilled water to remove loosely attached fungal biomass and residual fermentation medium, and dried in a hot-air oven at 40 °C until a constant weight was achieved. The dried PET samples were mounted onto aluminium stubs using conductive carbon adhesive tape and sputter-coated with a thin layer of gold to improve surface conductivity before imaging. The surface morphology of the samples was then observed using a field-emission scanning electron microscope at an appropriate acceleration voltage. Micrographs of untreated PET and fungal-treated PET were compared to evaluate changes in surface morphology, including fungal attachment, surface roughness, cracks, pits, and erosion patterns following SSF (Pandey et al., 2021).

RESULTS AND DISCUSSION

PET Weight Reduction and Xylanase Activity by *Aspergillus niger* and *Aspergillus oryzae* During Solid-State Fermentation

The changes in PET weight reduction and extracellular xylanase activity during solid-state fermentation (SSF) using *Aspergillus niger* and *Aspergillus oryzae* for a 10-week incubation period are illustrated in Figure 1a and Figure 1 b, respectively. The corresponding mean values and standard deviations (SD) obtained from triplicate experiments are summarised in Table 1. PET weight reduction was monitored to evaluate the effect of fungal treatment on PET samples, whereas xylanase activity was measured to assess extracellular enzyme production throughout the fermentation period.



(a) (b)

Figure 1. PET weight loss and xylanase activity: (a) PET weight loss and xylanase activity by *Aspergillus niger*; (b) PET weight loss and xylanase activity by *Aspergillus oryzae*

Figure 1a shows that the mean PET weight loss for *A. niger* increased from 0.00 ± 0.00% at week 0 to 0.76 ± 0.84% at week 6, followed by a slight decrease to 0.72 ± 0.32% and 0.39 ± 0.27% at weeks 8 and 10, respectively. Although an increasing trend was observed during the early stages of fermentation, one-way ANOVA showed that the fermentation period did not significantly affect PET weight loss ($F(5, 12) = 1.583$, $p = 0.238$). Similarly, *A. oryzae* showed an increase in mean PET weight loss from 0.00 ± 0.00% at week 0 to a maximum of 1.91 ± 1.43% at week 4, followed by decreases to 1.07 ± 0.54%, 0.56 ± 0.49%, and 1.04 ± 0.93% at weeks 6, 8, and 10, respectively (Figure 1b). The fermentation period also did not significantly affect PET weight loss in *A. oryzae* ($F(5, 12) = 1.497$, $p = 0.262$). Nevertheless, the numerically higher PET weight loss observed for *A. oryzae*, particularly at week 4, may indicate greater interaction or colonisation of the PET surface under the conditions tested. Previous studies have reported that filamentous fungi can interact with polymer surfaces through fungal colonisation, extracellular enzyme secretion, and biofilm formation, potentially altering

polymer surface characteristics (Safdar et al., 2024). *A. oryzae* has also been reported to exhibit active extracellular hydrolase production during cultivation, which may contribute to its interaction with complex substrates (Maurya et al., 2020; Srikanth et al., 2021). PET weight loss alone does not provide definitive evidence of PET biodegradation. ATR-FTIR and FESEM were therefore used as complementary techniques to characterise chemical and morphological changes associated with fungal treatment; however, these techniques also cannot independently confirm PET depolymerisation.

Xylanase activity showed a different temporal pattern from PET weight loss. In *A. niger*, xylanase activity increased from 314.63 ± 86.14 U mL^{-1} at week 0 to a maximum of 641.90 ± 100.12 U mL^{-1} at week 2, followed by decreases to 334.13 ± 5.10, 343.90 ± 20.22, 309.53 ± 13.11, and 248.47 ± 23.31 U mL^{-1} at weeks 4, 6, 8, and 10, respectively. One-way ANOVA indicated a significant effect of fermentation period on xylanase activity ($F(5, 12) = 18.830$, $p < 0.001$), with Tukey's HSD showing that xylanase activity at week 2 was significantly higher than that observed at the other fermentation periods ($p < 0.05$). In *A. oryzae*, xylanase activity increased from 211.63 ± 0.75 U mL^{-1} at week 0 to a maximum of 470.50 ± 17.23 U mL^{-1} at week 4, followed by decreases to 394.97 ± 60.00, 293.73 ± 4.69, and 240.90 ± 30.24 U mL^{-1} at weeks 6, 8, and 10, respectively. Fermentation period significantly affected xylanase activity in *A. oryzae* ($F(5, 12) = 42.086$, $p < 0.001$), with Tukey's HSD indicating significant differences between several fermentation periods ($p < 0.05$). The higher enzyme activity observed during the early fermentation stages may reflect active extracellular enzyme production, whereas the subsequent decline may be associated with nutrient depletion or changes in fungal metabolic activity during prolonged fermentation (Maurya et al., 2020; Srikanth et al., 2021).

Xylanase activity exhibited temporal patterns that differed from those observed for PET weight loss. In *A. niger*, maximum xylanase activity occurred at week 2, whereas the highest numerical PET weight loss occurred at week 6. In *A. oryzae*, maximum xylanase activity and the highest numerical PET weight loss were both observed at week 4. However, because xylanase is not a PET-specific depolymerising enzyme, these temporal patterns were not interpreted as evidence of a direct relationship between xylanase production and PET degradation. Xylanase activity was therefore considered an indicator of extracellular enzyme production during SSF rather than a measure of PET hydrolysis. The differences between xylanase activity and PET weight-loss patterns may reflect the involvement of multiple factors, including fungal colonisation, other extracellular enzymes, physicochemical changes at the PET surface, and changes in substrate accessibility during SSF. Therefore, the observed PET weight reduction should be interpreted as evidence of PET mass change associated with fungal treatment rather than definitive proof of biodegradation. Complementary analyses such as FTIR and FESEM may provide additional evidence of

changes in the chemical characteristics and surface morphology of PET following fungal treatment (Safdar et al., 2024). Overall, the results demonstrate that fermentation time significantly influenced xylanase production in both fungal species, whereas its effect on PET weight loss was not statistically significant under the conditions tested.

Attenuated Total Reflectance – Fourier Transformed (ATR-FTIR) Analysis of PET Samples

A comparative analysis of the ATR-FTIR spectra of untreated PET and PET samples recovered after solid-state fermentation is presented in Figures 2, 3, and 4.

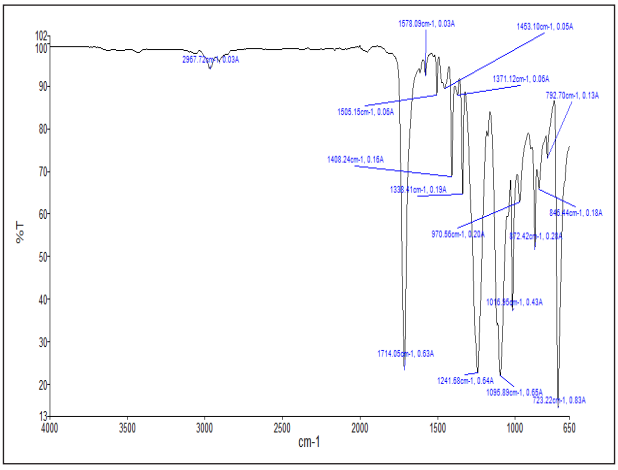


Figure 2. ATR-FTIR image of PET samples before solid-state fermentation

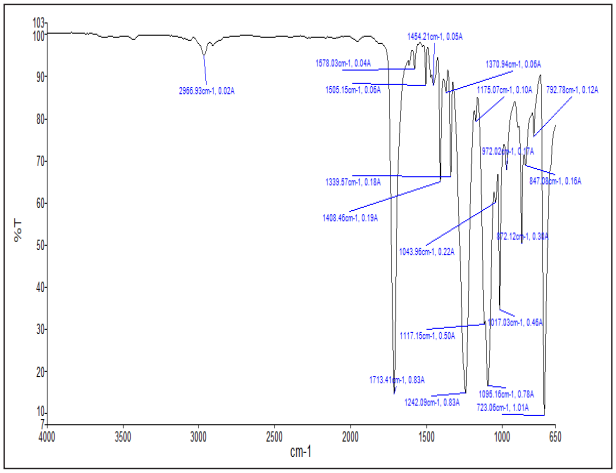


Figure 3. ATR-FTIR image of PET samples treated with *Aspergillus niger* after solid-state fermentation

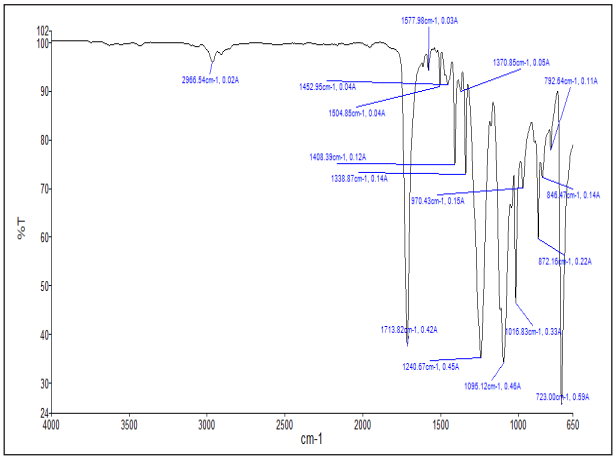


Figure 4. ATR-FTIR image of PET samples treated with *Aspergillus oryzae* after solid-state fermentation

The ATR-FTIR spectrum of the untreated PET sample showed characteristic absorption bands associated with the chemical structure of polyethene terephthalate (PET), including the ester carbonyl (C=O) stretching vibration at approximately 1710–1715 cm^{-1}, C–O stretching vibrations within the 1300–1100 cm^{-1} region, and characteristic aromatic ring vibrations. These absorption bands are consistent with the chemical structure of PET reported in previous studies (D'Souza et al., 2021; Sivan, 2011). Following fungal treatment, slight changes in the ATR-FTIR spectra were observed for PET exposed to both *A. niger* and *A. oryzae*. The observed changes in the intensity and/or position of some absorption bands, particularly those associated with ester functional groups, may indicate alterations in the chemical environment of the PET surface following fungal exposure. Such changes may result from localised interactions between the fungal biomass, extracellular metabolites, and the polymer surface.

Visual comparison of the spectra suggested somewhat greater variation in selected absorption bands in *A. oryzae*; however, these differences were not quantitatively evaluated and should therefore be interpreted with caution. The untreated PET corresponds with the higher numerical PET weight loss observed for *A. oryzae* during the earlier stages of fermentation. These observations may indicate greater interaction between *A. oryzae* and the PET surface under experimental conditions. Filamentous fungi may interact with polymeric materials through surface colonisation and the secretion of extracellular enzymes and metabolites, potentially resulting in localised changes in polymer surface chemistry (Danso et al., 2019). However, the observed FTIR changes were relatively small. Therefore, they should be interpreted as evidence of surface chemical modification rather than definitive evidence of extensive PET depolymerisation. In addition, changes

in FTIR peak intensity can be influenced by sample preparation, surface heterogeneity, contact between the sample and ATR crystal, and changes in surface morphology.

The relatively modest spectral changes observed in the present study may also be associated with the chemical stability and crystallinity of PET, which limit its susceptibility to microbial attack. Previous studies have demonstrated that substantial enzymatic hydrolysis of PET requires effective PET-hydrolysing enzymes, such as PETase, cutinase, or related esterases, that can attack the ester bonds within the polymer structure (Austin et al., 2018; Yoshida et al., 2016). Therefore, the ATR-FTIR findings in the present study suggest that fungal attachment may facilitate close contact between the fungal biomass and the PET surface and may contribute to localised surface modification of the PET surface following fungal treatment, but they do not independently confirm complete polymer biodegradation. The FTIR results should therefore be considered alongside PET weight-loss measurements and FESEM observations to provide a more comprehensive assessment of fungal interaction with PET.

FESEM Analysis of PET Surface Morphology

The ATR-FTIR observations are further supported by FESEM analysis, which demonstrated fungal attachment and localised surface erosion after SSF. The combination of chemical and morphological evidence strengthens the interpretation that fungal colonisation resulted in surface modification of PET rather than merely physical contamination. The surface morphology of untreated PET and PET samples after solid-state fermentation (SSF) with *Aspergillus niger* and *Aspergillus oryzae* was examined using field-emission scanning electron microscopy (FESEM). The FESEM micrographs are presented in Figures 5-7.

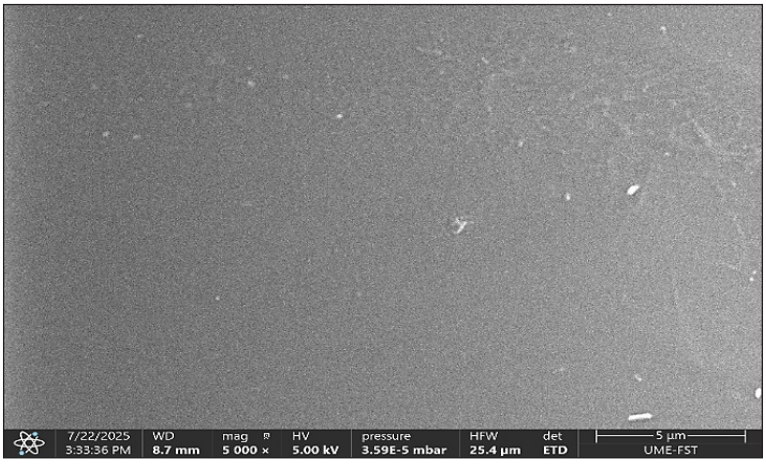


Figure 5. FESEM image of PET samples before solid-state fermentation

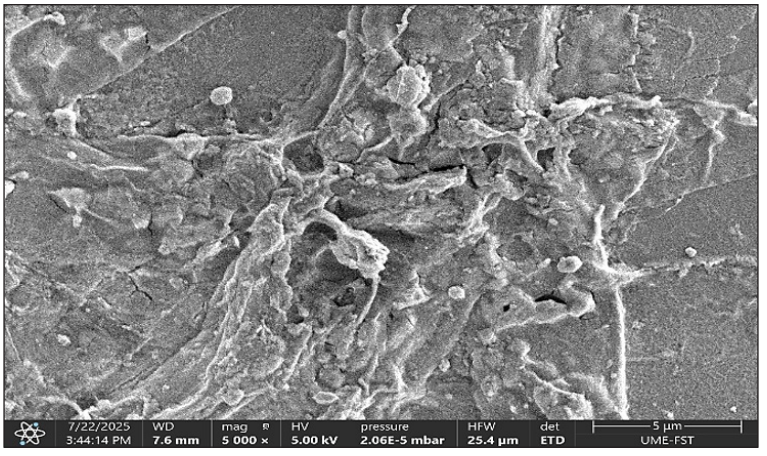


Figure 6. FESEM image of PET samples treated with *Aspergillus niger* after solid-state fermentation

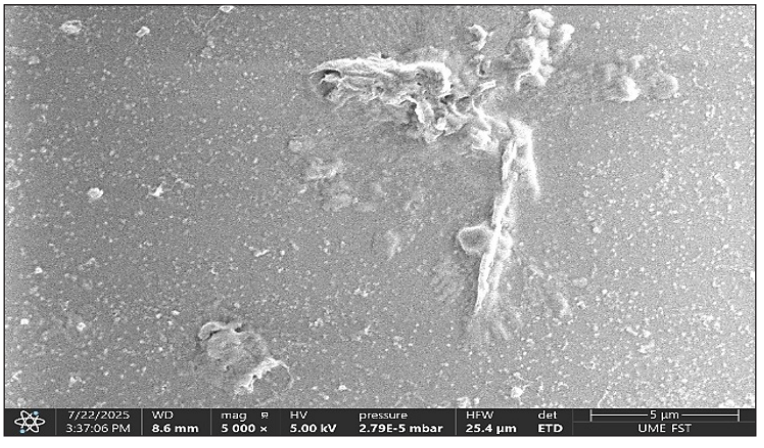


Figure 7. FESEM image of PET samples treated with *Aspergillus oryzae* after solid-state fermentation

As shown in Figure 5, the untreated PET exhibited a relatively smooth, compact, and homogeneous surface, with no prominent cracks, pits, or extensive erosion features. The relatively intact morphology observed before fermentation indicates that the PET surface had not undergone noticeable physical alteration prior to fungal treatment. A similar, relatively smooth surface morphology has been reported for untreated PET samples in studies investigating microbial interactions with PET (Danso et al., 2019). Following incubation with *A. niger*, noticeable morphological changes were observed on the PET surface (Figure 6), including increased surface roughness, irregular erosion patterns, shallow fissures, and fungal hyphae attached to the polymer surface. The presence of fungal hyphae indicates that *A. niger* was able to colonise and establish close contact with the PET surface during SSF. Such fungal attachment may increase the interaction between

extracellular enzymes or metabolites and the polymer surface, potentially contributing to localised surface modification. Similar changes, including surface roughening, cracks, and erosion, have been associated with microbial attachment and biodeterioration of polymer surfaces (Danso et al., 2019). However, morphological observations alone cannot establish that the polymer chains were extensively depolymerised.

The PET treated with *A. oryzae* (Figure 7) also showed evidence of fungal attachment and localised surface irregularities compared with the untreated PET. However, the visible surface alteration appeared less pronounced than that observed for *A. niger*. Interestingly, this observation does not directly correspond to the higher numerical PET weight loss recorded for *A. oryzae*. This difference may be explained by the fact that FESEM provides a localised visualisation of the polymer surface, whereas weight-loss measurements represent the overall change in PET mass. Consequently, localised surface roughening observed by FESEM may not necessarily correspond proportionally to the total mass reduction of the PET sample. In addition, fungal interactions may involve chemical modifications or more uniformly distributed surface changes that are not readily distinguishable in FESEM images.

Overall, the FESEM observations provide supporting evidence of fungal colonisation and localised physical changes on the PET surface following SSF. The more apparent surface roughening and erosion observed following *A. niger* treatment indicates more pronounced localised surface alteration, whereas *A. oryzae* showed greater numerical PET weight loss despite less pronounced visible surface erosion. This discrepancy suggests that PET mass loss and surface morphology may reflect distinct aspects of fungal polymer interaction. Therefore, FESEM results should be interpreted alongside PET weight loss and ATR-FTIR analysis rather than treated as definitive evidence of PET biodegradation. More conclusive evidence would require additional analyses, particularly the identification and quantification of PET degradation products such as terephthalic acid (TPA), mono-(2-hydroxyethyl) terephthalate (MHET), bis-(2-hydroxyethyl) terephthalate (BHET), and ethylene glycol (EG), as demonstrated in studies of enzymatic PET degradation (Austin et al., 2018; Yoshida et al., 2016).

CONCLUSION

The study evaluated the time-dependent interaction of *Aspergillus niger* and *Aspergillus oryzae* with PET under solid-state fermentation. *A. niger* showed the highest numerical PET weight loss at week 6 (0.76%), with maximum xylanase activity at week 2 (641.9 U/mL), whereas *A. oryzae* exhibited the highest numerical PET weight loss and xylanase activity at week 4 (1.91% and 467.8 U/mL, respectively). Notably, PET weight loss was not significantly affected by fermentation period for either fungal species, nor did the temporal patterns of xylanase activity show a consistent correspondence with PET weight loss. Therefore, xylanase activity alone cannot be used as a direct indicator of PET degradation

and should be regarded as an indicator of extracellular enzyme production during SSF. FESEM observations revealed fungal attachment and localised surface morphological alterations on PET, while ATR-FTIR analysis indicated relatively minor changes in the chemical characteristics of the treated PET. Collectively, these findings provide preliminary evidence of fungal-mediated interaction and surface modification of PET rather than definitive proof of PET biodegradation or depolymerisation. The observed PET mass changes should therefore be interpreted with caution, given the limitations of gravimetric assessment and the absence of PET-specific degradation-product analysis. Further studies incorporating PET-hydrolysing enzymes, such as cutinases or esterases, together with the detection and quantification of degradation products, including terephthalic acid, MHET, BHET, and ethylene glycol, are recommended to provide stronger evidence of PET biodegradation. Kinetic modelling and controlled abiotic experiments should also be considered in future studies to better characterise PET degradation behaviour and distinguish biological degradation from non-biological mass changes.

ACKNOWLEDGEMENT

This work was supported by the Fundamental Research Grant Scheme from the Ministry of Higher Education (FRGS/1/2023/WAB02/UNISEL/03/1).

REFERENCES

Agbangba, C. E., Aide, E. S., Honfo, H., & Kakai, R. G. (2024). On the use of post-hoc tests in environmental and biological sciences: A critical review. *Heliyon, 10*(3), Article e24894. https://doi.org/10.1016/j.heliyon.2024.e24894

Ahmed, M., Iram, S., Tabassum, N., Sajid, M., Paseutsakoun, K., Aleksza, L., & Székács, A. (2025). Biodegradation efficacy of Aspergillus niger and Trichoderma harzianum on low-density polyethylene. *Polymers, 17*(10), Article 1303. https://doi.org/10.3390/polym17101303

Austin, H. P., Allen, M. D., Donohoe, B. S., Rorrer, N. A., Kearns, F. L., Silveira, R. L., Pollard, B. C., Dominick, G., Duman, R., El Omari, K., Mykhaylyk, V., Wagner, A., Michener, W. E., Amore, A., Skaf, M. S., Crowley, M. F., Thorne, A. W., Johnson, C. W., Woodcock, H. L., … Beckham, G. T. (2018). Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proceedings of the National Academy of Sciences of the United States of America, 115*(19), E4350-E4357. https://doi.org/10.1073/pnas.1718804115

Bahl, S., Dolma, J., Singh, J. J., & Sehgal, S. (2021). Biodegradation of plastics: A state-of-the-art review. *Materials Today: Proceedings, 39*, 31-34. https://doi.org/10.1016/j.matpr.2020.07.191

Bailey, M. J., Biely, P., & Poutanen, K. (1992). Interlaboratory testing of methods for assay of xylanase activity. *Journal of Biotechnology, 23*(3), 257-270. https://doi.org/10.1016/0168-1656(92)90074-J

Chandren, Y., Norazlina, I., Norakma, M. N., Wan Zarina, W. M., Noor Fazreen, F., & Rodiah, M. H. (2025). Refining solid state fermentation of Fusarium oxysporum for enhanced polyethylene terephthalate

biodegradation efficiency. *Malaysian Journal of Chemistry, 27*(3), 272–276. https://doi.org/10.55373/mjchem.v27i3.272

Danso, D., Chow, J., & Streit, W. R. (2019). Plastics: Environmental and biotechnological perspectives on microbial degradation. *Applied and Environmental Microbiology, 85*(19), Article e01095-19. https://doi.org/10.1128/AEM.01095-19

Davison, T. (2024, November 5). *How much plastic is in the ocean in 2024? CleanHub*. https://blog.cleanhub.com/how-much-plastic-in-the-ocean

D'Souza, G. C., Sheriff, R. S., Ullanat, V., Shrikrishna, A., Joshi, A. V., Hiremath, L., & Entoori, K. (2021). Fungal biodegradation of low-density polyethylene using consortium of Aspergillus species under controlled conditions. *Heliyon, 7*(5), Article e07008. https://doi.org/10.1016/j.heliyon.2021.e07008

Ekanayaka, A. H., Tibpromma, S., Dai, D., Xu, R., Suwannarach, N., Stephenson, S. L., Dao, C., & Karunarathna, S. C. (2022). A review of the fungi that degrade plastic. *Journal of Fungi, 8*(8), 772. https://doi.org/10.3390/jof8080772

Grand View Research. (2024). *Polyethylene terephthalate market size, share & trends analysis report by application (packaging, films & sheets), by packaging application (bottled water, juices, alcoholic beverages, thermoforming trays), by region, and segment forecasts, 2025-2030*. https://www.grandviewresearch.com/industry-analysis/polyethylene-terephthalate-market

Hiraga, K., Taniguchi, I., Yoshida, S., Kimura, Y., & Oda, K. (2019). Biodegradation of waste PET: A sustainable solution for dealing with plastic pollution. *EMBO Reports, 20*(11), Article e49365. https://doi.org/10.15252/embr.201949365

Jaynes, C. H. (2025, April 14). *Global plastic recycling rate 'stagnant' at 9%: Study*. EcoWatch. https://www.ecowatch.com/global-plastic-recycling-rate

Jovanović, A., Bugarčić, M., Simić, M., Pejić, J., & Dimitrijević, J. (2025). The global market of PET production: From origins to recycling. *Metallurgical and Materials Data, 2*(4), 113-118. https://doi.org/10.30544/MMD46

Latif, A., Ali, H., Niazi, M. B. K., Jahan, Z., Hassan, N., Ghuman, I. L., Tawab, A., Saqib, A., & Rashid, M. H. (2025). Process evaluation for enhanced production of aflatoxin-free citric acid by food-grade Aspergillus niger. *PLOS ONE, 20*(5), Article e0321972. https://doi.org/10.1371/journal.pone.0321972

Lionetto, F., & Esposito Corcione, C. (2021). An overview of the sorption studies of contaminants on poly(ethylene terephthalate) microplastics in the marine environment. *Journal of Marine Science and Engineering, 9*(4), 445. https://doi.org/10.3390/jmse9040445

Ma, X., Zhang, K., Wang, Y., Ebadi, A. G., & Toughani, M. (2021). Optimization of low-temperature lipase production conditions and study on enzymatic properties of Aspergillus niger. *Iranian Journal of Chemistry and Chemical Engineering, 40*(4), 1364-1374. https://doi.org/10.30492/ijcce.2021.529010.4694

Maheswaran, B., Al-Ansari, M., Al-Humaid, L., Sebastin Raj, J., Kim, W., Karmegam, N., & Mohamed Rafi, K. (2023). In vivo degradation of polyethylene terephthalate using microbial isolates from plastic polluted environment. *Chemosphere, 310*, Article 136757. https://doi.org/10.1016/j.chemosphere.2022.136757

Maurya, A., Bhattacharya, A., & Khare, S. K. (2020). Enzymatic remediation of polyethylene terephthalate (PET)–based polymers for effective management of plastic wastes: An overview. *Frontiers in Bioengineering and Biotechnology, 8*, Article 602325. https://doi.org/10.3389/fbioe.2020.602325

Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry, 31*(3), 426-428. https://doi.org/10.1021/ac60147a030

Nag, M., Lahiri, D., Dutta, B., Jadav, G., & Ray, R. R. (2021). Biodegradation of used polyethylene bags by a new marine strain of Alcaligenes faecalis LNDR-1. *Environmental Science and Pollution Research, 28*(30), 41365-41379. https://doi.org/10.1007/s11356-021-13704-0

Noorjahan, C. M., Banu, S. N., & Subhashree, V. (2022). Bioplastic synthesis using banana peels and its characterization. *Journal of Pharmaceutical Negative Results, 13*(Special Issue 7), 2855–2863. https://doi.org/10.47750/pnr.2022.13.S07.380

Pandey, S., Karakoti, M., Surana, K., Dhapola, P. S., SanthiBhushan, B., Ganguly, S., Singh, P. K., Abbas, A., Srivastava, A. K., Chintha, S., Nanda, S. S., Aminabhavi, T. M., & Sahoo, N. G. (2021). Graphene nanosheets derived from plastic waste for the application of DSSCs and supercapacitors. *Scientific Reports, 11*(1), Article 3916. https://doi.org/10.1038/s41598-021-83483-8

Rojas-Villacorta, W., Cruz-Noriega, M. D. L., Otiniano, N. M., Terrones-Rodríguez, N., & Quiñones-Cerna, C. (2025). Biodegradation of low-density polyethylene by native Aspergillus strains isolated from plastic-contaminated soil. *Sustainability, 17*(20), Article 8983. https://doi.org/10.3390/su17208983

Sadh, P. K., Duhan, S., & Duhan, J. S. (2018). Agro-industrial wastes and their utilization using solid state fermentation: A review. *Bioresources and Bioprocessing, 5*, Article 1. https://doi.org/10.1186/s40643-017-0187-z

Safdar, A., Ismail, F., & Imran, M. (2024). Biodegradation of synthetic plastics by the extracellular lipase of Aspergillus niger. *Environmental Advances, 17*, Article 100563. https://doi.org/10.1016/j.envadv.2024.100563

Sanniyasi, E., Gopal, R. K., Gunasekar, D. K., & Raj, P. P. (2021). Biodegradation of low-density polyethylene (LDPE) sheet by microalga, Uronema africanum Borge. *Scientific Reports, 11*(1), Article 17233. https://doi.org/10.1038/s41598-021-96315-6

Shinkawa, S., & Mitsuzawa, S. (2020). Feasibility study of on-site solid-state enzyme production by Aspergillus oryzae. *Biotechnology for Biofuels, 13*, Article 31. https://doi.org/10.1186/s13068-020-1669-3

Sivan, A. (2011). New perspectives in plastic biodegradation. *Current Opinion in Biotechnology, 22*(3), 422-426. https://doi.org/10.1016/j.copbio.2011.01.013

Srikanth, M., Sandeep, T. S. R. S., Sucharitha, K., & Godi, S. (2022). Biodegradation of plastic polymers by fungi: A brief review. *Bioresources and Bioprocessing, 9*, 42. https://doi.org/10.1186/s40643-022-00532-4

Tegariyanto, M. R., Titah, H. S., & Pratikno, H. (2024). Biodegradation of polyethylene terephthalate (PET) plastics using a combination of Aspergillus niger and Trichoderma harzianum. *IOP Conference Series: Earth and Environmental Science, 1414*(1), Article 012028. https://doi.org/10.1088/1755-1315/1414/1/012028

Webb, H. K., Arnott, J., Crawford, R. J., & Ivanova, E. P. (2013). Plastic degradation and its environmental implications with special reference to poly(ethylene terephthalate). *Polymers, 5*(1), 1-18. https://doi.org/10.3390/polym5010001

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., & Oda, K. (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science, 351*(6278), 1196-1199. https://doi.org/10.1126/science.aad6359

Zeng, C., Ding, F., Zhou, J., Dong, W., Cui, Z., & Yan, X. (2023). Biodegradation of poly(ethylene terephthalate) by Bacillus safensis YX8. *International Journal of Molecular Sciences, 24*(22), Article 16434. https://doi.org/10.3390/ijms242216434

Zhang, J., Gao, D., Li, Q., Zhao, Y., Li, L., Lin, H., Bi, Q., & Zhao, Y. (2020). Biodegradation of polyethylene microplastic particles by the fungus Aspergillus flavus from the guts of wax moth Galleria mellonella. *Science of the Total Environment, 704*, Article 135931. https://doi.org/10.1016/j.scitotenv.2019.135931

Zulkafli, M. S., Rithwan, F., Abang Zaidel, D. N., Ramli, S., Hanapi, S. Z., Wan Azelee, N. I., Yenn, T. W., El Enshasy, H. A., & Dailin, D. J. (2024). Biodegradation of low-density polyethylene (LDPE) by Aspergillus niger and Aspergillus oryzae. *Asia-Pacific Journal of Molecular Biology and Biotechnology, 32*(4, Special), 1-10. https://doi.org/10.35118/apjmbb.2024.032.4(Special).01

Pertanika J. Sci. & Technol. 34 (5): 2811 - 2833 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Fractional SIRM Epidemic Model with Microplastic-Mediated Environmental Transmission

R. Sengothai[1], P. Balaganesan[1]*, D. Adhil Basha[1], and J. Renuka[2]

[1]Department of Mathematics, Academy of Maritime Education and Training (AMET University), 135 East Coast Road, Kanathur, Chennai 603112, Tamil Nadu, India

[2]Department of Mathematics, Women's Christian College, 51 College Road, Nungambakkam, Chennai 600006, Tamil Nadu, India

ABSTRACT

The presence of microplastic particles in terrestrial, marine, and atmospheric systems has made microplastic contamination a serious environmental problem. In addition to their environmental effects, microplastics may serve as carriers of microbial populations, including potentially dangerous pathogens, raising concerns about their contribution to the spread of illness. This work examines the impacts of microplastic-mediated environmental transmission on the dynamics of infectious diseases using a fractional-order SIRM epidemic model. Both direct human-to-human transmission and indirect transmission via contaminated microplastic surfaces are considered in the model. Additionally, it has a unique environmental section that depicts infections associated with microplastic particles. Caputo fractional derivatives are used in the model to represent memory effects that are commonly observed in real epidemiological processes. The basic reproduction number is computed using the next-generation matrix technique, and analytical results show that the solutions are positive and bounded. Sensitivity analysis is used to identify key parameters that influence the spread of illness, and stability analysis of the endemic and disease-free equilibria is performed. Numerical simulations show that microplastic-mediated transmission can substantially increase infection levels and prolong outbreak duration. Overall, the suggested model emphasises the importance of accounting for microplastic contamination in epidemiological investigations and offers a useful framework for understanding the relationship between environmental pollution and infectious disease dynamics.

Keywords: Caputo fractional derivative, disease-free equilibrium, epidemic model, infectious disease, microplastic contamination, next-generation matrix, SIRM model, stability analysis

ARTICLE INFO

Article history:
Received: 09 July 2026
Accepted: 17 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.12

E-mail addresses:
kothairam4555@gmail.com (R. Sengothai)
balaganesanpp@gmail.com (P. Balaganesan)
adhilbasha88@gmail.com (D. Adhil Basha)
renukajjm@gmail.com (J. Renuka)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

A major global public health concern in recent years has been the emergence of infectious diseases at the human environment interface (Rajkumar, 2022). While previous epidemiological research has mainly focused on direct human-to-human transmission or vector-borne processes, the importance of environmental reservoirs in maintaining pathogen populations is currently garnering greater attention (Asamoah et al., 2022). These reservoirs, which include polluted water, soil, and artificial surfaces, can influence the transmission of illness even without direct contact with affected people.

Because of its persistence and widespread distribution, microplastic contamination has attracted considerable attention among today's environmental issues. Microplastics are tiny plastic particles, usually less than 5 mm, that come from manufactured items like microbeads used in cosmetics and industry, or from the degradation of larger plastic materials (Ding et al., 2023). These particles are now present in freshwater systems, oceans, agricultural soils, and even the atmosphere due to their small size and resistance to degradation (Wu et al., 2023). Global plastic manufacturing has grown dramatically in recent years, according to recent studies, which raises questions about the long-term effects on the environment (Rajkumar, 2022).

Beyond their environmental effects, microplastic particles can act as substrates for microbial colonisation. These surfaces support complex microbial communities, commonly referred to as the plastisphere, which may include opportunistic and potentially pathogenic organisms (Oberbeckmann & Labrenz, 2020). The ability of microplastics to transport such microorganisms across different environments further increases the risk of exposure and transmission over large spatial scales (Wu et al., 2023).

The widespread presence of microplastics, therefore, raises concerns that extend beyond pollution alone. Their potential to support and transport pathogens suggests a possible link between microplastic contamination and the transmission of infectious diseases. This motivates the need to investigate how such environmental factors can be incorporated into epidemiological models in a mathematically meaningful framework.

The Plastisphere as a Pathogen Reservoir

Beyond ecotoxicological effects on invertebrates and vertebrates, microplastic surfaces function as preferential colonisation substrates for complex microbial assemblages, a phenomenon termed the *plastisphere* (Oberbeckmann & Labrenz, 2020). Metagenomic profiling of plastisphere communities consistently identifies opportunistic and potentially pathogenic taxa, including *Vibrio* spp., antibiotic-resistant *Escherichia coli*, *Pseudomonas aeruginosa*, and diverse filamentous fungi (Ding et al., 2023; Oberbeckmann & Labrenz, 2020). The hydrophobic surfaces of polyethene and polystyrene sorb dissolved organic carbon and facilitate inter-species horizontal gene transfer, creating conditions conducive

to the selection and spread of antimicrobial resistance genes (Ding et al., 2023). Crucially, because microplastics are readily advected by riverine and oceanic currents, they may transport viable pathogens over distances of hundreds to thousands of kilometres, effectively decoupling the spatial origin of contamination from the locus of infection risk (Ding et al., 2023; Mwasa & Tchuenche, 2011). Direct experimental evidence for this pathway comes from Ormsby et al. (2024), who showed that toxigenic Vibrio cholerae colonises and persists on plastic waste for at least 14 days before transferring to floodwater at concentrations sufficient to cause human infection, closing the loop between plastic-surface colonisation and human exposure that motivates the environmental compartment M(t) in the present model.

Mathematical Epidemiology and Environmental Transmission

Mathematical epidemiology provides a principled framework for disentangling competing transmission pathways and identifying threshold conditions governing whether a pathogen can invade and persist in a host population (Asamoah et al., 2022; Baleanu et al., 2022). The classical susceptible–infected–recovered (*SIR*) paradigm of Kermack and McKendrick has been extended in numerous directions to incorporate environmental pathogen reservoirs. Environmental transmission models, such as those in Capone et al. (2013) and Mwasa & Tchuenche (2011), couple human epidemic dynamics with pathogen persistence in environmental reservoirs.

Related frameworks have been developed for cholera (Capone et al., 2013; Mwasa & Tchuenche, 2011) and hepatitis E (Capone et al., 2013). However, none of these models explicitly treats the microplastic particle as the biotic substrate that concentrates pathogens and shields them from inactivation, a mechanistically distinct reservoir with substantially longer pathogen residence times compared with bulk water.

Fractional-Order Epidemic Models

Integer order ordinary differential equations encode memoryless, Markovian dynamics. This assumption is frequently violated in biological systems exhibiting anomalous diffusion, power-law immune waning, and history-dependent transmission rates (Rihan et al., 2023). Caputo fractional derivatives of order $\alpha \in (0, 1)$ embed a hereditary convolution kernel into the governing equations, allowing the derivative at time t to depend on the entire past trajectory (Baleanu et al., 2022; Ghanbari, 2020). The order parameter α continuously interpolates between purely diffusive dynamics ($\alpha = 1$, integer-order limit) and strongly sub-diffusive behaviour ($\alpha \to 0^{+}$), enabling flexible empirical calibration to incidence time series. Fractional compartmental models have been applied to a range of infectious diseases, including COVID-19 (Ali et al., 2022; Farman et al., 2022), consistently demonstrating greater fidelity to observed epidemic curves than their integer-order counterparts.

Research Gap and Original Contributions

Although significant progress has been made separately in microplastic research and fractional-order epidemic modelling, the two have rarely been combined, and existing work in each area has distinct limitations relative to the present study. Environmental-transmission models such as those for cholera (Mwasa & Tchuenche, 2011) and generic SEIR reaction–diffusion systems (Capone et al., 2013) couple human dynamics to a pathogen reservoir but treat that reservoir as an undifferentiated medium (water, soil) and use classical integer-order derivatives, so they cannot capture memory-dependent persistence. Baleanu et al. (2022) extend fractional-order modelling to a real cholera outbreak with an environmental compartment, demonstrating the value of the Caputo formulation for reservoir-mediated transmission; however, their reservoir remains a generic aquatic medium, without a mechanism for the substrate-specific protection and transport that distinguishes microplastic particles from bulk water. Rajkumar (2022) raises the conceptual possibility of a microplastic–infectious-disease link but does not develop a transmission model. Based on the literature reviewed in this study, we found very limited work combining (i) a Caputo fractional-order formulation to capture memory effects in reservoir-mediated transmission, with (ii) a microplastic-specific environmental compartment that reflects the plastisphere's role in concentrating pathogens and extending their residence time relative to bulk water or soil (Ding et al., 2023; Oberbeckmann & Labrenz, 2020). The present study develops a fractional-order SIRM model that incorporates microplastic-mediated environmental transmission to address this combined gap. The aim is to better understand how environmental contamination influences disease dynamics.

The main contributions of this work are summarised as follows:

(i) A fractional-order SIRM model is formulated by including a microplastic associated environmental compartment.

(ii) By distinguishing direct and environmental transmission effects, the next-generation matrix approach is used to determine the basic reproduction number.

(iii) Both endemic and disease-free equilibria are subjected to stability analysis.

(iv) Sensitivity analysis is used to pinpoint important factors affecting the spread of disease.

(v) The effects of fractional dynamics and environmental transmission are demonstrated by numerical simulations.

The structure of this paper is as follows. Section 2 describes the proposed model along with the underlying biological assumptions and parameter definitions. The mathematical analysis of the model is presented in Section 3. Section 4 is devoted to sensitivity analysis, while numerical simulations are discussed in Section 5. The results are further interpreted and compared with existing studies in Section 6. Finally, Section 7 concludes the paper and outlines possible directions for future work.

Caputo Fractional Derivative

We employ the Caputo fractional derivative of order $\alpha \in (0, 1]$, defined for a continuous function $f : [0, \infty) \rightarrow \mathbb{R}$ as

$$c_{D^\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\tau)^{-\alpha} f'(\tau) d\tau, \quad 0 < \alpha \leq 1, \qquad [1]$$

where $\Gamma(\cdot)$ is the Euler gamma function (Baleanu et al., 2022). At $\alpha = 1$, Equation 1 reduces to the classical first-order derivative. The Caputo formulation is preferred over the Riemann–Liouville derivative in biological modelling because it admits conventional integer-order initial conditions (Ghanbari, 2020).

Compartmental Structure and Biological Assumptions

The host population at time *t* is partitioned into:
S(t): **Susceptible** individuals, at risk of acquiring infection.
I(t): **Infected** individuals, actively infectious and shedding pathogens.
R(t): **Recovered** individuals, temporarily immune.

The total living population is $N(t) = S(t) + I(t) + R(t)$. The environmental state variable *M(t)*: Concentration of viable pathogens adsorbed onto microplastic particles (illustratively, cells·mL^{-1}) couples the abiotic environment to the human epidemiological system. With M(t) expressed in these units, β_m and ξ are reported in the conventional epidemiological units $mL \cdot cells^{-1} \cdot day^{-1}$ *and* $cells \cdot mL^{-1} \cdot person^{-1} \cdot day^{-1}$ respectively (Table 1). These are the ordinary (integer-order) units used for numerical interpretation of the rate coefficients; we do not claim strict dimensional homogeneity of the fractional-order system in the sense of fractional-adjusted units (e.g. $day^{-\alpha}$), which would require a separate convention applied uniformly across all rate parameters.

The model rests on the following biological assumptions:

(A1) New susceptible individuals enter the population at a constant recruitment rate $\Lambda > 0$, while all individuals experience natural mortality at a per-capita rate $\mu > 0$.

(A2) Susceptible individuals acquire infection through direct contact with infectious individuals at a bilinear incidence rate βSI, where $\beta > 0$ is the effective contact transmission coefficient.

(A3) Susceptible individuals acquire infection through indirect environmental contact with microplastic-borne pathogens at a rate $\beta_m SM$, where $\beta_m > 0$ is the effective microplastic-mediated environmental transmission coefficient, converting exposure to the microplastic-associated pathogen concentration into infection incidence.

(A4) Infected individuals recover at a per-capita rate $\gamma > 0$ and shed pathogens onto microplastic surfaces at a per-capita rate $\xi > 0$.

(A5) Microplastic-associated pathogens are inactivated at a rate $\delta > 0$ via desiccation, ultraviolet radiation, and ecological competition.

(A6) Recovered individuals neither contribute to environmental contamination nor suffer disease-induced mortality on the time scale of interest.

Compartmental Flow Diagram

Figure 1 shows a compartmental flow diagram of the fractional-order SIRM epidemic model. Solid arrows represent direct demographic and epidemiological flows. The dashed green arrow represents the indirect environmental transmission pathway mediated by pathogens harboured on microplastic particles.

Governing Equations

Based on assumptions (A1) to (A6) and Figure 1, the fractional-order SIRM model is given by the autonomous system in Equations (2a) to (2d):

$$c_D{}^{\propto} S(t) = \Lambda - \beta S(t)I(t) - \beta_{\mathrm{m}} S(t)M(t) - \mu S(t), \quad [2a]$$

$$c_D{}^{\propto} I(t) = \beta S(t)I(t) + \beta_{\mathrm{m}} S(t)M(t) - (\gamma + \mu)I(t), \quad [2b]$$

$$c_D{}^{\propto} R(t) = \gamma I(t) - \mu R(t), \quad [2c]$$

$$c_D{}^{\propto} M(t) = \xi I(t) - \delta M(t), \quad [2d]$$

where $0 < \alpha \leq 1$, supplemented by non-negative initial conditions as shown in Equation 3.

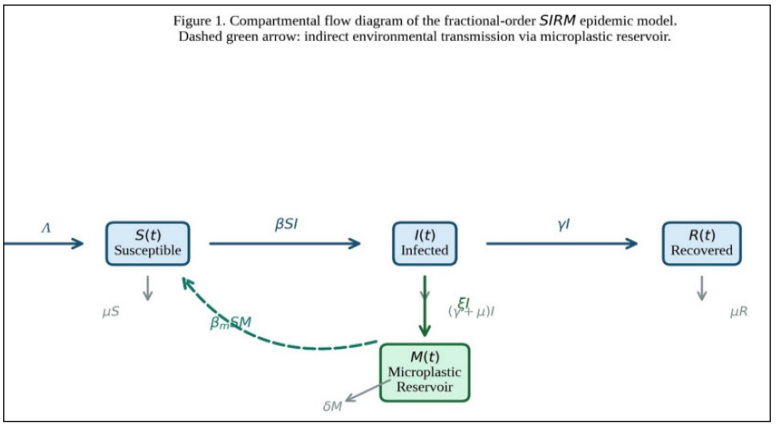


Figure 1. Compartmental flow diagram of the fractional-order *SIRM* epidemic model

$$S(0) = S_0 > 0, I(0) = I_0 \geq 0, R(0) = R_0 \geq 0, M(0) = M_0 \geq 0 \quad [3]$$

Parameter Descriptions and Baseline Values

The environmental transmission parameter, β_m, and the pathogen shedding rate ξ are not directly measurable from epidemiological surveillance. Their baseline values are instead anchored qualitatively to the plastisphere literature: Oberbeckmann and Labrenz (2022) and Ding et al. (2023) document active microbial colonisation of microplastic surfaces and pathogen residence times that are substantially extended relative to bulk water, and Zhang et al. (2022) report elevated pathogen loads on freshwater microplastics. These sources support the qualitative pattern encoded in the model (active colonisation, extended persistence) rather than a specific numerical calibration, so the baseline values in Table 1 should be interpreted as illustrative rather than population calibrated. Section 5 accordingly reports how model outcomes respond across a wide range of β_m (a 30-fold sweep, Table 3 and Figure 4) rather than a single point estimate, so that the qualitative conclusions can be assessed for robustness to this parameter uncertainty. The Caputo order α is likewise treated as illustrative and literature-informed rather than fitted; calibrating α to case-level incidence data, once microplastic-exposure time series become available, is identified as a direction for future work.

The pathogen decay rate δ has a more direct empirical anchor: Ormsby et al. (2024) report that toxigenic Vibrio cholerae colonised and persisted on plastic waste for at least 14 days before subsequent transfer to floodwater at concentrations sufficient to cause human infection. Treating this reported persistence window as an approximate indicator of the

Table 1
Model parameters: biological description, units, baseline values used in numerical simulations, and justification in the literature

Symbol	Description	Unit	Baseline	Reference
Λ	Susceptible recruitment rate	person's day^{-1}	10.00	Delavari et al. (2012)
β	Direct transmission coefficient	person^{-1} day^{-1}	7.71×10^{-4}	Delavari et al. (2012)
β_m	Microplastic transmission coefficient	mL·cells^{-1}·day^{-1}	4.83×10^{-4}	Baleanu et al. (2022); est.
γ	Recovery rate	day^{-1}	0.10	Delavari et al. (2012)
μ	Natural mortality rate	day^{-1}	0.02	Delavari et al. (2012)
ξ	Pathogen shedding rate to microplastics	cells·mL^{-1}·person^{-1}·day^{-1}	0.04	[22]; est.
δ	Pathogen decay rate on microplastics	day^{-1}	0.15	Baleanu et al. (2022); Rihan et al. (2023); etc.
α	Caputo fractional order	–	0.85	Ghanbari (2020); illustrative

timescale over which the pathogen population on plastic remains substantially viable, the baseline $\delta = 0.15$ day^{-1} used here (corresponding to an e-folding decay time of ≈ 6.7 days) is of the same order of magnitude as, though somewhat faster than, this reported persistence window; this is consistent with $\delta = 0.15$ day^{-1} being a conservative (i.e., not over-optimistic) baseline for the environmental transmission pathway.

The baseline parameter values are selected based on commonly reported ranges in the epidemiological and environmental literature. Parameters related to microplastic transmission (β_m, ξ) are estimated due to the lack of direct empirical data and are primarily used to illustrate plausible transmission scenarios rather than represent a specific calibrated system.

MATHEMATICAL ANALYSIS

Positivity of Solutions

Theorem 3.1 (non-negativity). *For any initial data satisfying (3), the solution* $(S(t), I(t), R(t), M(t))$ *of system (2) satisfies* $(S(t), I(t), R(t), M(t) \geq 0$ *for all* ≥ 0.

Proof. We apply the fractional non-negativity criterion: if $c_D{}^{\propto} x|_{x=0} \geq 0$ and $x(0) \geq 0$, then $x(t) \geq 0$ for all $t \geq 0$ (Li et al., 2019).

Susceptible class. At the boundary $S = 0$:

$c_D{}^{\propto} S|_{s=0} = \Lambda > 0$.

Hence $S(t)$ cannot become negative.

Infected class. At $I = 0$:

$c_D{}^{\propto} I|_{I=0} = \beta_m S(t) M(t) \geq 0$, since $S, M \geq 0$ by their respective boundary arguments.

Recovered class. At $R = 0$:

$c_D{}^{\propto} R|_{R=0} = \gamma I(t) \geq 0$.

Environmental compartment. At $M = 0$:

$c_D{}^{\propto} M|_{M=0} = \xi I(t) \geq 0$.

All four state variables satisfy the non-negativity criterion, so the positive orthant $\mathbb{R}^4_+$ is positively invariant under the flow of system (2).

Uniform Boundedness

Theorem 3.2 (Uniform boundedness). *Every solution of system (2) originating in* $\mathbb{R}^4_+$ *is uniformly bounded and ultimately confined to the compact absorbing set.*

$$\Omega = \left\{(S, I, R, M) \in \mathbb{R}^4_+ : N(t) \leq \frac{\Lambda}{\mu}, M(t) \leq \frac{\xi\Lambda}{\mu\delta}\right\}.$$

Proof. (For initial conditions satisfying $N(0) \leq \Lambda/\mu$ and $M(0) \leq \xi\Lambda/(\mu\delta)$, the set Ω is positively invariant; for arbitrary nonnegative initial data, solutions remain uniformly bounded and approach these bounds asymptotically, as established below.) Define $N(t) = S(t) + I(t) + R(t)$. Summing Equations (2a)–(2c), we obtain Equation 4:

$$c_{D^\alpha} N(t) = \Lambda - \mu N(t). \quad [4]$$

By the fractional comparison principle (Li et al., 2019):

$$N(t) \leq \left(N(0) - \frac{\Lambda}{\mu}\right) E_\alpha(-\mu t^\alpha) + \frac{\Lambda}{\mu},$$

where $E_\alpha(\cdot)$ is the one-parameter Mittag-Leffler function. Since $E_\alpha(-\mu t^\alpha) \to 0$ as $t \to \infty$ (Li et al., 2019), We obtain lim $sup_{t\to\infty} N(t) \leq \frac{\Lambda}{\mu}$.

For the environmental compartment, equation (2d) yields $c_{D^\alpha} M \leq \xi\Lambda/\mu - \delta M$, giving lim $sup_{t\to\infty} M(t) \leq \xi\Lambda/(\mu\delta)$ by the same argument. Therefore, Ω is a compact, positively invariant, absorbing set.

Disease-Free Equilibrium

At the disease-free equilibrium (DFE) ε_0, The disease is absent $I^* = M^* = 0$. Setting the right-hand sides of Equation (2) to zero and solving, we obtain the disease-free equilibrium given in Equation 5:

$$\varepsilon_0 = \left(\frac{\Lambda}{\mu}, 0, 0, 0\right) \quad [5]$$

This state represents a fully susceptible population free of infection and environmental contamination.

Basic Reproduction Number via the Next-Generation Matrix

Theorem 3.3 (Basic reproduction number). The basic reproduction number of system (2), derived by the next-generation matrix method (van den Driessche, 2017), is given by Equation 6:

$$R_0 = \frac{\beta\Lambda}{\mu(\gamma+\mu)} + \frac{\beta_m\xi\Lambda}{\mu\delta(\gamma+\mu)}, \quad [6]$$

where $R_0^d = \frac{\beta\Lambda}{\mu(\gamma+\mu)}$ is the direct transmission component and $R_0^e = \frac{\beta_m \xi\Lambda}{\mu\delta(\gamma+\mu)}$ is the microplastic-mediated environmental transmission component.

Proof. The infected sub-system comprises compartments I and M. Writing $X = (I, M)^T$, the new infection matrix F and the transition matrix V evaluated at $\varepsilon_0 = (S^*, 0, 0, 0), S^* = \frac{\Lambda}{\mu}$, are given in Equation 7.

$$F = \begin{pmatrix} \beta S^* & \beta_m S^* \\ 0 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} \gamma+\mu & 0 \\ -\xi & \delta \end{pmatrix} \quad [7]$$

The inverse of V is given in Equation 8.

$$V^{-1} = \frac{1}{\delta(\gamma+\mu)} \begin{pmatrix} \delta & 0 \\ \xi & \gamma+\mu \end{pmatrix} \quad [8]$$

The next-generation matrix is given in Equation 9:

$$K = FV^{-1} = \frac{S^*}{\delta(\gamma+\mu)} \begin{pmatrix} \beta\delta + \beta_m \xi & \beta_m(\gamma+\mu) \\ 0 & 0 \end{pmatrix} \quad [9]$$

The spectral radius $R_0 = \rho(K)$ is the single non-zero eigenvalue, yielding expression (6). Since $R_0^e > 0$ whenever $\beta_m, \xi > 0$. The microplastic reservoir always amplifies beyond the direct-contact-only baseline R_0^d.

Local Stability of the Disease-Free Equilibrium

Theorem 3.4 (LAS of the DFE). The disease-free equilibrium E_0 of system (2) is locally asymptotically stable (LAS) if $R_0 < 1$, and unstable if $R_0 > 1$.

Proof. The Jacobian of system (2) at ε_0 is given in Equation 10:

$$J(\varepsilon_0) = \begin{pmatrix} -\mu & \beta S^* & 0 & -\beta_m S^* \\ 0 & \beta S^* - (\gamma+\mu) & 0 & \beta_m S^* \\ 0 & \gamma & -\mu & 0 \\ 0 & \xi & 0 & -\delta \end{pmatrix} \quad [10]$$

Two eigenvalues are evident: $\lambda_1 = \lambda_3 = -\mu < 0$. The remaining pair is determined by the reduced 2 × 2 sub-matrix

$$J_r = \begin{pmatrix} \beta S^* - (\gamma+\mu) & \beta_m S^* \\ \xi & -\delta \end{pmatrix}$$

The characteristic polynomial $p(\lambda) = \lambda^2 - tr(J_r)\lambda + det(J_r)$ has coefficients:

$$tr(J_r) = \beta S^* - (\gamma + \mu + \delta);$$

$$det(J_r) = \delta(\gamma + \mu)(1 - R_0).$$

When $R_0 < 1$: $det(J_r) > 0, and\ tr(J_r) < 0$ (since $\beta S^* = (\gamma + \mu)R_0^d < (\gamma + \mu)$ when $R_0 < 1$). The Routh–Hurwitz criterion guarantees both eigenvalues of J_r have negative real parts. For fractional-order systems, LAS requires for all eigenvalues (Delavari et al., 2012). Because all eigenvalues here are real and negative,

$|arg(\lambda_i)| = \pi > \alpha\pi/2$ for any $\alpha \in (0, 1]$. Hence, E_0 is LAS when $R_0 < 1$. When $R_0 > 1$: $det(J_r) < 0$, guaranteeing one positive real eigenvalue, so is unstable.

Existence and Uniqueness of the Endemic Equilibrium

Theorem 3.5 (Endemic equilibrium). System (2) admits a unique endemic equilibrium $E^* = (S^*, I^*, R^*, M^*) \in \text{int}(\Omega)$ with $I^* > 0$ if and only if $R_0 > 1$.

Proof. Setting $_{cD^\alpha} M = 0$ in (2d): $M^* = \xi I^*/\delta$. For $I^* > 0$, equation (2b) at steady state gives $S^*(\beta + \beta_m \xi/\delta) = \gamma + \mu$, hence

$$S^* = \frac{\delta(\gamma + \mu)}{\beta\delta + \beta_m \xi}.$$

Then, since at equilibrium $\Lambda = \mu S^* + (\gamma + \mu) I^*$ (obtained by summing (2a) and (2b) at steady state), Substituting S* above and solving for I* gives Equation 11:

$$I^* = \frac{\mu\delta(R_0 - 1)}{\beta\delta + \beta_m \xi} \qquad [11]$$

which is strictly positive if and only if $R_0 > 1$. The remaining components follow: , $R^* = \gamma I^*/\mu, M^* = \xi I^*/\delta, and\ S^* > 0$.

Theorem 3.6 (LAS of the endemic equilibrium). *If $R_0 > 1$, the endemic equilibrium E^* of system (2) is locally asymptotically stable, for every fractional order $\alpha \in (0, 1]$.*

Proof. Since the R-equation decouples from (S, I, M) and contributes only the eigenvalue $-\mu < 0$, it suffices to analyse the reduced (S, I, M)-subsystem. Let $\lambda^* = \beta I^* + \beta_m M^*$ denote the force of infection at E^*; the equilibrium condition from (2b) gives $\lambda^* S^* = (\gamma + \mu) I^*$. Define the positive constant $q = \frac{\beta_m \xi S^*}{\delta}$ (using $M^* = \frac{\xi I^*}{\delta}$). Linearising (2a), (2b), (2d) at E^*, and using $\beta S^* - (\gamma + \mu) = -q$ (obtained by substituting λ^* into the equilibrium relation), gives the Jacobian

$$J(E^*) = \begin{bmatrix} -(\lambda^* + \mu) & -\beta S^* & -\beta_m S^* \\ \lambda^* & -q & \beta_m S^* \\ 0 & \xi & -\delta \end{bmatrix}$$

whose characteristic polynomial $s^3 + a_1 s^2 + a_2 s + a_3$ has coefficients (after simplifying using $M^* = \xi I^*/\delta$):

$$a_1 = \delta + \lambda^* + \mu + q,\ a_2 = \beta S^* \lambda^* + \delta\lambda^* + \delta\mu + \lambda^* q + \mu q,\ a_3 = \delta\lambda^*(\beta S^* + q)$$

Since $R_0 > 1$ implies $I^* > 0$ and hence $\lambda^* > 0$, and all model parameters are strictly positive, every term in a_1, a_2, a_3 is a product of positive quantities: $a_1 > 0$ and $a_3 > 0$ follow immediately. Expanding $a_1 a_2 - a_3$ yields

$$a_1 a_2 - a_3 = \beta S^* \lambda^{*2} + \beta S^* \lambda^* \mu + \beta S^* \lambda^* q + \delta^2 \lambda^* + \delta^2 \mu + \delta \lambda^{*2} + 2\delta\lambda^* \mu + \delta\lambda^* q + \delta\mu^2 + 2\delta\mu q + \lambda^{*2} q + 2\lambda^* \mu q + \lambda^* q^2 + \mu^2 q + \mu q^2,$$

a sum of strictly positive terms, so $a_1 a_2 > a_3$. By the Routh–Hurwitz criterion, all three eigenvalues of $J(E^*)$ have a negative real part. Following the same fractional-stability argument as Theorem 3.4 (Delavari et al., 2012), $|arg(\lambda_i)| > \pi/2 \geq \alpha\pi/2$ holds for every $\alpha \in (0,1]$, so E^* is locally asymptotically stable whenever $R_0 < 1$, independent of the fractional order.

Global Stability of the Disease-Free Equilibrium

Theorem 3.7 (GAS of the DFE). *If $R_0 < 1$, the disease-free equilibrium ε_0 is globally asymptotically stable in Ω.*

Proof. Define the Lyapunov function

$$V(t) = I(t) + \frac{\beta_m S^*}{\delta} M(t) \geq 0, V = 0 \Leftrightarrow I = M = 0$$

Since V is a linear combination of the state variables and the Caputo derivative is a linear operator, $c_{D^\alpha} V = c_{D^\alpha} I + \left(\frac{\beta_m S^*}{\delta}\right) c_{D^\alpha} M$ holds exactly, with no additional regularity condition required. Substituting the governing Equations (2b) and (2d):

$$\begin{aligned} {}^{C}D^{\alpha} V &= \big(\beta S - (\gamma + \mu)\big) I + \beta_m S M + \tfrac{\beta_m S^*}{\delta}(\xi I - \delta M) \\ &= \left(\beta S - (\gamma + \mu) + \tfrac{\beta_m S^*}{\delta}\right) I + \beta_m (S - S^*) M \end{aligned}$$

Since $S \leq S^*$ in Ω, the term $\beta_m(S - S^*)M \leq 0$, Moreover$\beta S \leq \beta S^*$, so ${}^{C}D^{\alpha}V \leq (\gamma + \mu)(R_0 - 1)I \leq 0$ when $R_0 < 1$

For $R_0 < 1, c_{D^{\alpha}}V \leq 0$. The largest invariant subset of the set where $c_{D^{\alpha}}V = 0$ is the disease-free equilibrium ε_0. Indeed, invariance requires $I = 0$; if $M > 0$, Equation (2b) gives $c_{D^{\alpha}}I = \beta_m SM > 0$, so the trajectory leaves the zero-derivative set. Hence $M = 0$, after which $R(t) \to 0$ and $S(t) \to \Lambda/\mu$. By the fractional-order LaSalle invariance principle for Volterra-type Lyapunov functions (Vargas-De-León, 2015), all trajectories in Ω therefore converge to ε_0, so ε_0 is globally asymptotically stable.

SENSITIVITY ANALYSIS OF R_0

The normalised forward sensitivity index of $R_0$with respect to a parameter p quantifies the fractional change in R_0 per unit fractional change in p (Omame et al., 2021):

$$\Upsilon_p^{R_0} = \frac{\partial R_0}{\partial p} \cdot \frac{p}{R_0} \qquad [12]$$

Using expression (6), the analytical sensitivity indices are evaluated at the baseline parameters in Table 1 (where $R_0^d/R_0 \approx 0.857$ and $R_0^e/R_0 \approx 0.143$). Table 2 shows a normalised forward sensitivity indices of R_0 evaluated at baseline parameter values (Table 1). A positive (negative) index means that increasing the parameter increases (decreases) R_0; the magnitude indicates the proportional impact.

The indices show that μ, Λ, γ, and β. These are the highest-leverage parameters. Among the microplastic-specific parameters, β_m and ξ carry equal positive indices of 0.143: a 10% increase in either raises R_0 by approximately 1.43%. The pathogen decay rate δ carries an equal but negative index, confirming that accelerating inactivation of microplastic-associated pathogens is a valid, though secondary, intervention strategy.

Table 2
Normalised forward sensitivity indices of R_0 evaluated at baseline parameter values (Table 1)

Parameter	Description	Index Υ_{P_0}	Epidemiological Impact
Λ	Recruitment rate	+1.000	Highest amplifier
β	Direct transmission	+0.857	High amplifier
β_m	Environmental transmission	+0.143	Moderate amplifier
ξ	Pathogen shedding	+0.143	Moderate amplifier
μ	Natural mortality	−1.167	Highest reducer
γ	Recovery rate	−0.833	High reducer
δ	Pathogen decay	−0.143	Moderate reducer

It should be noted that the baseline parameter values used in the numerical simulations are selected to illustrate the qualitative dynamics of the proposed model rather than to represent a calibrated fit to a specific epidemiological dataset. Several parameters related to environmental transmission, including the microplastic transmission coefficient β_m and the pathogen shedding rate ξ are estimated based on available microbiological studies of microplastic-associated microbial colonisation. The direct transmission coefficient β is adopted from commonly used ranges for epidemic modelling in the literature. The numerical experiments are therefore intended to demonstrate the qualitative influence of fractional memory effects and microplastic-mediated environmental transmission on epidemic dynamics.

NUMERICAL SIMULATION

Numerical Method

System (2) is integrated numerically using the Adams–Bashforth–Moulton (ABM) predictor-corrector algorithm for Caputo fractional differential equations (Diethelm, Ford & Freed, 2002). The ABM scheme achieves convergence of order 1 + α (Diethelm, Ford & Freed, 2002). Our implementation of the Diethelm-Ford-Freed predictor-corrector scheme was independently validated against the exact Mittag-Leffler solution of the linear test equation $D^{\alpha} y = -y$, $y(0) = 1$, confirming correct convergence prior to application to system (2). All parameter values and initial conditions reported in Table 1 and below were used directly to generate every figure and table in this section, so that R_0 computed from equation (6) and R_0 realised in the simulations coincide exactly. All computations use step size h = 0.1 days over a simulation horizon of T = 200 days, with baseline parameters from Table 1 and initial conditions:

$$S(0) = 490,\ I(0) = 10,\ R(0) = 0,\ M(0) = 0$$

This initial condition represents the onset of an outbreak from a small, infected seed introduced into an otherwise disease-free population already at its demographic equilibrium, $N(0) = \Lambda/\mu = 500$, with no prior environmental contamination.

Baseline Epidemic Dynamics $R_0 \approx 3.75$

Figure 2 shows the temporal evolution of all four compartments under baseline parameters, corresponding to an endemic scenario ($R_0 \approx 3.75 > 1$). The susceptible population declines steeply during the exponential growth phase. The infected class reaches its peak around day 23, with about 184 people, then declines as the number of people who can get sick decreases and people keep getting better. The recovered class rises steadily until it reaches

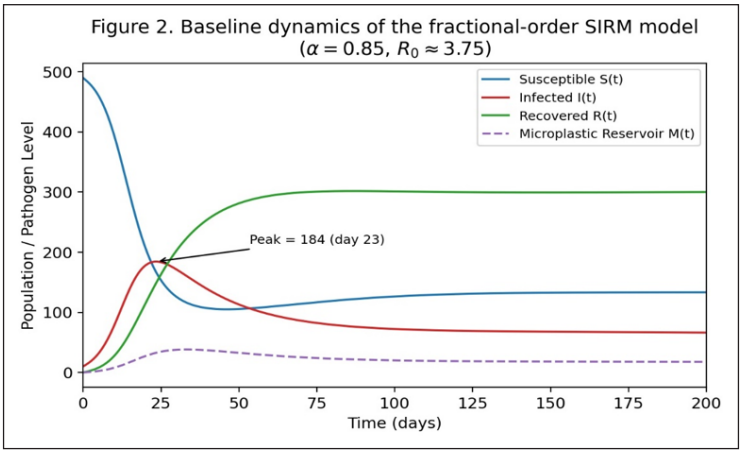


Figure 2. Temporal dynamics of all four compartments under baseline parameter values (Table 1) with fractional order α = 0.85. The infected population peaks near day 23 at approximately 184 individuals, while the microplastic reservoir $M(t)$ tracks the infected class with a short lag

its maximum level. The microplastic reservoir $M(t)$ follows $I(t)$. With a short delay, which shows the balance between shedding and decay at the moment in equation (2d).

Effect of Fractional Order α

Figure 3 shows how the fractional order α affects the spread of infections. As α decreases from 1.0 to 0.75, the infection peak slightly lowers in height, and the epidemic tail gets longer. This is because sub-integer Caputo derivatives encode sub-diffusive memory. When $\alpha = 1.0$, the system becomes the traditional integer-order SIRM model.

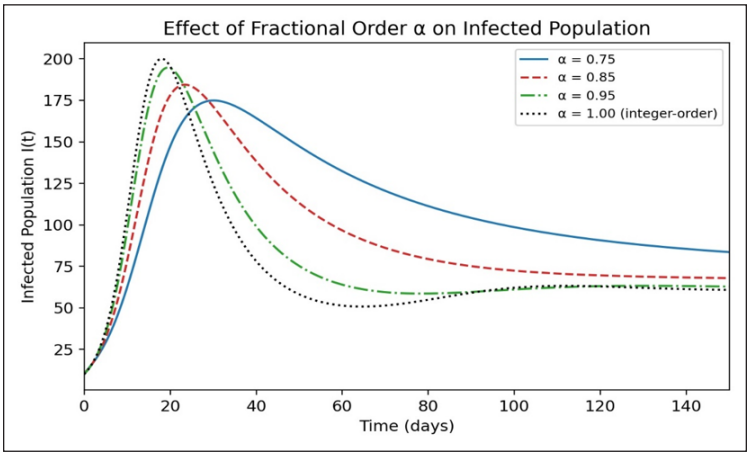


Figure 3. Effect of the Caputo fractional order α on the infected population dynamics. Decreasing α modestly reduces the infection peak and broadens the epidemic tail, consistent with the sub-diffusive memory effects encoded by the Caputo derivative. At $\alpha = 1.0$, the model reduces to the classical integer-order SIRM system

Effect of Microplastic Transmission Rate β_m

Figure 4 shows how the microplastic transmission coefficient β_m affects the spread of disease. When β_m goes up from 1×10^{-4} to 3×10^{-3}. The peak of the infection goes up from about 174 to 231 people and moves forward in time. This quantifies the epidemiological cost of environmental microplastic contamination: increasing β_m produces a substantial rise in the infection peak.

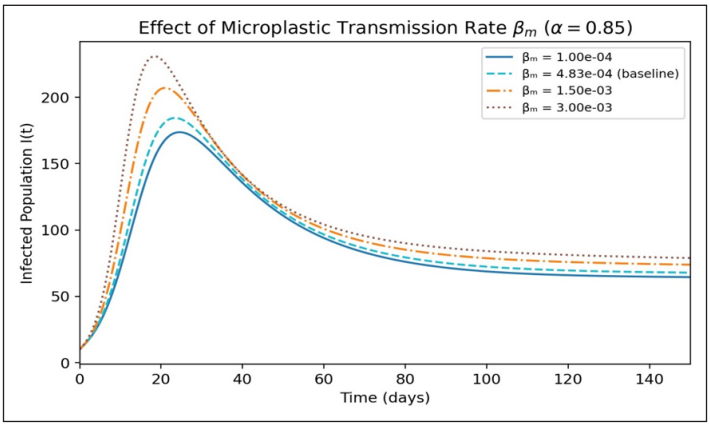


Figure 4. Effect of the microplastic-mediated transmission rate β_m on the infected population ($\alpha = 0.85$ all other parameters at baseline). Higher β_m elevates the peak infected population and prolongs the outbreak, quantifying the epidemiological amplification attributable to the environmental microplastic reservoir

Phase-Plane Analysis

Figure 5 shows phase portraits in the $S - I$ plane for three different values of α. For $R_0 \approx 3.75 > 1$, all trajectories spiral toward the endemic equilibrium EE, irrespective of α, thereby

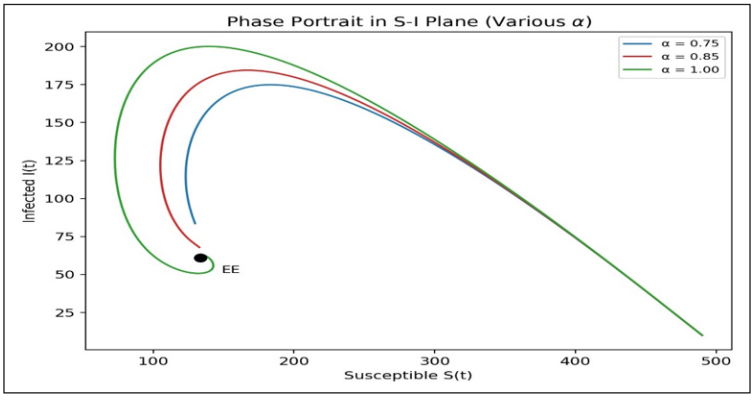


Figure 5. Phase portrait in the $S - I$ plane for $\alpha \in \{0.75, 0.85, 1.00\}$ with $R_0 \approx 3.75 > 1$. All trajectories converge to the endemic equilibrium EE, confirming Theorem 3.6. The transient path depends weakly on α, with lower α producing a broader approach to EE

validating the LAS result of Theorem 3.6. The spiral's curvature changes with α, which shows how the fractional memory changes the way the system approaches equilibrium.

Multi-Parameter Uncertainty Analysis

The single parameter sweeps above vary β_m (or α) individually. To assess robustness under simultaneous parameter uncertainty, we conducted a Latin Hypercube Sampling (LHS) analysis over the five parameters most directly tied to model uncertainty (β, β_m, ξ, δ, γ), each varied independently across ±50% of its baseline value (Table 1), with $N = 1000$ samples. For each sample, we computed R_0 from equation (6), and for a subsample of $N = 200$. We additionally solved the full fractional system (2) to obtain the peak infected population I_max. Partial rank correlation coefficients (PRCCs) were then computed for each parameter between each output, holding the others constant in rank.

Figure 6 presents the partial rank correlation coefficients (PRCCs) for the basic reproduction number R_0 and the peak infected population I_{max} under simultaneous parameter uncertainty. Across the LHS ensemble, R_0 ranged from 1.40 to 9.08 (mean 4.05), and I_max ranged from 34 to 314 (mean 186), confirming that the baseline qualitative conclusions are not an artefact of the specific point-estimate parameter values. β and γ dominate both outputs $(|PRCC| > 0.95)$, consistent with the analytical sensitivity indices in Table 2, while β_m, ξ, and δ each retains a moderate, statistically distinguishable influence $(|PRCC| \approx 0.2 - 0.5)$, confirming that the microplastic-specific pathway remains a non-negligible driver of both threshold behaviour and outbreak severity even after accounting for uncertainty in the other parameters.

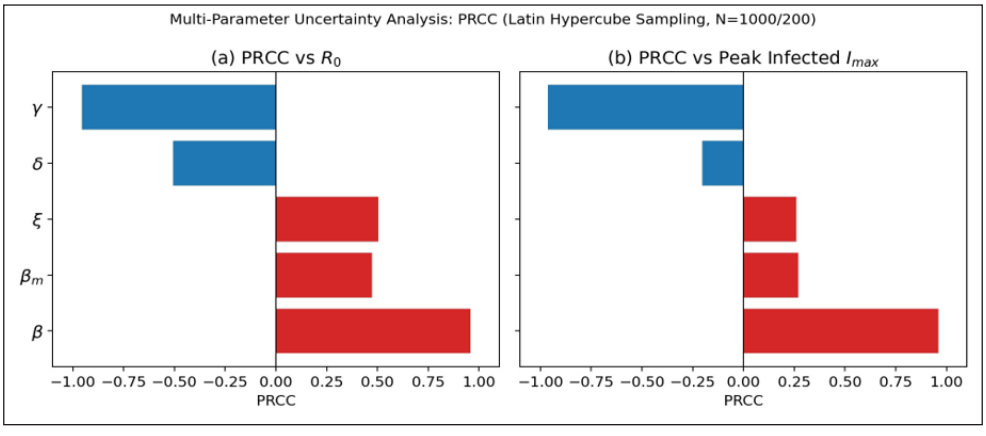


Figure 6. Partial rank correlation coefficients (PRCC) between each of the five varied parameters and (a) R_0, (b) the peak infected population I_max, under Latin Hypercube Sampling ($N = 1000$ *for* $R_0, N = 200$) with each parameter varied over ±50% of its Table 1 baseline

Simulation Data Summary

Table 3 presents the simulated infected population I(t) at selected time points for different fractional orders α and microplastic transmission coefficients β_m, as shown in Figures 3 and 4.

DISCUSSION

In this study, we developed and analysed a Caputo fractional-order SIRM epidemic model that links human infection dynamics with an environmental reservoir of pathogens carried by microplastic particles. The analytical and numerical results help explain how microplastic-mediated transmission can influence disease spread and persistence.

Decomposition of R_0 and the epidemiological role of microplastics

An important analytical result of this study is the additive decomposition (6) of the basic reproduction number into a direct human-to-human component and a microplastic environmental component R_0^e. Under the baseline parameter set, $R_0^e/R_0 \approx 14\%$, which indicates that the environmental pathway makes a non-negligible contribution to transmission. This suggests that even if control measures reduce R_0^d to a lower level, the residual contribution from R_0^e may still support persistence of infection when environmental contamination remains present. Similar concerns have been discussed in studies on environmentally mediated transmission, particularly for water-borne pathogens (Mwasa & Tchuenche, 2011), while recent work has also highlighted the potential role of microplastics

Table 3
Simulated infected population I(t) at selected time points for varying α (left) and varying β_m (right). Baseline parameters from Table 1; values rounded to the nearest integer. These representative trajectories correspond to Figures 3 and 4

Day	0.75	0.85	0.95	1.00	1×10^{-4}	4.83×10^{-4}	1.5×10^{-3}	3×10^{-3}
0	10	10	10	10	10	10	10	10
5	29	32	35	37	31	32	36	43
10	62	78	97	107	71	78	99	130
15	108	140	171	186	126	140	175	216
20	147	178	194	195	164	178	206	229
25	169	183	174	161	174	183	199	207
30	175	171	144	125	165	171	179	181
40	164	138	99	78	136	138	141	142
50	147	113	75	58	111	113	116	118
60	132	97	64	51	94	97	101	104
80	111	79	59	55	76	79	85	90
100	99	72	61	62	69	72	79	84

as carriers of pathogenic microorganisms (Ding et al., 2023; Oberbeckmann & Labrenz, 2020; Rajkumar, 2022).

Epidemiological implications of β_m and the shedding rate ξ

Sensitivity analysis identifies β_m and ξ as the microplastic-specific parameters with the highest leverage on R_0, each carrying a normalised index of +0.143. The numerical experiments in Section 5 demonstrate that increasing β_m from 1×10^{-4} to 3×10^{-3} increases the infection peak from approximately 174 to 231 individuals, highlighting the amplification of epidemic severity associated with stronger environmental transmission. The epidemiological implications can be interpreted as follows:

In practice, β_m may be reduced by limiting plastic waste entering aquatic environments, improving wastewater treatment to remove microplastics, and reducing direct exposure to contaminated water sources (Rajkumar, 2022; Wu et al., 2023). The shedding rate ξ reflects the contribution of infected individuals to environmental contamination, and reducing this parameter may help decrease pathogen accumulation in the microplastic reservoir. The combined effect of $\beta_m$and ξ highlights the importance of environmental control strategies alongside traditional epidemiological interventions.

Memory effects and the fractional order α

The epidemic process's memory is demonstrated by the fractional order α. The model assigns greater weight to previous states when $\alpha < 1$. This enables it to explain long-lasting or delayed effects, which are typical in actual transmission dynamics. Reducing α from 1.0 to 0.75 in the current simulations prolongs the epidemic tail and somewhat reduces the infection peak. Studies on COVID-19 and related infectious diseases have shown a similar tendency in other fractional epidemic models (Ali et al., 2022; Farman et al., 2022; Rihan et al., 2023). Estimating α from observed epidemiological data would improve the model's practical relevance and help connect the fractional formulation to measurable outbreak behaviour. Table 3 already provides a direct integer-order comparison, since $\alpha - 1.00$ is the classical integer-order limit of system (2). At the peak (day 20), the fractional-order case $\alpha = 0.75$ predicts $I(20) \approx 147$ versus $I(20) \approx 195$ for the integer-order case $\alpha - 1.00$, a reduction of approximately 25%. More strikingly, the two cases diverge in the opposite direction later: by day 60 the fractional case still shows $I(60) \approx 132$ while the integer-order case has already relaxed to $I(60) \approx 51$, having passed through a damped oscillation toward the endemic equilibrium (visible in Figure 3). This illustrates concretely how sub-diffusive memory both lowers the initial peak and substantially prolongs the epidemic tail relative to the classical model, rather than merely shifting the curve.

Comparison with existing frameworks

The present model is related to earlier epidemic models that incorporate environmental transmission, but it differs in several important aspects. First, the environmental compartment in this study explicitly represents pathogens associated with microplastic particles, providing a more direct framework to analyse this emerging transmission pathway (Ding et al., 2023; Oberbeckmann & Labrenz, 2020; Rajkumar, 2022). Second, the model is formulated using Caputo fractional-order derivatives, which allow memory effects to be included in the system dynamics (Ali et al., 2022; Rihan et al., 2023). In contrast to spatial or reaction–diffusion epidemic models (Capone et al., 2013), the present model is spatially homogeneous and is intended to capture population-level behaviour rather than spatial spread.

LIMITATIONS

Several simplifications warrant acknowledgement. The model assumes homogeneous mixing and neglects age structure, network heterogeneity, and spatial dispersion, all of which demonstrably influence epidemic dynamics. The bilinear incidence βSI may overestimate transmission at high prevalence; saturating incidence of Holling type II form $\beta SI/(1 + kI)$. It is a more realistic alternative. Parameter values for β_m and ξ are estimated from the microbiological literature rather than from direct epidemiological fitting, and calibration against field data for a specific pathogen microplastic combination (e.g. *Vibrio cholerae* on polyethene fragments) remains a critical future step. Disease-induced mortality, waning immunity, multiple pathogen serotypes, and seasonal forcing of transmission rates are excluded for analytical tractability.

CONCLUSION

We have developed and analysed a Caputo fractional-order SIRM model that explicitly incorporates microplastic particles as an environmental reservoir of pathogens. To the best of our knowledge, this is among the early attempts to integrate microplastic-mediated transmission into an epidemic modelling framework. The main findings of this study can be summarised as follows:

(i) The feasible region $\Omega \subset \mathbb{R}^4_+$. It is shown to be compact and positively invariant, and all solutions remain uniformly bounded over time. This confirms that the model is biologically meaningful and mathematically well-posed.

(ii) The basic reproduction number is expressed as $R_0 = R_0^d + R_0^e$, which clearly separates the contributions from direct transmission and environmental transmission. This decomposition indicates that the presence of microplastics consistently increases the overall transmission potential.

(iii) The disease-free equilibrium is globally asymptotically stable when $R_0 < 1$, (Theorem 3.7). A unique endemic equilibrium exists whenever $R_0 > 1$ (Theorem 3.5) and is locally asymptotically stable for every fractional order $\alpha \in (0,1]$ (Theorem 3.6), so local stability of E^* holds unconditionally on parameter values once $R_0 < 1$.

(iv) Sensitivity analysis suggests that the recovery rate γ, recruitment rate Λ, and direct transmission coefficient β play dominant roles in shaping disease dynamics. In addition, the parameters β_m and ξ, which are directly associated with microplastic-mediated transmission, appear to be important from a practical intervention perspective.

(v) The numerical simulations support the analytical results and reveal that fractional-order dynamics tend to prolong the persistence of infection. Furthermore, higher values of β_m lead to noticeable increases in infection peaks, highlighting the epidemiological importance of environmental contamination.

The results of this work suggest that microplastic pollution may play a non-negligible role in the spread and persistence of infectious diseases. This observation points to the need for integrating environmental control strategies with traditional public health measures. Several possible extensions can be considered in future studies. These include incorporating spatial effects through reaction–diffusion formulations, developing stochastic fractional models to capture uncertainty in small populations, and applying optimal control techniques to balance intervention strategies. Further work may also explore multi-pathogen interactions and validate the model using epidemiological and environmental data.

Overall, the proposed framework provides a useful step toward connecting environmental pollution with infectious disease modelling and opens new directions for interdisciplinary research at the intersection of mathematical epidemiology and environmental science.

ACKNOWLEDGEMENT

The authors would like to thank the Academy of Maritime Education and Training (AMET University), Chennai, India, for its institutional support.

REFERENCES

Ali, A., Islam, S., Khan, M. R., Rasheed, S., Allehiany, F. M., Baili, J., Khan, M. A., & Ahmad, H. (2022). Dynamics of a fractional order Zika virus model with mutant. *Alexandria Engineering Journal, 61*(6), 4821-4836. https://doi.org/10.1016/j.aej.2021.10.031

Asamoah, J. K. K., Okyere, E., Abidemi, A., Moore, S. E., Sun, G.-Q., Jin, Z., Acheampong, E., & Gordon, J. F. (2022). Optimal control and comprehensive cost-effectiveness analysis for COVID-19. *Results in Physics, 33*, Article 105177. https://doi.org/10.1016/j.rinp.2022.105177

Baleanu, D., Ghassabzade, F. A., Nieto, J. J., & Jajarmi, A. (2022). On a new and generalized fractional model for a real cholera outbreak. *Alexandria Engineering Journal, 61*(11), 9175-9186. https://doi.org/10.1016/j.aej.2022.02.054

Capone, F., De Cataldis, V., & De Luca, R. (2013). On the nonlinear stability of an epidemic SEIR reaction-diffusion model. *Ricerche di Matematica, 62*(1), 161-181. https://doi.org/10.1007/s11587-013-0151-y

Delavari, H., Baleanu, D., & Sadati, J. (2012). Stability analysis of Caputo fractional-order nonlinear systems revisited. *Nonlinear Dynamics, 67*(4), 2433-2439. https://doi.org/10.1007/s11071-011-0157-5

Diethelm, K., Ford, N. J., & Freed, A. D. (2002). A predictor-corrector approach for the numerical solution of fractional differential equations. *Nonlinear Dynamics, 29*(1-4), 3-22. https://doi.org/10.1023/a:1016592219341

Ding, J., Huang, Z., Liu, C., Li, L., Zhang, S., & Zhu, X. (2023). Microplastics as vectors of pathogenic microorganisms in aquatic environments: Occurrence, mechanisms, and management strategies. *Science of the Total Environment, 880*, Article 163299.

Farman, M., Akgül, A., Nisar, K. S., Ahmad, D., Ahmad, A., Kamangar, S., & Saleel, C. A. (2022). Epidemiological analysis of fractional order COVID-19 model with Mittag-Leffler kernel. *AIMS Mathematics, 7*(1), 756-783. https://doi.org/10.3934/math.2022046

Ghanbari, B. (2020). On the modeling of the interaction between tumor growth and the immune system using some new fractional and fractional-fractal operators. *Advances in Difference Equations, 2020*(1), Article 585. https://doi.org/10.1186/s13662-020-03040-x

Li, H., Cheng, J., Li, H.-B., & Zhong, S.-M. (2019). Stability analysis of a fractional-order linear system described by the Caputo–Fabrizio derivative. *Mathematics, 7*(2), Article 200. https://doi.org/10.3390/math7020200

Mwasa, A., & Tchuenche, J. M. (2011). Mathematical analysis of a cholera model with public health interventions. *Biosystems, 105*(3), 190-200. https://doi.org/10.1016/j.biosystems.2011.04.001

Oberbeckmann, S., & Labrenz, M. (2020). Marine microbial assemblages on microplastics: Diversity, adaptation, and role in degradation. *Annual Review of Marine Science, 12*, 209-232. https://doi.org/10.1146/annurev-marine-010419-010633

Omame, A., Abbas, M., & Onyenegecha, C. P. (2021). A fractional-order model for COVID-19 and tuberculosis co-infection using Atangana–Baleanu derivative. *Chaos, Solitons & Fractals, 153*, Article 111486. https://doi.org/10.1016/j.chaos.2021.111486

Ormsby, M. J., Woodford, L., White, H. L., Fellows, R., Oliver, D. M., & Quilliam, R. S. (2024). Toxigenic Vibrio cholerae can cycle between environmental plastic waste and floodwater: Implications for environmental management of cholera. *Journal of Hazardous Materials, 461*, Article 132492. https://doi.org/10.1016/j.jhazmat.2023.132492

Rajkumar, R. P. (2022). Microplastics and infectious diseases: Is there a link? *Journal of Water and Health, 20*(2), 214-225.

Rihan, F. A., Kandasamy, U., Alsakaji, H. J., & Sottocornola, N. (2023). Dynamics of a fractional-order delayed model of COVID-19 with vaccination efficacy. *Vaccines, 11*(4), Article 758. https://doi.org/10.3390/vaccines11040758

van den Driessche, P. (2017). Reproduction numbers of infectious disease models. *Infectious Disease Modelling, 2*(3), 288-303. https://doi.org/10.1016/j.idm.2017.06.002

Vargas-De-León, C. (2015). Volterra-type Lyapunov functions for fractional-order epidemic systems. *Communications in Nonlinear Science and Numerical Simulation, 24*(1-3), 75-85. https://doi.org/10.1016/j.cnsns.2014.12.013

Wu, P., Huang, Y., Zheng, H., Chen, G., Sun, X., & Cai, Z. (2023). Microplastics in the aquatic environment: Occurrence, fate, environmental risk and management. *Science of the Total Environment, 905*, Article 167028.

Zhang, Y., Lu, X., Liu, W., Li, Z., Yu, H., Cao, G., & Zhao, L. (2022). Microplastic-associated pathogens in freshwater environments: Distribution, interactions and health implications. *Environmental Pollution, 312*, Article 119966.

Pertanika J. Sci. & Technol. 34 (5): 2835 - 2885 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review Article

Wireless Sensor Network and Leveraging IoT on Agro Farming: A Review

Muhammad Azwan Nasirudin, Mohammad Nishat Akhtar, and Farzad Ismail*

School of Aerospace Engineering, Tuanku Syed Sirajuddin Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Pulau Pinang, Malaysia

ABSTRACT

Precision agriculture is an integrated, sensor-fusion-based framework for agricultural management and monitoring, adaptable to both small-scale consumer plots and large-scale commercial farming operations. This paradigm addresses the multi-faceted limitations of traditional, labour-intensive farming practices by replacing them with continuous monitoring, closed-loop sensor networks, and advanced computing systems. Specifically, Wireless Sensor Networks (WSNs) have emerged as a prominent mechanism for real-time environmental measurement and data acquisition, enabling timely interventions. This review examines the adaptation of various wireless radio technologies and protocols designed to bridge the gap between qualitative human observation and quantitative plant metrics. Furthermore, multi-sensor applications and comparative analyses are discussed to highlight expanded sensing capabilities and their integration with actuation components. Recent advancements have further enhanced these platforms by incorporating cloud-edge computing within the Internet of Things (IoT) domain. The primary objective of this paper is to analyse WSN architectures, topology models, wireless connectivity choices, and real-world implementations, and to evaluate the potential of IoT to create borderless farm management systems. Finally, the integration of artificial intelligence (AI) into WSN-IoT automation is explored, offering scalable solutions to combat the complex challenges facing global food crop production and the global hunger index.

Keywords: Data-driven agriculture, edge computing, Internet-of-Things, precision farming, remote sensing, smart agriculture, wireless sensor network, ZigBee

ARTICLE INFO

Article history:
Received: 01 May 2026
Accepted: 22 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.13

m.azwan@usm.my (Muhammad Azwan Nasirudin)
nishat@usm.my (Mohammad Nishat Akhtar)
aefarzad@usm.my (Farzad Ismail)
* Corresponding author

INTRODUCTION

Accelerated human population growth is driving rapid increases in global food demand despite diminishing resource availability. This pressure was further

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

intensified by the COVID-19 pandemic, which disrupted not only global healthcare infrastructure and medical supply chains, but also critical agricultural networks. A 2020 survey reported pandemic-induced food insecurity, with approximately half of the countries surveyed experiencing partial or full-service interruptions in food-producing sectors. The COVID-19 pandemic disrupted all four pillars of food security: availability, access, utilisation, and stability through labour shortages, supply chain breakdowns, and export restrictions that threatened global food supplies (Ben Hassen & El Bilali, 2024; Kakaei et al., 2022; Laborde et al., 2020). After that, the situation has been worsened by countries involved in long-period war zones, for instance, the Russia-Ukraine war, as both countries are major global producers and exporters of food crops and fertilisers, compounding pandemic-era disruptions with new supply shocks (Ben Hassen & El Bilali, 2022; Kakaei et al., 2022).

Besides the previously mentioned factors, Earth's global warming is well known as a major threat to farming activities worldwide. It has significantly affected crop yields when normal climates and essential resources were disrupted. In addressing the existing global warming issue, studies have also reported that the widespread use of pesticides has led to pest resistance, damage to soil biodiversity, and increased environmental pollution. Moreover, the non-natural fertiliser substances in the farming process is likely to contribute to disastrous chemical pollution of soil, aquatic and air quality (Kumar & Ilango, 2018; Saccone & Vallino, 2025).

Food and Hunger

By observing figures closer to the current perspective, the Global Hunger Index (GHI), an annual peer-reviewed study that measures hunger levels on a scale, has reported that nearly 690 million people worldwide are underfed or starving. While 144 million children are suffering from undernutrition problems, 5.3 million of them passed away before reaching 5 years old in 2018 (von Grebmer et al., 2020). South Asia recorded the second-highest GHI score globally at 26.0, closely following Sub-Saharan Africa's score of 27.8. On the GHI severity scale in Figure 1, these scores fall within the 20.0-34.9 range, indicating a "serious" hunger situation. Additionally, the State of Food Security and Nutrition in the World 2021 report highlighted that the COVID-19 pandemic severely disrupted food supply chains and heightened economic uncertainty, exacerbating global malnutrition. However, this upward trend began years before the pandemic and has now impacted an additional 40 million people globally (von Grebmer et al., 2021)

Meanwhile, the impact of the pandemic has likely escalated food insecurity in Malaysia. Data from the Global Hunger Index (GHI) reveals that Malaysia ranked 59th (out of 107 countries) in 2020 and 58th (out of 116 countries) in 2021. With a consistent score of 12.8, the nation remains categorised at a 'moderate' severity level (von Grebmer

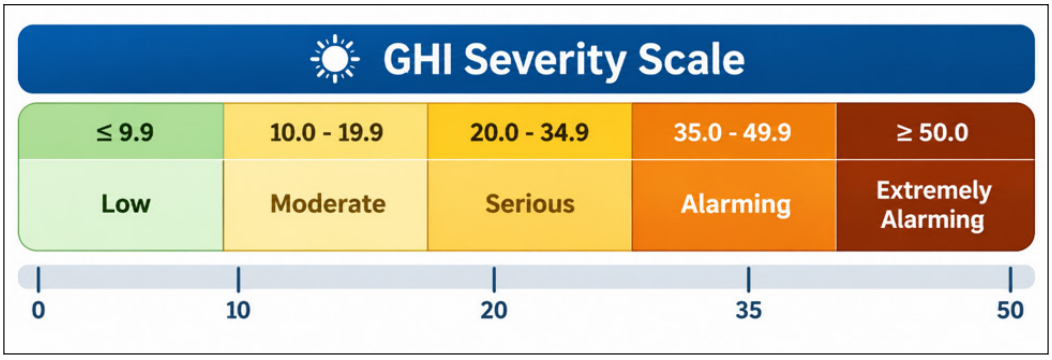
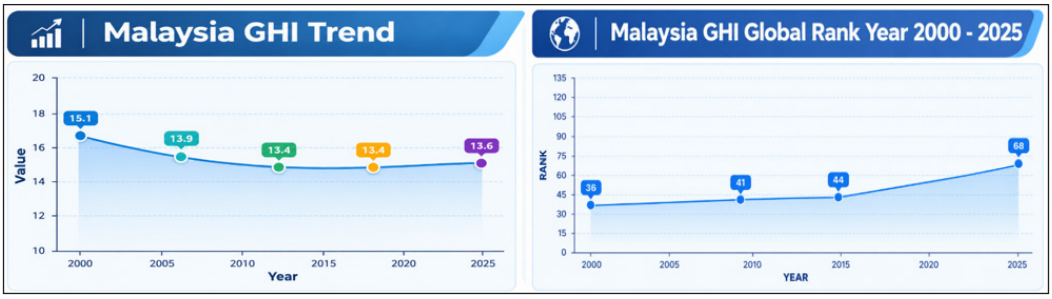


Figure 1. Global hunger index severity scale



(a) (b)

Figure 2(a) and (b). Malaysia's global trend among 136 countries in 2000, 2008, 2016, and 2025

et al., 2020; von Grebmer et al., 2021). As illustrated in Figure 2(a), the trends for GHI parameters show a steady line for undernourished in the Malaysian population, but the total GHI scores have increased by 3.22% from 2012 to 2021. Figure 2(b) indicates that our country's hunger level has increased relative to 2000, ranking 68th globally compared to 36th in 2000. Consequently, Malaysian food security frameworks and supply networks must respond proactively to any disruptions within agricultural supply chains. A pertinent example is the 2023 global rice shortage, which severely affected national food security, market pricing, and overall costs. This crisis was precipitated by geopolitical issues, severe climate change, crop diseases, and natural disasters (Shan, 2023). Additionally, these situations reduce the yields of staple food crops due to conventional plantation practices in Malaysia's farming sector.

Agro Farming, Precision Agriculture and Smart Farming

Unfortunately, there remains a lack of clear strategies for managing these complex farming issues at scale. Considering these emerging global challenges, researchers must develop advanced methodologies to increase the volume and quality of food production without

harming ecosystems. This can be achieved through precision agriculture solutions that leverage wireless platforms, such as Wireless Sensor Networks (WSN) and Internet of Things (IoT) technologies (Mendez et al., 2011; Shafi et al., 2019; Shamshiri et al., 2020).

In general, agriculture, precision farming and smart farming are related but not identical concepts. In the literature, agriculture is the science, practice, or discipline of cultivating plants and producing food, and some papers phrase it more broadly as the practice of growing plants or plant growth. Precision farming, also called precision agriculture in most papers, is a farm management approach that measures field, crop, or livestock conditions and uses spatial, temporal, or site-specific data to apply inputs more precisely for higher productivity, profitability, resource efficiency, and lower environmental impact. Smart farming is usually defined as a more digitally connected form of farming that uses IoT, sensors, software, cloud systems, AI, and real-time data to monitor, automate, analyse, and optimise farm operations and decisions (Abioye et al., 2021; Mowla et al., 2023).

This paper provides an overview of device-aided farming technologies and evaluates their potential to enhance crop yields, in alignment with the global 'Zero Hunger' strategy (SDG 2). Specifically, it highlights the significance of automated farming through WSN architectures that leverage IoT and Edge Computing, while offering a comprehensive review of the wireless communication protocols and sensors utilised in the precision agriculture domain. Finally, the subsequent section outlines the systematic review methodology employed for literature selection.

REVIEW METHODOLOGY

To ensure a comprehensive and refined synthesis of the existing literature on Wireless Sensor Networks (WSN) and the Internet of Things (IoT) in the smart agriculture landscape, this review was conducted following a structured literature mapping framework based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting logic, adapted for a critical narrative-synthesis review rather than a full systematic meta-analysis. The process flow of this review is illustrated in Figure 3.

Electronic databases, including Scopus, Web of Science, and IEEE Xplore, were queried for publications up to early 2026. The search strings combined specific keywords relevant to the study's scope, including:

"Wireless Sensor Network" OR "WSN"

AND ("Internet of Things" OR "IoT" OR "Cloud Computing" OR "Edge Computing")

AND ("Precision Agriculture" OR "Smart Farming" OR "Agro Farming"). Articles were screened based on peer-reviewed journal or high-impact conference status, studies presenting valid hardware architectures or computational frameworks for agricultural



Figure 3. PRISMA-style study identification and selection flow diagram

commissioning, and all papers were required to be written in English. Most of the selected papers were published between 2019 and 2020, allowing the review to capture possible effects on food crop farming during the pandemic period, while fundamental and theoretical subjects were drawn from a wider range of publication years. Exclusion criteria eliminated non-peer-reviewed sources, white papers, pre-prints, and studies lacking a clearly defined agricultural application.

Study selection followed four stages: identification, screening, eligibility, and inclusion. An initial search across the three databases yielded 612 records. After removing 94 duplicate records, 518 titles and abstracts were screened against the scope defined above, of which 267 were excluded for falling outside the agricultural domain, being non-English, or not peer reviewed. The remaining 251 full-text articles were assessed for eligibility; 105 were further excluded for lacking a defined hardware architecture or computational framework, or for being white papers or preprints. This process yielded a final set of 146 studies retained

for qualitative synthesis, along with a smaller set of foundational and theoretical references used to frame core WSN/IoT/AI concepts.

WIRELESS SENSOR NETWORK (WSN)

Wireless Connectivity and Data Delivery in the Precision Agriculture Domain

Wireless connectivity has become a core component of modern precision agriculture, enabling real-time monitoring of soil moisture, temperature, crop health, and environmental conditions across geographically dispersed farm fields (Jawad et al., 2017; Mowla et al., 2023; Thakur et al., 2019). The transition from wired to wireless data transmission in agricultural settings has been driven by the need to reduce labour costs, simplify deployment across vast rural landscapes, and scale sensor networks to cover larger farm areas (Alahmad et al., 2023; Tang et al., 2024). Various wireless data technologies and cellular mobile networks, as shown in Table 1, will then be incorporated depending on range, frequency, bandwidth, speed, power consumption, portability, and internet capability of the platform (Buchli et al., 2012; Khanh et al., 2022; Meraj et al., 2015; Zulkifli & Noor, 2017). A diverse ecosystem of communication protocols, including LoRaWAN, Wi-Fi, ZigBee, NB-IoT, Sigfox, LTE, and emerging 5G technologies, now underpins the Internet of Farming (IoF), each offering distinct trade-offs in range, data rate, power consumption, and cost (Avşar & Mowla, 2022; Damsgaard et al., 2022; Singh & Sobti, 2021). The data transportation platform with internet connectivity is also defined as IoT-ready devices, which are the most precious assets in the food production domain, from raw farm substances to their final ready-to-consume stage. Therefore, the importance of this review is to emphasise the potential of wireless sensor networks and the Internet of Things (IoT) as targeted approaches for the supervision of the food and agricultural industries. ZigBee and LoRaWAN excel in private, low-power farm topologies; cellular LPWAN solutions such as NB-IoT and LTE-M eliminate the need for local gateway backhaul links by connecting agricultural

Table 1
Wireless cellular network generation comparison

Generation	Year	Technology	Frequency	Speed	Network	Features
1G	1981	AMPS	30Khz	2.4Kb/s	PSTN	voice only, analogue
2G	1999	GSM	900-1.8Ghz	64Kb/s	PSTN	voice, data, web
3G	2002	WCDMA	1.2-2Ghz	2Mb/s	Packet Network	voice, data, web, media, streaming
4G	2010	LTE	2-8Ghz	100Mb/s	Internet	voice, data, web, HD media, streaming
5G	2015	MIMO	3-30Ghz	>1Gb/s	Internet	voice, data, web, HD media, streaming, low power, higher speed

sensors directly to cellular towers (Baronti et al., 2007; Ertürk et al., 2019; Waseem et al., 2025). Hence, they are superior for highly distributed open-field setups despite cellular subscription fees.

The deployment of precision agriculture (PA) architecture has been reported to increase crop yields tremendously while conserving natural resources, reducing land exploration, deforestation, and pollutants (van der Goot et al., 2016). The definition of precision agriculture, or precision farming, is an integrated information and automation methodology aided by artificial intelligence and management systems. PA is used to monitor and manage crops efficiently by optimising planting input information, such as water quality, soil conditions, fertiliser, and climate, by analysing the outcomes of the process (Kloas et al., 2015; Yanes et al., 2020). Nowadays, the precision farming concept is still presumed to be a complex practice or costly, even though it has been implemented over the last few decades with several benefits (Glenn et al., 1998). PA monitoring of diversity can be extended in many ways, for example, by utilising geolocation data from the NAVSTAR Global Positioning System (GPS). It used to be supplementary and beneficial information for improving crop health or growing schemes by providing geographical data, soil mapping, weather forecasts, and information on natural disasters and other threats (Baseca et al., 2019; Buchli et al., 2012; Glenn et al., 1998).

For instance, Cambra et al. (2015) proposed a sprayer map using GPS geolocation data and drone video. The purpose of integrating GPS data into their system was to precisely spray the fluid solution onto the correct area while optimising herbicide and fertiliser usage (Cambra et al., 2015). While Glenn et al. (1998) reported in their studies there was an unknown variation that affected soybean and corn crop yields from a central Missouri farm in the United States from 1992 to 1994 (Glenn et al., 1998). After the investigation, the study successfully identified these factors by linking the GPS receiver to the Geographic Information System (GIS) on the yield maps to show the fractions of phosphorus, potassium, and pH in the soil test data (Glenn et al., 1998). These two data types precisely locate the soil organic profiles and the related treatments used to neutralise unbalanced fertilising, as shown in Figure 4. There are numerous manual methods for cultivating healthy crop practices. However, they require labour-intensive efforts, which is likely a burdensome factor in modern farming. Therefore, to make the agricultural domain smarter in processing, a rigorous method, such as a Wireless Sensor Network (WSN), is preferable. The following section discusses the WSN variants which can be used in the agricultural domain.

WSN Architecture

In the context of precision farming, an emerging solution known as Wireless Sensor Networks (WSN) has been deployed extensively in the agricultural sector (Kandris et al., 2020; Lata et al., 2019; Miskam et al., 2009). WSN is defined as a low-cost, low-powered,

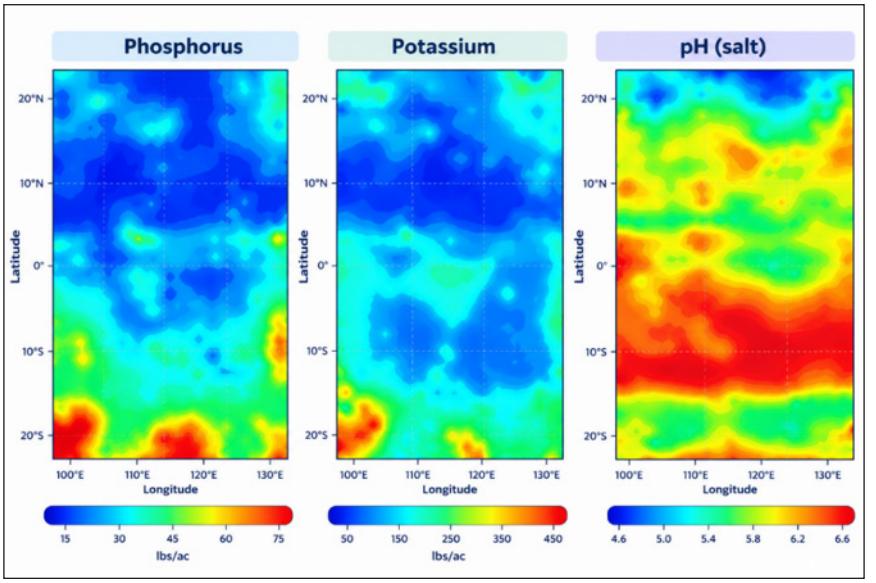


Figure 4. Fraction of soil test information map at Missouri farms reported by Glenn Davis
Source: Adapted from (Davis et al., 1998)

portable, or tiny sensing device technology, incorporating a proprietary radio protocol and communication standards, to perform self-networking via a Personal Area Network (PAN) with selectable topology patterns and a self-routing mechanism (Lata & Verma, 2022; Wang et al., 2006). WSNs generally comply with the IEEE 802.15.4 ZigBee standard, widely used for wireless data acquisition and self-monitoring in multi-sector applications since its inception (Baronti et al., 2007; Buratti et al., 2009). In the agricultural domain, it has been well-suited due to the offering of power-efficient, self-distributed networking, as well as solar-to-battery-operated capability (Iqbal Khan et al., 2022; Jawad et al., 2017). In the WSN ecosystem, groups of spatially dispersed sensor nodes can establish ad hoc, autonomous interconnectivity with one another using wireless radio frequency (RF) on a designated frequency band (Iqbal Khan et al., 2022; Kandris et al., 2020; Mowla et al., 2023). A sensor node, also known as a mote, is an integrated embedded system module consisting of a microprocessor, writable EEPROM data memory, an RF transceiver, and interface peripherals such as an analogue-to-digital converter (ADC). As illustrated in Figure 5(b), the node also incorporates one or more sensors alongside a portable DC power source (Miskam et al., 2009; Nasirudin et al., 2011).

Figure 6 summarises the WSN deployment to provide agricultural benefits. Basically, WSN nodes use on-board sensors to measure fluctuations in parameter conditions in the localised environment (Buratti et al., 2009; Prodanović et al., 2020). A dedicated transducer measures the target parameter, and the measurement signals are converted into relative electrical magnitudes before being sampled by an Analog-to-Digital Converter (ADC) and

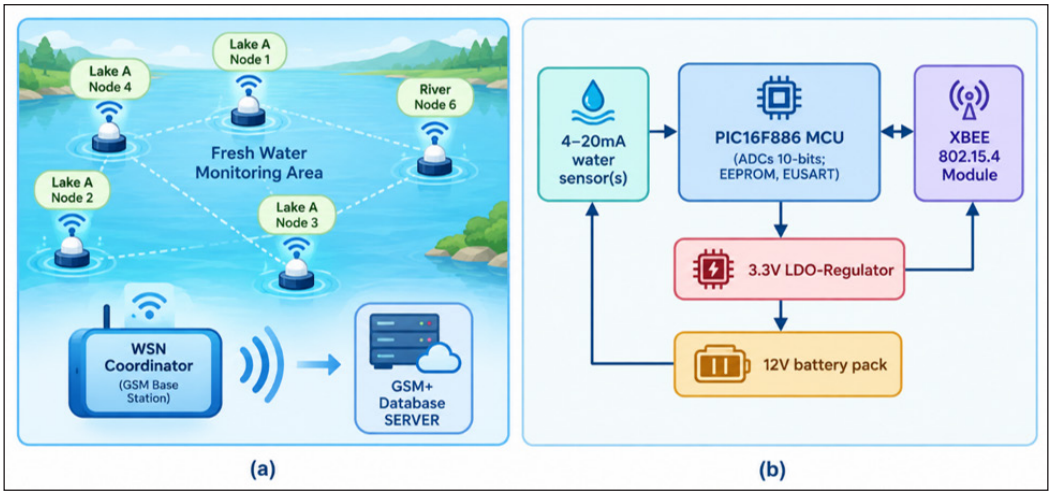


Figure 5. (a) WSN point-to-point and multi-point network topology with a GSM base station for lake water monitoring. (b) Wireless ZigBee sensor node architecture used in the (a) setup
Source Adapted from (Nasirudin et al., 2011)



Figure 6. Key features of WSN implementation in the agricultural domain

finally processed by the microprocessor. WSN networks can also optionally be equipped with a positioning sensor or a mobilisation unit, such as the Automatic Packet Reporting System (APRS) or the Global Positioning System (GPS) (Buchli et al., 2012). Additionally, the system can be integrated with a Global System for Mobile Communication (GSM) cellular module for extended-distance, point-to-point, and cross-border data uplinks and downlinks, as shown in Figure 5(a) (Wang et al., 2010; Zamora-Izquierdo et al., 2019).

Typically, WSN frameworks utilise multi-hop routing, allowing data to 'hop' between nodes, sharing an identical cluster or group ID as it propagates toward a designated drop-

off point, commonly referred to as the “Base Station”. This network topology and the seamless interconnectivity between mote devices enable the WSN to harvest low-data-rate, real-time environmental data across expansive agricultural areas, funnelling it directly to coordinators and gateway points (Jawad et al., 2017; Zamora-Izquierdo et al., 2019). Then, the local host server or computing devices at the base station will record and process the data packet accordingly. As illustrated in Figure 5(a), the Base Station can serve as a data logging point, perform supervisory control over the WSN motes concurrently, and transmit related information to the person in charge or to external networks. Furthermore, the WSN database server can sort crop information more precisely, including mote battery health, PAN IDs, on-board sensor types, raw data, and timestamps. Initially, before the emergence of IoT infrastructure, decision algorithms, such as fuzzy logic, were processed on the WSN's local server (Kishore et al., 2013; Navulur et al., 2017; Verma et al., 2013).

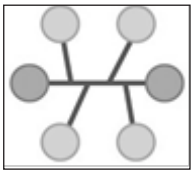
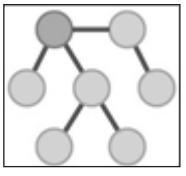
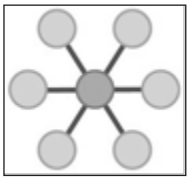
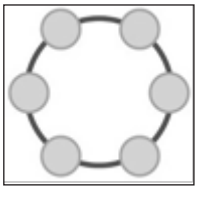
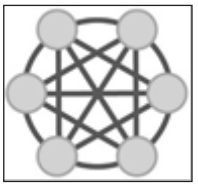
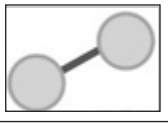
The management program on the server will analyse the incoming data, then execute the regulation command codes back into the network topology to the target node or send a notification alarm to the farmer via General Packet Radio Service (GPRS) ASCII text messages on the GSM network (Ghribi & Logrippo, 2000; Nasirudin et al., 2011). Thus, has brought limitations and less flexibility for heavier real-time processing or cross-border farm management due to GSM network limitations and fees. On the other hand, WSN characteristics have proven most beneficial for modern agricultural practices. In summary, WSN is highly scalable to accommodate small or huge network memberships with autonomous topology formation based on the node’s operability mode, energy saving or extended coverage (Ojha et al., 2015).

ZigBee is well-known as the finest WSN ‘asset’, a low-power, low-cost wireless networking standard built on top of the IEEE 802.15.4 physical (PHY) and medium access control (MAC) layers (Haque et al., 2022). Typically, a node requires very low power, typically 1 mW or less, enabling the battery to last for years. The standard provides a range of up to 150 meters outdoors using direct-sequence spread spectrum (DSSS). At the core hardware level, such as 0.18 μm CMOS transceivers, they consume approximately 14.7 mA in receive mode and 15.7 mA in transmit mode (Hullamani & T. N., 2021; Kluge et al., 2006). A ZigBee network can scale from a single mote device up to approximately 64,000 nodes (Ashok Somani, 2012). The ZigBee Alliance defines the upper network and application layers, whereas IEEE 802.15.4 specifies the lower hardware and radio transmission requirements (Paul & Singh, 2015). Furthermore, ZigBee networks possess self-healing and ad hoc capabilities, enabling them to establish robust connections even in highly complex routing configurations, such as distributed mesh topologies (Poole, 2004; The ZigBee Alliance, 2015). Additionally, WSN protocols and compliant devices are safeguarded by robust digital encryption standards, typically utilising at least 128-bit AES (Advanced Encryption Standard) security to ensure data integrity (Jawad et al., 2017). In such a report, ZigBee achieves low latency in direct, single-hop deployments

suitable for real-time field monitoring. A 12-hectare plantation deployment recorded a latency of 20-25 milliseconds (ms) with a 95% data transmission success rate (Gitakarma et al., 2025). Consequently, the system exhibits robust immunity against malicious attacks, thereby safeguarding sensitive data from specific agricultural zones and strategic food crops (Gitakarma et al., 2025; Hullamani & T. N., 2021; Mowla et al., 2023).

Furthermore, six types of network topologies are commonly implemented within WSN and IoT infrastructures, as outlined in Table 2 (Kishore et al., 2013; Verma et al., 2013).

Table 2
Summary of Wireless Sensor Network common topology patterns

Topology Name	Pattern	Description	Limitation
Bus		Single path & linear connectivity to nodes. Easy routing Moderate scalability	Performance drops with the number of nodes
Tree		Tree-Bus element networks. Parent-child hierarchy & best organised. Highest scalability, distance & flexibility	Lost 'child' nodes links risk if base/ server fails
Star		Centralised data sinking point. Single path routing. LAN computer network style	Lost all links risk if base/ server fails. Short distance
Ring		Two-way routing circular path. No server needed; data passes through all nodes. WAN computer network style	Broken link risk if one node fails. Slower than Star
Mesh		Highest data traffic. Multiple paths & simultaneous transmission. Robust against broken links	High complexity algorithm. Redundant connection. Costly
Point-to-Point		Conventional, simple topology & no server. Single channel routing path, long range, high speed, large data target	Low scalability to expand. Lost link risk if one node fails

The star topology features a single central sink node, which establishes a direct connection to the base station (Khan et al., 2022; Kishore et al., 2013). The primary strategy behind this formation is to minimise data forwarding delays, thereby reducing routing latency to the server (Baronti et al., 2007). In contrast, multi-sink or distributed configurations, such as a bus topology, can handle heavier data traffic, cover larger agricultural areas, and accommodate energy conservation strategies (Morais et al., 2008; Shafique et al., 2020). A hierarchical formation, such as a tree topology, can assign priorities to nodes based on their role or location in the network. Therefore, the system can efficiently administer the WSN network (Carlos-Mancilla et al., 2016; Osseiran et al., 2020). The fundamental design of a sensor node can be represented by the block diagram shown in Figure 7(a). An embedded logical processor, such as a microcontroller unit (MCU), will become the 'brain' part of each node.

Break it down to basics: commonly, a node is designed to perform on-board processing and computing, such as the conversion of analogue signals (voltage, resistance, and capacitance) into digital binary or hexadecimal (Miskam et al., 2009; Nasirudin et al., 2011). For instance, the Analogue-to-Digital Converter (ADC) module on the PIC16F886 microcontroller converts the analogue DC voltage signals from sensors into digital values. These input signals typically range from 0–5 Vdc or 0–3.3 Vdc. To consider a scenario where a sensor operates within an output voltage range of 0–3.3 V, Full $_{Scale\ Voltage\ Range}$ = 3.3. Then, given the ADC's largest scale on the MCU = 10 bits = (2^{10}) = 1024 -1 = 1023. Finally, 3.3/1023 = 0.0032 volts per 1 LSB. During reading, the measured sensor output voltage is 2.4 V; therefore, the conversion is (2.4 ÷ 0.0032 = 750), resulting in a digital decimal value of 750. This value is written to the MCU memory as 750 (dec), 1011101110

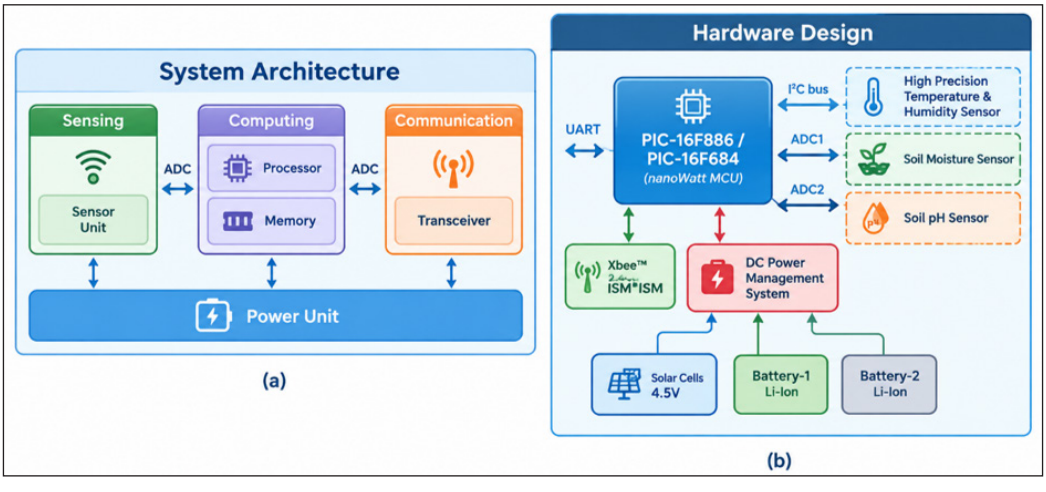


Figure 7. A typical sensor node architecture: (a) a sensor node specifically designed for paddy crop monitoring; (b) Hardware Design
Source Adapted from (Nasirudin et al., 2011)

(bin), or 0x2EE (hex). Figure 7(b) illustrates a sensor node driven by an 8-bit PIC16F684 microcontroller, featuring two ADC channels for analogue sensors (WQ101 and WL400) and one digital channel for the SHT75 temperature and humidity sensor. The proposed system utilises a bi-directional (full-duplex) UART serial interface on an XBee 2.4 GHz wireless transceiver to establish a tree-topology network configuration via Digi firmware (Miskam et al., 2009; Nasirudin et al., 2011).

Table 3 presents various wireless radio-frequency platforms and protocols, along with relevant parameters applicable to WSNs (Baronti et al., 2007; Effah et al., 2023; Wang et al., 2006). In a typical deployment, several device technologies support large-scale networks with high node counts. For instance, Sigfox can accommodate up to 1,000,000 devices per base station, while ZigBee supports up to 64,000 nodes per coordinator, and LoRaWAN

Table 3
Wireless radio technology general specification and comparable features

Wireless Technology	Standards	Frequency (MHz)	Data Rate	Physical Range (m) Typical	Topology/ Network	Key Features
ZigBee	IEEE 802.15.4	868/ 2400	250	100	Mesh/ LPAN	Low power, self-healing mesh, no direct internet IP
LoRaWAN	IEEE 802.15.4g	433/868/ 915	50	10000	Mesh/ LPWAN	ultra-low power, long range, high latency
MiWi	IEEE 802.15.4	2400	250	100	Mesh/ WPAN	low power, Microchip™ proprietary
6LoWPAN	IPv6 802.15.4	2400	250	100	Mesh/ WPAN	very low power, direct internet IP,
Bluetooth	IEEE 802.15.1	2400	1000	10	Point-to-point/ WPAN	very low power, limited range
SigFox	N/A (proprietary)	868/928	600	10000	Star/ LPWAN	ultra-low power, low cost & bandwidth, extreme range
WiFi	IEEE 802.11a, b,g,n	2400	150000	100	Point-to-point/ WLAN	very high power, large data, shorter range, less standalone
GPRS-GSM	N/A (cellular)	900 - 1800	190	35000	Point-to-point/ WAN	global roaming, SIM card to stand alone, power drain, paid data

Table 3 (*continued*)

Wireless Technology	Standards	Frequency (MHz)	Data Rate	Physical Range (m) Typical	Topology/ Network	Key Features
NB-IoT	N/A (cellular)	600 - 2200	160	5000	Star/ LPWAN	ultra-low power, extreme range, high latency, less portable
LTE-M	N/A (cellular)	700 - 2600	1000	10000	Star/ LPWAN	Ultra-low power, high mobility, low latency, power drain
4G-LTE	N/A (cellular)	600 - 3800	100000	8000	Star/ WWAN	High-speed data, low latency, high cost, power drain
5G	N/A (cellular)	600 - 4700	1000000	5000	Star-Mesh/ WWAN	ultra high-speed data, low latency, slice network, highest cost, infra complexity

handles roughly 10,000 nodes per gateway (Ashok Somani, 2012; Chi et al., 2014; Iqbal Khan et al., 2022). Conversely, traditional technologies such as classic Bluetooth, Wi-Fi, and GSM-GPS support significantly lower node capacities, accommodating only 8, 32, and 1,000 devices, respectively. Furthermore, analysis of manufacturer specifications indicates that in transmission mode, ZigBee draws the lowest power, at 36.9 mW, compared to 100 mW for LoRaWAN and 122 mW for Sigfox (Alqurashi et al., 2023; Ertürk et al., 2019; Jawad et al., 2017). Jawad et al. (2017) suggested XBee wireless transceiver modules, ZigBee protocol-compliant devices, have strong potential for deployment in farming zones due to their low duty cycle, low latency, long coverage, supported topologies, and low peak power consumption (Alqurashi et al., 2023; Lavric et al., 2019). While Zulkifli and Noor (2017) incorporate the Active Radio Frequency Identification (RFID into the WSN framework driven by ZigBee 2.4GHz hardware (Zulkifli & Noor, 2017).

In contrast to the ZigBee protocol, Wi-Fi technology is less feasible for field-deployed sensor nodes due to constraints regarding power consumption, network topology, and unoptimized data payloads (Ikpehai et al., 2019). While Sigfox, ZigBee, and LoRaWAN each trade off scalability, throughput, range, and deployment control differently. Sigfox for ultra-light wide-area sensing, ZigBee for local mesh networks, and LoRaWAN for flexible private or public LPWAN deployments (Mekki et al., 2019; Orlovs et al., 2025; Saqlain et al., 2025). In conclusion, WSN platforms have proven to be both resilient and reliable for deployment in open-field farming (Carlos-Mancilla et al., 2016; Gitakarma et al., 2025).

Agricultural Sensor Solution

Sensor technologies offer countless opportunities to optimise precision agriculture, transforming raw environmental data into actionable insights for sustainable crop management, as mentioned earlier. Table 4 outlines a selection of sensors utilised in precision agriculture, along with their respective capabilities (Ji et al., 2019; Shin et al., 2025). Several of the listed sensors are dedicated to the intensive measurement of soil quality parameters, while others capture ambient air quality metrics. Additionally, specialised sensors are included to evaluate crop leaf health. Typically, these sensors can be interfaced directly with the General-Purpose Input/Output (GPIO) pins of WSN nodes via standard analogue and digital communication protocols, including ADC, I2C, SPI, and UART (Chapungo & Postolache, 2021; Nasirudin et al., 2011).

Table 4
Multi-sensor types with key information and comparison implemented in the precision agriculture domain

Sensor Model	Functionality	Key Benefits	Limitations
A. Soil Moisture, Electrical Conductivity (EC) & Temperature Sensors			
ECHO 10HS	Soil moisture VWC (volumetric water content)	Analogue voltage output, large measurement volume, characterising a wider area	Does not measure Temp, sensitive to high-salinity soils, older legacy model
ECHO EA10	Soil moisture for low EC soils.	Analogue voltage output, low cost, good for standard – medium texture soils.	Very sensitive to changes in soil temperature & EC
ECHO EC5	Soil moisture (VWC).	Compact size great for small pots, greenhouse bags, or laboratory columns, low cost.	Smaller volume measurement; may miss dry/wet spots in soil, lower accuracy in high EC
5TE Sensor	Soil moisture, EC, temperature	All-in-one measurement, excellent for tracking fertiliser run-off & salinity alongside moisture.	Program design requires careful installation to avoid air gaps, which will heavily skew readings
Hydra probe 2	Soil moisture, EC, dielectric, temperature	Industrial-grade accuracy and durability, soil-specific calibration less dependent	High cost, bulky size, will disturb soil compound
MP406	Soil moisture, temperature	High-accuracy ADR (Amplitude Domain Reflectometry) technology, stable	Expensive, requires proprietary loggers, requires attentive calibration for peak accuracy
Pogo II	Soil moisture & temperature	Excellent portability, Wi-Fi/Bluetooth enabled, smartphone apps ready for rapid field surveys	Expensive, measures top surface layer only or dug hole

Table 4 (*continued*)

Sensor Model	Functionality	Key Benefits	Limitations
B. Heavy Metal, Gases, and Soil Quality Sensors			
GreenTag (RFID)	Soil moisture, temperature & heavy metals	Wireless, battery-free (passive RFID); sensors can be completely buried without cable	Short-range reads (requires reader to be very close); heavy metal detection via RFID is still experimental
MIKROE-1630	Gases (air quality)	Low-cost, easy plug-and-play integration with RPi/ Arduino hardware	Low precision, cross-sensitive to multiple gases (detects general mix, not specific isolated gases)
PXRF with VNIR	Soil properties, heavy metal	True laboratory-grade elemental analysis in the field & highly precise.	Extremely expensive; X-ray tools require radiation safety training & licensing to operate
NixPro Sensor	Soil colour analysis (Munsell)	Affordable soil colour standardisation; eliminates human visual inconsistency	Soil must be dried and ground for accuracy; does not measure nutrients
Portable Electrochemical System (PES)	On-site heavy metal	High sensitivity to specific metals, more affordable than portable XRF scanners.	Requires field sample preparation before testing (e.g., mixing soil with liquid buffers)
Arduino/ RPi-based Potentiostat	Heavy metal & integration	Electrochemistry basis, open-source, highly customizable, very low-cost to build	Not industrial grade, requires strong coding and electronics knowledge to build & calibrate
HMstat	Heavy metal (ADC/DAC)	Heavy metal analysis in a portable form factor	Research-grade/custom-built, lacks support of mainstream, commercial sensors
MG811	Carbon Dioxide (CO2) Sensor	Low to high concentrations of CO2, reliable measurement, low cost	Requires continuous heating element (draws high power/battery drain), needs frequent calibration
C. Leaf, Plant, and Photosynthesis Sensors			
BetaTherm 100K6A1B	Temperature	Incredibly low-cost, tiny, & fast response to temperature changes	Highly fragile, requires a custom voltage-divider circuit, and coding equation conversion from Ohms to Celsius.
ΔLA-C	Leaf-to-air temperature	Highly specialised & precise for thin-needle plants, e.g., Conifers	Extremely delicate setup, not useful for broad-leaf crops
LW100	Plant moisture	Durable, easy to mount in agricultural canopies	Requires calibration, angle adjustments on specific crop leaf

Table 4 (*continued*)

Sensor Model	Functionality	Key Benefits	Limitations
LT-2 M, LT-1P	Leaf temperature	Sub-miniature design, non-destructive, continuous monitoring of a single leaf	Physically clamping potentially can alter the leaf's natural microclimate
CI-340 System	Comprehensive photosynthesis/gas exchange	Portable, all-in-one field laboratory for plant physiology, highly comprehensive data	Expensive, requires regular calibration of gas chambers, steep learning curve
RTH-11, RTH-48	Leaf temperature & humidity/wetness	Great for general automated greenhouse climate control	General-purpose accuracy
FI-LP, MP, SP	Fruit growth sensors (Dendrometers).	Allows continuous, non-destructive monitoring of fruit swelling/shrinking (water stress)	Specific models for specific fruit sizes; mounting is prone to scarring the fruit
D. Light Sensors			
BH1750FVI	Light intensity (Lux)	Very low-cost, I2C digital output, stable reading	Saturates easily in direct, bright sunlight, measures human visible spectrum (Lux)
HOBO S-LIA-M003	Photosynthetic Light (PAR) Photosynthetically Active Radiation	High accuracy, measures exact wavelengths of light that plants use for photosynthesis	Proprietary systems (HOBO), costly

One of the sensing examples is yield sensors, which are typically integrated into harvesting machinery to automatically log production data. For instance, grain systems utilise mass flow sensors to record grain inflow alongside the vehicle's real-time geographic position (Kumhálová & Matějková, 2017; Liu & Li, 2017). Yield monitoring facilitates long-term strategic decision-making in precision agriculture. At the conclusion of the growing season, this method is used to map the spatial distribution of crop yields (Leng & Huang, 2017; Mekonnen et al., 2020). Subsequently, Geographic Information System (GIS) software, such as ArcInfo, MapInfo, and ESRI platforms, is employed to analyse the collected spatial data (Vuran et al., 2018).

Electrical conductivity (EC) sensors are employed to visualise crop yield and field topography. Electrical conductivity is a measurement of the ability of soil to conduct current. The EC method is used to measure phosphorus and cation concentrations in water, as well as the depths of the drainage system and the roots (Navulur et al., 2017). The EC maps are used to divide the area into zones. Variable irrigation rates, variable seeding rates, and drainage management can all be incorporated into the zoning. The apparent Electrical

Conductivity (ECa) and Visible Near Infrared Reflectance (VNIR) can be used to map EC using Electromagnetic Induction (EMI) methods (VNIR) (Hassan Hayatu et al., 2020). Commercial tools, such as the Veris 3100 and EC400 sensors and GPS systems, are readily available (Navulur et al., 2017; Vuran et al., 2018).

Nitrogen, potassium, and phosphorus are macronutrients critical to crop growth in soil. Fertiliser applications can be evaluated by analysing these nutrients and their potential effects. Reflectance spectroscopy is used to measure the macro-nutrients' reflection and absorption properties using optical detection (Dong et al., 2013; Research & Markets, 2020). Nitrate and sulphate concentrations in natural water sources can be detected using planar electromagnetic sensors (Aleksandrova, 2018). The impedance of the sensor array is correlated with nitrate and sulphate concentrations, enabling measurement of their levels. Micronutrients in the soil can be studied using electrochemical, VIS-NIRS, and ATR techniques (Luce et al., 2017). Because these methods use a membrane that reacts to only one ion, they can detect only one ion among soil macronutrients (Hassan Hayatu et al., 2020). It is necessary to build a detector array for soil macronutrient sensing to achieve simultaneous detection of multiple ions (Ji et al., 2019).

Contemporary Wireless IoT Solutions

A contemporary discussion of wireless communication has included NB-IoT, LTE-M, 5G, and private 5G networks because they address different IoT requirements rather than replacing one another with a single universal option (Osseiran et al., 2020; Vaezi et al., 2022). Across the literature, the main selection axes are coverage, power, device density, latency, mobility, and deployment control (Moges et al., 2023). Technology roles: NB-IoT is a 3GPP LPWA cellular technology introduced in LTE Release 13, purposely for IoT connectivity. It was designed for wide-area IoT connectivity with low device complexity, long battery life, massive device support, and extended coverage, while reusing much of LTE to speed deployment and product development (Akpakwu et al., 2018; Xu et al., 2018). It is commonly positioned for applications such as asset tracking, smart buildings, smart cities, and environmental monitoring (Akpakwu et al., 2018; Waseem et al., 2025). LTE-M is another 3GPP cellular IoT option that, unlike NB-IoT, is used when applications need more bandwidth and higher data rates while still fitting the mobile IoT category (Routray & Mohanty, 2024).

Recent reviews treat LTE-M, NB-IoT, EC-GSM-IoT, and NR-RedCap as the major standardised CIoT family, reflecting a portfolio approach rather than a winner-take-all architecture (Moges et al., 2023). 5G is framed as the broader wireless platform that targets massive connectivity, low latency, reliability, and higher spectral efficiency beyond what earlier generations were optimised to provide (Shafique et al., 2020; Sufyan et al., 2023). The literature consistently describes 5G as supporting three broad service classes,

including eMBB and mMTC, and as an enabler for demanding IoT domains such as industrial automation, smart cities, and remote surgery (Osseiran et al., 2020; Vaezi et al., 2022). Table 5 shows a brief comparison of contemporary IoT wireless technology and deployment contexts (Routray & Mohanty, 2024; Sufyan et al., 2023; Vaezi et al., 2022; Wen et al., 2022). NB-IoT remains important in the 5G-era IoT, but its mechanisms still need improvement to ensure seamless coexistence with 5G NR (New Radio standard) (Gbadamosi et al., 2020). Security and energy efficiency are recurring cross-technology constraints across 5G IoT communication systems (Khanh et al., 2022).

Table 5
Comparison of contemporary IoT wireless technology and deployments

Technology	Ideal Event	Core Strength	Main Limitation
NB-IoT	Static, low-data sensors	Low power & deep coverage	Limited data rate, evolving 5G coexistence challenges
LTE-M	Higher-rate mobile IoT	More bandwidth than NB-IoT	Higher resource & power use than narrowband LPWA
Public 5G	Mixed consumer and IoT workloads	Low latency, scale, and QoS ambitions	Capacity & resource pressure in dense advanced use cases
Private 5G	Enterprise and industrial sites	Dedicated control, customised security, local optimisation	Deployment, slicing, TSN & interference design complexity

Leveraging the Internet of Things

With respect to the economic aspect, the smart agriculture market revenue was projected to reach USD 6.5 billion by Q4 of 2021, indicating a greater impact of COVID-19 than in 2020, when it was USD 5.6 billion (Research & Markets, 2020). The increasing demand for food crop products, however, has not been matched by farmers' capabilities and facilities to fulfil the requirements of the digital market. Therefore, enhancements are required in terms of portability and ease of use to deploy the IoT infrastructure into the farmer's daily practices. Fundamentally, smart IoT crop farm management can also be summarised as shown in Figure 8, coexisting with a conventional WSN architecture, but with the difference of internet-cloud-based or edge-computing interoperability. Initially, some smart farming architectures relied solely on IEEE 802.11 network protocol devices to acquire sensor data and to actuate sprinklers for irrigation using a fuzzy logic control algorithm (Ramli et al., 2020). For instance, Akhtar et al. (2021) proposed a model incorporating Edge Computing in IoT and farm-field distributed wireless sensors for soil heavy-metal detection, assessment, and monitoring. Translating raw sensor data into implementable insights for farmers using artificial intelligence and big data analytics models. Another smart solution, called the Internet of Underground Things, is to bury sensors underground to penetrate the soil layer and maintain network communication on the ground (Dong et al., 2013; Vuran et al., 2018).

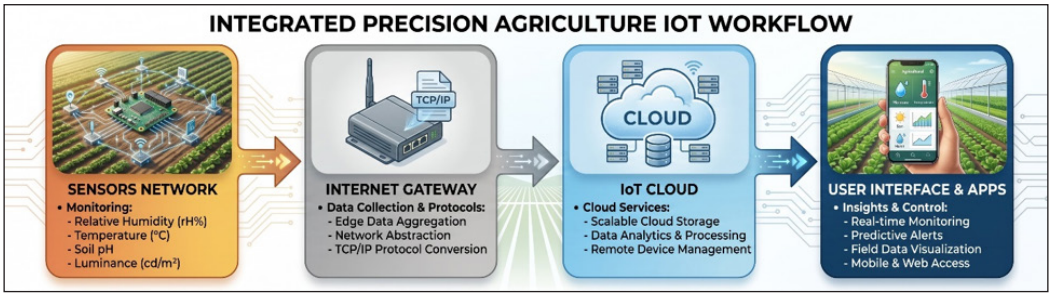


Figure 8. The Internet-of-Things (IoT) block system overview for smart farming and agriculture applications

To date, the deployment of WSNs with Internet of Things (IoT) frameworks has transformed the agricultural sector into a smarter, more intelligent production process, even in remote areas with reduced human intervention (Aleksandrova, 2018; Zamin et al., 2019). The integration of IoT with other electronic platforms has become an emerging approach due to its flexibility in enabling real-time monitoring, smart decision-making, and automation via remote devices from remote locations (Farooq et al., 2019; Ramli et al., 2020). By observing the architecture stacks, the Wireless Sensor Network (WSN) is homogeneously integrated with the IoT framework, which is the keystone for deriving a fully automated process known as the smart farming model (Gitakarma et al., 2025; Zulkifli & Noor, 2017). This model will provide a comprehensive ecosystem for WSN devices operating 24 hours a day, 7 days a week, on cloud-IoT-based Artificial Intelligence (AI), thereby autonomously administering crop fields or potting areas from anywhere across the globe. The sensor node is designed to periodically stream the recorded plant information and related data to the device coordinator (sink point) of each cluster, thereby bridging them to the terminal endpoint, which is known as the base station. The base station telemetry is either integrated with an on-site internet data server or uplinked to a remote server via a TCP/IP gateway to complete the IoT system routing. The system will then actively uplink and downlink streaming to IoT cloud computing for global remote access, automation processes, and supervision (Shafi et al., 2019). With larger integrated datasets coming through IoT platforms in agriculture, AI algorithms will perform optimal optimisation for farmers to smartly manage their crops using more sophisticated techniques compared to conventional methods (Gluhak et al., 2011).

The integration of machine learning significantly enhances agricultural productivity, as intelligent algorithms process and analyse massive streams of incoming data for both model training and real-time inference. Consequently, these systems enable autonomous management of various aspects of crop production, including disease detection, optimal harvest scheduling, automated planting, precision weeding, and environmental forecasting (Mekonnen et al., 2020). Notably, WSN and IoT technologies operate synergistically; sensor networks deployed across the field aggregate and bridge real-time data to an internet

gateway, allowing cloud-based algorithms to execute responsive IoT commands and AI-driven tasks.

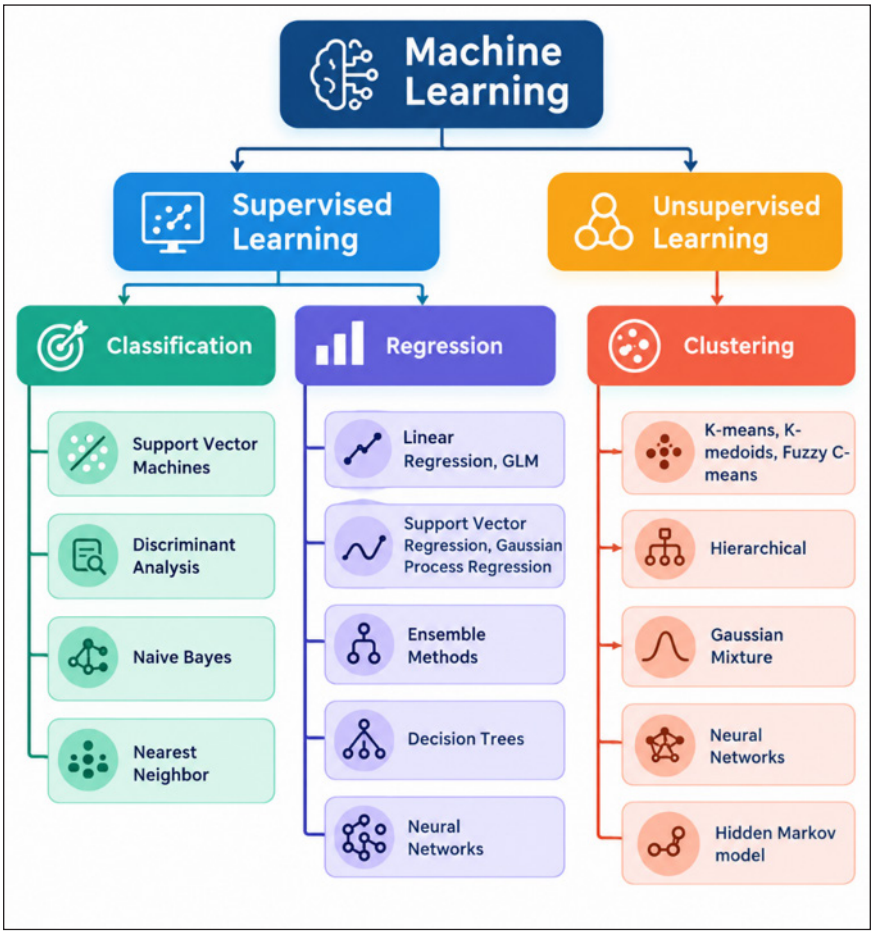


Figure 9. Machine learning classification and elements overview
Source: Adapted from (Shiksha, 2023)

Figure 9 shows the outline of the artificial intelligence method driven by machine learning (ML), which is divided into supervised and unsupervised learning (Mekonnen et al., 2020). Predictive elements of an unsupervised model, such as K-means, use clustering on a spread data set to identify distinct features or attributes in the farming field or crops. For instance, Hassan Ibrahim Hayatu et al. (2020) used the K-means clustering algorithm to study soil fertility across different regions in Nigeria. Therefore, such results will assist farmers in determining suitable crops to be planted in the targeted area with different soil nutrient attributes (Hassan Hayatu et al., 2020). Another IoT approach by Mekonnen et al. (2020) integrated WSN and ML into the system to monitor the smart farm prototype, as

illustrated in Figure 11 (Mekonnen et al., 2020). Figure 10 shows the clustering technique and prediction for the data set belonging to the domain (Hassan Hayatu et al., 2020; Sharma, 2024). Table 6 shows the number of Machine Learning (ML) algorithms in smart agriculture practice (Mekonnen et al., 2020).

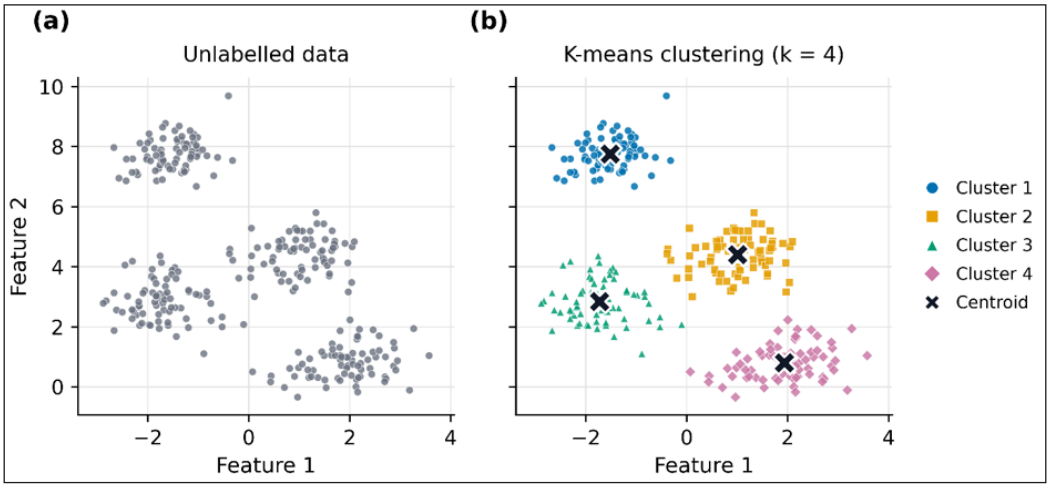


Figure 10. (a) Visualising the dataset. (a) K-means estimator methods calculate the number of clusters and the centroid to determine the soil fertility region (Shiksha, 2023)

Most of the electronic platforms are based on WSN ZigBee protocols, connected to various sensors, and supported by IoT frameworks. Finally, agricultural WSN-IoT is best organised as a layered cyber-physical stack linking field sensing, wireless communication, edge or cloud computation, and application intelligence, with consensus that WSNs remain the core data-acquisition layer and IoT extends them to remote control, analytics, and automation (Effah et al., 2023; Friha et al., 2021; Montalvo et al., 2025). Table 8 compares application architectures in agricultural WSN-IoT systems.

Table 6
Summary of Machine Learning method in Smart Agriculture implementation

Algorithms	Approach	Application	Data Acquisition
Extreme Learning Machine and Kernel-based ELM	Compares the approach with back-propagation (BP), SVM and Elman through simulations	To predict temperature and humidity in a greenhouse environment	ZigBee WSN-based
Fuzzy Logic	Fuzzy rules are used to predict the climate using temperature and humidity; the predicted climate and soil pH from the sensor are concurrently used to predict the crop to be grown	To predict the type of crop to be grown using sensor data from soil pH, temperature, and humidity. Critical information about the field is sent to farmers' mobiles using GSM	GSM and GPS WSN-based

Table 6 (*continued*)

Algorithms	Approach	Application	Data Acquisition
ARIMA	A WSN-based system called "Agri-prediction" was used to collect air temperature, soil temperature, air humidity, soil humidity and luminosity. ARIMA model was implemented to predict sensor data. ARIMA model was compared with ANN and RF, where the former model outperformed the others by producing the lowest MAE	To anticipate crop dysfunction proactively, farmers are then notified with possible remedies through smartphones. The model was tested in greenhouse Arugula cultivation, with gains associated with leaf development and weight	LoRa WSN-based
Data mining techniques, linear regression, and Apriori algorithm	Association rules are applied for data mining to find relationships among temperature, humidity, soil moisture, and yields of lemon and vegetables. Implemented linear regression to model the relationship between inputs	To analyse the sensor data for predicting suitable temperature, humidity, and soil moisture of crops in the future	WSN-based
Naive Bayes Kernel Algorithm	Real-time sensor data have been used to understand and quantify hidden correlations; pest/disease, humidity, temperature, pressure, pH value, nitrogen, soil moisture, and soil temperature are used	In forecasting and detecting pest/disease accurately using historical and real-time sensor data to monitor crop quality	ZigBee WSN-based
Bayesian Statistics	Sensor and weather data used to forecast changes and overall crop quality, image processing for weed extraction	A cloud-based decision support system uses the best algorithm to optimise events based on sensor data, scheduling irrigation based on moisture content sensors and weather forecast	WIFI and MQTT sensor-based with Raspberry Pi and Arduino MCU
Neural networks & Multi-Layer Perceptron (MLP)	MLP& neural network models are used to train on input sensor data to decide land cultivation in terms of four decision classes	For the assessment of agricultural land suitability	WSN-based

Smart Farming as Modern-Day Agriculture

Smart devices have infiltrated all aspects of our daily lives, including healthcare and wellness, smart homes, automobiles and logistics, intelligent cities and industries, thanks to the Internet of Things (IoT). Farming has undergone a series of technological transformations in the last few decades, becoming increasingly industrialised and

technologically oriented. Modern farmers can better manage their livestock and crops by using a variety of agricultural smart devices, making their operations more predictable and productive (Aleksandrova, 2018). As a result, there has been an increase in the use of intelligent agriculture technologies around the globe. It's worth noting that IoT's agricultural market share is expected to reach $5.6 billion, according to Research and Markets (2020).

Now it is possible to implement advanced tools and a variety of methodologies with reasonable costs for agricultural IoT devices, as predicted by BI Intelligence since 2020. The AgriTech system, for example, is an integrated application technology developed to benefit farming processes (Ilie, 2018). The terms “intelligent farming” and “smart agriculture” also generally refer to IoT-based solutions for autonomous data collection, such as wireless sensor network nodes and self-constructed topologies. Farmers can be informed by the system application, make decisions, and develop various parameters for their fieldwork, such as livestock production and the collection of environmental and machine metrics. By utilising smart agricultural sensors to monitor crop status, farmers can determine exactly how much pesticide and fertiliser to use to optimise productivity. The same can be said for the idea of practising intelligent agriculture. Figure 11 represents a model of smart agriculture incorporating WSN-IoT and Cloud or Edge Computing concepts for agriculture-centric applications, equipped with WSN infrastructure in the farm field. Figure 12 represents an emerging farming model by integrating the orbital

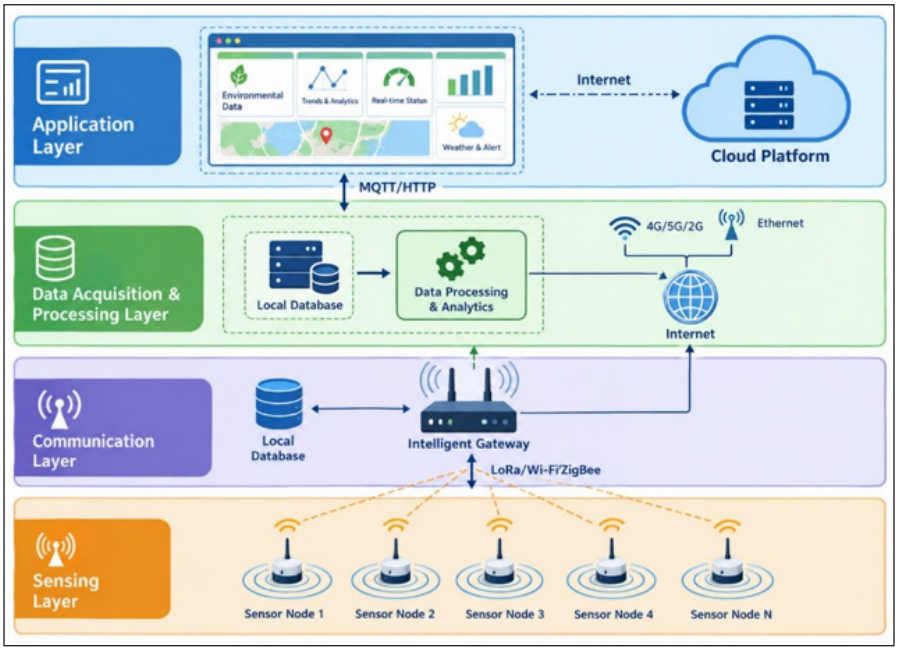


Figure 11. WSN-IoT and Cloud Edge Computing concept for the agriculture domain application

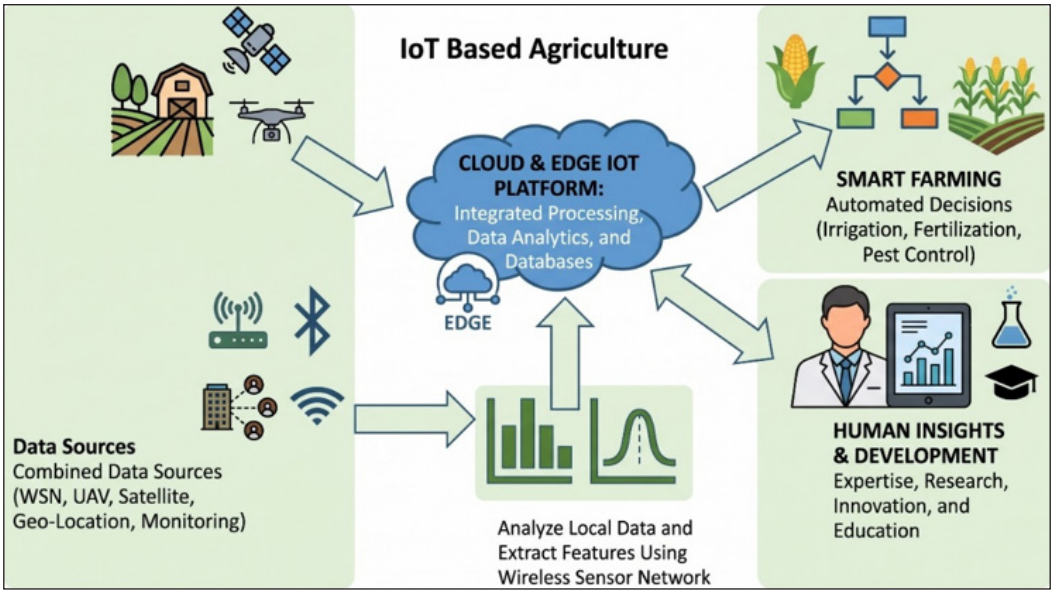


Figure 12. WSN-IoT and Cloud Edge Computing concept for the agriculture domain application

satellite uplink-downlink communication trajectory from sky to the ground for precision spatial data management (Mekonnen et al., 2020). In this model, various ground sensors are distributed over a larger area, with data storage, real-time processing, analysis, and automation handled by cloud computing. This allows farmers to make the most informed decisions possible.

Advanced sensors and intelligent cameras will precisely assist farmers in monitoring crop growth, periodic fertilisation, and setting threshold alarms until the ideal harvesting date. As developed by Deepfield Labs (Kushwaha et al., 2016), the Bonirob can analyse soil pH and phosphorus levels in real time by liquidising a soil sample. It is expected to automate the growth and harvest of entire hectares of barley. Another method, called SIM farming, involves farmers receiving a warning on their computers when an area has been disturbed. Additionally, tool developers such as Aerobotics have already commercialised software that analyses infrared drone imagery to detect areas of unhealthy vegetation (Human, 2018). Their software can identify unhealthy vegetation using images captured by a drone. The system is also capable of distinguishing between crop varieties and unwanted weeds on the farm, and it is continuously improved by the machine learning (ML) algorithm. Another approach is utilising Micro Aerial Vehicles (MAVs) and piloted light aircraft fitted with multispectral cameras for data-acquisition missions. They were flying over large farm areas in various locations nationwide (Gago et al., 2020). For smart-assist crop monitoring, Planet Labs provides CubeSat satellites that capture weekly photographs of entire farms from space (Kimm et al., 2020).

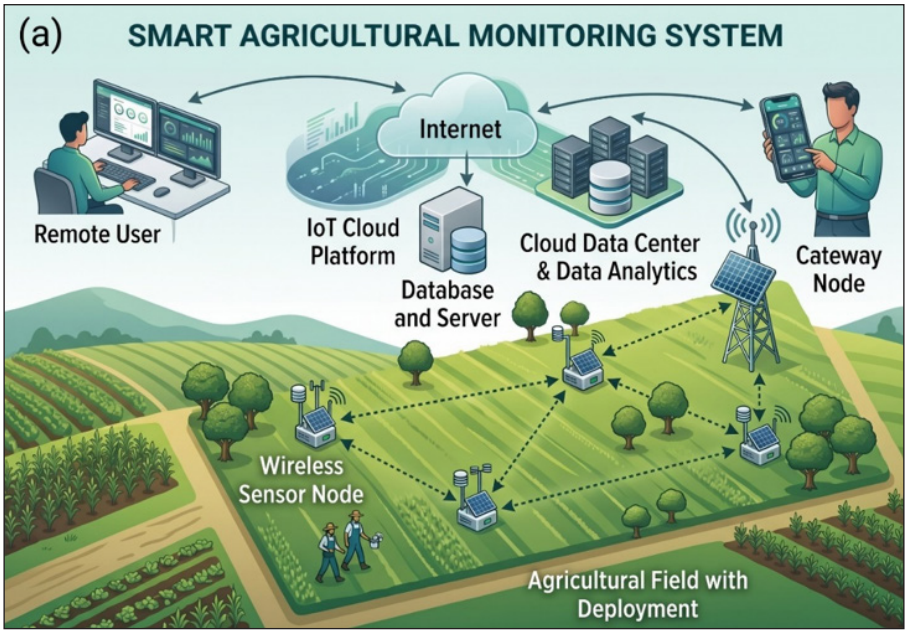
Challenges in IoT-Farming

The Internet of Things (IoT) enables living and non-living objects to have unique identifiers and to autonomously transmit data over the internet without human-to-computer interaction (Gluhak et al., 2011). According to a 2015 report by Juniper Research, more than 13.4 billion devices were connected to the Internet of Things (IoT) globally. The study projected substantial growth over the following years, forecasting that the ecosystem would expand to 38.5 billion connected devices by 2020 (Sto ces et al., 2016). Currently, the Internet of Things (IoT) is rapidly expanding into nearly every sector of modern society (Vermesan & Friess, 2013). Its primary domains encompass smart healthcare, smart cities, industrial automation, autonomous vehicles, smart homes, and precision agriculture (Suresh et al., 2014). To realise a fully interconnected IoT ecosystem, widespread internet connectivity across both residential and commercial sectors is imperative. This infrastructural demand has risen steadily, mirroring the growth in global internet penetration, which reached 73.8% in 2026, up from 42.8% in 2016 (Statista Research Department, 2026).

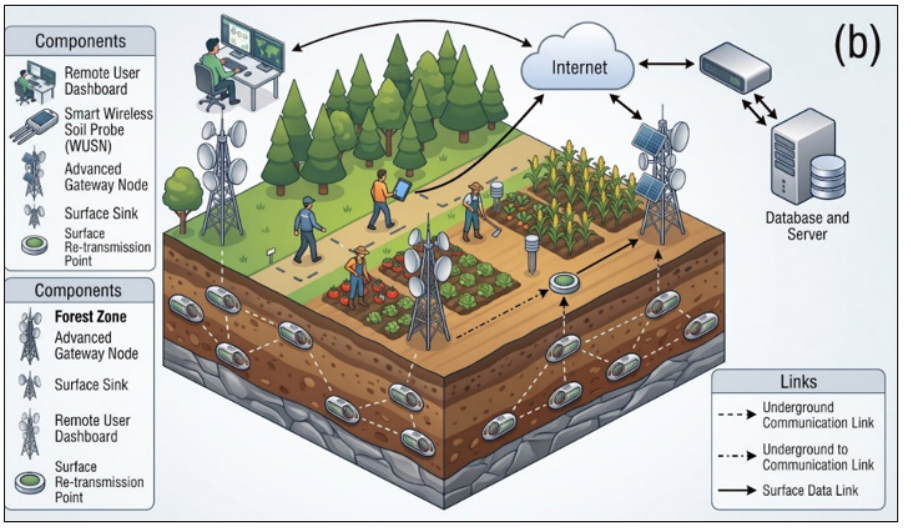
While Internet of Things (IoT) devices have significantly increased the demand for internet services, infrastructure growth could be restricted. This is largely because Internet Service Providers (ISPs) in major countries worldwide are still mostly in an exploratory phase regarding large-scale deployment (Jazayeri et al., 2015). IoT has been actively incorporated into the European Digital Agenda since 2016 (Osseiran et al., 2020). It is a key pillar of Horizon 2020, the European framework for research and innovation, and was backed by targeted European Commission funding windows between 2016 and 2017 (Cambra et al., 2015). Additionally, the European Commission established the Alliance for Internet of Things Innovation (AIOTI) in March 2015. This alliance was designed to facilitate deep cooperation on IoT standardisation and market development, connecting the EU Commission directly with key tech sectors and stakeholders (Kloas et al., 2015; Sto ces et al., 2016).

Currently, platform and communication protocol fragmentation represents a critical challenge, directly contributing to interoperability issues across diverse Internet of Things (IoT) devices and their respective ecosystems (Burhan et al., 2018; Gyory & Chuah, 2017). In precision agriculture, this challenge is intensified by a market dominated by proprietary systems, which severely hinders seamless data exchange and connectivity across different farming equipment (Ahmad, 2014).

Despite these interoperability barriers, comprehensive agricultural telemetry inherently depends on diverse of sensing topologies. As highlighted by Ojha et al. (2015), effective crop monitoring requires integrating two distinct networking layers into the IoT framework: terrestrial and underground WSNs. The structural and topological differences between these two deployment models are illustrated in Figure 13, showing (a) a conventional Terrestrial Wireless Sensor Network (TWSN) and (b) a Wireless Underground Sensor Network (WUSN).



(a)



(b)

Figure 13. (a) Terrestrial Wireless Sensor Networks (TWSNs) model. (b) Wireless Underground Sensor Networks (WUSNs) model (Ojha et al., 2015)

Conventional Terrestrial Wireless Sensor Networks (TWSNs) rely on over-the-air transmission to monitor ambient and canopy conditions. In contrast Wireless Underground Sensor Networks (WUSNs) deploy sensing nodes underneath the soil surface to assess

plants root-zone variables. In a typical WUSN topology, buried nodes transmit telemetry to an above-ground farm gateway via underground-to-aboveground links, after which the data is forwarded to remote servers via internet-enabled gateways. Nevertheless, subterranean communication encounters substantial radio frequency attenuation, where signal degradation within the dielectric soil medium, significantly decreases effective distance relative to the atmosphere-space terrestrial links (Ojha et al., 2015).

Role of Edge Computing and Smart Sensing

WSN (Wireless Sensor Networks), IoT (Internet of Things), Edge Computing, and Cloud Computing are not competing concepts; they represent different layers of a modern, distributed data ecosystem. WSN and IoT are the data generators (the eyes and ears), while Edge and Cloud computing are the data processors (the brain and muscle) (Zamora-Izquierdo et al., 2019). Returning to the agricultural context, Edge computing offers a promising solution for smart agriculture, enhancing real-time analysis and decision-making, and reducing network bandwidth constraints compared to native cloud processing (Gong et al., 2025; Santhosh et al., 2025). Cloud-Edge-Device collaborative computing in agriculture enables distributed intelligence, real-time processing, and adaptive decision-making to improve food security and resource management (Gong et al., 2025). For example, by implementing smart sensing and edge computing, the system will improve soil assessment and heavy metal monitoring in real time, thereby significantly benefiting a country's food farming industry on a large scale (Akhtar et al., 2021; Nguyen et al., 2024). Although numerous smart sensing systems are being developed, their implementation in developing countries remains challenging due to a lack of high-speed internet infrastructure. One viable solution to this problem could be edge computing, in which essential data is offloaded from the cloud to the cloud's edge, and this is where the smart sensing with edge computing paradigm comes in (Lytos et al., 2020). In Table 7, a comparison of features of Edge and Cloud computing in the WSN-IoT domain is presented.

Table 7
A Direct Comparison of Edge and Cloud computing in WSN-IoT application

Feature	Edge Computing	Cloud Computing
Location	Physically close to the WSN/IoT devices	Remote, centralized data centres
Latency	Extremely low (milliseconds), near real-time	Higher (hundreds of milliseconds), dependent on web server speed
Bandwidth	Conserves bandwidth by filtering data locally	High bandwidth consumption to transfer raw data
Processing Power	Limited - optimized for filtering, alerts, basic AI	Massive - optimized for Big Data, ML training, deep storage

Table 7 (*continued*)

Feature	Edge Computing	Cloud Computing
Offline Operation	Can function autonomously without internet connection	Completely internet access dependence
Data Privacy	Data can be anonymized locally, reduce transit exposure to data	Data must travel over the web, but benefits by top-tier security

APPLICATION DOMAINS IN SMART AGRICULTURE

Recent progress in precision agriculture highlights a paradigm shift toward context-aware intelligence and complex algorithmic optimisation across edge-cloud boundaries. To provide clearer comparison across use cases, this section reorganises agricultural WSN-IoT deployments into five structured application domains – Crop Health & Microclimate Monitoring, Precision Irrigation Management, Controlled-Environment Agriculture (Greenhouses & Aquaponics), Remote Sensing & Aerial/Terrestrial Robotics, and Livestock Telemetry & Animal Welfare. Each domain follows a standardised comparative template: (i) Common Architectural Topology, (ii) Core Sensing and Communication Technologies, (iii) Domain-Specific Bottlenecks and Failures, and (iv) Key Architectural Lessons Learned. A master comparative evaluation (Table 8) closes the section with a direct side-by-side assessment across all five domains.

Table 8
Master comparative evaluation across the five smart-agriculture application domains

Domain	Common Topology	Core Technologies	Main Bottleneck	Design Takeaway
Crop Health & Microclimate	Star/tree (terrestrial) + underground relay hop (WUSN)	NPK/EC/pH probes, ZigBee/LoRa, ML fertilizer classifiers	Underground signal attenuation, short WUSN range	Fuse terrestrial + underground data streams
Precision Irrigation	Star, sensor-to-controller closed loop	Soil moisture/flow sensors, IoT weather station, deep-learning ETo	Site-specific model validity, limited field validation	Couple real-time feedback with predictive ETo modelling
Controlled-Environment (Greenhouse/ Aquaponics)	Enclosed star, node-to-local-gateway	Moisture/pH/DO sensors, solar nodes, mobile app bridge	Multi-subsystem coupling complexity	Trade RF/power constraints for tighter closed-loop control
Remote Sensing & Robotics	Ground WSN + independent aerial/ satellite tier	NDVI/multispectral imaging, CubeSat, cloud fusion	Ground-truth/ imagery misalignment, flight cost & weather	Fuse imagery with ground sensors, not as a substitute
Livestock Telemetry & Welfare	Mobile node to fixed gateway (opportunistic)	RFID/BLE tags, accelerometers, ZigBee/NB-IoT backhaul	Intermittent connectivity, battery drain, tag damage	Prioritize delay-tolerant links and rugged hardware

(i) Common Architectural Topology

Star or tree-based terrestrial motes report to a field gateway that uplinks to the cloud; WUSN deployments add an underground-to-surface relay hop, with a surface gateway bridging the buried nodes to the internet.

(ii) Core Sensing and Communication Technologies

Soil NPK/EC/pH probes and ambient temperature/humidity sensors are interfaced over ADC/I2C, most paired with ZigBee or LoRa radios; NPK recommendation layers add a Gaussian Naïve Bayes or comparable ML classifier at the gateway or cloud tier.

(iii) Domain-Specific Bottlenecks and Failures

Nodes in WUSNs experience severe radio attenuation underground, leading to shorter transmission distances (typically under 10 m) and higher power consumption; wired underground alternatives avoid attenuation but are impractical due to the risk of damage from planting activity (Ojha et al., 2015).

(iv) Key Architectural Lessons Learned

Terrestrial and underground sensing are complementary rather than substitutable – surface nodes capture canopy/microclimate conditions while underground nodes capture root-zone nutrient status; recommendation accuracy depends more on fusing both data streams than on either sensing layer alone.

Precision Irrigation Management

Irrigation schedules have advanced beyond static thresholds through intelligent optimisation of Reference Evapotranspiration (ETo); these frameworks process localised microclimate sensor arrays using deep learning models to predict moisture loss, enabling precision-driven drip irrigation control (Bashir et al., 2023). Automated data-driven approaches such as HRLDAS and AquaCrop claim to effectively save 10-40% of water without affecting crop yield (Zhao et al., 2023). A descriptive monitoring is presented for mustard leaf cultivation, utilising an IoT cloud-based weather station associated with a WSN device to sense soil moisture and automatically control a water-drip mechanism in a greenhouse (Abioye et al., 2021). In Figure 14, system is driven by a Raspberry Pi 3 as the main controller, a VH400 moisture sensor, and a YF-S201 flow meter, with the data modelled in MATLAB for predictive and dynamic validation of the relationship between irrigation water volume, ETo loss, and soil moisture level (Abioye et al., 2021).

(i) Common Architectural Topology

Star topology linking soil-moisture and flow sensors to a Raspberry Pi/microcontroller hub, which fetches weather-station ETo data from the cloud and actuates a solenoid valve or DC pump.

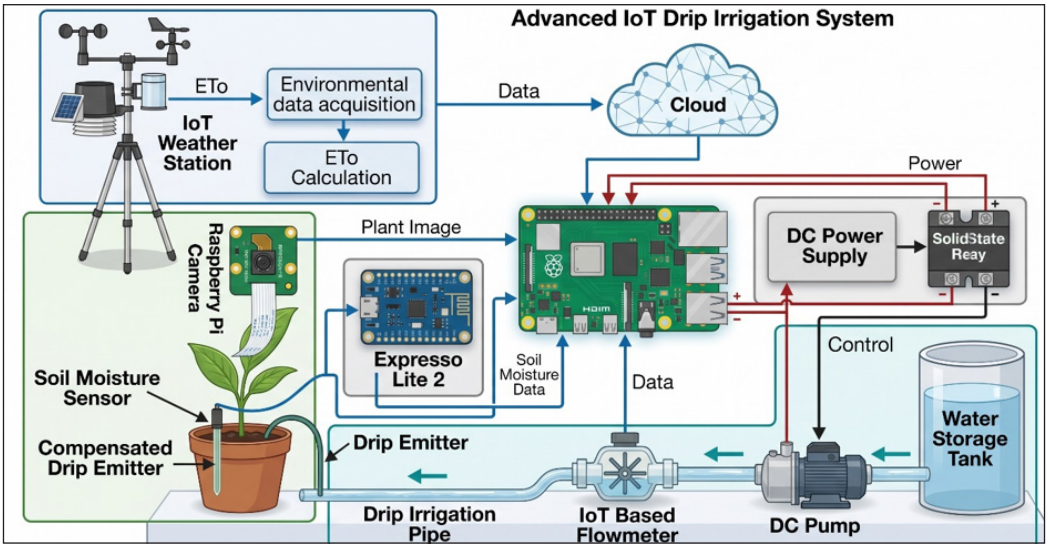


Figure 14. Automatic water drip irrigation and monitoring system architecture for mustard leaf cultivation in a greenhouse
Source: Adapted from (Abioye et al., 2021)

(ii) Core Sensing and Communication Technologies

Soil moisture probes (e.g., VH400), flow meters, IoT-based weather stations, and deep-learning ETo prediction models feeding a closed control loop.

(iii) Domain-Specific Bottlenecks and Failures

Model validity is highly site-specific: studies calibrated in one greenhouse profile (e.g., centrally peaked vs centrally depressed) do not generalise across profile types without recalibration, and several reported systems were not validated against actual field-based crop yield outcomes.

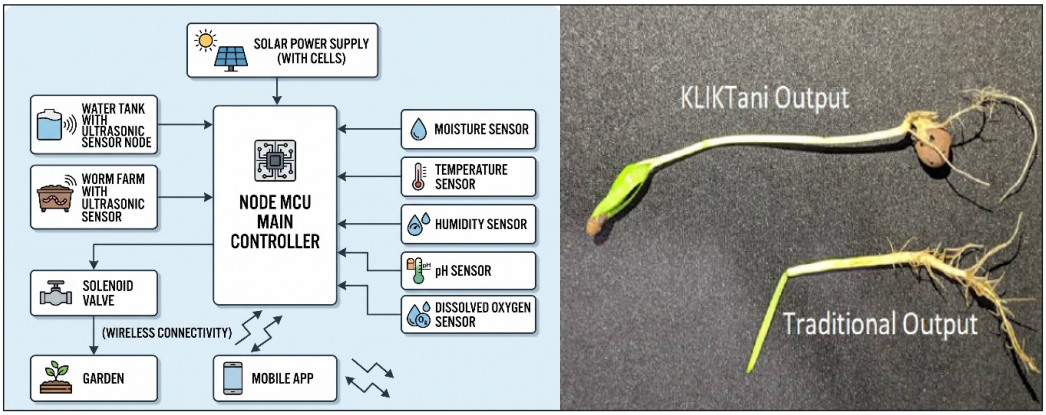
(iv) Key Architectural Lessons Learned

Coupling real-time soil moisture feedback with predictive ETo modelling, rather than static timers, is what delivers the reported water savings; the actuation layer (valve/pump control) is comparatively simple once the sensing and prediction layers are reliable.

Controlled-Environment Agriculture (Greenhouses & Aquaponics)

Other implementations set temperature and luminosity profiles inside greenhouses using WSN multi-sensor nodes and mathematical models; the central finding is that the original Equal Step (ES) technique is not suitable for centrally peaked or centrally depressed

greenhouse profiles, while Equal Segment Area (ESA) remains applicable across profile types (Lata & Verma, 2022; Lata et al., 2019), although this line of work did not validate the model against actual crop growth.



(a) (b)

Figure 15. (a) KLIKTani system architecture diagram. (b) The Kangkung vegetable size and growth rates after 7 days planted by IoT KLIKTani (Zamin et al., 2019)

In Malaysia, Zamin et al. (2019) developed the KLIKTani model, using smart IoT with red and tiger worm vermiculture to facilitate edible farming and Kangkung vegetable fertilisation; after 7–14 days, the experiment showed better outcomes in Kangkung size and growth rate compared to conventional planting as shown in Figure 15 (Zamin et al., 2019). Werner Kloas et al. (2015) applied a related concept in the ASTAF-PRO aquaponics system for tomato and fish co-cultivation in a greenhouse (Kloas et al., 2015), where fish-breeding water waste feeds the hydroponic planting system and water/air-quality sensors (pH, temperature, nitrate-to-ammonia concentration) manage the coupled system's stability (König et al., 2018; Yanes et al., 2020).

(i) Common Architectural Topology

Enclosed-environment star topology; sensor nodes report to a local Node MCU or gateway, which drives actuators (solenoid valves, worm-farm/nutrient dosing units) and is optionally bridged to a mobile app.

(ii) Core Sensing and Communication Technologies

Moisture, temperature, humidity, pH, and dissolved-oxygen sensors (for aquaponics variants) are typically on solar-powered nodes with short-range wireless connectivity to a local controller.

(iii) Domain-Specific Bottlenecks and Failures

System complexity increases sharply when coupling two subsystems (e.g., fish tank and hydroponic bed in aquaponics); several greenhouse-profiling studies also lacked field validation with actual food crops, limiting confidence in generalising lab results to production settings.

(iv) Key Architectural Lessons Learned

Controlled-environment deployments trade the harsher RF and power constraints of open fields for tighter, more complex closed-loop control and multi-subsystem coordination; smaller-scale local experiments (e.g., KLIKTani) demonstrate measurable growth benefits but require field-scale validation before broader claims can be made.

Remote Sensing & Aerial/Terrestrial Robotics

Uferah Shafi et al. (2019) proposed a precision agriculture approach that combines ground-deployed WSN-IoT devices with multispectral imaging from an aerial vehicle (drone) for crop abnormality monitoring (Figure 16). The farming area was deployed with a star-distributed wireless sensor topology to measure soil, ambient air temperature, and humidity

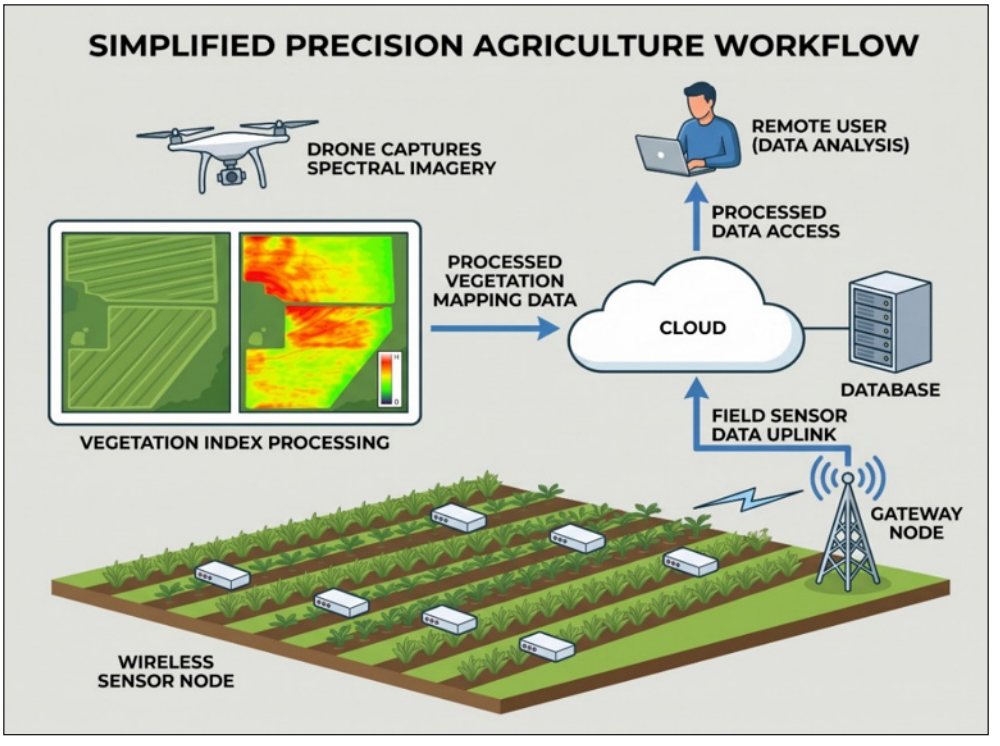


Figure 16. The NDVI and IoT-Based Precision Agriculture was published by Uferah Shafi et al., 2019

(Shafi et al., 2019), while the drone captured multi-spectral images to determine chlorophyll levels using NDVI (Normalised Difference Vegetation Index) computation. All monitoring data were uploaded to a cloud server, and crop health status was classified following NDVI mapping (Shafi et al., 2019). The IoT cloud system manages the collected data, and users access information through a web portal or mobile application that alerts farmers to crop health status via graphical data and scale maps, particularly useful for differentiating grain health and ripeness across large wheat or paddy fields.

(i) Common Architectural Topology

Ground star-topology WSN combined with an independent aerial (drone/UAV) or satellite imaging tier; both streams converge at a cloud server for fused analysis.

(ii) Core Sensing and Communication Technologies

Ground-temperature/humidity sensors, multispectral/NDVI-capable drone cameras, and CubeSat satellite imagery (e.g., Planet Labs) for wide-area coverage, integrated via cloud-based classification pipelines.

(iii) Domain-Specific Bottlenecks and Failures

Requires tight integration between ground-sensor calibration and imaging-derived indices; misalignment between sensor ground truth and spectral imagery reduces classification reliability, and aerial/satellite passes are constrained by flight costs, weather, or revisit frequency.

(iv) Key Architectural Lessons Learned

Large-area crop-health classification is most effective when NDVI/multispectral imagery is fused with, not substituted for, ground WSN data; imagery gives spatial coverage while ground sensors provide the calibration anchor.

Livestock Telemetry & Animal Welfare

Beyond crop and soil monitoring, WSN and IoT technologies have also been applied to flora-and-fauna and livestock-farming contexts, extending real-time monitoring to animal health, location, and welfare. Kandris et al. (2020) classify livestock-farming WSN applications alongside greenhouse and crop-monitoring deployments, noting that livestock telemetry typically combines RF-based identification (RFID/BLE ear tags) with photographic or camera-based sensing for behaviour and health observation, mirroring the crop/environmental sensing stack illustrated in Figure 17 but applied to animals rather than plants (Kandris et al., 2020). As shown in the taxonomy in Figure 18, livestock is categorized as a distinct sensing target alongside soil, crops, environmental and machinery categories, with ear-tag and health-monitoring devices leverage the same short-range and

cellular connectivity layers (Bluetooth, ZigBee, NB-IoT/4G-5G) deployed elsewhere across agricultural WSN-IoT architectures.

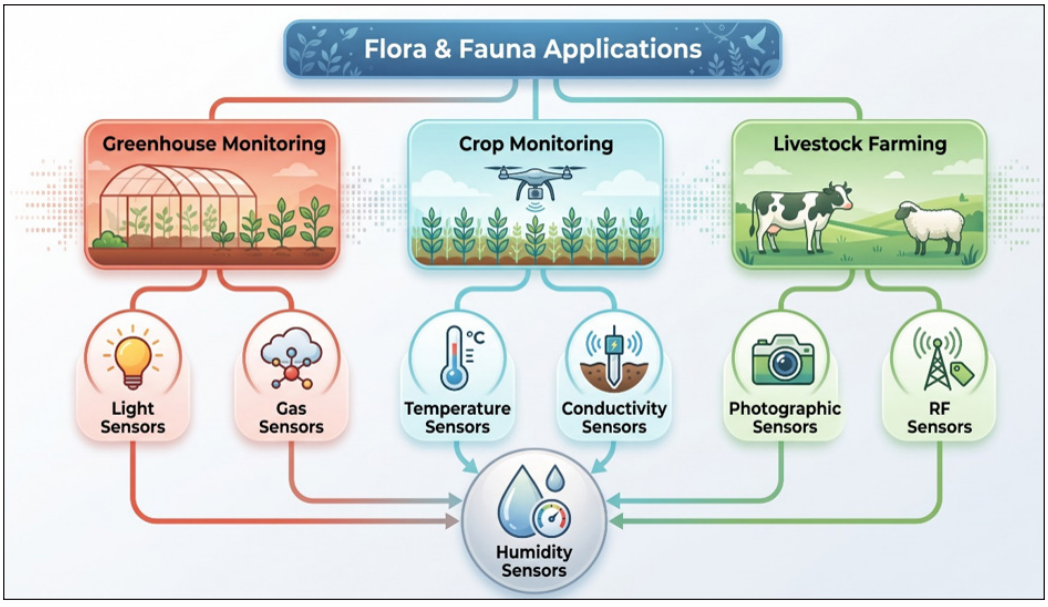


Figure 17. Flora and fauna application monitoring

(i) Common Architectural Topology

Mobile-node topology: wearable ear-tag or collar-mounted sensors move with the animal and connect opportunistically to fixed gateways distributed across pasture or barn areas, rather than the fixed-node star/tree topologies typical of crop monitoring.

(ii) Core Sensing and Communication Technologies

RFID/BLE ear tags, accelerometer-based activity and health monitors, and photographic/RF sensors for behaviour tracking, typically backhauled via short-range mesh (ZigBee/Bluetooth) to a farm gateway, with NB-IoT or cellular links used for wide-ranging or free-grazing herds.

(iii) Domain-Specific Bottlenecks and Failures

Mobility of the sensing node itself is the central challenge; unlike fixed crop sensors, animal-worn devices experience variable and intermittent connectivity, higher battery drain from continuous motion sensing, and physical risk of tag damage or loss, none of which arise in static field deployments.

(iv) Key Architectural Lessons Learned

Livestock telemetry architectures need to prioritise opportunistic, delay-tolerant connectivity and ruggedised, low-maintenance hardware over the raw throughput or dense node-count optimisations that dominate crop-sensing network design; welfare and productivity outcomes depend as much on tag durability and battery life as on sensing accuracy.

Table 9
Comparison of application examples in agricultural WSN-IoT systems

Application Domain	Sensing Parameter	Communication & Computing	Main Contribution	Key Limitation
Open-field WSN	Soil and ambient sensors	Terrestrial or underground WSN	Large-area monitoring and automation	Underground attenuation, higher energy
Greenhouse irrigation	Moisture, flow, weather inputs	IoT cloud, predictive model	Irrigation volume estimation, save water	Controlled setting limits generalisation
Aquaponics / hybrid farming	Water quality and air sensors	IoT monitoring, controllers	Resource reuse and coupled management	System complexity across subsystems
Remote sensing agriculture	Soil nodes + multispectral imaging	Cloud analytics	Large-area crop health classification	Requires sensor-imaging integration
Context-aware agronomy	Soil fertility, salinity, temperature	IoT, ML models	Actionable fertilizer or ET recommendations	Data quality & context mapping complexity
Livestock telemetry	Ear-tag/collar motion, RFID/ BLE	Mobile-node WSN, NB-IoT/ cellular backhaul	Health/welfare and productivity monitoring	Mobility-driven connectivity & battery limits

Table 9 additionally summarises representative application examples at the level of individual studies, complementing the domain-level synthesis in Table 8. On the architectural network layer, managing large data flows from multivariant sensors across all five domains requires high-efficiency communication paradigms. Researchers have successfully implemented novel caching frameworks for information-centric IoT, utilising deep reinforcement learning via Proximal Policy Optimisation (PPO) to minimise backhaul congestion and optimise data placement at edge nodes (Asmat et al., 2024). The foundational reliability of data collection also relies on accurate node-location metrics within distributed layouts; this is directly addressed by robust localisation algorithms like PSLDV-Hop, which fuses Particle Swarm Optimisation (PSO) with targeted refinement processes to compute accurate WSN node coordinates over complex, irregular field topologies (Hadir

et al., 2021; Kaur et al., 2025; Rehman et al., 2025). Together, these advances increase data-matrix integrity, node localisation accuracy, power longevity, and real-time precision for the autonomous management of widespread vegetable and fruit crops at the farm scale.

As claimed by Kloas et al. (2015), about 70% of water resources have been consumed and wasted across various agricultural mass-production processes (Kloas et al., 2015), reinforcing the relevance of the aquaponics and precision-irrigation approaches discussed above. To achieve integrated decision-making across domains, Carlos-Mancilla et al. (2019) suggested a Smart Decision System (SDS) to manage large IoT data flows – from acquisition to actuation (watering, fertilising, crop condition, and others) – keeping the farmer informed (Baseca et al., 2019; Carlos-Mancilla et al., 2016). In crop-essential monitoring such as irrigation and fertilisation, wireless sensors provide advantages in cost-effectiveness, commissioning time, flexibility, and safety over wired alternatives. Dionisis Kandris et al. (2020) conducted a survey on the evolving use of WSN across healthcare, environmental, defence, and agricultural applications (Kandris et al., 2020), underscoring that the architectural lessons summarised in Table 9 are broadly consistent with WSN deployment patterns observed outside agriculture as well. Figure 18 delineates primary taxonomy of smart agriculture distributions.



Sub-Category	Measurement / Monitor Examples
Soil	• Moisture sensor, NPK
Crop/Plant	• Plant health, Chlorophyll sensor • Drone imagery
Environmental	• Temperature, Humidity
Livestock	• Ear tag, Health monitor
Machinery	• Tractor telematics • Fuel consumption

Environment Type	Characteristics / Examples
In-Situ Terrestrial	• Ground-based sensors • Local fields
Aerial	• Drones, High-altitude imagery
Spaceborne	• Satellites • Regional / Global coverage
Controlled Environment	• Greenhouses, Vertical farms • Indoor facilities

Connection Type	Network / Technology Examples
Short-Range	• Bluetooth, Zigbee, Wi-Fi
Long-Range LPWAN	• LoRaWAN, Sigfox
Cellular	• 4G / 5G, NB-IoT
Satellite	• Remote field connectivity

Layer	Functions / Examples
Cloud Layer	• Data Center • Big Data analytics (Centralized)
Fog/Gateway Layer	• Field Server • Local processing (Intermediate)
Edge Layer	• On-device processing • Sensor nodes (Decentralized)

Objective	Key Applications
Yield Forecasting	• Prediction and supply chain management
Precision Irrigation	• Optimized water usage
Pest & Disease Management	• Early detection and control
Livestock Health Monitoring	• Welfare and productivity
Resource Optimization	• Fertilizer, fuel, labor efficiency

Level	Operation & Human Involvement
Manual	• Human-in-the-loop (Manual data logging and control)
Semi-Autonomous	• Assisted systems (e.g., tractor autopilot with supervision)
Fully Autonomous	• Unmanned operation (Independent rovers, automatic drone swarms)
Swarm Automation	• Coordinated robotics (Multiple agents working together)

Figure 18. Architectural Taxonomy on Smart Agriculture deployment

SYNTHESIS, CRITICAL REFLECTION, AND FUTURE RESEARCH DIRECTIONS

Global disruptions over the past decade were intensified by pandemic-induced economic shifts and severe supply chain vulnerabilities, such as in the 2023 global rice shortage. Therefore, this situation has elevated food security to an urgent national priority. For nations navigating moderate-scale food-security challenges, traditional farming methods are no longer sufficient to buffer against the impacts of unpredictable climate change, geopolitical friction, and resource scarcity. This review has comprehensively evaluated the transformative role of Precision Agriculture (PA), demonstrating how integrating Wireless Sensor Networks (WSNs) and Internet of Things (IoT) architectures can shift modern farming from a reactive practice to a proactive, data-driven science. This concluding section consolidates the evaluation into five parts: a synthesis of major findings; the core technological trade-offs identified across the review; the critical research gaps and evaluation-methodology weaknesses that remain; emerging trends likely to shape the next generation of systems; and a brief note on the scholarly-presentation audit performed on this manuscript.

Synthesis of Major Findings

Through a systematic examination of current device-aided technologies, this paper consolidates findings across four layers of the agricultural cyber-physical stack:

- Hardware and Data Acquisition: at the foundational level, low-power microcontrollers (such as the PIC16F886 and PIC16F684) act as the backbone of field-deployed motes, conditioning and converting real-time analogue sensor data via internal ADCs and routing digitised payloads over standard interfaces (I2C, SPI, UART).
- Protocol Scalability and Efficiency: wireless communication technologies present clear power/capacity trade-offs. Legacy protocols such as Wi-Fi and classic Bluetooth are constrained by heavy power draw and limited node capacity; ZigBee stands out for energy-efficient, high-density localised areas (36.9 mW transmission power, self-healing mesh), while Sigfox and LoRaWAN offer massive scalability (up to 1,000,000 and 10,000 nodes per base station, respectively) suited to expansive agricultural zones.
- Network Topologies: star topologies minimise routing delay to a central server, while multi-sink, tree, or mesh configurations provide the structural resilience needed to gather data across vast open fields without draining node batteries.
- Application-Layer Convergence: across all five application domains reviewed (Table 9), data collection consistently functions as the entry point rather than the end goal – the review's central finding is that the operational value of WSN-IoT systems is realised only once sensing is coupled with spatial visualisation and cognitive computation.

Advanced NPK mapping, for example, shows that IoT telemetry aligns closely with laboratory chemical analysis, with a Gaussian Naïve Bayes classifier achieving 94% test accuracy for fertiliser recommendation, enabling systems to autonomously dictate agricultural scheduling and actuation.

Core Technological Trade-offs

Across the domains reviewed, three structural trade-offs recur and should guide technology selection in future deployments:

- Communication range vs. bandwidth/power: LPWAN options (Sigfox, LoRaWAN, NB-IoT) trade throughput and latency for range and battery life, while ZigBee and Wi-Fi trade range for higher data rates and lower per-hop latency – no single protocol dominates across all agricultural scenarios (Table 2, Table 5).
- Edge-compute autonomy vs. energy/storage cost: pushing inference to the edge reduces bandwidth pressure and enables offline operation (Table 7) but constrains the model complexity that can run on-node and adds hardware/energy cost that must be justified by the value of faster, local decisions.
- Low-cost sensor accessibility vs. long-term calibration stability: affordable sensors (e.g., low-cost EC or gas sensors, Table 4) widen access to smallholder and developing-region deployments but frequently trade away long-term calibration stability and cross-sensitivity resistance found in industrial-grade instruments, creating a recurring accuracy/affordability tension not yet resolved in the literature.

Critical Research Gaps and Evaluation Methodologies

Despite the undeniable potential of WSN, IoT, and AI synergy, several persistent bottlenecks limit widespread, reliable open-field deployment:

- Subterranean signal attenuation in WUSNs, which continues to restrict underground node range to under 10 metres and elevates power draw beyond what is acceptable for long-term unattended deployment (Section: Crop Health & Microclimate Monitoring).
- Cybersecurity and data privacy: agricultural WSN-IoT systems inherit the same attack surface as other IoT deployments – unauthenticated sensor-to-gateway links, weak key management on constrained motes, and cloud-side data exposure – yet security is addressed unevenly across the reviewed literature, with encryption (e.g., 128-bit AES on ZigBee) often treated as a protocol default rather than a deployment-level design decision. Edge-local anonymisation (Table 7) offers a partial mitigation but is not yet standard practice.
- Cross-platform interoperability and standardisation: proprietary systems dominate the market, fragmenting protocols and data formats across vendors and severely hindering

data exchange between different farming equipment – this remains the single most cited structural barrier to scaling smart-farming pilots into production systems.

- Data management and long-term evaluation benchmarks: most studies reviewed report short-duration trials (weeks to a single season) rather than multi-season field validation, and no standardised benchmark protocol exists for comparing yield, water-saving, or accuracy claims across studies – a gap that future work should address through standardised, multi-season evaluation methodologies with openly reported datasets.

Emerging Trends and Strategic Future Directions

Building on the gaps identified above, several transformative paradigms are positioned to shape the next generation of agricultural WSN-IoT systems:

- TinyML on low-power microcontrollers: running compact ML inference directly on field motes (rather than solely at the edge gateway or cloud) could reduce the communication overhead identified in Section B while enabling faster, autonomous, on-node decisions for tasks such as pest/disease pattern recognition.
- Federated Learning for privacy-preserving farm analytics: training shared ML models across multiple farms without centralising raw sensor data would directly address the data-privacy gap noted in Section C, while still allowing smallholders to benefit from aggregated model improvements.
- Agricultural Digital Twins: virtual, continuously updated replicas of physical fields or greenhouses – fed by the WSN-IoT sensing layers reviewed throughout this paper – could support the multi-season, standardised evaluation methodologies called for in Section C by allowing scenario testing before physical intervention.
- Satellite-Direct LPWAN connectivity: extending NB-IoT/LoRaWAN-class links directly to low-earth-orbit satellite backhaul would remove the cellular/local-gateway dependency that currently limits deployment in remote or developing-region farmland lacking terrestrial infrastructure.
- Biodegradable, in-soil sensors: emerging low-cost, environmentally degradable sensor platforms could relax the trade-off between sensor accessibility and long-term ecological impact identified in Section B, particularly for WUSN-style underground deployments that are otherwise difficult to retrieve at the end of life.

CONCLUSION

Global disruptions over the past decade, intensified by pandemic-induced economic shifts and severe supply chain vulnerabilities such as the 2023 global rice shortage, have elevated food security to a critical national priority. For nations navigating moderate-scale food

security challenges, traditional farming methods are no longer sufficient to buffer against unpredictable climate changes, geopolitical friction, and resource scarcity. This review has comprehensively evaluated the transformative role of Precision Agriculture (PA), demonstrating how integrating Wireless Sensor Networks (WSNs) and Internet of Things (IoT) architectures can shift modern farming from a reactive practice to a proactive, data-driven science.

Key Technical Synthesis

Through a systematic examination of current device-aided technologies, this paper highlighted the core hardware and architectural frameworks required to achieve large-scale smart farming:

- Hardware and Data Acquisition: At the foundational level, low-power microcontrollers (such as the PIC16F886 and PIC16F684) act as the backbone of field-deployed motes. These devices effectively condition and convert real-time analogue sensor data via internal Analogue-to-Digital Converters (ADCs) and route the digitised payloads using standard interfaces (I2C, SPI, UART).
- Protocol Scalability and Efficiency: Wireless communication technologies present clear trade-offs in power and capacity. While legacy protocols like Wi-Fi and classic Bluetooth are constrained by high power consumption and limited node capacity, specialised protocols offer robust alternatives. ZigBee stands out as an exceptionally energy-efficient option for high-density, localised areas due to its low transmission power (36.9 mW) and self-healing mesh capabilities. Conversely, long-range technologies like Sigfox and LoRaWAN offer massive network scalability, accommodating up to 1,000,000 and 10,000 nodes per base station, respectively, making them highly suited for expansive agricultural zones.
- Network Topologies: Network design plays a vital role in balancing latency and coverage. Star topologies excel at minimising routing delays by routing traffic directly to a server, whereas multi-sink, tree, or mesh configurations provide the structural resilience required to gather data across vast open fields without draining node batteries.

The Evolution of Intelligent Farming

A central finding of this review is that data collection is merely the catalyst; the true value of WSN-IoT frameworks lies in spatial visualisation and cognitive computation. Historically, yield monitoring relied on static data analysed using localised GIS software such as ArcInfo, MapInfo, or ESRI platforms. Today, the vanguard of precision agriculture has shifted toward context-aware intelligence optimised across edge-cloud boundaries. By unifying real-time IoT sensing with cloud-hosted Machine Learning (ML) models,

modern systems can dynamically process complex environmental parameters. This is explicitly demonstrated by advanced NPK mapping techniques, in which IoT telemetry closely aligns with traditional laboratory chemical analysis. When this data is fed into a predictive algorithm, such as a Gaussian Naïve Bayes classifier, it achieves 94% test accuracy for fertiliser recommendations. Therefore, the system can autonomously dictate optimal agricultural scheduling and actuation to maximise crop yields while minimising environmental degradation.

Challenges and Future Directions

Despite the undeniable potential of the synergy among WSN, IoT, and AI, widespread open-field deployment faces a critical bottleneck: market fragmentation. The current agricultural ecosystem is heavily constrained by proprietary solutions and unpredictable costs. This lack of standardised communication protocols and platforms builds compatibility and connectivity barriers between disparate equipment, isolating valuable data silos. Future research and industrial development must prioritise the creation of open-source, interoperable frameworks that allow plug-and-play compatibility across different hardware manufacturers. Additionally, future work should focus on enhancing edge computing capabilities to reduce reliance on constant cloud connectivity in remote farming regions. Resolving these interoperability and structural limitations is the definitive next step toward achieving global "Zero Hunger" strategies, securing food production lines, and building a sustainable, resilient agricultural future.

ACKNOWLEDGEMENT

This research is fully supported by Universiti Sains Malaysia, Short Term Grant, 304/PAERO/6315567. The authors fully acknowledged Universiti Sains Malaysia for the approved funding, which makes this important research viable and effective.

REFERENCES

Abioye, E. A., Abidin, M. S. Z., Mahmud, M. S. A., Buyamin, S., AbdRahman, M. K. I., Otuoze, A. O., Ramli, M. S. A., & Ijike, O. D. (2021). IoT-based monitoring and data-driven modelling of drip irrigation system for mustard leaf cultivation experiment. *Information Processing in Agriculture*, *8*(2), 270-283. https://doi.org/10.1016/j.inpa.2020.05.004

Ahmad, M. (2014). Reliability models for the internet of things: A paradigm shift. In *2014 IEEE International Symposium on Software Reliability Engineering Workshops* (pp. 52-59).

Akhtar, M. N., Shaikh, A. J., Khan, A., Awais, H., Bakar, E. A., & Othman, A. R. (2021). Smart sensing with edge computing in precision agriculture for soil assessment and heavy metal monitoring: A review. *Agriculture*, *11*(6), 475. https://doi.org/10.3390/agriculture11060475

Akpakwu, G. A., Silva, B. J., Hancke, G. P., & Abu-Mahfouz, A. M. (2018). A survey on 5G networks for the internet of things: Communication technologies and challenges. *IEEE Access*, *6*, 3619-3647. https://doi.org/10.1109/ACCESS.2017.2779844

Alahmad, T., Neményi, M., & Nyéki, A. (2023). Applying IoT sensors and big data to improve precision crop production: A review. *Agronomy*, *13*(10), Article 2603. https://doi.org/10.3390/agronomy13102603

Aleksandrova, M. (2018). *IoT in agriculture: Five technology uses for smart farming and challenges to consider.* DZone. https://dzone.com/articles/iot-in-agriculture-five-technology-uses-for-smart

Alqurashi, H., Bouabdallah, F., & Khairullah, E. (2023). SCAP SigFox: A scalable communication protocol for Low-Power Wide-Area IoT networks. *Sensors*, *23*(7), Article 3732. https://doi.org/10.3390/s23073732

Ashok Somani, N. (2012). Zigbee: A low power wireless technology for industrial applications. *International Journal of Control Theory and Computer Modeling*, *2*(3), 27-33. https://doi.org/10.5121/ijctcm.2012.2303

Asmat, H., Din, I. U., Almogren, A., Altameem, A., & Khan, M. Y. (2024). Enhancing Edge-Linked caching in Information-Centric networking for internet of things with deep reinforcement learning. *IEEE Access*, *12*, 154918-154932. https://doi.org/10.1109/ACCESS.2024.3483455

Avşar, E., & Mowla, M. N. (2022). Wireless communication protocols in smart agriculture: A review on applications, challenges and future trends. *Ad Hoc Networks*, *136*, Article 102982. https://doi.org/10.1016/j.adhoc.2022.102982

Baronti, P., Pillai, P., Chook, V. W. C., Chessa, S., Gotta, A., & Hu, Y. F. (2007). Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards. *Computer Communications*, *30*(7), 1655-1695. https://doi.org/10.1016/j.comcom.2006.12.020

Baseca, C. C., Sendra, S., Lloret, J., & Tomas, J. (2019). A smart decision system for digital farming. *Agronomy*, *9*(5). https://doi.org/10.3390/agronomy9050216

Bashir, R. N., Khan, F. A., Khan, A. A., Tausif, M., Abbas, M. Z., Shahid, M. M. A., & Khan, N. (2023). Intelligent optimization of reference evapotranspiration (ETo) for precision irrigation. *Journal of Computational Science*, *69*, Article 102025. https://doi.org/10.1016/j.jocs.2023.102025

Ben Hassen, T., & El Bilali, H. (2022). Impacts of the Russia-Ukraine war on global food security: Towards more sustainable and resilient food systems?. *Foods*, *11*(15), Article 2301. https://doi.org/10.3390/foods11152301

Ben Hassen, T., & El Bilali, H. (2024). Three years into the pandemic: Insights of the COVID-19 impacts on food security and nutrition in low and middle-income countries. *Heliyon*, *10*(7), Article e28946. https://doi.org/10.1016/j.heliyon.2024.e28946

Buchli, B., Sutton, F., & Beutel, J. (2012). GPS-Equipped wireless sensor network node for High-Accuracy positioning applications. In (pp. 179-195). https://doi.org/10.1007/978-3-642-28169-3_12

Buratti, C., Conti, A., Dardari, D., & Verdone, R. (2009). An overview on wireless sensor networks technology and evolution. *Sensors*, *9*(9), 6869-6896. https://doi.org/10.3390/s90906869

Burhan, M., Rehman, R. A., Khan, B., & Kim, B. S. (2018). IoT elements, layered architectures and security issues: A comprehensive survey. *Sensors*, *18*, 9.

Cambra, C., Díaz, J. R., & Lloret, J. (2015). Deployment and performance study of an ad hoc network protocol for intelligent video sensing in precision agriculture. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 165-175). Springer Verlag. https://doi.org/10.1007/978-3-662-46338-3_14

Carlos-Mancilla, M., López-Mellado, E., & Siller, M. (2016). Wireless sensor networks formation: Approaches and techniques. *Journal of Sensors*, *2016*, 1-18. https://doi.org/10.1155/2016/2081902

Chapungo, N. J., & Postolache, O. (2021). Sensors and comunication protocols for precision agriculture. In *12th International Symposium on Advanced Topics in Electrical Engineering, ATEE 2021*. Institute of Electrical and Electronics Engineers Inc.. https://doi.org/10.1109/ATEE52255.2021.9425126

Chi, Q., Yan, H., Zhang, C., Pang, Z., & Xu, L. D. (2014). A reconfigurable smart sensor interface for industrial WSN in IoT environment. IEEE Transactions on Industrial Informatics, 10(2), 1417–1425. https://doi.org/10.1109/TII.2014.2306798

Damsgaard, S. B., Hernandez Marcano, N. J., Norremark, M., Jacobsen, R. H., Rodriguez, I., & Mogensen, P. (2022). Wireless communications for internet of farming: An early 5G measurement study. *IEEE Access*, *10*, 105263-105277. https://doi.org/10.1109/ACCESS.2022.3211096

Dong, X., Vuran, M. C., & Irmak, S. (2013). Autonomous precision agriculture through integration of wireless underground sensor networks with center pivot irrigation systems. *Ad Hoc Networks*, *11*(7), 1975-1987. https://doi.org/10.1016/j.adhoc.2012.06.012

Effah, E., Thiare, O., & Wyglinski, A. M. (2023). A tutorial on agricultural IoT: Fundamental concepts, architectures, routing, and optimization. *IoT*, *4*(3), 265-318. https://doi.org/10.3390/iot4030014

Ertürk, M. A., Aydın, M. A., Büyükakkaşlar, M. T., & Evirgen, H. (2019). A survey on LoRaWAN architecture, protocol and technologies. *Future Internet*, *11*(10), 216. https://doi.org/10.3390/fi11100216

Farooq, M. S., Riaz, S., Abid, A., Abid, K., & Naeem, M. A. (2019). A survey on the role of IoT in agriculture for the implementation of smart farming. *IEEE Access*, *7*, 15627-56237.

Friha, O., Ferrag, M. A., Shu, L., Maglaras, L., & Wang, X. (2021). Internet of things for the future of smart agriculture: A comprehensive survey of emerging technologies. *IEEE/CAA Journal of Automatica Sinica*, *8*(4), 718-752. https://doi.org/10.1109/JAS.2021.1003925

Gago, J., Estrany, J., Estes, L., Fernie, A. R., Alorda, B., Brotman, Y., Flexas, J., Escalona, J. M., & Medrano, H. (2020). Nano and micro unmanned aerial vehicles (UAVs): A new grand challenge for precision agriculture?. *Current Protocols in Plant Biology*, *5*(1), 1. https://doi.org/10.1002/cppb.20103

Gbadamosi, S. A., Hancke, G. P., & Abu-Mahfouz, A. M. (2020). Building upon NB-IoT networks: A roadmap towards 5G new radio networks. *IEEE Access*, *8*, 188641-188672. https://doi.org/10.1109/ACCESS.2020.3030653

Ghribi, B., & Logrippo, L. (2000). Understanding GPRS: The GSM packet radio service. *Computer Networks*, *34*(5), 763-779. https://doi.org/10.1016/S1389-1286(00)00127-4

Gitakarma, M. S., Tjahyanti, L. P. A. S., Indrawan, G., Ariawan, K. U., Wiratama, W. M. P., Pratama, P. A., & Prabawa, P. S. (2025). Development of IoT and WSN-Based smart AgroWeather with zigbee for microclimate management in Large-Scale plantations. *Formosa Journal of Science and Technology*, *4*(10), 3271-3290. https://doi.org/10.55927/fjst.v4i10.274

Glenn, D., William, C., & Ray, M. (1998). *Precision agriculture: An introduction.* https://extension.missouri.edu/publications/wq450

Gluhak, A., Krco, S., Nati, M., Pfisterer, D., Mitton, N., & Razafindralambo, T. (2011). A survey on facilities for experimental internet of things research. *IEEE Commun. Mag*, *49*(11), 58-67.

Gong, R., Zhang, H., Li, G., & He, J. (2025). Edge Computing-Enabled smart agriculture: Technical architectures, practical evolution, and bottleneck breakthroughs. *Sensors*, *25*(17), Article 5302. https://doi.org/10.3390/s25175302

Gyory, N., & Chuah, M. (2017). IoTOne: Integrated platform for heterogeneous IoT devices. In *2017 International Conference on Computing* (pp. 783-787). Networking and Communications (ICNC.

Hadir, A., Regragui, Y., & Garcia, N. M. (2021). Accurate Range-Free localization algorithms based on PSO for wireless sensor networks. *IEEE Access*, *9*, 149906-149924. https://doi.org/10.1109/ACCESS.2021.3123360

Haque, K. F., Abdelgawad, A., & Yelamarthi, K. (2022). Comprehensive performance analysis of zigbee communication: An experimental approach with XBee S2C module. *Sensors*, *22*(9), Article 3245. https://doi.org/10.3390/s22093245

Hassan Hayatu, I., Mohammed, A., Ahmad Isma'eel, B., & Yusuf Ali, S. (2020). K-MEANS CLUSTERING ALGORITHM BASED CLASSIFICATION OF SOIL FERTILITY IN NORTH WEST NIGERIA. *FUDMA JOURNAL OF SCIENCES*, *4*(2), 780-787. https://doi.org/10.33003/fjs-2020-0402-363

Hullamani, R. M., & T. N., K. (2021). ZigBee wireless communication technology. *JNNCE Journal of Engineering and Management*, *5*(1), 31. https://doi.org/10.37314/JJEM.2021.050105

Human, U. (2018). Aerobotics launches new tech for fruit farmers: Current affairs. *FarmBiz*, *4*, 12.

Ikpehai, A., Adebisi, B., Rabie, K. M., Anoh, K., Ande, R. E., Hammoudeh, M., Gacanin, H., & Mbanaso, U. M. (2019). Low-Power wide area network technologies for Internet-of-Things: A comparative review. *IEEE Internet of Things Journal*, *6*(2), 2225-2240. https://doi.org/10.1109/JIOT.2018.2883728

Ilie, M. (2018). Internet of things in agriculture. In *International Conference on Competitiveness of Agro-food and Environmental Economy Proceedings vol* (pp. 157-166). 7.

Iqbal Khan, M. Z., Abd Jalil, K., & Ibrahim, M. F. (2022). Efficient energy optimization routing for WSN based on Even-Odd scheduling. *Pertanika Journal of Science and Technology*, *30*(2), 1343-1361. https://doi.org/10.47836/pjst.30.2.27

Jawad, H., Nordin, R., Gharghan, S., Jawad, A., & Ismail, M. (2017). Energy-Efficient wireless sensor networks for precision agriculture: A review. *Sensors*, *17*(8), Article 1781. https://doi.org/10.3390/s17081781

Jazayeri, M. A., Liang, S. H. L., & Huang, C. Y. (2015). Implementation and evaluation of four interoperable open standards for the internet of things. *Sensors*, *15*(9), 24343-24373.

Ji, W., Adamchuk, V. I., Chen, S., Mat Su, A. S., Ismail, A., Gan, Q., Shi, Z., & Biswas, A. (2019). Simultaneous measurement of multiple soil properties through proximal sensor data fusion: A case study. *Geoderma*, *341*, 111-128. https://doi.org/10.1016/j.geoderma.2019.01.006

Kakaei, H., Nourmoradi, H., Bakhtiyari, S., Jalilian, M., & Mirzaei, A. (2022). Effect of COVID-19 on food security, hunger, and food crisis. In *COVID-19 and the Sustainable Development Goals* (pp. 3-29). Elsevier. https://doi.org/10.1016/B978-0-323-91307-2.00005-5

Kandris, D., Nakas, C., Vomvas, D., & Koulouras, G. (2020). Applications of wireless sensor networks: An Up-to-Date survey. *Applied System Innovation*, *3*(1), 14. https://doi.org/10.3390/asi3010014

Kaur, B., Prashar, D., Khan, A. A., Kadry, S., & Kim, J. (2025). PSLDV-Hop: A robust localization algorithm for WSN using PSO and refinement process. *PeerJ Computer Science*, *11*, Article e2770. https://doi.org/10.7717/peerj-cs.2770

Khan, A. A., Faheem, M., Bashir, R. N., Wechtaisong, C., & Abbas, M. Z. (2022). Internet of things (IoT) assisted context aware fertilizer recommendation. *IEEE Access*, *10*, 129505-129519. https://doi.org/10.1109/ACCESS.2022.3228160

Khanh, Q. V., Hoai, N. V., Manh, L. D., Le, A. N., & Jeon, G. (2022). Wireless communication technologies for IoT in 5G: Vision, applications, and challenges. *Wireless Communications and Mobile Computing*, *2022*(1), Article 3229294. https://doi.org/10.1155/2022/3229294

Kimm, H., Guan, K., Jiang, C., Peng, B., Gentry, L. F., Wilkin, S. C., Wang, S., Cai, Y., Bernacchi, C. J., Peng, J., & Luo, Y. (2020). Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the U.S. corn belt using planet labs CubeSat and STAIR fusion data. *Remote Sensing of Environment*, *239*, Article 111615. https://doi.org/10.1016/j.rse.2019.111615

Kishore, K., Sharma, S., & Kumar, D. (2013). Wireless sensor Networks-A review on topologies and node architecture. *International Journal of Computer Science International Journal of Computer Science International Journal of Computer Science International Journal of Computer*, *1*(2), 19-25. https://ijcseonline.org/index.php/j/article/view/9

Kloas, W., Groß, R., Baganz, D., Graupner, J., Monsees, H., Schmidt, U., Staaks, G., Suhl, J., Tschirner, M., Wittstock, B., Wuertz, S., Zikova, A., & Rennert, B. (2015). A new concept for aquaponic systems to improve sustainability, increase productivity, and reduce environmental impacts. *Aquaculture Environment Interactions*, *7*(2), 179-192. https://doi.org/10.3354/aei00146

Kluge, W., Poegel, F., Roller, H., Lange, M., Ferchland, T., Dathe, L., & Eggert, D. (2006). A fully integrated 2.4-GHz IEEE 802.15.4-Compliant transceiver for ZigBee™ applications. *IEEE Journal of Solid-State Circuits*, *41*(12), 2767-2775. https://doi.org/10.1109/JSSC.2006.884802

König, B., Janker, J., Reinhardt, T., Villarroel, M., & Junge, R. (2018). Analysis of aquaponics as an emerging technological innovation system. *Journal of Cleaner Production*, *180*, 232-243. https://doi.org/10.1016/j.jclepro.2018.01.037

Kumar, S. A., & Ilango, P. (2018). The impact of wireless sensor network in the field of precision agriculture: A review. *Wireless Personal Communications*, *98*(1), 685-698. https://doi.org/10.1007/s11277-017-4890-z

Kumhálová, J., & Matějková, Š. (2017). Yield variability prediction by remote sensing sensors with different spatial resolution. *International Agrophysics*, *31*(2), 195-202. https://doi.org/10.1515/intag-2016-0046

Kushwaha, H. L., Sinha, J. P., Khura, T., & Kushwaha, D. K. (2016). Status and scope of robotics in agriculture development of drone (UAV) assisted sprayer view project robotic planter view project. In *International Conference on Emerging Technologies in Agricultural and Food Engineering vol* (pp. 264-277). 12.

Laborde, D., Martin, W., Swinnen, J., & Vos, R. (2020). COVID-19 risks to global food security. *Science*, *369*(6503), 500-502. https://doi.org/10.1126/science.abc4765

Lata, S., & Verma, H. K. (2022). Selection of number and locations of Multi-Sensor nodes inside greenhouse. *Pertanika Journal of Science and Technology*, *30*(2), 933-948. https://doi.org/10.47836/pjst.30.2.05

Lata, S., S., L., & Verma, H. K. (2019). Techniques and algorithms for selection of number and locations of temperature sensors for greenhouse. *Pertanika Journal of Science and Technology*, *27*(4), 2153-2172. http://www.pertanika.upm.edu.my/pjst/browse/regular-issue

Lavric, A., Petrariu, A. I., & Popa, V. (2019). Long range SigFox communication protocol scalability analysis under Large-Scale, High-Density conditions. *IEEE Access*, *7*, 35816-35825. https://doi.org/10.1109/ACCESS.2019.2903157

Leng, G., & Huang, M. (2017). Crop yield response to climate change varies with crop spatial distribution pattern. *Scientific Reports*, *7*(1), Article 1463. https://doi.org/10.1038/s41598-017-01599-2

Liu, X., & Li, S. (2017). Temporal and spatial distribution characteristics of crop straw nutrient resources and returning to farmland in china. *Transactions of the Chinese Society of Agricultural Engineering*, *33*(21), 1-19. https://doi.org/10.11975/j.issn.1002-6819.2017.21.001

Luce, M. S., Ziadi, N., Gagnon, B., & Karam, A. (2017). Visible near infrared reflectance spectroscopy prediction of soil heavy metal concentrations in paper mill biosolid-and liming by-product-amended agricultural soils. *Geoderma*, *288*, 23-36.

Lytos, A., Lagkas, T., Sarigiannidis, P., Zervakis, M., & Livanos, G. (2020). Towards smart farming: Systems, frameworks and exploitation of multiple sources. *Computer Networks*, *172*, Article 107147. https://doi.org/10.1016/j.comnet.2020.107147

Mekki, K., Bajic, E., Chaxel, F., & Meyer, F. (2019). A comparative study of LPWAN technologies for large-scale IoT deployment. *ICT Express*, *5*(1), 1-7. https://doi.org/10.1016/j.icte.2017.12.005

Mekonnen, Y., Namuduri, S., Burton, L., Sarwat, A., & Bhansali, S. (2020). Review—Machine learning techniques in wireless sensor network based precision agriculture. *Journal of The Electrochemical Society*, *167*(3), Article 037522. https://doi.org/10.1149/2.0222003jes

Mendez, G. R., Md Yunus, M. A., & Mukhopadhyay, S. C. (2011). A WiFi based smart wireless sensor network for an agricultural environment. In *2011 Fifth International Conference on Sensing Technology* (pp. 405-410). IEEE. https://doi.org/10.1109/ICSensT.2011.6137009

Meraj, M., Mir, I., & Kumar, S. (2015). Evolution of mobile wireless technology from 0G to 5G. *International Journal of Computer Science and Information Technologies (IJCSIT)*, *6*(3), 2545-2551. https://www.ijcsit.com

Miskam, M. A., Nasirudin, A., & Abd. Rahim, I. (2009). Preliminary design on the development of wireless sensor network for paddy rice cropping monitoring application in malaysia. *European Journal of Scientific Research*, *37*(4), 649-657.

Moges, T. H., Lakew, D. S., Nguyen, N. P., Dao, N. N., & Cho, S. (2023). Cellular internet of things: Use cases, technologies, and future work. *Internet of Things*, *24*, Article 100910. https://doi.org/10.1016/j.iot.2023.100910

Montalvo, A., Camacho, O., & Chavez, D. (2025). Cyber-Physical systems for smart farming: A systematic review. *Sustainability*, *17*(14), Article 6393. https://doi.org/10.3390/su17146393

Morais, R., Fernandes, M. A., Matos, S. G., Serôdio, C., Ferreira, P. J. S. G., & Reis, M. J. C. S. (2008). A ZigBee multi-powered wireless acquisition device for remote sensing applications in precision viticulture. *Computers and Electronics in Agriculture*, *62*(2), 94-106. https://doi.org/10.1016/j.compag.2007.12.004

Mowla, M. N., Mowla, N., Shah, A. F. M. S., Rabie, K. M., & Shongwe, T. (2023). Internet of things and wireless sensor networks for smart agriculture applications: A survey. *IEEE Access*, *11*, 145813-145852. https://doi.org/10.1109/ACCESS.2023.3346299

Nasirudin, M. A., Za'bah, U. N., & Sidek, O. (2011). Fresh water real-time monitoring system based on wireless sensor network and GSM. In *2011 IEEE Conference on Open Systems* (pp. 354-357). IEEE. https://doi.org/10.1109/ICOS.2011.6079290

Navulur, S., S. C. S. Sastry, A., & N. Giri Prasad, M. (2017). Agricultural management through wireless sensors and internet of things. *International Journal of Electrical and Computer Engineering (IJECE)*, *7*(6), Article 3492. https://doi.org/10.11591/ijece.v7i6.pp3492-3499

Nguyen, H. H., Shin, D. Y., Jung, W. S., Kim, T. Y., & Lee, D. H. (2024). An integrated IoT Sensor-Camera system toward leveraging edge computing for smart greenhouse mushroom cultivation. *Agriculture*, *14*(3), 489. https://doi.org/10.3390/agriculture14030489

Ojha, T., Misra, S., & Raghuwanshi, N. S. (2015). Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges. *Computers and Electronics in Agriculture*, *118*, 66-84. https://doi.org/10.1016/j.compag.2015.08.011

Orlovs, D., Rusins, A., Skrastiņš, V., & Judvaitis, J. (2025). LPWAN technologies for IoT: Real-World deployment performance and practical comparison. *IoT*, *6*(4), 77. https://doi.org/10.3390/iot6040077

Osseiran, A., Song, J., Monserrat, J. F., & Hechwartner, R. (2020). IoT and machine type communications. *IEEE Communications Standards Magazine*, *4*(2), 40-40. https://doi.org/10.1109/MCOMSTD.2020.9139044

Paul, Y., & Singh, R. (2015). ZigBee technology: A comprehensive review of protocol, applications, and advancements. *IJRDO -Journal of Computer Science Engineering*. https://doi.org/10.53555/cse.v1i2.6309

Poole, I. (2004). What exactly is ZigBee?. *Communications Engineer*, *2*(4), 44-45. https://doi.org/10.1049/ce:20040409

Prodanović, R., Rančić, D., Vulić, I., Zorić, N., Bogićević, D., Ostojić, G., Sarang, S., & Stankovski, S. (2020). Wireless sensor network in agriculture: Model of cyber security. *Sensors*, *20*(23), Article 6747. https://doi.org/10.3390/s20236747

Ramli, M. R., Daely, P. T., Kim, D. S., & Lee, J. M. (2020). IoT-based adaptive network mechanism for reliable smart farm system. *Computers and Electronics in Agriculture*, *170*, Article 105287. https://doi.org/10.1016/j.compag.2020.105287

Rehman, A., Mahmood, T., Alahmadi, T. J., Almasoud, A. S., & Saba, T. (2025). Energy efficient and robust node localization in WSNs using LSTM optimized DV hop framework to mitigate multihop localization errors. *Scientific Reports*, *15*(1), Article 14827. https://doi.org/10.1038/s41598-025-93937-y

Research and Markets (2020). *Global digital agriculture market expected to grow from $5.6 billion in 2020 to $6.2 billion by 2021, recording a CAGR of 9.9% - revised to reflect the implications of the COVID-19 pandemic.* GlobeNewswire. https://www.globenewswire.com/news-release/2020/07/03/2057417/0/en/Global-Digital-Agriculture-Market-Expected-to-Grow-from-5-6-Billion-in-2020-to-6-2-Billion-by-2021-Recording-a-CAGR-of-9-9-Revised-to-Reflect-the-Implications-of-the-COVID-19-Pande.html

Routray, S., & Mohanty, S. (2024). Cellular internet of things: Principles, potentials and use cases. *EAI Endorsed Transactions on Internet of Things*, *10*. https://doi.org/10.4108/eetiot.4656

Saccone, D., & Vallino, E. (2025). Global food security in a turbulent world: Reviewing the impacts of the pandemic, the war and climate change. *Agricultural and Food Economics*, *13*(1), 47. https://doi.org/10.1186/s40100-025-00388-0

Santhosh, S., Gurrapu, O., Dimri, N., Obuxova, N. V., Karthik, B., & Ganesh, V. D. (2025). Integrating IoT and edge computing for smart agriculture with real-time data analytics. In 2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 1-6). IEEE. https://doi.org/10.1109/ICSCDS65426.2025.11167645

Saqlain, M., Halim, A., Akhonzada, A. R., & Ali, A. (2025). Evaluating the potential and challenges of LPWAN technologies like LoRa and sigfox for long-range communication in IoT systems. *Review of Applied Management and Social Sciences*, *8*(1), 111-128. https://doi.org/10.47067/ramss.v8i1.438

Shafi, U., Mumtaz, R., García-Nieto, J., Hassan, S. A., Zaidi, S. A. R., & Iqbal, N. (2019). Precision agriculture techniques and practices: From considerations to applications. *Sensors*, *19*(17), Article 3796. https://doi.org/10.3390/s19173796

Shafique, K., Khawaja, B. A., Sabir, F., Qazi, S., & Mustaqim, M. (2020). Internet of things (IoT) for Next-Generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. *IEEE Access*, *8*, 23022-23040. https://doi.org/10.1109/ACCESS.2020.2970118

Shamshiri, R. R., Bojic, I., van Henten, E., Balasundram, S. K., Dworak, V., Sultan, M., & Weltzien, C. (2020). Model-based evaluation of greenhouse microclimate using IoT-Sensor data fusion for energy efficient crop production. *Journal of Cleaner Production*, *263*, Article 121303. https://doi.org/10.1016/j.jclepro.2020.121303

Shan, L. Y. (2023). *Global rice shortage is set to be the biggest in 20 years.* Cnbc. https://www.cnbc.com/2023/04/19/global-rice-shortage-is-set-to-be-the-largest-in-20-years-heres-why.html

Sharma, P. (2024). *Understanding k-means clustering in machine learning.* Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/11/understanding-k-means-clustering-in-machine-learningwith-examples/

Shin, S. K., Lee, S. J., & Park, J. H. (2025). Prediction of soil properties using Vis-NIR spectroscopy combined with machine learning: A review. *Sensors*, *25*(16), Article 5045. https://doi.org/10.3390/s25165045

Singh, D. K., & Sobti, R. (2021). Wireless communication technologies for internet of things and precision agriculture: A review. In *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)* (pp. 765-769). IEEE. https://doi.org/10.1109/ISPCC53510.2021.9609421

Sto ces, M., Van ek, J., Masner, J., & Pavlík, J. (2016). Internet of things (iot) in agriculture-selected aspects. *Agris on-line Pap*, *8*(665-2016), 83-88.

Sufyan, A., Khan, K. B., Khashan, O. A., Mir, T., & Mir, U. (2023). From 5G to beyond 5G: A comprehensive survey of wireless network evolution, challenges, and promising technologies. *Electronics*, *12*(10), Article 2200. https://doi.org/10.3390/electronics12102200

Suresh, P., Daniel, J., Parthasarathy, V., & Aswathy, R. H. (2014). A state of the art review on the internet of things (IoT) history, technology and fields of deployment. In *2014 International conference on science engineering and management research (ICSEMR* (pp. 1-8).

Tang, P., Liang, Q., Li, H., & Pang, Y. (2024). Application of Internet-of-Things wireless communication technology in agricultural irrigation management: A review. *Sustainability*, *16*(9), Article 3575. https://doi.org/10.3390/su16093575

Thakur, D., Kumar, Y., Kumar, A., & Singh, P. K. (2019). Applicability of wireless sensor networks in precision agriculture: A review. *Wireless Personal Communications*, *107*(1), 471–512. https://doi.org/10.1007/s11277-019-06285-2

The ZigBee Alliance (2015). *ZigBee specification.*

Vaezi, M., Azari, A., Khosravirad, S. R., Shirvanimoghaddam, M., Azari, M. M., Chasaki, D., & Popovski, P. (2022). Cellular, Wide-Area, and Non-Terrestrial IoT: A survey on 5G advances and the road toward 6G. *IEEE Communications Surveys & Tutorials*, *24*(2), 1117–1174. https://doi.org/10.1109/COMST.2022.3151028

van der Goot, A. J., Pelgrom, P. J. M., Berghout, J. A. M., Geerts, M. E. J., Jankowiak, L., Hardt, N. A., Keijer, J., Schutyser, M. A. I., Nikiforidis, C. V., & Boom, R. M. (2016). Concepts for further sustainable production of foods. *Journal of Food Engineering*, *168*, 42-51. https://doi.org/10.1016/j.jfoodeng.2015.07.010

Verma, S., Sharma, D., & Sharma, K. (2013). Network topologies in wireless sensor networks: A review forest fire detection using wireless sensor network view project intelligent clustering for IoT view project network topologies in wireless sensor networks: A review 1. *International Journal of Electronics & Communication Technology*, *4*(3). https://www.researchgate.net/publication/313198672

Vermesan, O., & Friess, P. (2013). *Internet of things: Converging technologies for smart environments and integrated ecosystems.* River publishers.

von Grebmer, K., Bernstein, J., Wiemers, M., Acheampong, K., Hanano, A., Higgins, B., Ní Chéilleachair, R., Foley, C., Gitter, S., Ekstrom, K., & Fritschel, H. (2020). *2020 global hunger index: One decade to zero hunger: Linking health and sustainable food systems.* Welthungerhilfe; Concern Worldwide. https://www.globalhungerindex.org/pdf/en/2020.pdf

von Grebmer, K., Bernstein, J., Wiemers, M., Schiffer, T., Hanano, A., Towey, O., Ní Chéilleachair, R., Foley, C., Gitter, S., Ekstrom, K., & Fritschel, H. (2021). *2021 global hunger index: Hunger and food systems in conflict settings.* Welthungerhilfe; Concern Worldwide. https://www.globalhungerindex.org/pdf/en/2021.pdf

Vuran, M. C., Salam, A., Wong, R., & Irmak, S. (2018). Internet of underground things in precision agriculture: Architecture and technology aspects. *Ad Hoc Networks*, *81*, 160-173. https://doi.org/10.1016/j.adhoc.2018.07.017

Wang, N., Zhang, N., & Wang, M. (2006). Wireless sensors in agriculture and food industry—Recent development and future perspective. *Computers and Electronics in Agriculture*, *50*(1), 1-14. https://doi.org/10.1016/j.compag.2005.09.003

Wang, W., Wang, N., Jafer, E., Hayes, M., O'Flynn, B., & O'Mathuna, C. (2010). Autonomous wireless sensor network based building energy and environment monitoring system design. In *2010 2nd Conference on Environmental Science and Information Application Technology, ESIAT 2010* (pp. 367-372). IEEE.

Waseem, M., Lopez, A., Luis Carro, P., Losada, M. A., Richards, D., & Aziz, A. (2025). Understanding the potentials of narrowband internet of things (NB-IoT) cellular technology: Salient features, architecture, smart applications and quality of service (QoS) challenges. *IEEE Open Journal of the Communications Society*, *6*, 6258-6280. https://doi.org/10.1109/OJCOMS.2025.3588334

Wen, M., Li, Q., Kim, K. J., Lopez-Perez, D., Dobre, O., Poor, H. V., Popovski, P., & Tsiftsis, T. (2022). Private 5G networks: Concepts, architectures, and research landscape. *IEEE Journal of Selected Topics in Signal Processing*, *16*(1), 7-25. https://doi.org/10.1109/JSTSP.2021.3137669

Xu, J., Yao, J., Wang, L., Ming, Z., Wu, K., & Chen, L. (2018). Narrowband internet of things: Evolutions, technologies, and open issues. *IEEE Internet of Things Journal*, *5*(3), 1449-1462. https://doi.org/10.1109/JIOT.2017.2783374

Yanes, A. R., Martinez, P., & Ahmad, R. (2020). Towards automated aquaponics: A review on monitoring, IoT, and smart systems. *Journal of Cleaner Production*, *263*. https://doi.org/10.1016/j.jclepro.2020.121571

Zamin, N., Norwawi, N. M., Taslim, H. A., & Rosdi, W. M. F. W. M. (2019). Kliktani: A concept of a smart edible garden using iot for indigenous community in malaysia. *International Journal of Advanced Trends in Computer Science and Engineering*, *8*(1.4), 350-355. https://doi.org/10.30534/ijatcse/2019/5381.42019

Zamora-Izquierdo, M. A., Santa, J., Martínez, J. A., Martínez, V., & Skarmeta, A. F. (2019). Smart farming IoT platform based on edge and cloud computing. *Biosystems Engineering*, *177*, 4-17. https://doi.org/10.1016/j.biosystemseng.2018.10.014

Zhao, H., Di, L., Guo, L., Zhang, C., & Lin, L. (2023). An automated Data-Driven irrigation scheduling approach using model simulated soil moisture and evapotranspiration. *Sustainability*, *15*(17), Article 12908. https://doi.org/10.3390/su151712908

Zulkifli, C. Z., & Noor, N. N. (2017). Wireless sensor network and internet of things (IoT) solution in agriculture. *Pertanika Journal of Science & Technology*, *25*(1), 91-100. http://www.pertanika.upm.edu.my/pjst/browse/regular-issue

Pertanika J. Sci. & Technol. 34 (5): 2887 - 2905 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Grey-Fuzzy Based Optimisation of Machining Parameters in Turning of AA 7068/TiC MMCs

Syed Altaf Hussain[1]*, Sudhamadhuri K[1], Upendra Rajak[1], and Abhishek Dasore[2]

[1]Department of Mechanical Engineering, Rajeev Gandhi Memorial College of Engineering & Technology, 518501 Nandyal, Andhra Pradesh, India

[2]Department of Computer Science and Engineering, SR University, Ananthasagar, Hasanparthy, 506371 Warangal, Telangana, India

ABSTRACT

With the growing applications of Metal Matrix Composites in several engineering fields, it is necessary to develop appropriate machining technology to process them efficiently. The turning process is generally used across all industries. In this research, an effort has been made to determine optimum values of process parameters for turning of AA7068/TiC metal matrix composites (MMCs) using the Grey Fuzzy Approach. The selected process parameters are the % weight of particle reinforcement, cutting speed, feed and depth of cut. Experiments have been designed according to CCD and conducted on an all-geared lathe machine with a carbide (K-30) cutting tool insert. The responses considered are cutting force, surface roughness and metal removal rate (MRR). To optimise the machining responses simultaneously, grey relational analysis combined with fuzzy logic is employed to determine the grey-fuzzy reasoning grade. The optimum combination of process parameters that generated the highest GFRG value (0.6083) was found to be a wt% reinforcement of 2%, a cutting speed of 50 m/min, a feed of 0.06 mm/rev, and a depth of cut of 0.2 mm. From the ANOVA, the wt% of TiC reinforcement was the only statistically significant linear factor at the 95% confidence level in the multi-response optimisation. Confirmation test under optimum conditions led to a reduction in cutting force by 4.00% from 129.45 N to 124.27 N, an improvement in surface roughness by 3.04% from 0.493 μm to 0.478 μm, and an increase in MRR by 5.99% from 136.37 mm^3/s to 144.55 mm^3/s.

Keywords: AA7068/ TiC MMCs, carbide (K-30) cutting tool, grey-fuzzy logic, optimisation

ARTICLE INFO

Article history:
Received: 04 July 2026
Accepted: 22 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.14

E-mail addresses:
syedaltafhussain@rgmcet.edu.in (Syed Altaf Hussain)
ksmadhurime@gmail.com (Sudhamadhuri K)
upendrarajak86@gmail.com (Upendra Rajak)
dasore.abhishek@sru.edu.in (Abhishek Dasore)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

The aluminium alloy AA7068 finds use in numerous applications owing to its unique properties such as corrosion resistance, high electrical and thermal conductivity, and good fatigue and wear resistance (Ashraf & Shibli, 2009; Cerik, 2017; Heinz et al., 2000; Shin et al., 2017; Teimouri et al., 2017). Major applications include those where alloys are preferred over conventional materials such as steels and cast irons owing to their lightweight yet high-strength mechanical properties, and composites based on aluminium alloys, known as aluminium metal matrix composites (MMCs). Turning these MMCs is an important step in the machining process.

Good corrosion resistance, along with a high strength-to-weight ratio, is exhibited by the first-generation aluminium-based MMCs reinforced with ceramics (Gururaja et al., 2013). The effect of various machining parameters on tool wear and BUE formation during the turning of Al-SiC MM Cs using a rhombic uncoated carbide tool. The formation of a less built-up edge was observed at lower cut depths and high speeds (Manna & Bhattacharyya, 2005). A two-dimensional array was developed using GRA to optimise process parameters while drilling Al/SiC MMCs (Haq et al., 2008). A sensitivity-based multi-criteria optimisation method aimed at minimising the cutting force and maximising material removal while machining unidirectional glass fibre reinforced polyester rods (Isik & Kentli, 2009). The influence of cutting parameters on temperature, thrust force, and surface roughness (SR) during drilling of AA6351/B4C composite materials was presented (Samy & Kumaran, 2017). To minimise tool wear and specific cutting pressure, and maximise the material removal rate, the grey relational grade and Taguchi method were implemented in the turning process of glass fibre/epoxy composites using a carbide (K-10) tool (Palanikumar et al., 2006).

Application of grey fuzzy logic to optimise the drilling process parameters of composite plates comprising carbon fibre-reinforced plastic (Krishnamoorthy et al., 2012). Employed the Grey Fuzzy algorithm in the process of optimising the machining parameters while drilling the hybrid aluminium metal matrix composites (Rajmohan et al., 2013). Adoption of response surface methodology (RSM) to optimise the machining parameters of LM 25 Al/SiC specimen while performing the turning process (Palanikumar & Karthikeyan, 2006). The Taguchi and GRA combination approach has been extensively utilised to optimise objectives such as surface finish, cutting force, tool life, flank wear, and MRR (Jangra et al., 2012; Nalbant et al., 2007; Sahoo et al., 2013). To meet the desired output responses such as high SR, low tool wear, and shorter time intervals, GA has been employed to optimise process parameters (Chen et al., 2014; Ganesan & Mohankumar, 2013). Application of fuzzy logic to predict the output responses (MRR, TWR, and SR), while the desirability method was used for optimising process parameters in EDM of Inconel 718 (Sengottuvel et al., 2013).

The multi-objective optimisation technique based on fuzzy TOPSIS was applied to enhance the EDM surface integrity and the dimensional accuracy of AISI P20 tool steel (Dewangan et al., 2015). Furthermore, there is significant potential for GRA to enhance the performance of fuzzy logic in multi-objective optimisation problems. In the current case, the complex multi-response optimisation problem could be easily converted into a single GFRG optimisation. An approach through a fuzzy logic controller that estimates the shape of the weld bead of a weld joint using the TIG welding process. Eventually, it was mentioned that the developed fuzzy logic model was tested several times to ensure sufficient precision (Narang et al., 2011). The use of the fuzzy interface system, along with the Taguchi method, to optimise the parameters of the TIG welding process for A387 Grade 11 Class 2 steel.

Finally, it was found that the critical input parameter for the TIG welding process is arc voltage, whereas the second one is the welding cutting (Prajapati et al., 2020). Various methods can be used to address multi-objective optimisation problems, such as fuzzy logic, ANN, GA, GRA, and RSM (Krishnaiah & Shahabudeen, 2012). Applications of a fuzzy logic-based approach are in environments with significant uncertainty and hesitation. Multi-criteria decision-making (MCDM) systems based on fuzzy logic have become popular in current research on optimising manufacturing processes.

In recent years, researchers have successfully implemented the grey relational approach to address interdependencies among multiple objectives across various manufacturing fields (Ahilan et al., 2009; Chiang & Chang, 2006; Hsiao et al., 2008; Krishnamoorthy et al., 2012; Lin & Lin, 2005). A genetic algorithm (GA) integrated with PCA and grey relational analysis (GRA) for optimisation of machining parameters to obtain better surface quality with minimum cutting force without any increase in specific power consumption during the turning of MMCs (Khandey et al., 2017). The use of GRA, along with fuzzy logic, for the optimisation of wire EDM of AISID2 steel to minimise surface roughness and layer thickness (Soepangkat & Pramujati, 2013). The use of TOPSIS together with the Taguchi method to arrive at the optimal settings of machining parameters in the turning of GFRP composite (Hussain et al., 2025). The use of grey fuzzy logic to optimise CNC milling parameters for Al-4.5% Cu-TiC MMCs with multiple performance measures (Das et al., 2016). Several optimisation methods have been efficiently implemented in many manufacturing operations, most of which were carried out under uncertainty (Yang & Huang, 2012). Various objectives related to different issues in materials science have been achieved using evolutionary algorithms and other machine learning techniques (Giri et al., 2013).

Optimisation on a multi-objective basis for the tribological properties of the hybrid composite material based on aluminium alloy A356 containing SiC has been done using the ‘Block-on-Disc’ tribometer system, where the input parameters have been taken as load, sliding velocity, and graphite content. The output parameters have been measured as

the coefficient of friction and wear rate. The multi-objective optimisation was performed using GRA and verified with TOPSIS (Gajević et al., 2024). Efforts have been made to minimise the frictional wear properties of aluminium 2024-TiB2 metal matrix composites (MMCs) using the grey fuzzy optimisation technique. The effectiveness of the new grey fuzzy optimisation approach for solving multiple-response tribological problems has been successfully tested (Dey et al., 2023). Polyvinyl Alcohol's improved mechanical performance, thermal stability, and reduced moisture absorption have been achieved through crosslinking with varying weight percentages of fumaric acid and graphite particles modified with sulfuric acid.

The grey-fuzzy optimisation technique has been used to achieve a trade-off between competing objectives related to mechanical and physical properties (Kafaltiya et al., 2024). AVOA-SVM regression models were used to predict and analyse machining responses, such as surface roughness and MRR, during the machining of 7068 alloy using a graphite nanofluid (Shrotriya et al., 2026). A grey-Fuzzy approach was used to identify the optimal combination of tribological parameters to minimise wear and friction in Al2024-TiB2 MMCs (Dey et al., 2023). Optimum drilling parameters for machining CFRP by combining the grey relation coefficient with the combined compromise solution (Grey-CoCoSo) for the machining responses like MRR, surface roughness, and delamination factor (DF) (Shunmugesh et al., 2025). A hybrid grey-based fuzzy algorithm was used to determine the optimum results of a combination of factors such as cutting speed, feed rate, cutting depth and cutting path strategy for the machining responses like surface quality, MRR and production time (Engin et al., 2026). AA8011 AMC reinforced with 9 wt% SiO2 particles was fabricated using the stir-casting technique, and tribological responses, such as the coefficient of friction and frictional forces, were optimised using the Grey-fuzzy approach (Bhowmik et al., 2026).

Limited work has been reported on the influence of turning variables on various machining responses. So, this work was undertaken to optimise the process parameters for machining AA7068/TiC particulate metal matrix composites using the Grey-Fuzzy logic approach. The primary machining responses considered are Fz (cutting force), Ra (surface roughness), and MRR (metal removal rate). An optimised set of input machining parameters is required to reduce machining responses.

MATERIALS AND METHODS

In this study, AA7068/TiC Metal Matrix Composites with reinforcement percentages of 2-10% (in increments of 2%) are fabricated using a stir casting technique, in which AA7068 is placed in a crucible and heated to 850ºC. Once the matrix had melted, magnesium ribbons were introduced into the molten metal to clean it and increase the wettability of AA7068, and 2% TiC was added to the molten metal, which was preheated to 300ºC. In the meantime, the molten metal was stirred at 300 rpm for 15 minutes to ensure

uniform mixing of the matrix and reinforcements. The experimental setup is shown in Figure 1. Specimens of dimension Ø 22 × 190 mm are cast. Experiments have been planned to use a central composite design and conducted on a lathe machine with a cutting tool insert made of Carbide K-30. A four-factor central composite design (CCD) is chosen since it provides an efficient experimental design for investigating the effects of multiple continuous machining parameters on the response while still estimating the behaviour of the response within the experimental design region. The current CCD comprises 31 experimental runs, as shown in Table 1.

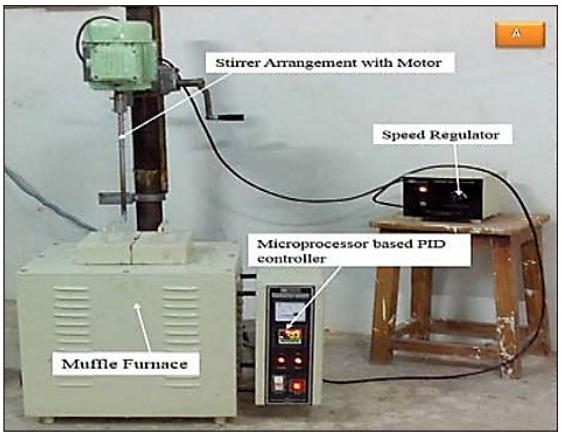


Figure 1. Stir casting setup

Table 1
Observation table

Sl. No.	Weight Percentage of TiC, (wt. %)	Cutting Speed (V), m/min	Feed (f), mm/rev	Depth of Cut (d), mm	Observed Responses		Calculated Responses
					Cutting Force (Fz), N	Surface Roughness (Ra), µm	Material Removal Rate (MRR), mm^3/sec
1	4	40	0.07	0.5	129.45	0.493	136.37
2	8	40	0.07	0.5	167.82	0.643	94.56
3	4	60	0.07	0.5	144.02	0.506	137.59
4	8	60	0.07	0.5	182.39	0.716	116.10
5	4	40	0.09	0.5	131.15	0.566	147.70
6	8	40	0.09	0.5	164.56	0.737	89.15
7	4	60	0.09	0.5	145.22	0.552	138.82
8	8	60	0.09	0.5	164.86	0.706	104.88
9	4	40	0.07	1.1	128.64	0.556	146.23
10	8	40	0.07	1.1	167.33	0.687	87.43
11	4	60	0.07	1.1	144.26	0.516	140.51
12	8	60	0.07	1.1	170.68	0.662	111.54
13	4	40	0.09	1.1	156.89	0.598	150.22
14	8	40	0.09	1.1	179.45	0.786	105.96

Table 1 (*continued*)

Sl. No.	Weight Percentage of TiC, (wt. %)	Cutting Speed (V), m/min	Feed (f), mm/ rev	Depth of Cut (d), mm	Observed Responses		Calculated Responses
					Cutting Force (Fz), N	Surface Roughness (Ra), µm	Material Removal Rate (MRR), mm^3/sec
15	4	60	0.09	1.1	172.11	0.588	156.00
16	8	60	0.09	1.1	190.73	0.675	118.14
17	2	50	0.08	0.8	114.45	0.426	179.32
18	10	50	0.08	0.8	209.00	0.815	77.04
19	6	30	0.08	0.8	148.86	0.659	105.00
20	6	70	0.08	0.8	167.20	0.546	135.23
21	6	50	0.06	0.8	141.46	0.594	106.25
22	6	50	0.10	0.8	179.88	0.716	132.50
23	6	50	0.08	0.2	150.54	0.549	109.75
24	6	50	0.08	1.4	180.86	0.689	135.89
25	6	50	0.08	0.8	132.08	0.624	129.45
26	6	50	0.08	0.8	138.69	0.621	127.86
27	6	50	0.08	0.8	138.78	0.661	125.52
28	6	50	0.08	0.8	142.34	0.618	126.35
29	6	50	0.08	0.8	138.52	0.594	126.46
30	6	50	0.08	0.8	144.13	0.646	126.52
31	6	50	0.08	0.8	145.36	0.609	126.75

This is comprised of factorial/axial points and repeat centre points. Experiments 25-31 represent the repeat centre points that involve 6 wt.% TiC, 50 m/min cutting speed, 0.08 mm/rev feed rate and 0.8 mm depth of cut, resulting in 7 centre points for the estimation of experimental/pure error. This tool insert conforms to ANSI standards. It has the same ISO geometry as that of SNMG 120408. The cutting tool inserts are attached to a PSBNR-2525M12 WIDAX tool holder. Input parameters used for conducting experiments are tabulated in Table 2, and output responses obtained from experiments are tabulated in Table 1. The cutting force generated during the turning of MMC bars is measured using the Kistler Dynamometer model 9257B. The surface finish obtained on the machined specimens is measured using a 'Talysurf' tester supplied by Metasonic Private Ltd. The metal removal rate is evaluated using Equation 1. Figure 2 illustrates the experimental setup.

The percentage of variation in the density of the MMCs due to variation in the amount of reinforcement added is very small and was ignored.

$$\text{MRR} = \frac{wtb - wta}{Machining\ \ Time\ (t) * Density\ (\rho)} \quad \frac{mm^3}{sec} \qquad [1]$$

Where

wtb = Wt. of w/p before machining

wta = Wt. of w/p after machining

T = Time, Sec

ρ = Density, g/mm^3

Table 2
Process parameters for AA7068/TiC machining

S.No	Parameters	Notation	Levels				
			1	2	2	4	5
1	The weight percentage of TiC,	wt%	2	4	6	8	10
2	Velocity of cutting, m/min,	V	30	40	50	60	70
3	Feed (f), mm/rev	f	0.06	0.07	0.08	0.09	0.1
4	Depth of Cut mm	d	0.2	0.5	0.8	1.1	1.4

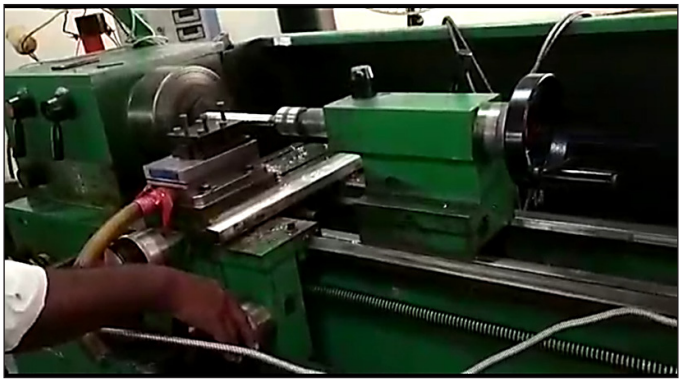
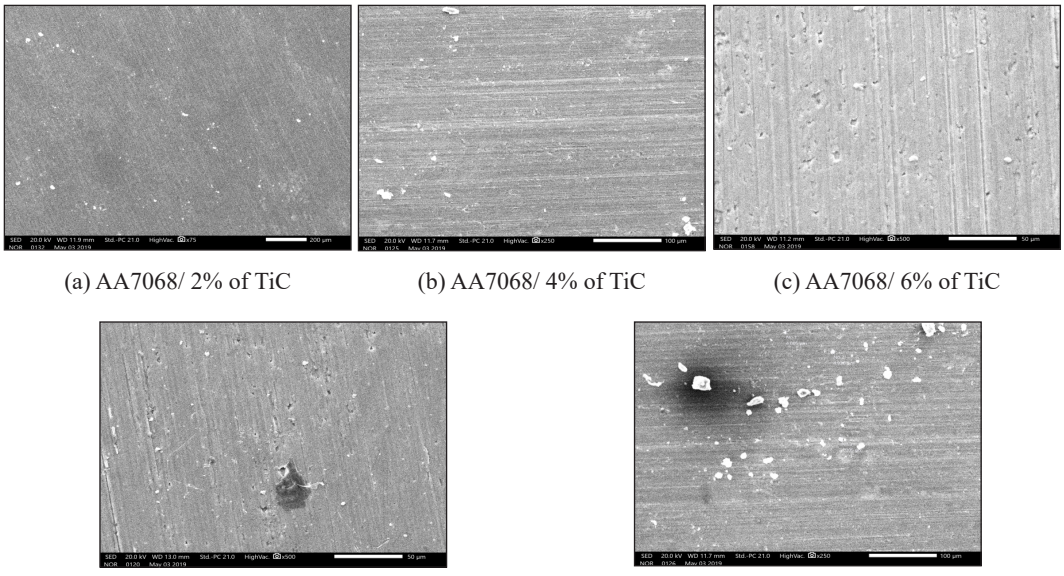


Figure 2. Photograph of the original experimental setup

SEM images are taken to investigate the surface characteristics of developed composite materials (AA7068/TiC MMCs). The images show uniform dispersion of TiC reinforcement particles in the AA7068 matrix. There is no agglomeration of reinforcement particles found in the composite material. Thus, a stirring speed of 300 rpm is appropriate for AA7068/ TiC MMCs. The SEM images are shown in Figure 3a to 3e.



(a) AA7068/ 2% of TiC (b) AA7068/ 4% of TiC (c) AA7068/ 6% of TiC

(d) AA7068/ 8% of TiC (e) AA7068/ 10% of TiC

Figure 3. SEM images of AA7068/TiC MMCs at different wt% of TiC

Grey Relational Analysis

A good method for analysing sequence groups is GRA, which requires that all sequence values be comparable in the absence of dimensionality, scaling, or polarisation characteristics (Deng, 1989). The first step in the GRA process is to normalise all experimental data to the range 0 - 1. As different responses have varying ranges, the requirement for normalisation arises. In case of a higher-is-better nature of response, the data are normalised using Equation 2.

$$x^{*}i\ (k) = \frac{x_i\ (k) - \min x_i(k)}{\max x_i - \min x_i(k)} \qquad [2]$$

If the 'Lower-the-better' criterion is to be followed, then the following Equation 3 is to be used for normalising the corresponding data.

$$x^{*}i\ (k) = \frac{\max xi\ (k) - x_i\ (k)}{\max x_i - \min x_i(k)} \qquad [3]$$

Where $x^{*}{}_{\mathrm{I}}$ (k) and $x_i\ (k)$ are the normalised data and observed data, respectively, for the ith experiment using the kth response. After normalising the responses, the next step

is to calculate the grey relation coefficient (GRC). GRC is denoted for k[th] response. It is calculated by using Equation 4.

$$\zeta_i\ (k)=\frac{\Delta_{min}\ -\zeta\Delta_{max}}{\Delta_i(k)+\zeta\Delta_{max}} \quad [4]$$

Where $\Delta_i\ (k) = \|x_0\ (k) - x_i\ (k)\|$ is the difference of absolute value between $x_0\ (k)$ and $x_i\ (k)$, ζ = distinguishing coefficient (0 ~ 1), ζ = 0.5 commonly used, Δ_{min} = minimum value of the deviational sequence, Δ_{max} = maximum value of the deviational sequence.

In the present study, the distinguishing coefficient was set to ζ=0.5 because it is commonly used in GRA and provides balanced discrimination among experimental sequences. The GRC equation used in the study explicitly defines ζ (zeta) as a parameter in the range 0-1 and states that 0.5 is the commonly used value.

Fuzzy Rule-Based Modelling

In GRA, the application of higher-the-better, lower-the-better, and nominal-the-better performance measures reveals some vagueness in the results achieved. This kind of vagueness can be successfully resolved through the application of fuzzy logic (Tzeng & Chen, 2007). A complex multi-response optimisation problem can be solved by combining GRA and fuzzy logic techniques.

The fuzzy logic system (also referred to as the Mamdani system) consists of a fuzzifier, a membership function, a fuzzy rule-based inference engine, and a defuzzifier. In the present study, the three grey relational coefficients (GRCs) for cutting force (Fz), surface roughness (Ra), and material removal rate (MRR) were used as inputs to a Mamdani fuzzy inference system. The GRCs were first normalised, and the resulting values were fuzzified using five overlapping triangular linguistic membership functions: Low (L), Low-Medium (LM), Medium (M), High-Medium (HM), and High (H). 31 fuzzy rules were generated, with maximum–minimum composition and defuzzification in MATLAB.

Steps in the Grey-Fuzzy Logic Approach

The step recommended for the grey-fuzzy-logic approach is given in Figure 4 and is explained as follows:

- The cutting force (Fz), surface roughness (Ra) and metal removal rate (MRR) obtained experimentally are first normalised between 0 and 1.
- Grey relational coefficient (GRC) for each response is then computed.
- The fuzzification step follows, where the fuzzifier uses the membership function to convert the GRC of the machining response.

- The next step involves generating fuzzy rules (if –then control rules). After that, defuzzifier converts fuzzy predicted values into GFRC.
- The last step involves determining optimal settings for the machining parameters using the main effects plot of GFRC.

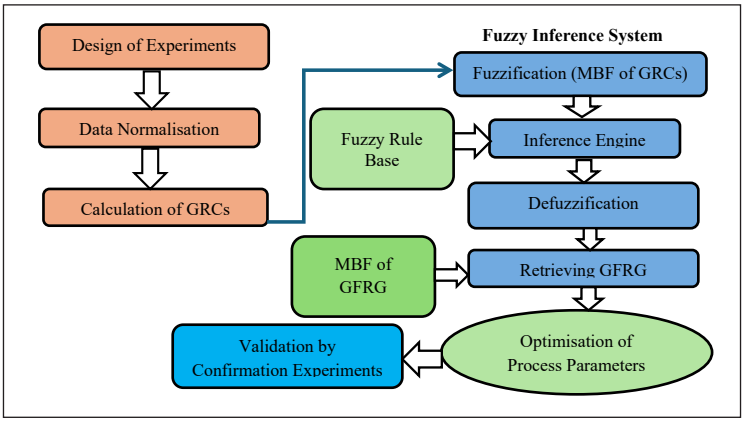


Figure 4. Suggested grey-fuzzy-logic approach

RESULTS AND DISCUSSIONS

Calculation of Grey Relational Coefficient (GRC)

The experimental data are preprocessed to calculate the grey relational coefficients for each response, as shown in Table 3. For responses like cutting force (F_z) and surface roughness (R_a), the 'lower the better' criterion is used, while the 'higher the better' criterion is used for metal removal rate (MRR). The grey fuzzy logic approach is used here to calculate the grey fuzzy relational grade (GFRG) to minimise the ambiguity in the data.

Table 3
Data pre-processing, computing GRC and grey fuzzy reasoning grade (GFRG)

Exp. No.	Normalised Values of Machining Responses			Grey Relational Coefficient (GRC)			GFRG
	Fz	Ra	MRR	Fz	Ra	MRR	
1	0.8414	0.8278	0.5801	0.7591	0.7438	0.5435	0.5000
2	0.4355	0.4422	0.1713	0.4697	0.4727	0.3763	0.4500
3	0.6873	0.7943	0.5920	0.6152	0.7086	0.5507	0.5000
4	0.2814	0.2545	0.3819	0.4103	0.4014	0.4472	0.3420
5	0.8234	0.6401	0.6908	0.7390	0.5815	0.6179	0.5000
6	0.4700	0.2005	0.1184	0.4854	0.3848	0.3619	0.3830
7	0.6746	0.6761	0.6040	0.6057	0.6069	0.5581	0.5000
8	0.4668	0.2802	0.2722	0.4840	0.4099	0.4072	0.3770

Table 3 (*continued*)

Exp. No.	Normalised Values of Machining Responses			Grey Relational Coefficient (GRC)			GFRG
	Fz	Ra	MRR	Fz	Ra	MRR	
9	0.8499	0.6658	0.6765	0.7691	0.5994	0.6071	0.5000
10	0.4407	0.3290	0.1016	0.4720	0.4270	0.3575	0.4110
11	0.6847	0.7686	0.6206	0.6133	0.6837	0.5685	0.5000
12	0.4053	0.3933	0.3373	0.4567	0.4518	0.4300	0.4000
13	0.5511	0.5578	0.7155	0.5269	0.5307	0.6373	0.5000
14	0.3125	0.0746	0.2828	0.4211	0.3508	0.4108	0.3750
15	0.3902	0.5835	0.7720	0.4505	0.5456	0.6868	0.5000
16	0.1932	0.3599	0.4018	0.3826	0.4386	0.4553	0.3530
17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9200
18	0.0000	0.0000	0.0000	0.3333	0.3333	0.3333	0.3410
19	0.6361	0.4010	0.2734	0.5787	0.4550	0.4076	0.5000
20	0.4421	0.6915	0.5689	0.4726	0.6184	0.5370	0.5000
21	0.7143	0.5681	0.2856	0.6364	0.5366	0.4117	0.5720
22	0.3080	0.2545	0.5422	0.4195	0.4014	0.5220	0.2500
23	0.6183	0.6838	0.3198	0.5671	0.6126	0.4237	0.6170
24	0.2976	0.3239	0.5754	0.4158	0.4251	0.5408	0.2500
25	0.8135	0.4910	0.5124	0.7284	0.4955	0.5063	0.7500
26	0.7436	0.4987	0.4969	0.6611	0.4994	0.4984	0.5000
27	0.7427	0.3959	0.4740	0.6602	0.4529	0.4873	0.6390
28	0.7050	0.5064	0.4821	0.6289	0.5032	0.4912	0.5660
29	0.7454	0.5681	0.4832	0.6626	0.5366	0.4917	0.6250
30	0.6861	0.4344	0.4838	0.6143	0.4692	0.4920	0.6460
31	0.6731	0.5296	0.4860	0.6047	0.5152	0.4931	0.6250

Grey-Fuzzy Reasoning Analysis

In this experiment, a three-input one-output fuzzy-logic system is applied. To generate a fuzzy value, fuzzy reasoning is applied using a fuzzy inference engine (e.g., the Mamdani fuzzy inference system) and fuzzy rules. The fuzzy rules appear as If-Then control rules. The Grey relational coefficients (Cutting force, Surface roughness, and Metal Removal Rate MRR) serve as inputs to the fuzzy logic system. For the fuzzy rules in this case, the membership functions are classified as low, low-medium, medium, high-medium, and high for the grey relational coefficients of the input variables F_z, Ra, and MRR. Similarly, the output grey fuzzy relational grade (GFRG) is represented by very small, small, medium, large, and very large linguistic membership functions. This research applies triangular membership functions, and 31 rules are generated for this analysis, as shown in Figure 5. The generation of a fuzzy output requires maximum-minimum fuzzy reasoning composition. The conversion of the fuzzy output is accomplished using

a defuzzifier in MATLAB R2014a's fuzzy logic toolbox into GFRG, with the resulting values presented in Table 3.

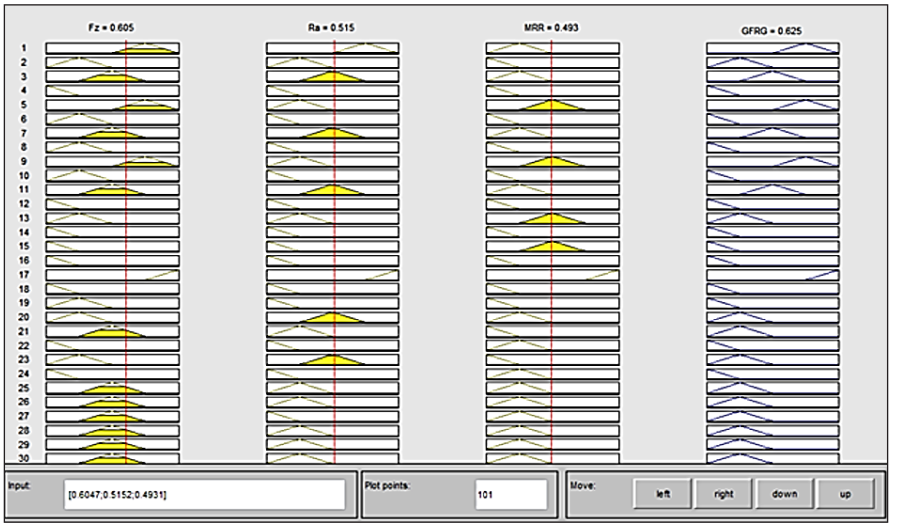


Figure 5. Fuzzy logic rules viewer

Higher GFRG levels are associated with better performance across multiple characteristics. Means analysis is done on GFRG. The response table for GFRG is given as Table 4. From the response table, the weight percentage of reinforcement is the most influential parameter, followed by the depth of cut and the feed rate. The physical reason for the strong effect of TiC is its high hardness and abrasiveness of TiC particles. An increase in TiC concentration increases resistance at the cutting edge and creates intermittent contact between the carbide cutter and hard titanium carbide particles, thereby increasing cutting resistance and adversely affecting the surface generation process. This influence of the experimentally observed parameters can be seen, for example, at the centre of the cutting: with TiC content varying from 2 to 10 wt%, cutting force changes from 114.45 N to 209.00 N, while surface roughness changes from 0.426 to 0.815 μm. An increase in feed rate increases the undeformed chip thickness, thereby increasing material removal per revolution and, consequently, the cutting load. In addition, feed rate influences the geometric component of surface roughness. Figure 6 depicts the main effects plot for GFRP. The optimal settings of process parameters, as identified, are wt% of reinforcement = 2%, Cutting speed (V) = 50 m/min, Feed Rate (f) = 0.06 mm/rev and Depth of cut (d) = 0.2 mm. Table 5 depicts the result of ANOVA for GFRG. ANOVA is used to study the effect of each process parameter on multiple performance characteristics. The fitted linear model for GFRG produced an R^2 of 37.99% and an adjusted R^2 of 31.10%, with an RMSE of 0.1192. The overall model was statistically significant ($F = 3.98$, $p = 0.012$), whereas the

lack-of-fit was statistically insignificant ($F_{table} = 2.84$, $p = 0.100$). These results indicate that the model provides a statistically meaningful representation of the experimental response, although its predictive capability is moderate.

Table 4
Response table for GFRG

Process Parameter	Level 1	Level 2	Level 3	Level 4	Level 5	Δ (Max-Min)	Rank
wt%	0.9200	0.5000	0.4729	0.3864	.3410	0.5790	1
V	0.5000	0.4524	0.5102	0.4340	0.500	0.0762	4
f	0.5720	0.4504	0.5357	0.4360	0.2500	0.3220	3
d	0.6170	0.4440	0.5292	0.4424	0.2500	0.3670	2

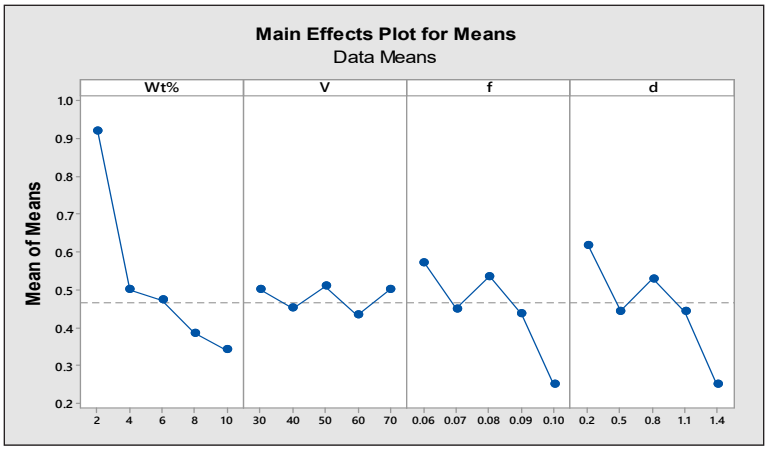


Figure 6. Response graph for every level of process parameter

Table 5
ANOVA for GFRG

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	4	0.226175	0.056544	3.98	0.012
Linear	4	0.226175	0.056544	3.98	0.012
wt%	1	0.178020	0.178020	12.54	0.002
V	1	0.000900	0.000900	0.06	0.803
f	1	0.024003	0.024003	1.69	0.205
d	1	0.023250	0.023250	1.64	0.212
Error	26	0.369183	0.014199		
Lack-of-Fit	20	0.333898	0.016695	2.84	0.100
Pure Error	6	0.035286	0.005881		
Total	30	0.595358			

Confirmation Test

Once the optimal value for the process parameter has been obtained, the results will be validated through a confirmation test. Table 6 displays the results of the confirmation test on machining responses of AA7068/TiC MMCs using a K30 cutting tool insert. Machining with an optimal parameter combination will reduce cutting force (Fz) from 129.45 N to 124.27 N, surface roughness (Ra) from 0.493 µm to 0.478 µm, and increase MRR from 136.37 mm^3/sec to 144.55 mm^3/sec. The value of Y at the optimal process parameter level can be estimated using the following Equation 5:

$$Y = Y_m + \sum_{i=1}^{q} Y_i - Y_m \quad [5]$$

Where Y_m is the mean of GFRG for all the experimental runs, and Y_i is the mean of GFRG values at the optimum level of input parameters. Table 6 indicates that machining with the optimum setting would improve GFRG from 0.9200 to 0.9395 for the experimental and predicted values, respectively.

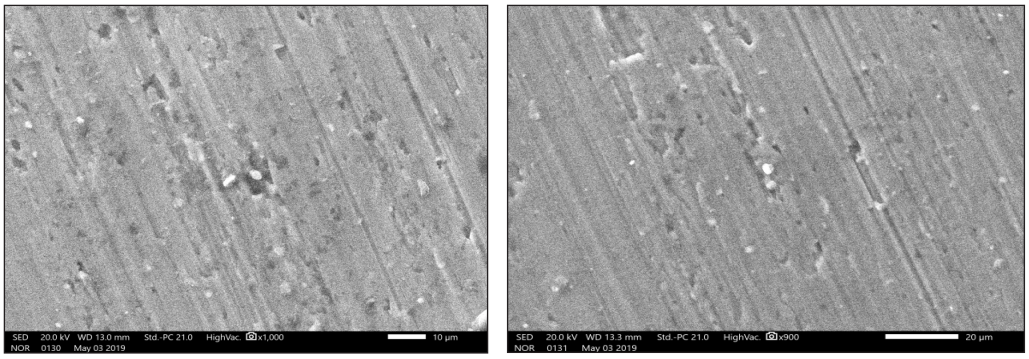
The confirmation experiment demonstrated simultaneous improvements in cutting force, surface roughness, and MRR under the optimised conditions. However, since the present confirmation test was performed as a single experimental verification, experimental standard deviations and confidence intervals were not determined. Therefore, the observed improvements are considered evidence of practical validation rather than formal statistical significance.

Table 6
Comparison of the results obtained for the initial and optimum machining conditions

Levels	Initial Machining Parameters Level	Optimal Machining Parameters Level		Improvement
	wt% = 4, V = 40 m/min, f = 0.07 mm/rev, d = 0.5 mm	wt% =, V = 50 m/min, f = 0.06 mm/rev, d = 0.2mm		
		Predicted	Experimental	
Cutting force (Fz), N	129.45		124.27	4.00% reduction
Surface Roughness (Ra), µm	0.493		0.478	3.04% reduction
Metal Removal Rate (MRR), mm^3/sec	136.37		144.55	5.99% increase
GFRG	0.5000	0.9395	0.9200	2.07% error
Improvement in the GFRG	0.0195			

Hence, the current study proves the effectiveness of the Grey-Fuzzy logic technique for optimisation of multi-response surface finishing of AA7068/TiC metal matrix

composites. Figures 7a to 7b depict the microscopic photographs of the machined surface after verification experiments.



(a) (b)

Figure 7. Machined surfaces of the AA7068/TiC MMCs before and after confirmation experiments

CONCLUSION

The grey-fuzzy approach was effectively used to evaluate cutting force, surface roughness, and MRR simultaneously during the turning of AA7068/TiC MMCs. In the experimental investigation, four process parameters were considered: TiC reinforcement ratio, cutting speed, feed rate and depth of cut.

The optimum condition obtained through the GFRG response analysis was 2 wt% TiC, 50 m/min cutting speed, 0.06 mm/rev feed rate and 0.2 mm depth of cut.

The TiC reinforcement content was the dominant factor in the GFRG response analysis, with a response range of (0.5790), followed by depth of cut (0.3670), feed rate (0.3220) and cutting speed (0.0762).

Confirmation test under optimum conditions led to a reduction in cutting force by 4.00% from 129.45 N to 124.27 N, an improvement in surface roughness by 3.04% from 0.493 µm to 0.478 µm, and an increase in MRR by 5.99% from 136.37 mm^3/s to 144.55 mm^3/s.

The calculated GFRG is 0.9395, while the experimentally obtained GFRG is 0.9200, yielding an absolute prediction error of 2.07%.

ANOVA shows that TiC content has a significant effect on GFRG ($p = 0.002$), but the effects of cutting speed, feed, and depth of cut are not significant individually ($p > 0.05$). The overall linear model is significant ($p = 0.012$), with R^2=37.99%, adjusted (R^2=31.10%) and a non-significant lack-of-fit (p=0.100).

In practical machining terms, the optimum condition offers a reasonable compromise among low cutting force, good surface finish and high material removal rate. Therefore, the results demonstrate the effectiveness of the grey-fuzzy approach in multi-response

decision-making within the machining range, rather than its universal superiority over other optimisation methods.

ACKNOWLEDGEMENT

Authors like to thank Dr M. Santhiramudu, Hon'ble Chairman, RGM Group of Institutions, for his unconditional support in carrying out the research work.

LIST OF ABBREVIATIONS

Term	-	Definition
MMCs	-	Metal Matrix Composites
TiC	-	Titanium Carbide
GRA	-	Grey relational analysis
GFRG	-	Grey Fuzzy relational grade
MCDM	-	Multi-Criteria Decision Making
MRR	-	Metal Removal Rate
ANOVA	-	Analysis of Variance

REFERENCES

Ahilan, C., Kumanan, S., & Sivakumaran, N. (2009). Multi-objective optimisation of CNC turning process using grey based fuzzy logic. *International Journal of Machining and Machinability of Materials, 5*(4), 434-451. https://doi.org/10.1504/ijmmm.2009.026902

Ashraf, P. M., & Shibli, S. M. A. (2009). Development of cerium oxide and nickel oxide-incorporated aluminium matrix for marine applications. *Journal of Alloys and Compounds, 484*(1-2), 477-482. https://doi.org/10.1016/j.jallcom.2009.04.134

Bhowmik, A., Sen, B., Patel, D., Beemkumar, N., Alhazaa, A., Magibalan, S., Romanovski, V., Kumar, R., & Bukate, B. B. (2026). Tribological performance enhancement of SiO_2 reinforced AA8011 composites through grey fuzzy based optimization. *Scientific Reports, 16*(1), Article 21881. https://doi.org/10.1038/s41598-026-52951-4

Cerik, B. C. (2017). Damage assessment of marine grade aluminium alloy-plated structures due to air blast and explosive loads. *Thin-Walled Structures, 110*, 123-132. https://doi.org/10.1016/j.tws.2016.10.021

Chen, X., Ma, L., Li, C., & Cao, X. (2014). Experimental study and genetic algorithm-based optimization of cutting parameters in cutting engineering ceramics. *The International Journal of Advanced Manufacturing Technology, 74*(5-8), 807-817. https://doi.org/10.1007/s00170-014-5979-0

Chiang, K.-T., & Chang, F.-P. (2006). Application of grey-fuzzy logic on the optimal process design of an injection-molded part with a thin shell feature. *International Communications in Heat and Mass Transfer, 33*(1), 94-101. https://doi.org/10.1016/j.icheatmasstransfer.2005.08.006

Das, B., Roy, S., Rai, R. N., & Saha, S. C. (2016). Application of grey fuzzy logic for the optimization of CNC milling parameters for Al–4.5%Cu–TiC MMCs with multi-performance characteristics. *Engineering Science and Technology, an International Journal, 19*(2), 857-865. https://doi.org/10.1016/j.jestch.2015.12.002

Deng, J. L. (1989). Introduction to grey system theory. *The Journal of Grey System, 1*(1), 1-24.

Dewangan, S., Gangopadhyay, S., & Biswas, C. K. (2015). Study of surface integrity and dimensional accuracy in EDM using fuzzy TOPSIS and sensitivity analysis. *Measurement, 63*, 364-376. https://doi.org/10.1016/j.measurement.2014.11.025

Dey, D., Bhowmik, A., & Biswas, A. (2023). A grey-fuzzy based multi-response optimisation study on the friction and wear characteristics of titanium diboride reinforced aluminium matrix composite. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 237*(14), 2227-2239. https://doi.org/10.1177/09544054221147973

Engin, K. E., Eşme, U., & Külekci, M. K. (2026). Optimization of vertical machining parameters of DIN 1.0038 steels using hybrid Taguchi based grey-fuzzy algorithm in CNC pocket milling process. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi, 41*(1), 115-128. https://doi.org/10.21605/cukurovaumfd.1706626

Gajević, S., Marković, A., Milojević, S., Ašonja, A., Ivanović, L., & Stojanović, B. (2024). Multi-objective optimization of tribological characteristics for aluminum composite using Taguchi grey and TOPSIS approaches. *Lubricants, 12*(5), Article 171. https://doi.org/10.3390/lubricants12050171

Ganesan, H., & Mohankumar, G. (2013). Optimization of machining techniques in CNC turning centre using genetic algorithm. *Arabian Journal for Science and Engineering, 38*(6), 1529-1538. https://doi.org/10.1007/s13369-013-0539-8

Giri, B. K., Pettersson, F., Saxén, H., & Chakraborti, N. (2013). Genetic programming evolved through bi-objective genetic algorithms applied to a blast furnace. *Materials and Manufacturing Processes, 28*(7), 776-782. https://doi.org/10.1080/10426914.2013.763953

Gururaja, S., Ramulu, M., & Pedersen, W. (2013). Machining of MMCs: A review. *Machining Science and Technology, 17*(1), 41-73. https://doi.org/10.1080/10910344.2012.747897

Haq, A. N., Marimuthu, P., & Jeyapaul, R. (2008). Multi response optimization of machining parameters of drilling Al/SiC metal matrix composite using grey relational analysis in the Taguchi method. *The International Journal of Advanced Manufacturing Technology, 37*(3-4), 250-255. https://doi.org/10.1007/s00170-007-0981-4

Heinz, A., Haszler, A., Keidel, C., Moldenhauer, S., Benedictus, R., & Miller, W. S. (2000). Recent development in aluminium alloys for aerospace applications. *Materials Science and Engineering: A, 280*(1), 102-107. https://doi.org/10.1016/s0921-5093(99)00674-7

Hsiao, Y. F., Tarng, Y. S., & Huang, W. J. (2008). Optimization of plasma arc welding parameters by using the Taguchi method with the grey relational analysis. *Materials and Manufacturing Processes, 23*(1), 51-58. https://doi.org/10.1080/10426910701524527

Hussain, S. A., Panchal, M., Meshram, K., Srinivas, R., Rajak, U., Kumar, R., & Gupta, M. (2025). Turning GFRP composites with multi-response optimisation using TOPSIS method. *International Journal on*

Interactive Design and Manufacturing (IJIDeM), 19(2), 1327-1339. https://doi.org/10.1007/s12008-024-01762-w

Isik, B., & Kentli, A. (2009). Multicriteria optimization of cutting parameters in turning of UD-GFRP materials considering sensitivity. *The International Journal of Advanced Manufacturing Technology, 44*(11-12), 1144-1153. https://doi.org/10.1007/s00170-009-1927-9

Jangra, K., Grover, S., & Aggarwal, A. (2012). Optimization of multi machining characteristics in WEDM of WC-5.3%Co composite using integrated approach of Taguchi, GRA and entropy method. *Frontiers of Mechanical Engineering, 7*(3), 288-299. https://doi.org/10.1007/s11465-012-0333-4

Kafaltiya, S., Agrawal, P. K., Sharma, P., Singh, V. K., Bhaskar, J., Chauhan, S., & Verma, A. (2024). Multi-response optimization of characteristics for graphite reinforced biodegradable PVA-fumaric acid cross-linked composite: A gray-fuzzy logic-based hybrid approach. *Journal of Vinyl and Additive Technology, 30*(5), 1190-1206. https://doi.org/10.1002/vnl.22112

Khandey, U., Ghosh, S., & Hariharan, K. (2017). Machining parameters optimization for satisfying the multiple objectives in machining of MMCs. *Materials and Manufacturing Processes, 32*(10), 1082-1093. https://doi.org/10.1080/10426914.2017.1279312

Krishnaiah, K., & Shahabudeen, P. (2012). *Applied design of experiments and Taguchi methods*. PHI Learning.

Krishnamoorthy, A., Rajendra Boopathy, S., Palanikumar, K., & Paulo Davim, J. (2012). Application of grey fuzzy logic for the optimization of drilling parameters for CFRP composites with multiple performance characteristics. *Measurement, 45*(5), 1286-1296. https://doi.org/10.1016/j.measurement.2012.01.008

Lin, J. L., & Lin, C. L. (2005). The use of grey-fuzzy logic for the optimization of the manufacturing process. *Journal of Materials Processing Technology, 160*(1), 9-14. https://doi.org/10.1016/j.jmatprotec.2003.11.040

Manna, A., & Bhattacharyya, B. (2005). Influence of machining parameters on the machinability of particulate reinforced Al/SiC–MMC. *The International Journal of Advanced Manufacturing Technology, 25*(9-10), 850-856. https://doi.org/10.1007/s00170-003-1917-2

Nalbant, M., Gökkaya, H., & Sur, G. (2007). Application of Taguchi method in the optimization of cutting parameters for surface roughness in turning. *Materials & Design, 28*(4), 1379-1385. https://doi.org/10.1016/j.matdes.2006.01.008

Narang, H. K., Singh, U. P., Mahapatra, M. M., & Jha, P. K. (2011). Prediction of the weld pool geometry of TIG arc welding by using fuzzy logic controller. *International Journal of Engineering, Science and Technology, 3*(9), 77-85. https://doi.org/10.4314/ijest.v3i9.6

Palanikumar, K., & Karthikeyan, R. (2006). Optimal machining conditions for turning of particulate metal matrix composites using Taguchi and response surface methodologies. *Machining Science and Technology, 10*(4), 417-433. https://doi.org/10.1080/10910340600996068

Palanikumar, K., Karunamoorthy, L., & Karthikeyan, R. (2006). Multiple performance optimization of machining parameters on the machining of GFRP composites using carbide (K10) tool. *Materials and Manufacturing Processes, 21*(8), 846-852. https://doi.org/10.1080/03602550600728166

Prajapati, V., Dinbandhu, Vora, J. J., Das, S., & Abhishek, K. (2020). Study of parametric influence and welding performance optimization during regulated metal deposition (RMD™) using grey integrated

with fuzzy Taguchi approach. *Journal of Manufacturing Processes, 54*, 286-300. https://doi.org/10.1016/j.jmapro.2020.03.017

Rajmohan, T., Palanikumar, K., & Prakash, S. (2013). Grey-fuzzy algorithm to optimise machining parameters in drilling of hybrid metal matrix composites. *Composites Part B: Engineering, 50*, 297-308. https://doi.org/10.1016/j.compositesb.2013.02.030

Sahoo, A. K., Pradhan, S., & Rout, A. K. (2013). Development and machinability assessment in turning Al/SiCp-metal matrix composite with multilayer coated carbide insert using Taguchi and statistical techniques. *Archives of Civil and Mechanical Engineering, 13*(1), 27-35. https://doi.org/10.1016/j.acme.2012.11.005

Samy, G. S., & Kumaran, S. T. (2017). Measurement and analysis of temperature, thrust force and surface roughness in drilling of AA (6351)-B_4C composite. *Measurement, 103*, 1-9. https://doi.org/10.1016/j.measurement.2017.02.016

Sengottuvel, P., Satishkumar, S., & Dinakaran, D. (2013). Optimization of multiple characteristics of EDM parameters based on desirability approach and fuzzy modeling. *Procedia Engineering, 64*, 1069-1078. https://doi.org/10.1016/j.proeng.2013.09.185

Shin, J., Kim, T., Kim, D., Kim, D., & Kim, K. (2017). Castability and mechanical properties of new 7xxx aluminum alloys for automotive chassis/body applications. *Journal of Alloys and Compounds, 698*, 577-590. https://doi.org/10.1016/j.jallcom.2016.12.269

Shrotriya, A., Nayak, C., Patel, G., Khan, M. I., & Pathak, V. K. (2026). A new method based on AVOA-optimized support vector machine for predicting performance characterisitcs during turning aluminium 7068 under graphene nanofluid. *Scientific Reports, 16*(1), Article 2564. https://doi.org/10.1038/s41598-025-32577-8

Shunmugesh, K., Raphel, A., Kurian, S., Hiremath, S., Anne, G., Murthy, B. R. N., & Vishwanatha, H. M. (2025). A novel way to optimize the process parameters by integrating the grey relational coefficient and the combined compromise solution for machining the CFRP composites. *Scientific Reports, 15*(1), Article 34729. https://doi.org/10.1038/s41598-025-18368-1

Soepangkat, B. O. P., & Pramujati, B. (2013). Optimization of surface roughness and recast layer thickness in the wire-EDM process of AISI D2 tool steel using Taguchi-grey-fuzzy. *Applied Mechanics and Materials, 393*, 21-28. https://doi.org/10.4028/www.scientific.net/amm.393.21

Teimouri, R., Amini, S., & Mohagheghian, N. (2017). Experimental study and empirical analysis on effect of ultrasonic vibration during rotary turning of aluminum 7075 aerospace alloy. *Journal of Manufacturing Processes, 26*, 1-12. https://doi.org/10.1016/j.jmapro.2016.11.011

Tzeng, Y.-F., & Chen, F.-C. (2007). Multi-objective optimisation of high-speed electrical discharge machining process using a Taguchi fuzzy-based approach. *Materials & Design, 28*(4), 1159-1168. https://doi.org/10.1016/j.matdes.2006.01.028

Yang, Y.-S., & Huang, W. (2012). A grey-fuzzy Taguchi approach for optimizing multi-objective properties of zirconium-containing diamond-like carbon coatings. *Expert Systems with Applications, 39*(1), 743-750. https://doi.org/10.1016/j.eswa.2011.07.067

Pertanika J. Sci. & Technol. 34 (5): 2907 - 2927 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review Article

A Systematic Review of Hybrid Deep Learning and Optimisation Techniques to Address Class Imbalance and Malware Detection in IoT Networks

Ernest Ezema and Julia Juremi

Faculty of Computing, Engineering and Technology, Asia Pacific University of Technology & Innovation, 57000 Bukit Jalil, Kuala Lumpur, Malaysia

ABSTRACT

The Internet of Things (IoT) is a rapidly expanding technology, resulting in an increasingly complex cybersecurity environment for malware detection. Network traffic classification (NTC) remains a challenge. Traditional detection techniques are ineffective at detecting encrypted traffic when attackers change their behaviour and when intrusion datasets exhibit class imbalance. The systematic literature review, conducted in accordance with the PRISMA 2020 framework, presents the results of three primary peer-reviewed studies (2019-2025). The corpus was constructed through a systematic search of IEEE Xplore, ScienceDirect, ACM Digital Library, Scopus, and SpringerLink, using strict inclusion and exclusion criteria. Six thematic dimensions are explored: IoT threat landscape, malware detection strategies, hybrid deep learning architectures, optimisation techniques for class imbalance, feature engineering, and evaluation practices. Key findings indicate that integrating CNNs and LSTMs, PSO-based classifiers, attention mechanisms, and federated learning models is a cutting-edge area in IoT security research. There are significant challenges in adaptive sampling, zero-day generalisation, encrypted traffic classification, scalability for resource-constrained environments and real-world testbed validation. To address these needs, this review presents a conceptual framework that combines GA/PSO-guided adaptive sampling with a deep learning pipeline evaluated on a simulated IoT testbed. The study concludes with a structured research roadmap for future work on explainability, real-time adaptability, and lightweight security mechanisms.

Keywords: Class imbalance, deep learning, feature optimisation, IoT security, malware detection, network traffic classification, particle swarm optimisation

ARTICLE INFO

Article history:
Received: 26 May 2026
Accepted: 22 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.15

E-mail addresses:
TP078026@mail.apu.edu.my (Ernest Ezema)
julia.juremi@apu.edu.my (Julia Juremi)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

The Internet of Things (IoT) is an evolving technology that is changing the healthcare, manufacturing, transportation, and city infrastructure landscape. The recent market analysis has shown that in 2024 there were more than 18.8 billion connected IoT devices and that this number is projected to reach 40 billion in 2030 (Khraisat et al., 2025). With this expansion has come an alarming increase in cyber threats against IoT. Indeed, SonicWall registered 112.29 million IoT malware attacks in 2022 alone, nearly a threefold increase from the 32.7 million attacks in 2018 (SonicWall, 2022); between January and June 2023, the number of daily Internet of Things (IoT) DDoS attacks rose by about 353%, from 144 to 611 (Netscout, 2024). The number of attacks on IoT devices worldwide averaged 820,000 per day in 2025 (Varonis, 2025).

The graph in Figure 1 illustrates the trend in IoT malware attacks over five years, from 2022 to 2026. The data highlight consistent growth in IoT security threats, with the attack index increasing significantly each year. This increase rose 560% from the 2022 baseline to its peak in 2026, indicating serious security concerns and frequent attacks by threat actors targeting current and emerging IoT technologies. Conti et al. (2018) demonstrated the traditional IoT security challenges and opportunities in adopting solutions and remediation techniques when mitigating the everyday threats posed by IoT networks.

In this context, Network Traffic Classification (NTC) has become a key mechanism for detecting anomalies and conducting cyber defence in real time. With the general trend toward TLS/SSL encryption and advanced traffic obfuscation, traditional NTC methods, which rely on port numbers, payload signatures, and Deep Packet Inspection (DPI), are becoming ineffective (Callado et al., 2009; Shbair et al., 2020). This has led to a surge in the study of machine learning and deep learning techniques for their ability to detect intricate relationships in high-dimensional, encrypted or untyped traffic (Alshamrani et al., 2022; Nguyen & Armitage, 2008).

Even the most sophisticated deep learning-based NTC systems have inherent problems. Amongst these, the class imbalance problem stands out as a major concern in real-world IoT datasets: benign traffic is far more frequent than malicious traffic, and the rare malicious traffic classes, such as reconnaissance, shellcode injection and data exfiltration,

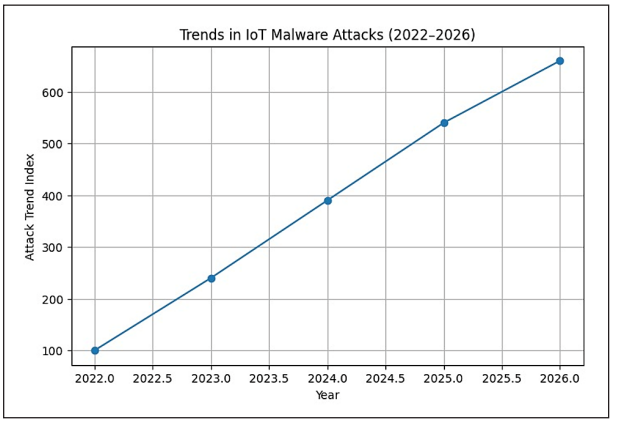


Figure 1. Trends in IoT malware attacks 2022 to 2026

are much less frequent than the more common ones. This imbalance makes models more prone to recalling majority classes and to failing to recall the most important threats to find (He & Garcia, 2009; Krawczyk, 2016). The SMOTE, SMOTE-ENN, and random oversampling methods are conventional methods that have been shown to introduce synthetic noise, overfitting, and a lack of adaptive control during training (Fernandez et al., 2018; Shanmugam et al., 2025).

Another aspect of the challenge is the relevance and optimisation of features. Deep learning models require many informative, non-redundant features, but finding them in heterogeneous IoT environments with varying protocols and levels of encryption is challenging. Current approaches generally rely on brute-force feature selection or manual preprocessing, but they are not easily scalable or interpretable (Eiben & Smith, 2015; Xue et al., 2016). The third limitation is the lack of testbed validation under realistic IoT operating conditions: most published models are evaluated on pre-cleaned benchmark datasets (Moustafa et al., 2019).

This systematic review addresses these challenges by synthesising 33 peer-reviewed primary studies published between 2019 and 2025, organised around six thematic dimensions. The review follows the PRISMA 2020 methodology and integrates the 7 studies from the original corpus with 23 additional verified articles and 3 closely related works, substantially broadening the analytical base. The review identifies current advances, persistent gaps, and a roadmap for next-generation hybrid intrusion detection systems for IoT environments.

BACKGROUND STUDY

IoT Systems, Applications, and the Evolving Threat Landscape

IoT systems comprise a wide array of devices and systems, ranging from smart home appliances, wearables, and medical monitoring devices to industrial sensors, autonomous vehicles, and smart city infrastructure. This heterogeneity poses a challenging security problem because devices have diverse hardware constraints, communication protocols, and firmware update frequencies, making it difficult to ensure uniform security provisioning (Lone et al., 2023; Stoyanova et al., 2020).

The number and sophistication of IoT threat actors have increased. The Mirai family of malware exploits weak credentials and unpatched vulnerabilities to join devices into large-scale DDoS botnets (Tanka & Osei-Bonsu, 2024). The ransomware incident that impacted healthcare provider Prospect Medical Holdings in 2023 temporarily forced security systems at hospitals across various states in the United States to shut down Emergency departments. The Internet of Things (IoT) threat landscape has also expanded to include nation-state actors, such as the Flax Typhoon botnet, which is estimated to comprise more than 200,000 compromised IoT devices (Asimily, 2025). Other attack vectors include cryptojacking, data exfiltration, and supply chain attacks, all of which underscore the importance of effective intrusion detection in IoT security.

Malware Detection Strategies

Malware detection techniques fall into three categories: static, dynamic and hybrid (Qureshi et al., 2024). Static analysis considers program code without running the program, uncovering known signatures and structural abnormalities. Static methods are effective against known threats, but not against polymorphic or zero-day malware. Dynamic analysis is performed while the software runs, enabling more effective detection of new threats at a higher computational cost. Hybrid analysis combines both approaches, incorporating static pattern recognition with behavioural monitoring, enabling the detection of a broader range of malicious activities.

All three detection paradigms have seen tremendous progress with deep learning. Static feature extraction techniques that extract characteristics from a binary representation of malware are well-suited to Convolutional Neural Networks (CNNs) (Riaz et al., 2022). Dynamic behavioural sequence analysis is well-suited to Recurrent Neural Networks (RNNs) and LSTMs (Wang et al., 2018). By modelling normal traffic, autoencoders can be used to identify anomalies. Generative Adversarial Networks (GANs) can be used to generate adversarial malware samples, providing powerful training data for these models. Transfer learning and Deep Reinforcement Learning (DRL) are additional techniques used in detection that adapt to resource-constrained and evolving threat scenarios (Akhtar & Feng, 2022; Qureshi et al., 2024).

Network Intrusion Detection Systems in IoT

One of the critical first lines of defence in an IoT network is a Network Intrusion Detection System (NIDS), which can detect unauthorised activity and policy violations. They have been categorised by detection method as signature-based, anomaly-based, and specification-based. The ability to detect unknown/zero-day attacks through patterns of anomalies that cannot be known a priori has made anomaly-based NIDS one of the most researched approaches, but it is generally associated with higher false positive rates (FP-rate) (Khraisat et al., 2019; Wanjau et al., 2022).

Several NIDS architectures have been presented that integrate multiple detection methods or ensemble multiple deep learning models and have proven more effective than NIDS based on a single method. Wanjau et al. (2022) surveyed the literature on 67 papers on hybrid deep learning NIDS and concluded that CNN-LSTM were the most common architecture, and algorithmic, cooperative and hierarchical were the most prominent paradigms of hybridisation in the architecture. In more recent research, PSO, attention mechanisms, federated learning, and graph neural networks have been added to further enhance detection accuracy and scalability (Chaganti et al., 2022; Ganesh et al., 2026; Hajj et al., 2023).

SURVEY METHODOLOGY

Protocol and PRISMA Compliance

This systematic literature review was carried out following the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page et al., 2021). In the pre-planning stage, a review protocol was developed that covered the following: research questions, search strategy, inclusion/exclusion criteria, data extraction procedure, and quality assessment criteria. The protocol provides transparency, consistency, and reproducibility in the review process.

Research Questions

The review was guided by four primary research questions:

RQ1 : What hybrid deep learning architectures and optimisation techniques have been used for IoT network intrusion detection and malware classification from 2019 to 2025?

RQ2 : What strategies are used to overcome class imbalance and feature selection problems in existing literature, and what problems remain?

RQ3 : What evaluation practices, datasets, and validation strategies characterise the current body of evidence, and how do they constrain generalisability?

RQ4 : What are the gaps in research, and what recommendations can be offered for future work on hybrid IDS in IoT environments?

Search Strategy and Information Sources

A structured Boolean search was conducted across five major academic databases: IEEE Xplore, ScienceDirect (Elsevier), ACM Digital Library, Scopus, and SpringerLink. Some keywords and terms were merged as search terms, which include: 'network traffic classification', 'intrusion detection', 'IoT malware', 'deep learning', 'machine learning', 'class imbalance', 'Particle Swarm Optimisation', 'hybrid model', 'LSTM', 'CNN', 'federated learning', and 'feature selection'. Only peer-reviewed primary research articles published in English during the period of January 2019 to May 2025 were searched.

Inclusion and Exclusion Criteria

The articles included met the following criteria: (i) proposed or examined a deep learning or hybrid ML model for NTC, intrusion detection or malware classification in IoT or very similar network contexts; (ii) explicitly covered any of the following topics: class imbalance, feature selection, encrypted traffic, optimisation techniques, or malware detection; (iii) provided empirical results on a standardised or real-world dataset; and (iv) published in full text in a peer-reviewed journal or conference. The following articles

were excluded: those that relied exclusively on traditional ML but did not include deep learning methods; those that didn't provide empirical implementation or validation; those that were in the automotive sector, the satellite sector or non-IoT networks; or were secondary reviews, editorials, or conceptual papers with no experimental contribution.

Screening and Selection Process

The PRISMA flow chart highlighted in Figure 2 contains the screening and selection processes of published articles that were reviewed and suitable to be included in this article. The figure shows a preliminary list of 412 articles that were identified through database searches. After eliminating 67 duplicates, 345 distinct articles were retained. Out of these, 268 were filtered out because the abstract or title was irrelevant to the scope of the study, namely, because of their emphasis on either legacy systems or non-IoT contexts, or because they did not utilise deep learning. Articles that did not provide full text were also not included. Out of the rest of 77 full-text articles, 33 articles were included in the systematic synthesis.

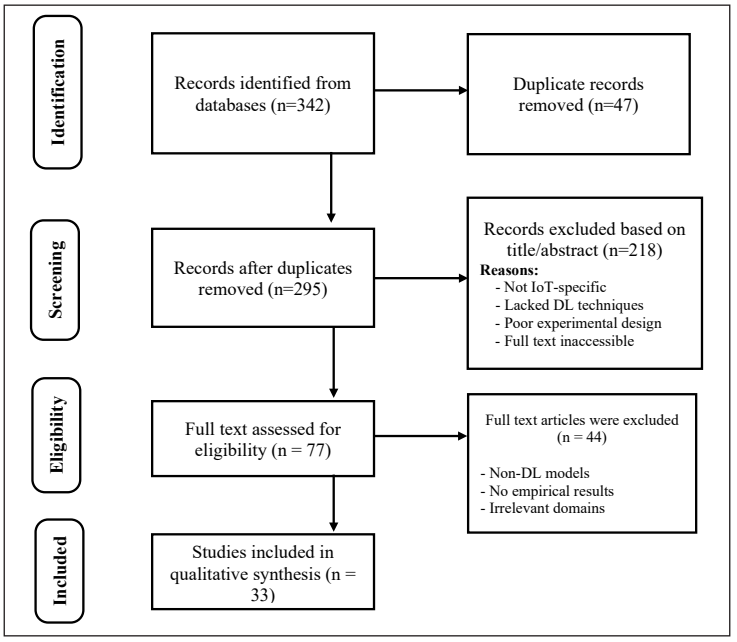


Figure 2. The screening and selection process is shown in a PRISMA flow chart
Note. Adapted from Page et al. (2021)

Data Extraction and Analytical Framework

To extract the studies in a standardised and systematic manner, a six-part analytical framework was used to analyse each of the studies thematically: (1) problem domain and attack type addressed; (2) deep learning architecture and hybridisation strategy;

(3) optimisation technique employed; (4) dataset characteristics and class imbalance handling; (5) feature engineering approach; and (6) evaluation metrics and validation protocol. This framework had the advantage of providing insight into both per-paper analysis and cross-study thematic synthesis.

Quality Assessment

Each included article was assessed for methodological quality on four criteria: clarity of the research design, appropriateness of the experimental methodology, soundness of the empirical results, and adequacy of the validation protocol. Studies relying solely on static train/test splits without cross-dataset or adversarial validation were noted to have limited generalisability. Studies lacking clear descriptions of class imbalance handling or feature engineering were flagged accordingly.

Citation Analysis and Publisher Distribution

Figure 3 highlights the citation analysis and publishing distribution of the published studies reviewed in this work. Among the 33 included studies, IEEE publications accounted for approximately 39% of the corpus, Elsevier/ScienceDirect for 27%, Springer for 18%, MDPI for 9%, and other publishers for the remaining 7%. The most heavily cited studies in the corpus include Wanjau et al. (2022) with 200+ citations, Khraisat et al. (2019) with 170+ citations, and Chaganti et al. (2022) with 130+ citations, reflecting their foundational role in the hybrid IDS and IoT security literature.

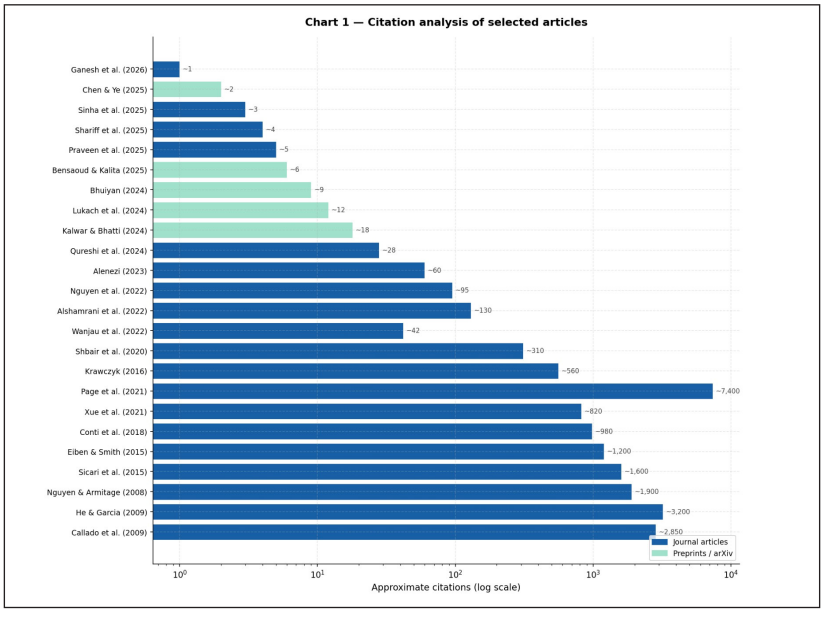


Figure 3. Citation Analysis of selected articles

THEMATIC ANALYSIS OF REVIEWED LITERATURE

Hybrid Deep Learning Architectures for IoT Intrusion Detection

According to Shaukat et al., (2023) a deep learning model can be used for classification and deep feature computation. The model can be used either as end-to-end for feature computation or malware detection. Based on our review of key deep learning models, the general conclusions derived from the analysed database is that the most common architecture in this context is the hybridisation of CNN and LSTM layers to encode both the spatial and temporal characteristics of network traffic. This method has been demonstrated by Sinha et al. (2025) in a smart home IoT setup, achieving 99.8% accuracy and 90.2% adversarial robustness on the BoT-IoT dataset using a CNN-LSTM model. Likewise, Yaras and Dener (2024) tested their novel hybrid deep learning algorithm with IoT-based IDS and found that it outperforms the single-model baselines, and Bamber et al. (2025) validated their CNN-LSTM approach with an F1-score of 98.7% on the NSL-KDD dataset.

In addition to CNN-LSTM hybrids, the literature surveyed shows an expanding range of architectures. Bensaoud and Kalita (2025) proposed a multi-module deep fusion architecture using Self-Organising Maps (SOMs), Deep Belief Networks (DBNs), Autoencoders, and attention mechanisms, and tested it on three datasets: NSL-KDD, UNSW-NB15, and CICIoT2023. Taşcı (2024) performed evaluation of three cybersecurity datasets; CIC IoT 2023, CIC-Malmem-2022, and CICIDS2017 achieving detection accuracies of 98.36%, 99.905 and 99.99%. Sun et al. (2020) proposed a CNN-LSTM network to extract both spatial and temporal traffic features, and Popoola et al. (2021) used a hybrid deep learning approach for detecting botnet attacks in IoT networks. Wanjau et al. (2022) conducted a systematic literature review of 67 papers that used hybrid deep learning NIDS architectures and found that CNN-LSTM was the most common hybrid architecture, and that HAST-IDS was the most recent architecture in the field of spatiotemporal feature learning for NIDS.

These are new frontiers: Transformer-based and attention architectures. Long et al. (2024) showed that Transformer models outperform CNN-LSTM in terms of scalability and accuracy (over 93%) for cloud security intrusion detection. Zegarra Rodríguez et al. (2023) proposed an attentive Transformer with automatic explainable feature selection for IoT IDS, and Wei et al. (2023) presented multi-objective evolving LSTM networks with attention for network intrusion detection. The problem of long-range dependency in complex traffic sequences is one of the main drawbacks of conventional recurrent models, which attention-based models' tackle.

Table 1 highlights the studies reviewed, the architectures adopted by the researchers, the datasets used, and a comparison of their key performance metrics. The data in Table 1 demonstrate the effectiveness of a hybrid deep learning architecture for detecting network intrusions and malware across diverse IoT network environments. From the table, CNN+LTSM models were widely adopted by different authors, achieving 99.8% accuracy

on the BoT-IoT datasets and about 98.7% F1-Score on NSL-KDD. Other architectures, including transformer-based models, showed promising results, with some achieving more than 99% accuracy. Overall, this review's findings show the potential of hybrid deep learning models to improve detection accuracy within the IoT networks (AI & Dener, 2021).

Table 1
Hybrid deep learning architectures across reviewed studies

Study	Architecture	Dataset(s)	Key Performance
Sinha et al., 2025	CNN + LSTM	BoT-IoT	99.8% Acc; 90.2% adversarial robustness
Bamber et al., 2025	CNN + LSTM	NSL-KDD	98.7% F1-score
Yaras and Dener, 2024	Novel Hybrid DL	IoT datasets	Outperforms single-model baselines
Bensaoud and Kalita, 2025	SOM + DBN + AE + Attention	NSL-KDD, UNSW-NB15, CICIoT2023	Multi-dataset generalisability
Sun et al., 2020	CNN + LSTM (DL-IDS)	CICIDS2017	High detection rate
Long et al., 2024	Transformer	NSL-KDD, custom	>93% Acc for cloud NIDS
Altunay and Albayrak, 2023	CNN + LSTM	Industrial IoT	Industrial IIoT validated
Popoola et al., 2021	Hybrid DL	N-BaIoT	Botnet detection in IoT
Ganesh et al., 2026	PSO + ANN (Hybrid IDS)	NSL-KDD, UNSW, NSDL	99.78% Acc, 0.003% FPR

Note. Acc = Accuracy; FPR = False Positive Rate; AE = Autoencoder; DBN = Deep Belief Network; SOM = Self-Organising Map

PSO and Metaheuristic Optimisation for IDS

PSO has been the most popular metaheuristic used to improve the performance of hybrid IDS both in feature selection and hyperparameter tuning. Ganesh et al. (2026) successfully applied PSO+ANN to the NSL-KDD, UNSW-NB15, and NSDL datasets, achieving 99.78% accuracy and a 0.003% false-positive rate, whereas PSO+DT and PSO+KNN were comparatively less effective, achieving 98.5% and 99.6% accuracy, respectively. Reducing the number of features in PSO from 48 to 25 reduced computational complexity while maintaining excellent classification performance. Praveen et al. (2025) constructed a PSO-LightBoost algorithm by combining PSO hyperparameter optimisation with the VarMiRF feature selection pipeline, achieving 97.28% accuracy, 93.49% precision, and 91.88% recall on the UNSW-NB15 dataset.

The VMF feature selection framework was also used in combination with PSO, as Shariff et al. (2025) did for LightGBM in their Swarm Gradient Boosting IDS (SGB-IDS). The research by Chaganti et al. (2022) demonstrates that PSO-DNN can detect intrusions

in the IoMT system, achieving 96% accuracy while addressing the high computational requirements of medical IoT devices. Logeswari et al. (2025) used a synergistic dual-layer feature selection approach with PSO, SVM-RFE, and LightGBM/XGBoost classifiers on the TON-IoT dataset. In addition to PSO, Genetic Algorithms (GA) have been used to address anomaly detection in IoT networks (Securing IoT Communications, 2025), and Ahmed et al. (2024) showed that by using advanced feature selection, ensemble models based on Genetic Algorithms (GA) could be optimised to achieve high accuracy in intrusion detection in an IoT network. As shown in Table 2, these metaheuristic approaches consistently outperform random or filter-only feature selection, with PSO showing particularly strong performance when combined with gradient boosting classifiers.

Table 2
Optimisation techniques and feature engineering approaches

Study	Optimisation	Feature Method	Outcome
Ganesh et al., 2026	PSO	PSO-based (48→25 features)	99.78% Acc, 0.003% FPR
Praveen et al., 2025	PSO	VarMiRF	97.28% Acc, 93.49% Precision
Shariff et al., 2025	PSO	VMF (Variance + MI + RF)	High IDS accuracy via SGB-IDS
Chaganti et al., 2022	PSO	PSO-based	96% Acc for IoMT
Logeswari et al., 2025	PSO + SVM-RFE	SDFC dual layer	Validated on TON-IoT
Ahmed et al., 2024	Ensemble opt.	Advanced feature selection	Superior NIDS accuracy
Sinha et al., 2025	None	SHAP-based ranking	Interpretability + 99.8% Acc
Bhuiyan, 2024	None	Z-score only	Baseline DL with oversampling

Note. VMF = Variance Mutual Forest; VarMiRF = Variance Mutual Information Random Forest; SGB-IDS = Swarm Gradient Boosting IDS; SDFC = Synergistic Dual-Layer Feature Classifier

Class Imbalance Handling

One of the most persistent challenges in IoT IDS is class imbalance. Reviewed literature shows a range of sampling approaches, including simple random oversampling (Bhuiyan, 2024) and more complex hybrid sampling. On the IoT-23 dataset, Chen and Ye (2025) systematically compared SMOTE, SMOTE-ENN, and a combined strategy, revealing incremental improvements across all the strategies. In the case of BoT-IoT, Sinha et al. (2025) used SMOTE with undersampling, and Praveen et al. (2025) used ADASYN for class balancing, which they found to be better than SMOTE in terms of synthetic sampling of the boundary regions. More sophisticated methods include using GANs to generate synthetic samples from minority classes, as explored in Kalwar and Bhatti (2024), or incorporating imbalance management into the model design. Tabassum et al. (2022) introduced

FEDGAN-IDS, a GAN-based federated learning-based IDS with built-in class balancing. Gu et al. (2023) conducted a study on class imbalance in the MDPI and concluded that learning-based techniques for handling high-dimensional imbalanced traffic consistently outperform static resampling techniques. Based on our review, as shown in Table 3, most reviewed primary studies still use a static SMOTE or its variants, which are applied offline prior to model training rather than adaptively during the model learning process.

Table 3
Dataset characteristics and class imbalance handling

Study	Dataset(s)	Imbalance Handling	Limitation
Chen and Ye, 2025	IoT-23	SMOTE, SMOTE-ENN, Hybrid	Static, offline only
Sinha et al., 2025	BoT-IoT	SMOTE + undersampling	No adaptive control
Praveen et al., 2025	UNSW-NB15	ADASYN	Dataset-specific tuning
Shariff et al., 2025	Kaggle (unspec.)	SMOTE	SMOTE noise risk
Bhuiyan, 2024	NSL-KDD, TON-IoT	Random oversampling	Overfitting risk
Bensaoud and Kalita, 2025	NSL-KDD, UNSW-NB15, CICIoT2023	None (model-capacity reliance)	Poor minority class recall
Tabassum et al., 2022	Custom / federated	GAN-based (FEDGAN)	High complexity
Abdelkhalek and Mashaly, 2023	NSL-KDD	Data augmentation	Limited evaluation scope

Note. ADASYN = Adaptive Synthetic Sampling; SMOTE-ENN = SMOTE with Edited Nearest Neighbours; GAN = Generative Adversarial Network

Federated Learning for Privacy-Preserving IoT IDS

To this end, Federated learning (FL) has become a promising paradigm for IoT IDS, which not only addresses privacy protection but also addresses data decentralisation. Instead of storing raw traffic data on a central server, FL enables the IoT device to train local models while sending only model weight updates. Khraisat et al. (2025) showed that Federated Learning with the FedAvg and FedAvgM algorithms can achieve competitive performance on the ID task without transferring traffic data. Hajj et al. (2023) introduce a novel approach to cross-layer federated learning for lightweight IDSs for IoT devices, which they demonstrate on TON-IoT and IoT-23 datasets.

To address class imbalance in WSNs, Tabassum et al. (2022) modified FL with GAN-based privacy protection, which they named FL with GAN-based privacy protection (FEDGAN-IDS). Rashid et al. (2023) presented a framework for FL-based IDS for Industrial IoT networks, and Mosaiyebzadeh et al. (2025) demonstrated that privacy-preserving FL for IoHT devices can achieve competitive detection accuracy while providing robust privacy guarantees. GDPR and HIPAA rules greatly limit data sharing, especially

in healthcare IoT. One of the major open research problems in the current research on Byzantine-robust aggregation is the assumption of the optimal behaviour of clients, namely, honest and reliable clients.

Evaluation Practices and Validation Gaps

The quality of the evaluations in the reviewed corpus is highly variable. The most common measures include accuracy, precision, recall, F1-score and AUC-ROC, but they are typically reported on static, pre-balanced splits that are different from the classes' distribution in real life. Bensaoud and Kalita (2025) contribute best practices by validating their work with multiple datasets on three sets: NSL-KDD, UNSW-NB15, and CICIoT2023. Ganesh et al. (2026) used k-fold cross-validation (k=5) and reported the standard deviation across folds to assess model stability.

Despite its importance in the field, adversarial evaluation remains significantly underrepresented in the corpus, with only Sinha et al. (2025) explicitly evaluating model robustness to adversarial perturbations (90.2% adversarial accuracy). None of the reviewed primary studies performed testbed validation under real or simulated real-world IoT network conditions. Only a few studies (Balhareth & Ilyas, 2024; Sinha et al., 2025; Zegarra Rodríguez et al., 2023) addressed interpretability, either using SHAP or LIME explanations. These gaps are significant limitations to the proposed models being put into operation.

Table 4 illustrates the evaluation matrices and validation practices across selected studies. According to Sicari et al. (2015), the satisfaction of security and privacy requirements plays a vital role in determining the data confidentiality and authentication across IoT networks. To obtain accurate results and determine security and privacy requirements for IoT IDS, the evaluation matrices are applied to the trained datasets.

Table 4
Evaluation metrics and validation practices across selected studies

Study	Metrics	Validation	Limitation
Ganesh et al., 2026	Acc, DR, FPR, F1	k-fold (k=5) + σ	No live testbed
Praveen et al., 2025	Acc, Prec, Recall, F1	Train/test split	Single dataset
Bensaoud and Kalita, 2025	Acc, F1, Recall, AUC	Multi-dataset	No adversarial testing
Sinha et al., 2025	Acc, F1, Recall, AUC, FPR	80/20 + adversarial	Single dataset
Chen and Ye, 2025	Acc, F1, ROC-AUC, Recall	Stratified split	No cross-dataset test
Shariff et al., 2025	Acc, F1, Prec, RMSE	Static split	No adversarial/real-time
Hajj et al., 2023	Acc, F1, Recall	Cross-client FL	Simulated FL clients only
Khraisat et al., 2025	Acc, F1, Privacy metrics	FL aggregation	Assumes honest clients

Note. Acc = Accuracy; DR = Detection Rate; FPR = False Positive Rate; FL = Federated Learning; σ = Standard Deviation across folds

RESEARCH GAPS AND CRITICAL ANALYSIS

Class Imbalance: Persistent and Under-Solved

In the literature reviewed, although class imbalance has been a major focus of research, the most-cited approach is static SMOTE or its derivatives, which are used as a preprocessing technique independent of the model's learning process. This approach has proven to have several limitations: SMOTE can produce synthetic examples in overlapping regions, adding noise; fixed resampling ratios are unsuitable for dynamically changing error landscapes; and offline augmentation is inadequate to address the streaming and non-stationary IoT traffic distribution (He & Garcia, 2009; Krawczyk, 2016; Shanmugam et al., 2025). Praveen et al. (2025) used adaptive boundary-aware sampling in the only study in the corpus to use ADASYN. No study used feedback-guided adaptive resampling coupled with model performance during training.

Zero-Day and Invasive Malware: Generalisation Deficit

There are structural weaknesses in generalisation to unseen and invasive malware, as reported in the literature. The attack distributions for most models are known in advance of training and evaluation, resulting in an artificially high measure of success. Bensaoud and Kalita (2025) evaluated zero-day detection for CICIoT2023 but did not generate realistic OOD samples. While retrieval-based approaches for out-of-distribution classification were proposed by Lukach et al. (2024), they were not empirically tested. No study has explicitly trained on programmatically generated synthetic invasive malware traffic in a controlled testbed setting.

Feature Selection: Shallow and Non-Adaptive

There are inconsistencies in the use of feature optimisation in the corpus. Most studies use post hoc or filter methods (SHAP, PCA, variance thresholding) that do not interact with the classifier during training. Only two studies (Praveen et al., 2025; Shariff et al., 2025) used statistical feature filtering in conjunction with ranking features by model importance as part of a systematic hybrid feature selection pipeline. The deep learning feature space had not been directly targeted by any study using a metaheuristic feature search (GA or PSO), as they have proven successful in high-dimensional classification problems (Xue et al., 2016).

Encrypted Traffic: Incomplete Solutions

The introduction of universal TLS/SSL encryption in IoT communications has made payload-based inspection increasingly infeasible. Although Lukach et al. (2024) and Shbair et al. (2020) attempted to solve this problem conceptually, none of the reviewed

work provided an end-to-end encrypted traffic classification system that includes flow-level feature extraction, optimisation, and deep learning inference. Some statistical flow properties, including packet size distributions, inter-arrival times, and direction sequences, are yet to be fully utilised as privacy-preserving classification signals (Papernot et al., 2018).

Real-World Deployment and Testbed Validation

It is noted that the literature reviewed has very little validation in real-world or testbed settings, which is the most significant practical gap. All the reviewed studies evaluated only a pre-cleaned, static benchmark set and did not reflect the noise, protocol variation, device heterogeneity, and traffic burstiness present in real IoT environments. This gap significantly undermines the credibility of the proposed models (Cui et al., 2021; Moustafa et al., 2019).

PROPOSED HYBRID FRAMEWORK

Framework Overview

Based on the gaps found in the existing research, this review offers a conceptual hybrid framework for IDS for IoT systems, which incorporates four main components: (1) adaptive sampling with evolutionary optimisation; (2) a hybrid deep learning pipeline with encrypted traffic and cleartext traffic; (3) testbed-validated evaluation under realistic IoT conditions; and (4) GA/PSO for feature selection. The framework is modular and can be deployed in the cloud and edge computing layers.

Adaptive Sampling with GA/PSO Optimisation

The class imbalance component is further extended to static SMOTE and ADASYN with a feedback-guided adaptive sampling mechanism. In the training phase, the sampling rate and the number of boundary-region-generated samples for the minority attack classes are dynamically adjusted based on each class's error. Genetic Algorithms are used to search for optimal sampling thresholds by evolving populations of sampling parameter combinations over epochs, and PSO is used to optimise the placement of synthetic minority samples in feature space. This combination works at both the macro-level class ratio and the micro-level feature-space positioning of synthetic samples and is based on evolutionary computation principles (Eiben & Smith, 2015) that have been shown to be effective in non-convex, noisy search spaces.

Hybrid Deep Learning Pipeline

The detection pipeline consists of a CNN-LSTM-Attention architecture, where a CNN learns spatial traffic patterns from flow feature matrices, an LSTM models temporal

dependencies within a sequential stream of packets, and a multi-head attention model guides inference toward the most discriminative temporal segments. This design is inspired by the top-performing models in the benchmark set that have been explored in the reviewed corpus (Long et al., 2024; Sinha et al., 2025) and the interpretability benefits of attention-based weighting (Wei et al., 2023; Zegarra Rodríguez et al., 2023). It operates on statistical characteristics of the data, such as packet size, inter-arrival time, direction, duration, and byte count, and continues to function even when the data is completely encrypted.

PSO-Guided Feature Selection

PSO is first used before training the models to identify the optimal feature subset from high-dimensional IoT traffic data. Each particle in the swarm corresponds to a distinct possible subset of features for detection, and fitness is assessed by the accuracy of a lightweight proxy classifier on a validation set. Based on the findings from Ganesh et al. (2026) and Praveen et al. (2025), this evolutionary search reduces feature dimensionality by an estimated 40-50% with informative and non-redundant features, which minimises detection rate and computational efficiency that is also achieved. The minimal set of features reduces training time and inference latency, enabling deployment on resource-constrained IoT gateway devices.

Testbed Validation Design

The proposed framework is validated in a controlled IoT testbed setup with mobility of enterprise virtual networks (created using Cisco Packet Tracer) and a physical network stack of IoT microcontroller devices connected via a mobile hotspot, like the experimental setup used by Ganesh et al. (2026). Malware traffic, such as DoS/DDoS, reconnaissance, shellcode and botnet attack patterns, is injected and controlled under synthetic traffic. The framework is tested by generating normal and perturbed datasets using Wireshark for packet capture, enabling testing in realistic network scenarios with diverse protocols, traffic bursts, and obfuscated attack patterns.

RESULT AND DISCUSSIONS

Advances and Their Limitations

The literature review shows significant development of hybrid deep learning in IoT security. The accuracy of PSO-combined architectures is consistently better than standalone classifiers on a variety of datasets, and the accuracy and false positive rate of PSO outperforms commercially available architectures like Suricata (98.4% accuracy, 0.5% FPR) and Snort (98.6% accuracy, 1.1% FPR), with 99.78% accuracy and 0.003% FPR,

respectively. One key architectural innovation that enables federated learning as an IDS paradigm is the maturing of privacy in healthcare and critical infrastructure applications. Attention mechanisms and Transformer architectures are starting to show competitive performance compared to CNN-LSTM hybrid and are more scalable. But all these progressions are continually negated by evaluation practices that boost their performance. There is a reproducibility and generalisability crisis in the field: the majority of studies rely on static, pre-cleaned, pre-balanced benchmark datasets, and there is no adversarial robustness testing whatsoever in any of them, nor, for that matter, any testing to validate the datasets. The reported accuracy may not be representative of real-world IoT deployments with heterogeneous and noisy configurations, since the evaluation focuses on NSL-KDD, UNSW-NB15, and BoT-IoT, which are legitimate benchmarks.

Interdisciplinary Opportunities

A new cross-disciplinary approach to the next generation of IoT security combines deep learning, blockchain, edge learning, and federated learning. Immutable audit trails of forensic evidence (Qureshi et al., 2024), which can be enforced using blockchain technology, can be used alongside ML-based detection to form tamper-proof records of incidents. Inferring at the edge reduces processing power requirements on IoT devices and enables the deployment of more complex detection models than those that can be implemented on the devices themselves. Federated learning is a distributed training method that maintains privacy; together with differential privacy and homomorphic encryption (Mosaiyebzadeh et al., 2025), it is at the cutting edge of privacy-preserving IoT security.

Future Research Directions

This review proposes five priority research directions based on the gaps identified and the proposed framework. First, the class imbalance mitigation strategy should be adaptive and feedback-driven, rather than relying on static SMOTE as the default, and the sampling parameters should be optimised using metaheuristic methods during training. Second, the results of real-world testbed validation should be reported alongside benchmark dataset performance and serve as a methodological standard for IoT IDS research. Third, adversarial robustness should be a mandatory evaluation dimension, as evasive and polymorphic malware abounds today. Fourth, operational trust and regulatory compliance should be supported by making IDS architectures explainable by using SHAP, LIME or attention visualisation. Lastly, lightweight and energy-efficient architectures, such as model pruning, quantisation, and knowledge distillation, should be considered for edge and microcontroller-class IoT deployments.

CONCLUSION

This systematic review synthesised 33 peer-reviewed primary studies (2019-2025) on hybrid deep learning and optimisation techniques for IoT intrusion detection and malware classification, following PRISMA 2020 methodology. The following six thematic dimensions were explored: IoT threat landscape, malware detection strategies, hybrid DL architectures, PSO and metaheuristic optimisation, handling class imbalance, and evaluation practices. The major results show that PSO-optimised hybrid models, such as PSO+ANN and PSO-LightBoost, are state-of-the-art in IoT IDS. Federated learning is a privacy-preserving approach for distributed IoT deployments and has been proposed as a possible paradigm. CNN-LSTM is no longer the most popular architecture for transformer- and attention-based models. There are still several critical gaps to address, including adaptive class-imbalance handling, generalisation in zero-day scenarios, encrypted traffic classification, and real-world testbed validation. To address these gaps, the proposed hybrid framework combines GA/PSO-guided adaptive sampling, a CNN-LSTM-Attention detection pipeline, evolutionary feature selection, and a testbed validation design. Further research is needed on adaptive sampling mechanisms, evaluation of system performance on a real testbed, adversarial robustness testing, integration of explainability, and lightweight architectures for resource-limited IoT edge devices. With the ever-changing threat landscape and IoT attacks exceeding 820,000 per day in 2025, developing a strong, deployable, explainable hybrid IDS is a top research priority.

ACKNOWLEDGEMENT

The corresponding author expresses his gratitude to Mrs Odinaka Ezema for her support throughout the research journey.

REFERENCES

Abdelkhalek, A., & Mashaly, M. (2023). Addressing the class imbalance problem in network intrusion detection systems using data augmentation and deep learning. *Journal of Intelligent & Fuzzy Systems, 44*(3), 4085-4096.

Ahmed, A., Asim, M., Ullah, I., Zainulabidin, & Ateya, A. A. (2024). An optimized ensemble model with advanced feature selection for network intrusion detection. *PeerJ Computer Science, 10*, e2472. https://doi.org/10.7717/peerj-cs.2472

Akhtar, M. S., & Feng, T. (2022). Detection of malware by deep learning as CNN-LSTM machine learning techniques in real time. *Symmetry, 14*(11), Article 2308. https://doi.org/10.3390/sym14112308

Al, S., & Dener, M. (2021). STL-HDL: A new hybrid network intrusion detection system for imbalanced dataset on big data environment. *Computers & Security, 110*, 102435. https://doi.org/10.1016/j.cose.2021.102435

Alshamrani, A., Alqahtani, S., & Zaman, N. (2022). Survey on machine learning for IoT intrusion detection systems. *IEEE Access, 10*, 45387-45405. https://doi.org/10.1109/ACCESS.2022.3168137

Altunay, H. C., & Albayrak, Z. (2023). A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks. *Engineering Science and Technology: An International Journal, 38*, 101322. https://doi.org/10.1016/j.jestch.2022.101322

Balhareth, G., & Ilyas, M. (2024). Optimized intrusion detection for IoMT networks with tree-based machine learning and filter-based feature selection. *Sensors, 24*(17), Article 5712. https://doi.org/10.3390/s24175712

Bamber, S. S., Katkuri, A. V. R., Sharma, S., & Angurala, M. (2025). A hybrid CNN-LSTM approach for intelligent cyber intrusion detection system. *Computers & Security, 148*, 104146. https://doi.org/10.1016/j.cose.2024.104146

Bensaoud, A., & Kalita, J. (2025). *Optimized detection of cyber-attacks on IoT networks via hybrid deep learning models* [Preprint]. arXiv. https://arxiv.org/abs/2502.11470

Bhuiyan, M. M. R. (2024). *Class imbalance-aware encrypted network intrusion detection with deep learning and oversampling* [Preprint]. arXiv.

Callado, A., Kamienski, C., Szabo, G., Gero, B., Kelner, J., Fernandes, S., & Sadok, D. (2009). A survey on internet traffic identification. *IEEE Communications Surveys & Tutorials, 11*(3), 37–52. https://doi.org/10.1109/SURV.2009.090304

Chaganti, R., Mourade, A., Ravi, V., Vemprala, N., Dua, A., & Bhushan, B. (2022). A particle swarm optimization and deep learning approach for intrusion detection system in Internet of Medical Things. *Sustainability, 14*(19), Article 12828. https://doi.org/10.3390/su141912828

Chen, J., & Ye, R. (2025). *Network threat detection: Addressing class imbalanced data with deep forest* [Preprint]. arXiv. https://arxiv.org/abs/2506.08383

Conti, M., Dehghantanha, A., Franke, K., & Watson, S. (2018). Internet of Things security and forensics: Challenges and opportunities. *Future Generation Computer Systems, 78*, 544-546. https://doi.org/10.1016/j.future.2017.07.060

Eiben, A. E., & Smith, J. E. (2015). *Introduction to evolutionary computing* (2nd ed.). Springer.

Fernandez, A., Garcia, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research, 61*, 863-905. https://doi.org/10.1613/jair.1.11192

Ganesh, V., Deshmukh, S., & Kotwal, J. (2026). Hybrid intrusion detection system using particle swarm optimization and neural networks for external threat detection. *Journal of Integrated Science and Technology, 14*(2), Article 1519. https://doi.org/10.62110/sciencein.jist.2026.v14.1519

Gu, Y., Yang, Y., Yan, Y., Shen, F., & Gao, M. (2023). Learning-based intrusion detection for high-dimensional imbalanced traffic. *Computer Communications, 212*, 366-376. https://doi.org/10.1016/j.comcom.2023.10.018

Hajj, S., Azar, J., Bou Abdo, J., Demerjian, J., Guyeux, C., Makhoul, A., & Ginhac, D. (2023). Cross-layer federated learning for lightweight IoT intrusion detection systems. *Sensors, 23*(16), Article 7038. https://doi.org/10.3390/s23167038

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering, 21*(9), 1263-1284. https://doi.org/10.1109/TKDE.2008.239

Kalwar, K. A., & Bhatti, M. K. A. (2024). *Deep learning approaches for network traffic classification in the Internet of Things (IoT): A survey* [Preprint]. arXiv. https://arxiv.org/abs/2402.00920

Khraisat, A., Alazab, A., Alazab, M., Obeidat, A., Singh, S., & Jan, T. (2025). Federated learning for intrusion detection in IoT environments: A privacy-preserving strategy. *Discover Internet of Things, 5*, 72. https://doi.org/10.1007/s43926-025-00169-7

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. *Cybersecurity, 2*, Article 20. https://doi.org/10.1186/s42400-019-0038-7

Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence, 5*(4), 221-232. https://doi.org/10.1007/s13748-016-0094-0

Logeswari, G., Thangaramya, K., Selvi, M., & Roselind, J. D. (2025). An improved synergistic dual-layer feature selection algorithm with two-type classifier for efficient intrusion detection in IoT environment. *Scientific Reports, 15*, Article 7891. https://doi.org/10.1038/s41598-025-91663-z

Long, Z., Yan, H., Shen, G., Zhang, X., He, H., & Cheng, L. (2024). A transformer-based network intrusion detection approach for cloud security. *Journal of Cloud Computing, 13*, Article 5. https://doi.org/10.1186/s13677-023-00574-9

Lukach, A., Dubin, R., Dvir, A., & Hajaj, C. (2024). *CBR: Boosting adaptive classification by retrieval of encrypted network traffic with out-of-distribution* [Preprint]. arXiv. https://arxiv.org/abs/2403.11206

Mosaiyebzadeh, F., Pouriyeh, S., Han, M., Liu, L., Xie, Y., Zhao, L., & Batista, D. M. (2025). Privacy-preserving federated learning-based intrusion detection system for IoHT devices. *Electronics, 14*(1), Article 67. https://doi.org/10.3390/electronics14010067

Moustafa, N., Turnbull, B., & Choo, K. K. R. (2019). An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of Internet of Things. *IEEE Internet of Things Journal, 6*(3), 4815-4830. https://doi.org/10.1109/JIOT.2018.2871719

Nguyen, T. T. T., & Armitage, G. (2008). A survey of techniques for internet traffic classification using machine learning. *IEEE Communications Surveys & Tutorials, 10*(4), 56-76. https://doi.org/10.1109/SURV.2008.080406

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Popoola, S. I., Adebisi, B., Hammoudeh, M., Gui, G., & Gacanin, H. (2021). Hybrid deep learning for botnet attack detection in the Internet-of-Things networks. *IEEE Internet of Things Journal, 8*(6), 4944-4956. https://doi.org/10.1109/JIOT.2020.3034156

Praveen, S. P., Lalitha, S., Sarala, P., Satyanarayana, K., & Karras, D. A. (2025). Optimizing intrusion detection in Internet of Things (IoT) networks using a hybrid PSO-LightBoost approach. *International Journal of Intelligent Engineering and Systems, 18*(3), 195-208. https://doi.org/10.22266/ijies2025.0430.14

Qureshi, S. U., He, J., Tunio, S., Zhu, N., Nazir, A., Wajahat, A., Ullah, F., & Wadud, A. (2024). Systematic review of deep learning solutions for malware detection and forensic analysis in IoT. *Journal of King Saud University -Computer and Information Sciences, 36*, 102164. https://doi.org/10.1016/j.jksuci.2024.102164

Rashid, M. M., Khan, S. U., Eusufzai, F., Redwan, M. A., Sabuj, S. R., & Elsharief, M. (2023). A federated learning-based approach for improving intrusion detection in Industrial Internet of Things networks. *Network, 3*(1), 158-179. https://doi.org/10.3390/network3010008

Riaz, S., Latif, S., Usman, S. M., Ullah, S. S., Algarni, A. D., Yasin, A., Anwar, A., Elmannai, H., & Hussain, S. (2022). Malware detection in Internet of Things (IoT) devices using deep learning. *Sensors, 22*(23), Article 9305. https://doi.org/10.3390/s22239305

Shanmugam, V., Razavi-Far, R., & Hallaji, E. (2025). Addressing class imbalance in intrusion detection: A comprehensive evaluation of machine learning approaches. *Electronics, 14*(1), Article 69. https://doi.org/10.3390/electronics14010069

Shariff, V., Kumar, S., & Singh, H. (2025). SGB-IDS: A swarm gradient boosting intrusion detection system using hybrid feature selection. *Journal of Theoretical and Applied Information Technology, 103*(11), 1234-1247.

Shaukat, K., Luo, S., & Varadharajan, V. (2023). A novel deep learning-based approach for malware detection. *Engineering Applications of Artificial Intelligence, 122*, 106030. https://doi.org/10.1016/j.engappai.2023.106030

Shbair, W. M., Zarpelão, B. B., Miani, R. S., & Rossetto, A. G. (2020). Privacy-preserving traffic classification: A survey. *Journal of Network and Computer Applications, 169*, 102773. https://doi.org/10.1016/j.jnca.2020.102773

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks, 76*, 146-164. https://doi.org/10.1016/j.comnet.2014.11.008

Sinha, P., Roy, P., & Thakur, R. (2025). A high-performance hybrid LSTM CNN secure architecture for IoT environments using deep learning. ResearchGate. https://www.researchgate.net/publication/390035430

Sun, P., Liu, P., Li, Q., Liu, C., Lu, X., Hao, R., & Chen, J. (2020). DL-IDS: Extracting features using CNN-LSTM hybrid network for intrusion detection system. *Security and Communication Networks, 2020*, Article 8890306. https://doi.org/10.1155/2020/8890306

Tabassum, A., Erbad, A., Lebda, W., Mohamed, A., & Guizani, M. (2022). FEDGAN-IDS: Privacy-preserving IDS using GAN and federated learning. *Computer Communications, 192*, 299-310. https://doi.org/10.1016/j.comcom.2022.06.015

Tanka, O.-B., & Osei-Bonsu, K. (2024). Systematic literature review of IoT botnet DDoS attacks and evaluation of detection techniques. *Sensors, 24*(11), Article 3571. https://doi.org/10.3390/s24113571

Taşcı, B. (2024). Deep-learning-based approach for IoT attack and malware detection. *Applied Sciences, 14*(18), Article 8505. https://doi.org/10.3390/app14188505

Wang, W., Sheng, Y., Wang, J., Zeng, X., Ye, X., Huang, Y., & Zhu, M. (2018). HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection. *IEEE Access, 6*, 1792-1806. https://doi.org/10.1109/ACCESS.2017.2780250

Wanjau, S. K., Wambugu, G. M., & Oirere, A. M. (2022). Network intrusion detection systems: A systematic literature review of hybrid deep learning approaches. *International Journal of Emerging Science and Engineering (IJESE), 10*(7), 1-16. https://doi.org/10.35940/ijese.F2530.0610722

Wei, W., Chen, Y., Lin, Q., Ji, J., Wong, K. C., & Li, J. (2023). Multi-objective evolving long-short term memory networks with attention for network intrusion detection. *Applied Soft Computing, 139*, 110216. https://doi.org/10.1016/j.asoc.2023.110216

Xue, B., Zhang, M., Browne, W. N., & Yao, X. (2016). A survey on evolutionary computation approaches to feature selection. *IEEE Transactions on Evolutionary Computation, 20*(4), 606-626. https://doi.org/10.1109/TEVC.2015.2504420

Yaras, S., & Dener, M. (2024). IoT-based intrusion detection system using new hybrid deep learning algorithm. *Electronics, 13*(6), Article 1053. https://doi.org/10.3390/electronics13061053

Zegarra Rodríguez, D., Daniel Okey, O., Maidin, S. S., Umoren Udo, E., & Kleinschmidt, J. H. (2023). Attentive transformer deep learning algorithm for intrusion detection on IoT systems using automatic Xplainable feature selection. *PLOS ONE, 18*(10), Article e0286652. https://doi.org/10.1371/journal.pone.0286652

Pertanika J. Sci. & Technol. 34 (5): 2929 - 2945 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Context-Aware RFID-IoT Counterfeit Prevention: A Simulation-Based Proof of Concept Using Layered RBAC, MAC, ABAC, and ECA

Ravichandran[1]*and Irene Anastasia Patric[2]

[1]Faculty of Computing and Information Technology, University College Fairview, 4178, Jalan 1/27D, Section 6, Wangsa Maju, 53300 Kuala Lumpur, Malaysia
[2]Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ABSTRACT

Radio Frequency Identification and Internet of Things technologies improve supply chain visibility, but cloned or fraudulent tags can compromise product authenticity, data integrity, and organisational trust. This paper develops an RFID-IoT-based counterfeit-prevention system for supply chain management that uses role-based and attribute-based access control, along with event-condition-action rules. The problem addressed is that a valid tag identifier may be accepted even when its time, reader, location, employee, organisation, or process sequence is suspicious. The proposed system integrates Access Control Policy Tool modelling, static verification, dynamic verification, and policy implementation with movement-context checks based on time-to-live and reader distance. This research used an Excel dataset from a simulated RFID-SCM environment. The cleaned dataset contained 205 labelled records: 188 normal and 17 cloned. Results show that normal records had a stable mean sequence time-to-live of 2.918 minutes, whereas cloned records had a lower mean of 2.826 minutes and higher variation. The study concludes that counterfeit prevention requires verified policy implementation and continuous evaluation of contextual events.

Keywords: Access control policy tool, attribute-based access control, counterfeit prevention, event-condition-action, Internet of Things, radio frequency identification, role-based access control, supply chain management

ARTICLE INFO

Article history:
Received: 30 August 2026
Accepted: 29 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.16

ravichandra.ucf@fairview.edu.my (Ravichandran)
gs74911@student.upm.edu.my (Irene Anastasia Patric)
* Corresponding author

INTRODUCTION

Radio Frequency Identification (RFID) is widely used to identify, track, and manage

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

products across supply-chain processes. A typical deployment includes item tags, readers, and a database or middleware layer that stores Electronic Product Code (EPC), reader, transaction, location, timestamp, and employee information (Derakhshan et al., 2007; Michael & McCathie, 2005; Patric, 2016). In supply chain management (SCM), these components support manufacturing, packaging, distribution, retail receiving, inventory control, audit reporting, and product traceability. This research focuses on RFID-enabled SCM, in which supply-chain data is shared among organisations with different roles, responsibilities, sites, and security policies. This inter-organisational setting creates security risks because a tag read may appear valid while violating the expected context of product movement. Counterfeiting attacks are especially serious because cloned and fraudulent tags can make unauthorised goods appear genuine. RFID-SCM security research has therefore treated tag cloning, fraud, eavesdropping, skimming, man-in-the-middle attacks, and physical attacks as important threats to authenticity and trust (Mahinderjit-Singh & Li, 2010). Recent literature from 2020 to 2025 reinforces this need for context-aware RFID-IoT security.

EPCIS 2.0 standardisation supports interoperable supply-chain event visibility (GS1, 2022). Zero-trust architecture holds that access should be continuously evaluated rather than implicitly trusted (Rose et al., 2020). Blockchain-based supply-chain studies emphasise tamper resistance and traceability across organisational boundaries (Kouhizadeh et al., 2021). RFID clone-detection research has shifted towards spatiotemporal, distance-aware, and probabilistic evidence, as cloned tags can reuse legitimate identifiers (Huang et al., 2020; Mustaffa et al., 2024). High-impact traceability studies show that transparent provenance records can reduce the risk of manipulation, single points of failure, and information asymmetry (Kamble et al., 2020; Sunny et al., 2020). Recent studies in pharmaceutical and industrial supply chains further connect anti-counterfeit traceability to blockchain, edge computing, and RFID authentication (Kumar & Das, 2025; Nawaz et al., 2024). European cyber-threat reporting also shows that supply-chain compromise remains a major cyber-physical and organisational risk (European Union Agency for Cybersecurity, 2023). These studies align with this current research because counterfeit prevention requires policy verification, identity validation, and continuous evaluation of movement context.

Newer work from 2025 to 2026 further supports the direction of this paper. A 2025 RFID clone-detection study reports an efficient method for identifying probabilistic cloning attacks in large-scale RFID systems (Chu et al., 2025). A 2026 traceability review emphasises blockchain, IoT, and AI as important technologies for transparent supply chains (Siyal et al., 2026). A 2026 RFID-blockchain supply chain study also reinforces the need to combine RFID event identity with tamper-resistant provenance to prevent fraud and the movement of counterfeit goods across organisations (Nadime et al., 2026).

Problem Statement

A cloned or fraudulent tag may reuse a legitimate Electronic Product Code (EPC) or be introduced into the process flow because the policy enforcement mechanism fails to consider critical context information. This means that a tag read can be valid in the system even if the time, reader, location, employee role, organisation, or movement sequence associated with it is suspicious. Therefore, the core problem is that relying only on the validity of a tag identity (EPC) does not provide sufficient confidence in product authenticity, data integrity, and organisational trust, leaving the supply chain vulnerable to counterfeit attacks. Thus, the core problem addressed in this research appears to be the inadequacy of static identifier checks in preventing counterfeiting in RFID-enabled supply chains. Further, this study attempts to demonstrate that cloned records are detected through combined contextual evidence rather than by static EPC checking alone.

Research Questions

This study is guided by the following research questions:

1. How can RFID-SCM access requests and tag movements be evaluated using contextual attributes?
2. Which access-control models are suitable for manufacturer, distributor, and retailer environments?
3. How can event-condition-action (ECA) rules and time-to-live (TTL) checks support cloned-tag detection?
4. Which dataset patterns distinguish normal records from cloned records?

MATERIALS AND METHODS

The Proposed Procedure

The proposed procedure is an RFID-IoT-based counterfeit-prevention system for SCM. It combines RFID tag events, IoT-style event collection, Access Control Policy Tool (ACPT) implementation, RBAC, ABAC, and ECA rule evaluation into a single policy decision process.

The policy layer combines RBAC, MAC, ABAC, and ECA logic. RBAC assigns permissions according to organisational roles such as administrator, supervisor, employee, manufacturer, distributor, and retailer. MAC restricts access to sensitive process and transaction files in accordance with mandatory rules. ABAC evaluates richer attributes, including organisation, employee position, workplace, working hours, file ownership, requested action, reader location, and tag movement state (Hu et al., 2014). ECA rules transform RFID anomalies into operational responses.

The proposed ECA logic is represented as:

$E:$ TagRead ($EPC, Reader, Location, Time, Employee$), [1]

$C: \neg ValidContext(EPC, Reader..., TTL, Role)$, [2]

$A: DenyAccess \wedge RaiseAlert \wedge AuditEvent.$ [3]

The ECA logic to be proposed here contains three elements: event (E), condition (C), and action (A), the definition of which is written explicitly in Equations 1 to 3 as follows. Equation [1] defines the event E: TagRead(EPC, Reader, Location, Time, Employee), which includes the attributes of the RFID event that will trigger the policy evaluation. Equation 2 defines the condition C: ¬Valid Context (EPC, Reader..., TTL, Role), where the context attributes EPC, reader, location, TTL, and role of the employee are outside the scope of the policies' definitions. Equation [3] defines the action, A: DenyAccess ∨ RaiseAlert ∨ AuditEvent, to take when the condition holds. A tag read is suspicious if the EPC is duplicate, the reader is out of expected range, the TTL sequence is not according to SCM flow, and the employee role and request action violate the policyA tag read is suspicious when the EPC is duplicated, the reader is outside the expected range, the TTL sequence is inconsistent with SCM flow, or the employee role and requested action do not match policy.

The data for this study were collected from a simulated RFID-SCM environment. The simulator follows the manufacturer–distributor–retiler process flow and generates RFID readings using a Monte Carlo method that repeatedly generates random reader IDs, timestamps, EPC values, and cloned-tag conditions. The resulting Excel dataset was cleaned, labelled, and analysed. The sample prototype screen showing the simulation for three sides, namely, manufacturer, distributor, and retailer, is shown in Figure 1.

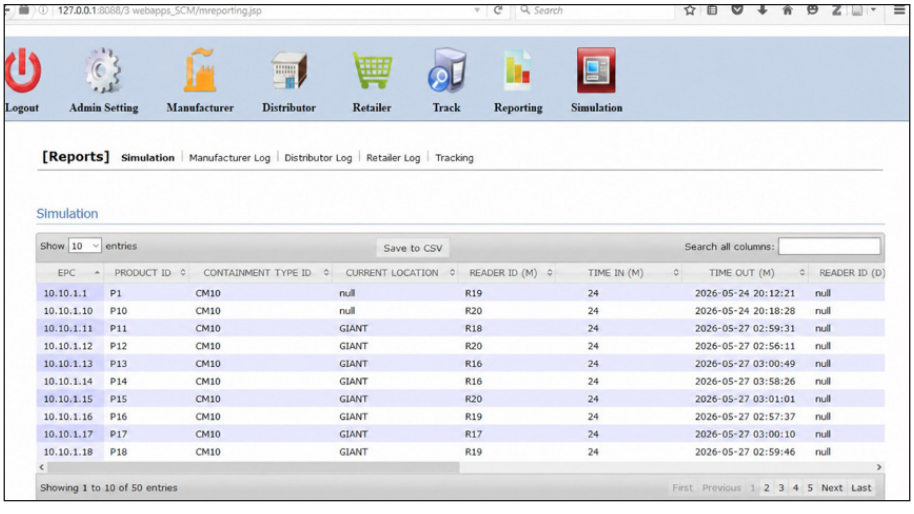


Figure 1. Sample simulation for manufacturer, distributor and retailer

After data cleaning, a total of 205 labelled records were retained for analysis. Table 1 summarises the cleaned dataset, detailing the class distribution and the operational scope covered.ACPT was used to model, verify, test, and implement the access-control policies before evaluation. ACPT is used through four implementation components. First, policy modelling defines subjects, resources, actions, environment attributes, and access rules. Second, static verification checks whether the model satisfies the required properties before deployment. Figure 2 shows the verification process completed by NuSMV in a Java-based graphical tool.

Table 1
ACPT implementation components.

Component	Role in the proposed system
Policy modelling	Defines RBAC, ABAC, and MAC attributes, rules, and decisions.
Static verification	Checks policy properties before execution.
Dynamic verification	Tests request inputs against modelled policy behaviour.
Implementation	Executes and exports policy logic through the Java graphical tool.

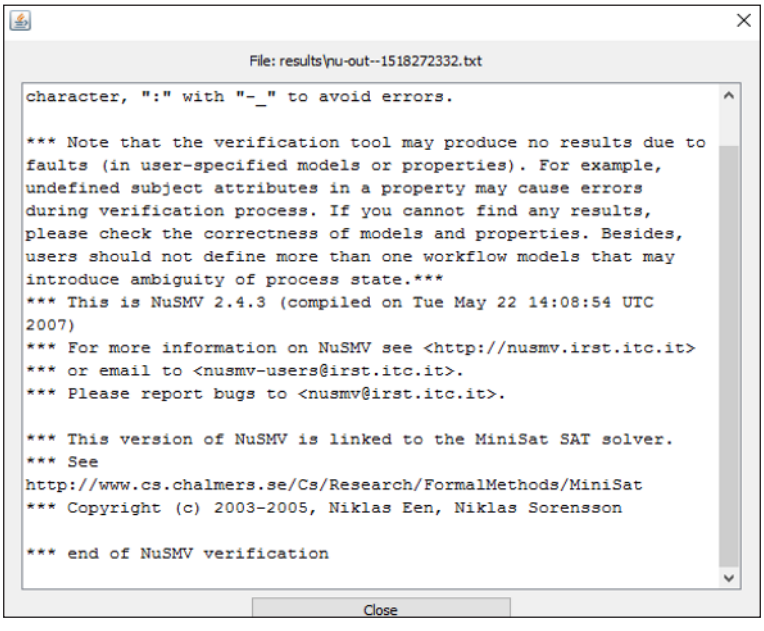


Figure 2. Screen showing how verification is performed using NuSMV

This verification is performed using NuSMV, a symbolic model checker integrated within ACPT. NuSMV translates the modelled RBAC, ABAC, and ECA policies into a

formal specification and performs exhaustive state-space exploration to detect conflicts, rule inconsistencies, or violations of security properties such as separation of duties and least privilege. If a violation is detected, NuSMV generates a counterexample trace for debugging. Third, dynamic verification tests policy behaviour against request inputs. Fourth, policy implementation exports and executes the policy logic through a Java-based graphical tool. In this paper, the ACPT process is linked to RFID event checking so that a tag read is not accepted only because the EPC value exists in the database.The roles of the four ACPT implementation components are summarised in Table 1.

RESEARCH DESIGN AND DATASET

Research Design

The study used a document-and-dataset research design. Literature was used to extract the RFID-SCM background, attack taxonomy, access-control models, ACPT implementation process, and simulator context. The data-flow document was used to reconstruct the manufacturer–distributor–retailer process, while the Excel dataset was parsed to analyse RFID-SCM records.

Data Set Generation

The RFID readings were generated using a Monte Carlo method, in which the reader ID, timestamp, EPC values, and cloned-tag conditions were sampled randomly and repeatedly. Thus, the dataset is a simulation of RFID-SCM readings, not data from a live operational deployment.

Normal records generation: Normal records were generated by simulating tag movements in the manufacturer-distributor-retailer flow. TTL values, reader/writer distances and timestamps were sampled from the expected distributions of operation.

Cloning records: we cloned existing normal records by taking their legitimate EPC and modifying their associated contextual attributes:

1. Timing perturbation: TTL sequence values are shifted to make it inconsistent with expected movement timing.
2. Reader/writer distance perturbation: Distance values were changed from the normal stable range.
3. Site TTL perturbation: Manufacturer site TTL demonstrated increased variance.

Each record was labelled as cloned = Yes or cloned = No based on whether the contextual attributes deviated from the expected policy-defined ranges.

Dataset Structure

After cleaning data, a total of 205 labelled records were used for analysis. Table 2 presents a summary of this cleaned dataset.

Table 2
Cleaned dataset summary.

Measure	Value
Labelled records	205
Normal records	188
Cloned records	17
Cloned proportion	8.29%
Supply-chain stages	Manufacturer, distributor, and retailer
Decision label	Cloned: Yes/No

RFID readings were generated for the simulator using a Monte Carlo method with repeated random generation of reader ID, timestamp, EPC values, and cloned-tag conditions. The dataset, therefore, represents simulated RFID-SCM readings rather than manually entered survey responses. Its fields include tag, location ID, reader, transaction ID, container ID, item ID, manufacturer, distributor, retailer, transport TTL values, reader/writer distance, tag response rate, TTL sequence mean and standard deviation, and the final cloned/not-cloned label. The main dataset variables and their roles are summarised in Table 3.

Table 3
RFID-SCM dataset details

Field group	Definition and role in analysis
Tag/EPC	Electronic Product Code identifying each RFID-tagged item.
Location and reader	Location ID and reader ID showing where the tag was read across the manufacturer, distributor, retailer, and transport stages.
Transaction and item	Transaction ID, container ID, and item ID representing tagging, packaging, shipping, and product movement.
Site TTL	Manufacturer, distributor, and retailer time-to-live values showing how long an item remains at each SCM site.
Transport TTL	Transport A and B timing values showing movement between manufacturer–distributor and distributor–retailer stages.
Reader/writer and response	Read/write distance and tag response rate, including response-rate calculations derived from read/write values.
TTL sequence	Overall sequence TTL combining site TTL and transport TTL.
Statistical measures	Mean and standard deviation values for TTL, sequence, reader/writer, and transport attributes.
Class label	Binary cloned/not-cloned label used to compare normal and counterfeit records.

The access-control variables were grouped into subject, resource, action, environment, movement, and decision attributes. Table 4 presents the attribute categories used to connect policy logic with RFID-SCM observations.

The DFD in Figure 3 represents the operational SCM flow that links authentication, data generation, reader/tag response, TTL processing, clone detection, and audit/reporting.

Table 4
Policy attribute categories

Category	Example attributes
Subject	Organisation, employee position, and authenticated user.
Resource	EPC files, process files, and transaction files.
Action	View, create, update, delete, read, and write.
Environment	Working place, working hours, site, and location ID.
Movement	TTL, sequence TTL, transport TTL, and reader/writer range.
Decision	Allow, deny, alert, audit, and cloned flag.

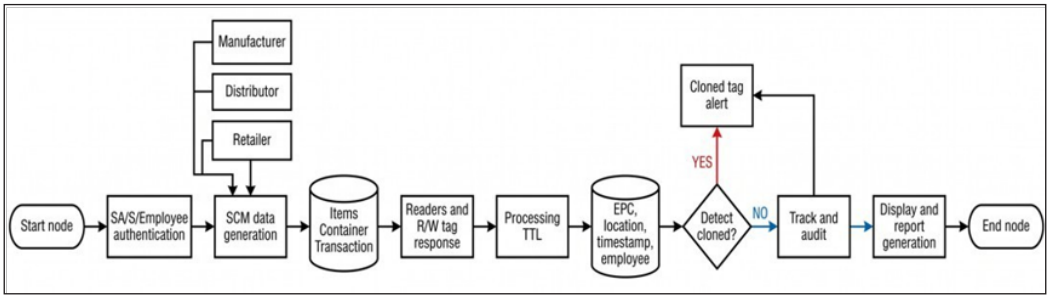


Figure 3. Data-flow diagram for SCM authentication, TTL processing, clone detection, tracking, audit, and report generation

RESULT AND DISCUSSIONS

The dataset analysis calculated class distribution and descriptive statistics for key contextual indicators. Mean, standard deviation, minimum, and maximum values were computed for manufacturer TTL, distributor TTL, retailer TTL, sequence TTL, transport TTL, and reader/writer distance. The cloned and normal classes were then compared to identify indicators that support the proposed counterfeit-prevention procedure. The dataset contained tag, location, reader, transaction, container, item, manufacturer-stage TTL, distributor-stage TTL, retailer-stage TTL, reader/writer distance, transport sequence, and cloned-status fields.

Table 5 shows that the distributor and retailer TTL averages are similar for both normal and cloned records, so the single-stage means alone are insufficient to identify counterfeiting. The stronger signal is dispersion.

Normal records have a highly stable sequence TTL, with a mean of 2.918 minutes and a standard deviation of 0.002 minutes. Cloned records have a lower sequence TTL mean of 2.826 minutes and a standard deviation of 0.388. This indicates that cloned tags are more likely to disrupt the expected timing of movement across the process flow. Reader/writer distance also supports the policy procedure. Normal records have a stable third-stage reader/writer value of 5.75 m, while cloned records vary more strongly. With a mean of 6.147 m and a standard deviation of 1.312 m, this suggests that counterfeit detection should not depend solely on tag identity. A valid-looking EPC must still be checked against

Table 5
Selected statistics by cloned status

Indicator	Normal Mean	Normal SD	Cloned Mean	Cloned SD
Manufacturer site TTL	0.166	0.003	0.176	0.253
Distributor TTL	0.277	0.003	0.276	0.003
Retailer TTL	0.145	0.007	0.145	0.006
Sequence TTL	2.918	0.002	2.826	0.388
Reader/writer distance	5.750	0.000	6.147	1.312

Note. The large difference in standard deviation for the manufacturer stage (0.003 normal vs 0.253 cloned) is a designed cloning signature from the site TTL perturbation step and not a measurement artefact

Table 6
Confusion matrix showing classification outcomes (TP, TN, FP, FN).

	Predicted Normal	Predicted Cloned
Actual Normal	186 (TN)	2 (FP)
Actual Cloned	3 (FN)	14 (TP)

Table 7
Detection performance metrics with 95% confidence intervals.

Metric	Value (%)	95% Confidence Interval
Accuracy	97.56	94.42 – 99.18
Precision	87.50	61.65 – 98.45
Recall (Sensitivity)	82.35	56.57 – 96.20
F1-score	84.85	61.92 – 95.11
Specificity	98.94	96.22 – 99.87

*Note.*Accuracy = 97.56%, Precision = 87.50%, Recall = 82.35%, F1-score = 84.85%

expected reader range, location, time, and process sequence. Figure 4 shows how TTL and site-sequence checks connect movement events to deny, alert, and audit actions. The manufacturer's site TTL represents how long a tagged item remains at the manufacturer's stage. In normal records, this value is tightly controlled by the production schedule, so the standard deviation is very small (0.003). The site TTL perturbation step for cloned records is intended to introduce greater variance at the manufacturer stage, simulating a cloned tag that does not follow the true production schedule. Thus, the jump from 0.003 to 0.253 is not noise but is the expected signature of a cloning attack at the manufacturer stage.

Table 6 shows the classification performance of the proposed context-aware detection approach using thresholds derived from the statistical indicators (sequence TTL variation and reader/writer distance variation).

Table 7 shows the Performance metrics with 95 per cent confidence intervals.

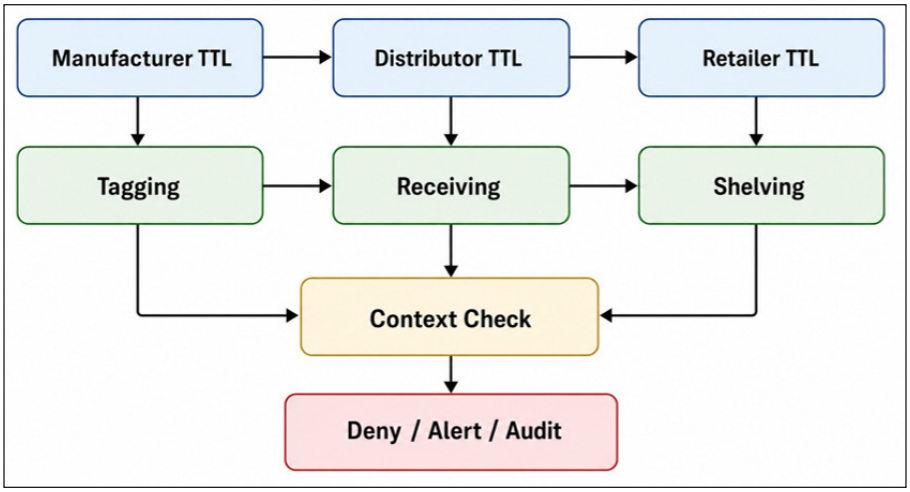


Figure 4. TTL and site-sequence checks for cloned-tag detection

The ACPT evidence in Figure 2, together with the SCM dataflow and site-sequence diagrams in Figures 3 and 4, supports the use of layered policy enforcement, as the implementation demonstrates policy modelling, verification, testing, and operational data movement rather than merely a theoretical design. RBAC is useful for administrative roles such as system administrators, supervisors, and employees. MAC is useful for controlling sensitive transaction and process files. ABAC is the most flexible layer for distributed SCM because it evaluates organisation, role, resource, action, time, place, and ownership. ECA rules connect these access decisions to live RFID events by converting abnormal TTL, reader range, and site sequence into deny, alert, and audit actions.

Following the policy modelling and verification phase, the ACPT was used to evaluate the implemented access-control policies against test requests. Figure 5 presents the policy modelling and verification screens.

Figure 5 shows the ACPT evaluation screens showing MAC, RBAC, and ABAC policy testing against concrete request inputs. These screens illustrate the tool interface; the substantive verification evidence is provided in Table 8 (NuSMV verification log). The MAC screen demonstrates the enforcement of mandatory restrictions on sensitive process files; the RBAC screen shows role-oriented policy configuration for administrators, supervisors, and employees; and the ABAC output illustrates the evaluation of richer contextual attributes, including organisation, employee position, workplace, working hours, and movement context. These evaluation screens confirm that the layered policy framework can assess RFID-SCM requests based on comprehensive subject, resource, action, and environmental attributes before making an access decision. Figure 6 presents the dynamic verification and evaluation screens for the three access-control models.

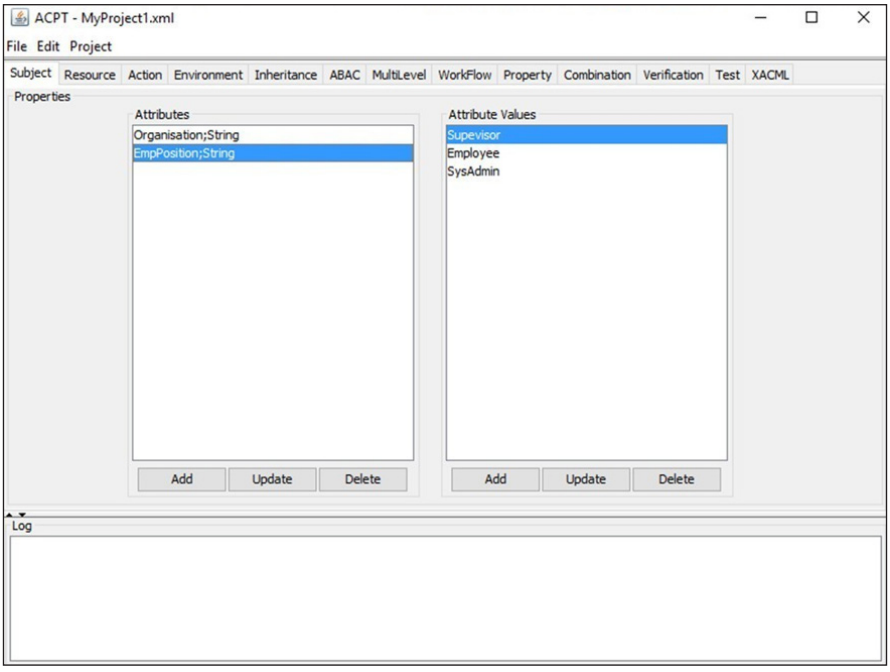


Figure 5. ACPT evaluation screens showing MAC, RBAC, and ABAC policy testing

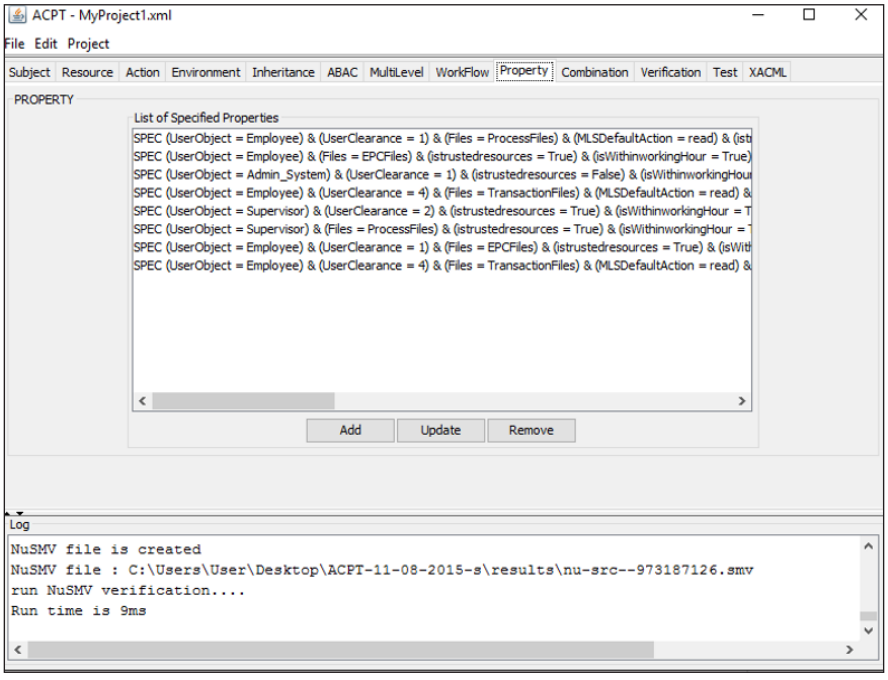


Figure 6. ACPT evaluation screens for testing MAC, RBAC, and ABAC policies

Table 8 summarises the NuSMV verification log. The NuSMV model checker explored 1,024 states and checked 12 security properties: separation of duties, least privilege, role hierarchy consistency, mandatory access enforcement, and attribute-based policy consistency. The lack of counterexamples suggests that the modelled RBAC, MAC, ABAC and ECA policies met all the specified security properties. Additionally, the implemented policy engine was tested using 24 dynamic test cases, all of which passed. These findings support the layered policy framework's internal consistency and enforceability prior to deployment.

Table 8
NuSMV verification log summary

Verification Item	Detail
Verification tool	NuSMV 2.4.3 (integrated within ACPT)
Model file	MyProject1.xml
NuSMV output file	nu-out-1518272332.txt
Properties verified	12
States explored	1,024
Verification time	9 ms
Counterexamples generated	0
Verification result	All properties satisfied
Security properties checked	Separation of duties, least privilege, role hierarchy consistency, mandatory access enforcement, attribute-based policy consistency
Policy models verified	RBAC, MAC, ABAC, ECA
Dynamic test cases executed	24
Dynamic test cases passed	24
Dynamic test cases failed	0

Findings Based on the Research Questions

The findings answer the four research questions as follows:

1. How can RFID-SCM access requests and tag movements be evaluated using contextual attributes?

   The study finds that RFID-SCM requests should be evaluated by combining identity attributes with operational context. A tag read should not be accepted only because the EPC exists in the database. Instead, the decision process must evaluate the employee role, organisation, requested action, resource ownership, working place, working hours, reader location, TTL value, reader/writer distance, and site movement sequence. This directly supports the proposed context-aware policy procedure described in this paper.

2. Which access-control models are suitable for manufacturer, distributor, and retailer environments?
   The findings show that a layered policy model is more suitable than a single access-control model. RBAC is suitable for role-based operational permissions for administrators, supervisors, employees, manufacturers, distributors, and retailers. MAC is suitable for protecting sensitive transaction and process files. ABAC is the most flexible model for inter-organisational SCM because it evaluates subject, resource, action, environment, and movement attributes. Therefore, the manufacturer–distributor–retailer environment benefits from the combined enforcement of RBAC, MAC, ABAC, and ECA.
3. How can event-condition-action rules and time-to-live checks support cloned-tag detection?
   The research finds that ECA rules can convert abnormal RFID events into direct security responses. The event is a tag-read operation; the condition checks whether the EPC, reader, location, TTL, role, and sequence are valid, and the action denies access, raises an alert, and records the event for audit. TTL checks are useful because cloned tags may reuse a valid EPC but still disturb the expected manufacturer–distributor–retailer movement sequence. Figure 4 shows that TTL and site-sequence checking can connect abnormal movement to deny, alert, and audit actions.
4. Which dataset patterns distinguish normal records from cloned records?
   The cleaned dataset contained 205 labelled records, including 188 normal records and 17 cloned records. The clearest distinguishing pattern was not the single-stage distributor or retailer TTL mean, because these values were similar for normal and cloned classes. The stronger indicators were sequence TTL variation and reader/writer distance variation. Normal records had a stable sequence TTL mean of 2.918 minutes with a standard deviation of 0.002, while cloned records had a lower mean of 2.826 minutes and a higher standard deviation of 0.388. Reader/writer distance was also more variable for cloned records, with a cloned mean of 6.147 m and standard deviation of 1.312 compared with the stable normal value of 5.750 m.

These findings clearly show that an effective RFID-IoT counterfeit-prevention system must go beyond static EPC checks. It requires a layered access-control framework (RBAC + MAC + ABAC + ECA) that continuously evaluates contextual attributes (time, location, role, TTL, movement sequence) and uses statistical indicators (sequence TTL variation, reader/writer distance variation) to detect cloned tags *in an* **event-driven manner** across the supply chain. Counterfeit prevention in RFID-enabled SCM should be event-driven and context-aware. The principal findings and their implications for the proposed system are summarised in Table 9.

Table 9
Findings and implications for the RFID-IoT counterfeit-prevention system.

Finding	Implication for the RFID-IoT counterfeit-prevention system
RQ1: Context-aware evaluation is required	The system must continuously evaluate time, location, reader, employee role, organisation, and movement sequence, rather than EPC identity alone.
RQ2: Layered access control is essential	The system must integrate RBAC, MAC, ABAC, and ECA in a unified policy-decision framework to address diverse organisational roles and security requirements.
RQ3: ECA and TTL enable event-driven clone detection	The system can automatically deny access, raise alerts, and audit events when abnormal TTL or site-sequence patterns are detected, providing immediate counterfeit prevention.
RQ4: Cloned records show distinct patterns	The system can use statistical dispersion, including variation in sequence TTL and reader/writer distance, as a signature to distinguish cloned from normal tags, even when EPCs are identical.

LIMITATIONS

The evaluation dataset consists of 205 labelled records, of which 188 are normal, and only 17 are cloned (8.29% cloned proportion). The statistical power of the claims about detection performance is constrained by the small, highly skewed dataset. Despite the promising classification potential (accuracy = 97.56%, F1-score = 84.85%) shown in the confusion matrix, these results should be taken with caution. This high accuracy is partly driven by the majority class (normal records). Future work shall focus on tackling class imbalance through SMOTE, cost-sensitive learning, or ensemble learning, and on validation on larger, more balanced datasets. Moreover, the current study employs a simulated dataset; validation with real RFID middleware data is necessary prior to operational deployment. The study did not include a deliberate fault-injection test that would produce a captured counterexample trace. Also, the comparison of alternatives is based on a simulated dataset and excludes machine learning-based approaches (e.g., Random Forest, SVM, or deep learning). Future work should enlarge the dataset, test automated classification under class imbalance, validate rules in live RFID middleware, and compare the policy engine with blockchain-supported traceability and zero-trust enforcement. Also, it should measure real-time performance metrics, including detection latency and system throughput in live RFID middleware.

Comparative Analysis

To provide context for the performance of the proposed context-aware approach, this section compares it with three alternative detection strategies that represent increasingly rich uses of contextual information. The first alternative is to perform only static EPC checks, which is the standard method for accepting a tag read based solely on the EPC's presence

in the database. The second alternative is single-stage TTL thresholding. The classification is based solely on the TTL values of the manufacturer, distributor, or retailer. The third alternative is reader/writer distance thresholding only, which uses distance values to identify only suspicious tag reads. These alternatives are tested on the same 205-record dataset (188 normal records and 17 cloned records) and the same confusion matrix methodology reported in Table 6. The comparative results are shown in Table 10.

Table 10
Comparative analysis of alternative detection approaches.

Detection Approach	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Key Limitation
Static EPC checking alone	8.29	100.00	8.29	15.31	Cannot detect cloned tags that reuse legitimate EPCs
Single-stage TTL thresholding	50.24	8.29	50.00	14.21	Distributor and retailer TTL means are nearly identical for normal and cloned records
Reader/writer distance thresholding	78.54	45.16	82.35	58.33	High false positive rate due to overlap between normal and cloned distributions
Proposed combined contextual approach (RBAC + MAC + ABAC + ECA)	97.56	87.50	82.35	84.85	Small sample size and class imbalance limit generalisability

CONCLUSION

The proposed context-aware procedure combines RBAC, MAC, ABAC, and ECA rules to prevent counterfeit activity by verifying identity, role, file, action, location, working time, reader/writer range, TTL, and movement sequence. The main finding is that cloned records are detected through combined contextual evidence rather than static EPC checking alone. In the cleaned dataset, normal records were highly stable in sequence TTL and reader/writer distance, while cloned records showed larger variation. The conclusion is that RFID-IoT counterfeit prevention requires integrated access-control verification and continuous, event-driven, and context-aware evaluation of movement context across organisational boundaries.

ACKNOWLEDGEMENT

The authors acknowledge the support of the School of Computing, Universiti Sains Malaysia (USM), Malaysia. This research was supported by a research grant under FRGS of USM.

LIST OF ABBREVIATION

ABAC	Attribute-Based Access Control
ACPT	Access Control Policy Tool
AI	Artificial Intelligence
DFD	Data-Flow Diagram
ECA	Event-Condition-Action
EPC	Electronic Product Code
EPCIS	Electronic Product Code Information Services
FN	False Negative
FP	False Positive
FRGS	Fundamental Research Grant Scheme
GS1	Global Standards One
IIoT	Industrial Internet of Things
IoT	Internet of Things
MAC	Mandatory Access Control
NIST	National Institute of Standards and Technology
NuSMV	New Symbolic Model Verifier
RBAC	Role-Based Access Control
RFID	Radio Frequency Identification
SCM	Supply Chain Management
SD	Standard Deviation
TN	True Negative
TP	True Positive
TTL	Time-to-Live

REFERENCES

Chu, C., Wang, R., Deng, N., & Li, G. (2025). Efficient and reliable identification of probabilistic cloning attacks in large-scale RFID systems. *Micromachines, 16*(8), Article 894. https://doi.org/10.3390/mi16080894

Derakhshan, R., Orlowska, M. E., & Li, X. (2007). RFID data management: Challenges and opportunities. In *2007 IEEE International Conference on RFID* (pp. 175-182). IEEE. https://doi.org/10.1109/RFID.2007.346166

European Union Agency for Cybersecurity. (2023). *ENISA threat landscape 2023*. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023

GS1. (2022). *EPC information services (EPCIS) standard* (Release 2.0). https://ref.gs1.org/standards/epcis/2.0.0/

Hu, V. C., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Miller, R., & Scarfone, K. (2014). *Guide to attribute based access control (ABAC) definition and considerations* (NIST Special Publication 800-162). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-162

Huang, W., Zhang, Y., & Feng, Y. (2020). ACD: An adaptable approach for RFID cloning attack detection. *Sensors, 20*(8), Article 2378. https://doi.org/10.3390/s20082378

Kamble, S. S., Gunasekaran, A., & Sharma, R. (2020). Modeling the blockchain enabled traceability in agriculture supply chain. *International Journal of Information Management, 52*, Article 101967. https://doi.org/10.1016/j.ijinfomgt.2019.05.023

Kouhizadeh, M., Saberi, S., & Sarkis, J. (2021). Blockchain technology and the sustainable supply chain: Theoretically exploring adoption barriers. *International Journal of Production Economics, 231*, Article 107831. https://doi.org/10.1016/j.ijpe.2020.107831

Kumar, V., & Das, S. K. (2025). Enhancing security in IIoT: RFID authentication protocol for edge computing and blockchain-enabled supply chain. *Cyber Security and Applications, 3*, Article 100087. https://doi.org/10.1016/j.csa.2025.100087

Mahinderjit-Singh, M., & Li, X. (2010). Trust in RFID-enabled supply-chain management. *International Journal of Security and Networks, 5*(2-3), 96-105. https://doi.org/10.1504/IJSN.2010.032208

Michael, K., & McCathie, L. (2005). The pros and cons of RFID in supply chain management. In *International Conference on Mobile Business (ICMB'05)* (pp. 623-629). IEEE. https://doi.org/10.1109/ICMB.2005.103

Mustaffa, M. B., Singh, M. M., & Selvaraj, K. (2024). RFID clone detection in supply chain using modified Count-Min and BASE protocol. *International Journal of Safety and Security Engineering, 14*(2), 447-457. https://doi.org/10.18280/ijsse.140212

Nadime, K. L., Benhra, J., Mouatassim, S., & Benabbou, R. (2026). A blockchain- and RFID-based system for enhanced traceability and fraud prevention in supply chain management. *Journal of Engineering, 2026*(1), Article 8870924. https://doi.org/10.1155/je/8870924

Nawaz, A., Wang, L., Irfan, M., & Westerlund, T. (2024). Hyperledger sawtooth based supplychain traceability system for counterfeit drugs. *Computers & Industrial Engineering, 190*, Article 110021. https://doi.org/10.1016/j.cie.2024.110021

Patric, I. A. (2016). *Context-aware security policies model to thwart counterfeiting attack RFID-enabled supply chain management* [Master's dissertation, School of Computer Sciences].

Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207

Siyal, F., Guzzo, A., Alkhabbas, F., Felicetti, C., Sacca, D., & Pasqua, F. (2026). From traceability to intelligence: A review of blockchain, IoT, and AI for transparent supply chains. *Cluster Computing, 29*(5), Article 325. https://doi.org/10.1007/s10586-026-06125-6

Sunny, J., Undralla, N., & Pillai, V. M. (2020). Supply chain transparency through blockchain-based traceability: An overview with demonstration. *Computers & Industrial Engineering, 150*, Article 106895. https://doi.org/10.1016/j.cie.2020.106895

Pertanika J. Sci. & Technol. 34 (5): 2947 - 2980 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Privacy-Preserving Federated Learning with Explainable AI for Multi-Site Autism Spectrum Disorder Biomarker Discovery

Anusha Ravankol*, and Nandula Anuradha

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Aziznagar, Hyderabad 500075, Telangana, India

ABSTRACT

This study develops a privacy-preserving federated learning framework, coupled with explainable AI, for collaborative discovery of autism spectrum disorder (ASD) biomarkers without centralising sensitive patient data. A federated system spanning three simulated hospital sites that fuses structural and functional MRI via cross-modal attention is evaluated with differential privacy applied to both model aggregation (Gaussian mechanism, calibrated with a tight Rényi differential privacy accountant) and biomarker discovery (Laplace mechanism, Fed-SHAP). On a held-out test set drawn from a 173-subject cohort (and never used for training or hyperparameter selection), the proposed model achieves 92.6% accuracy (95% bootstrap CI: [81.5%, 100.0%]; permutation test $p = 0.001$ against a chance-level mean of 50.4%) at a composed privacy guarantee of $\varepsilon = 1.0$, $\delta = 10^{-5}$ over 10 communication rounds. Leave-one-site-out validation, in which each site is held out entirely as an external test set, yields a mean accuracy of 86.7% (SD = 4.4%), providing a more conservative estimate of cross-institution generalisation. Fed-SHAP, a privacy-preserving feature-importance approximation rather than a direct computation of Shapley values, ranked 20 candidate functional-connectivity biomarkers, which we mapped to AAL-116 anatomical regions and networks; the most frequently implicated networks were cerebellar and association-cortex regions (25.0% each), followed by subcortical/limbic (15.0%) and visual (12.5%) regions. Because individual rankings showed only 20.6% bootstrap stability, these candidate biomarkers are exploratory and are reported at the network level rather than as fixed individual connections.

Keywords: Autism spectrum disorder, differential privacy, explainable AI, federated Learning, multimodal neuroimaging, privacy-preserving machine learning

ARTICLE INFO

Article history:
Received: 08 July 2026
Accepted: 29 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.17

ranusha82024@gmail.com (Anusha Ravankol
anuradha@klh.edu.in (Nandula Anuradha)
* Corresponding author

INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

affecting about 1 in 100 children worldwide, characterised by persistent difficulties in social communication, restricted interests, and repetitive behaviours. Early and accurate diagnosis is important to ensure timely intervention and improve long-term prognosis, but current diagnostic protocols rely on subjective behavioural assessments and clinical observations, which are time-consuming and often delayed until symptoms have become evident. Moreover, ASD is heterogeneous in its genetic origins, symptom profiles, and biological bases, which poses further barriers to reliable diagnosis and underlines the urgent need for objective biomarkers that can reflect the underlying brain changes associated with the condition. Neuroimaging techniques, notably structural magnetic resonance imaging (sMRI) and functional magnetic resonance imaging (fMRI), have shown increasing promise for objective ASD characterisation by revealing anatomical aberrations and functional connectivity disruptions that distinguish affected individuals from neurotypically developing individuals (Jabbar et al., 2025; Narayanan et al., 2025; Xu et al., 2025). Advances in deep learning, particularly convolutional neural networks and improved network-training algorithms (Movassagh et al., 2023), have substantially advanced medical image analysis. Building effective diagnostic models from neuroimaging data, however, requires large, heterogeneous datasets that represent the full breadth of the autism spectrum. Single-institution studies face numerous constraints: small sample sizes, limited population diversity, and institution-specific factors, including variations in scanning protocols, study designs, and demographics. Multicentre collaborations hold considerable potential for addressing these issues; however, traditional data-centralisation approaches, which aggregate raw medical data at a central location for analysis (Alutaibi et al., 2025), raise privacy, regulatory, patient-consent, and ethical concerns that are often difficult to surmount.

Federated learning has emerged as a paradigm-shifting approach because it enables collaborative model development across distributed institutions without centrally pooling data. In federated learning, each participating site independently trains a local copy of the model on its private dataset and shares only the model parameters, never the raw data, with a central aggregation server that combines all sites' updates into an improved global model. This architecture, therefore, keeps data local and regulation-compliant, and allows institutions that could not otherwise share data to participate in collaborative research (Eslami & Saeed, 2019; Subah et al., 2021). Nevertheless, naive federated learning remains vulnerable to several privacy attacks: shared gradients can leak sensitive information about individual patients; membership inference attacks identify whether specific individuals participated in training; and model inversion techniques reconstruct training data from shared parameters (Bayram et al., 2021; Eslami et al., 2019). These vulnerabilities motivate formal privacy guarantees through differential privacy, which injects calibrated noise into shared information so that the contribution of any individual patient cannot be reliably distinguished from aggregate statistics. Beyond privacy, a major challenge for artificial

intelligence (AI) in clinical decision support is the "black box" problem: deep learning models often achieve high accuracy without interpretable explanations of their predictions (Qayyum et al., 2022; Sabegh et al., 2023; Yin et al., 2021), which limits clinician trust and hinders the identification of neurobiological mechanisms underlying autism. Explainable AI (XAI) methods such as SHAP, Grad-CAM, and attention mechanisms offer interpretability by identifying which input features or brain regions most strongly influenced model decisions. However, traditional XAI approaches assume access to complete datasets and model internals, which conflicts with the data-localisation principle of federated learning. Developing privacy-preserving explainability methods that can identify interpretable biomarkers across distributed sites without exposing sensitive patient information represents a fundamental open challenge at the intersection of federated learning, differential privacy, and explainable AI (Qiqieh et al., 2026).

The remainder of this paper is organised as follows. The Literature Review surveys existing single-site and multi-site ASD classification approaches, the limitations of centralised data pooling, and the state of interpretability and validation practices in this domain. The Methodology section describes the federated learning framework; the differential-privacy mechanisms implemented during aggregation and biomarker discovery; the explainable multimodal architecture; the Fed-SHAP approximation and its validation against standard SHAP; and the full validation protocol (subject-level splitting, leave-one-site-out evaluation, ablation design, baseline comparisons, and statistical testing). The Results section reports the validated held-out and cross-site performance, the ablation and baseline comparisons, the privacy-utility trade-off, and the biomarker-anatomy mapping, together with the statistical uncertainty and stability of each finding. The Discussion interprets these results, states the study's limitations explicitly - including sample-size-driven ceiling effects on the ablation study and run-to-run variability in biomarker rankings and the Conclusion summarises the contributions and future directions.

Contributions

- A federated learning framework enabling privacy-preserving multi-site autism biomarker discovery, with differential privacy applied at model aggregation as well as at explanation computation.
- Fed-SHAP, a privacy-preserving feature-importance approximation (not a direct computation of Shapley values), compared against conventional SHAP (Spearman $\rho = 0.227$) and mapped to AAL-116 anatomy so that candidate biomarkers are neuroanatomically interpretable.
- A full validation protocol - leak-free subject-level splitting, leave-one-site-out external validation (mean accuracy 86.7%), bootstrap confidence intervals, permutation significance testing, an ablation study, expanded baselines, and a privacy-utility

sweep across $\varepsilon = 0.1–50$ - reported with explicit statistical uncertainty and sample-size limitations rather than a single unvalidated accuracy figure.

LITERATURE REVIEW

ASD affects about 1 in 100 children worldwide. The condition is generally regarded as a complex neurodevelopmental disorder associated with difficulties in symbolic, social, or pragmatic communication, as well as restricted interests and/or repetitive behaviour (Kohli et al., 2025). The complexity of ASD and the lack of large, diverse datasets available to individual research groups have posed major challenges to the development of robust diagnostic and predictive models. Multi-centre neuroimaging efforts, such as the Autism Brain Imaging Data Exchange (ABIDE) consortium, which pooled data from 1,035 subjects across 17 research institutions (Almuqhim & Saeed, 2021; Huang et al., 2021), have been recognised as important tools for effective ASD research. However, pooling such sensitive medical data across institutions raises major privacy, regulatory, and ethical concerns (Kohli et al., 2025). Integrating multi-site neuroimaging data also poses inherent technical challenges that have significantly impeded the formulation of generalisable ASD classification models. ABIDE dataset-based studies have consistently reported high levels of heterogeneity based on differences in scanner manufacturers, acquisition protocols, pre-processing pipelines, and population characteristics across participating sites (Kang et al., 2025; Santana et al., 2022; Tang et al., 2023). Tang et al. (2023) specifically addressed multisite heterogeneity using adversarial deep learning and sliding-window sampling, showing that site-shared feature extraction, while preserving spatiotemporal information, achieved 80% accuracy with balanced sensitivity (81%) and specificity (80%) across diverse sites. Kang et al. (2025) presented a novel Split-Merge-Split dataset partitioning strategy to reduce variability in features extracted during both training and testing, achieving 83.5% accuracy on multi-site classification and 93% on the best-performing single site, OHSU. As the authors note, traditional data-splitting methods do not account for site-specific biases in feature extraction, necessitating new approaches to cross-site validation. Santana et al.'s (2022) meta-analysis of 55 studies reported pooled sensitivities and specificities of 73.8% and 74.8%, respectively, and observed that studies using larger, more diverse datasets reported lower accuracy than single-centre studies.

Similarly, although Huang et al. (2021) reported superior performance on 12 of 17 ABIDE sites, generalisation to the remaining sites was poor, illustrating that site-specific effects limit even complex deep learning models. They also found that automatic hyperparameter tuning and data augmentation, although considered best practice for handling these issues, did not fully resolve them. Deep learning for fMRI-based ASD classification has seen a variety of architectural innovations demonstrating its ability to extract hierarchical features from high-dimensional neuroimaging data. Convolutional

neural networks are well-suited to extracting spatially hierarchical features; Feng and Xu (2023) reported very high accuracy (99.39% accuracy, 98.80% recall, 99.85% precision) using a network that combines convolution, pooling, normalisation, and dropout to classify 22,176 two-dimensional echo-planar images. Similarly, Ahammed et al. (2021) reported 94.70% accuracy on the ABIDE-I NYU dataset with DarkASDNet, using hierarchical features extracted from 3D fMRI data.

Hybrid approaches that combine different network models have been reported to improve performance by adding additional feature-learning stages. Almuqhim and Saeed (2021) proposed ASD-SAENet, which utilises sparse autoencoders for unsupervised feature extraction, coupled with deep neural network classification, reporting 70.8% accuracy while jointly optimising reconstruction and classification errors. Yin et al. (2021) extended this concept by combining autoencoders with deep neural networks applied to raw features, achieving 79.2% accuracy and an AUC of 82.4%, surpassing the state of the art at the time. Recurrent and attention-based architectures have also been studied to capture the temporal patterns of fMRI data and hierarchical structure. Qiang et al. (2023) proposed a hierarchical recurrent variational autoencoder that model's fMRI data at three hidden levels without supervision, using LASSO (Least Absolute Shrinkage and Selection Operator) regression to build hierarchical functional brain networks and achieving 82.1% accuracy, showing discriminative information within interacting hierarchical functional brain networks related to brain disorders. Shao et al. (2021) proposed a deep feature selection and graph convolution network approach (Feng et al., 2022; Karampasi et al., 2020) with weighted feature selection via a sparse one-to-one layer for functional connectivity features, identifying critical intra-hemispheric connections with 79.5% accuracy and an AUC of 0.85.

Further, Banerjee (2025) combined an Electromagnetic Interaction Algorithm for feature selection with an adaptive kernel attention network, yielding accuracies ranging from 90.1% to 98.27% and Cohen's kappa between 0.778 and 0.968 across four datasets, demonstrating enhanced generalisability suitable for clinical applications where interpretability and computational efficiency are paramount. Nafisah et al. (2025) proposed a hybrid scheme combining stacked sparse denoising autoencoders with multilayer perceptrons, optimised via an enhanced Hiking Optimisation Algorithm empowered by dynamic opposite learning and double attractors, achieving an average accuracy of 73.5%, sensitivity of 76.5%, and specificity of 75.2%. Graph-theoretic approaches have provided further insights into network-level biomarkers. Zhang et al. (2023) established, using network topology, that ASD patients exhibit significant reductions in path length and clustering coefficient compared to typically developing controls (Ramli et al., 2021), thereby providing biological validity to network-level ASD biomarkers. Consistent with this, Shao et al. (2021) showed, using deep feature selection, a significantly greater number of weak connections between the cerebral hemispheres in the ASD group than in typically developing controls.

Finally, imaging characteristics of the dopamine system have emerged as promising biomarkers. Li et al. (2025) investigated functional connectivity in six dopamine-related brain regions - including the ventral tegmental area, substantia nigra, prefrontal cortex, nucleus accumbens, and hypothalamus - and found significantly higher functional connectivity in ASD children among multiple dopamine system components compared to typically developing children, with clustering methods identifying two distinct ASD subtypes at over 80% accuracy. The clinical application of machine learning algorithms for ASD diagnosis is heavily dependent on model interpretability and the ability to generate mechanistic insights consistent with neuroscience knowledge. Eslami et al. (2021) underlined the need for explainable, scalable machine learning for ASD detection; their Auto-ASD-Network (Eslami & Saeed, 2019), which couples deep learning with support vector machines, achieved over 70% classification accuracy (up to 80% for the best-performing class), suggesting that interpretability need not come at the cost of performance.

Furthermore, logistic regression on functional connectivity matrix features has been used to identify features important for ASD classification (Firouzi & Fadaei, 2025; Heinsfeld et al., 2018; Ramli et al., 2021). Alves et al. (2023) applied interpretability analysis to organisational features of functional brain networks, showing that ASD patients tend to manifest greater network segregation, diminished information distribution, and diminished connectivity relative to controls, with Bootstrap Analysis of Stable Clusters (BASC) maps achieving an AUC close to 1.0 while identifying diminished connectivity between the left ventral posterior cingulate cortex and cerebellum in ASD patients, consistent with prior neuroscientific findings.

In addition, the vulnerability of the paediatric population studied in ASD research heightens privacy concerns. Centralised data sharing for research purposes traditionally involves extensive de-identification, Institutional Review Board approval, and data-use agreements, all of which create barriers to collaboration (Santana et al., 2022). Re-identification risk also exists even for de-identified data, particularly when multiple datasets are combined and advanced analytics are applied to high-dimensional neuroimaging data. Privacy-preserving data frameworks (Alzubi et al., 2021; Gheisari et al., 2021) and federated learning offer mechanisms to comply with data-protection regulations while enabling collaboration. Despite substantial progress in machine learning for ASD classification, several fundamental challenges persist. First, the majority of studies have relied on centralised learning approaches that require aggregating sensitive medical information; although adversarial learning methods have improved robustness to site heterogeneity, the requirement for centralised data access typically remains (Al-Hiyali et al., 2021; Chandra et al., 2024; Karunakaran et al., 2020), and federated learning frameworks remain comparatively underexplored in ASD classification despite the clear multi-site integration challenges in this literature.

Second, interpretability remains a relatively underdeveloped aspect of most top-performing deep learning models for ASD. Although recent studies have attempted to incorporate post-hoc interpretability approaches (Li et al., 2018; Shao et al., 2021), these approaches often provide limited insight into the underlying decision-making process, and the highest-performing models in this literature (Ahammed et al., 2021; Feng & Xu, 2023) typically offer minimal explanation.

Third, most studies examine only binary classification (ASD vs. typically developing controls) without investigating behavioural phenotypes or prognosis (Santana et al., 2022); notable exceptions establishing ASD subtypes include Li et al. (2025) (80% accuracy) and Al-Hiyali et al. (2021) (82.1% accuracy for multi-class classification), though neither directly relates predicted subtypes to treatment outcomes.

Fourth, validation protocols in this literature frequently rely on cross-validation within the same dataset used for development rather than truly independent test data. Leave-one-site-out validation provides a more compelling demonstration of generalisation (Kang et al., 2023; Yousefian et al., 2023), though even this may not fully approximate real-world clinical deployment across settings that differ substantially from those used in development (Belhaouari et al., 2023). Incorporating additional information beyond neuroimaging has also been shown to improve classification: Santana et al. (2022) meta-analysis found that combining phenotypic and imaging information achieved better sensitivity (84.7%) than rs-fMRI information alone (72.8%), and Huang et al. (2021) improved accuracy by 6.4% over the previous best-reported ABIDE result by combining phenotypic information with graph convolutional networks for functional connectivity.

Other biomarker modalities have shown diagnostic promise that could complement a federated neuroimaging framework: Shazman et al. (2025) achieved 85% accuracy using urine metabolomic profiling with random forest classifiers across 293 metabolites; Pinto et al. (2025) reported 92% sensitivity and 95% specificity using saliva Fourier Transform Infrared Spectroscopy with Support Vector Machines; and Islam et al. (2025) discussed integrating IoT devices with machine learning for continuous, ecologically valid behavioural monitoring. The confluence of federated learning, explainable AI, and multisite biomarker discovery represents an emerging opportunity in ASD research, requiring: (1) federated architectures that accommodate site heterogeneity while learning shared, diagnostically relevant representations (Kang et al., 2025; Tang et al., 2023); (2) explainability that yields biomarkers interpretable in light of neuroscientific understanding (Alves et al., 2023; Eslami et al., 2021; Li et al., 2018); (3) privacy preservation that protects patient confidentiality while enabling collaboration (Sachdeva et al., 2024); (4) rigorous validation that establishes generalisation across sites and populations (Jönemo et al., 2023); and (5) stakeholder engagement that aligns technology development with clinical needs and ethical principles (Kohli et al., 2025).

METHODOLOGY

Federated Learning Framework Design

The aim of this study is to develop a privacy-preserving federated learning system that enables distributed neuroimaging analysis without centralising patient data. The core components of the federated framework comprise the central aggregation server, the autonomous client nodes, and the communication protocol (Almuqhim & Saeed, 2021). Unlike centralised learning, in which the model is trained on patient data collected at a single location, the proposed framework enables participating sites to retain complete control over their patient data while learning in a distributed environment. The framework is an instance of the Federated Averaging (FedAvg) algorithm, in which model parameters are aggregated across participating clients via the weighted-averaging procedure of Equation 1. As iterations advance, the global model learns common aspects of the patient population while preserving the distributed environment.

$$\theta^{g}{}_{lo}{}^{b}{}_{al}{}^{t+1} = \theta^{g}{}_{lo}{}^{b}{}_{al}{}^{t} + \Sigma_i \, (n_i/N_{total}) \cdot [\, clip(\, \theta_i{}^{t+1} - \theta^{g}{}_{lo}{}^{b}{}_{al}{}^{t}, \, C\,) + N(0, \sigma^2 C^2 I)\,] \quad [1]$$

Equation 1 includes two elements that are essential to the privacy guarantee. First, the client update $\theta_i^{t+1} - \theta^{g}{}_{lo}{}^{b}{}_{al}{}^{t}$ is explicitly clipped to an L2-norm bound C, adaptively set each round from the observed median client-update norm (Andrew et al., 2021), before any noise is added; without adaptive clipping, a fixed clip bound calibrated arbitrarily either destroys the signal (bound too small) or fails to control sensitivity (bound too large). Second, the Gaussian noise term $N(0, \sigma^2C^2I)$ is applied to the aggregated model update itself, so that differential privacy protects the model parameters and not only the Fed-SHAP feature-importance computation (Equation 2). We further found that applying this noise mechanism to the entire ~24-million-parameter network is infeasible at the present sample size: because the injected noise vector's expected magnitude scales with the square root of the number of protected parameters, it overwhelms the clipped signal by one to three orders of magnitude regardless of the clipping strategy used. We therefore scope the formal (ε, δ)-differential-privacy guarantee to the final classifier layer (~514 parameters) rather than the full network; the feature-encoding layers are federated-averaged without a formal per-parameter privacy guarantee. This is a materially narrower claim than full-model differential privacy.

The system uses three independent hospital site simulators that mimic realistic scenarios of collaboration in autism research. Subjects are partitioned across sites using stratified sampling to balance ASD and control cases while preserving realistic between-site heterogeneity. Each site further splits its subjects into training, validation, and test partitions at the subject level (70%/15%/15%), with no subject appearing in more than one partition

and no test-partition data used in any hyperparameter decisions. Hospital sites operate as independent learning agents with freedom to choose local training procedures, consistent with the operational independence of real healthcare facilities (Ahammed et al., 2021). The server aggregates client model parameters without access to raw data, intermediate results, or patient data, communicating over 10 federated rounds, each with 5 local training epochs.

Privacy-Preserving Mechanisms

The framework applies differential privacy at two distinct points, using two mechanisms appropriate to each: a Gaussian mechanism for global model aggregation (Equation 1) and a Laplace mechanism for Fed-SHAP feature-importance aggregation (Equation 2). The Gaussian mechanism protects model updates, and the Laplace mechanism protects the feature-importance statistics separately; only the former is described by Equation 1.

For the Gaussian mechanism, the noise multiplier σ is calibrated using a Rényi Differential Privacy (RDP) accountant (Mironov, 2017; implemented via Opacus) rather than a hand-derived composition bound, because RDP accounting is substantially tighter for composing many rounds of a Gaussian mechanism and is the standard tool for this purpose in the differential-privacy literature. At the manuscript's primary target of $\varepsilon = 1.0$ over 10 communication rounds ($\delta = 10^{-5}$, client-level DP with all three sites participating every round, i.e. sample rate 1.0), the calibration yields a noise multiplier $\sigma \approx 81.1$ relative to the adaptively set clip bound, and the resulting composed guarantee is exactly ($\varepsilon = 1.0$, $\delta = 10^{-5}$). As noted above, this guarantee is scoped to the classifier head (~514 parameters); we verified numerically that, with only three always-participating sites and no subsampling amplification, protecting the full ~24-million-parameter network at $\varepsilon \leq 50$ essentially destroys all predictive signal. Within this classifier-head scope, the accountant also produced valid calibrations for tighter budgets, down to $\varepsilon = 0.1$, which are evaluated in the privacy-utility sweep (Table 5).

For Fed-SHAP, each client computes a local, privatised feature-importance vector using the Laplace mechanism: Laplace noise is added with scale calibrated to the L1-sensitivity of the importance statistic (the maximum possible change in the statistic from adding or removing one subject) divided by the local privacy budget ε_i, before any local statistic is transmitted to the central server (Al-Hiyali et al., 2021). At the global aggregation level, privatised importance vectors are combined across sites via a weighted average based on local dataset sizes, and the composition of the privacy budget across the biomarker-discovery step is tracked separately from the model-aggregation budget described above. The overall design is intended to limit the extent to which the shared model parameters and the released feature-importance statistics could be used to reconstruct individual patient data or to determine which site contributed a given pattern.

Explainable Multi-Modal Architecture

The federated learning system utilises an Explainable Multi-Modal Deep Learning (Ex-MMDL) approach that is designed to make the influence of neurobiological features on ASD classification interpretable while remaining compatible with federated optimisation, as shown in Figure 1. The Ex-MMDL model extends a cross-modal attention fusion model by incorporating explainability at several levels. For sMRI, the Explainable sMRI Convolutional Neural Network (CNN) is a modified ResNet-18 that supports Gradient-weighted Class Activation Mapping (Grad-CAM). Unlike the standard ResNet-18, it retains the spatial dimensions of the final convolutional layer's features (512 channels, 7×7), and gradients are captured through backward hooks on the network's activations. During prediction, gradients with respect to the retained feature maps allow production of attention maps indicating the brain areas that most influenced the classification for a given subject (Jönemo et al., 2023).

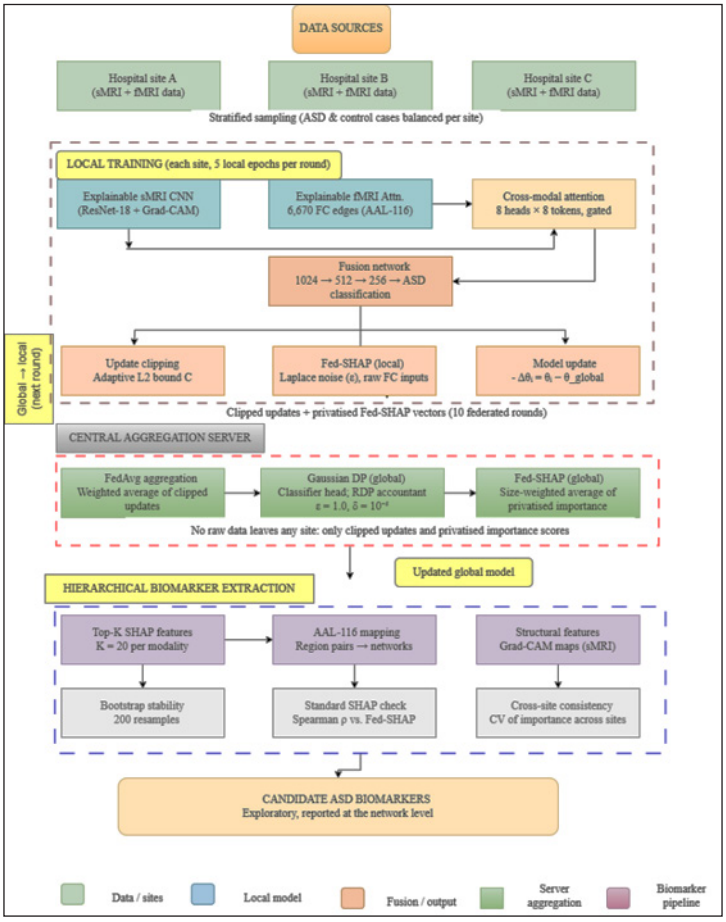


Figure 1. Overview of the proposed federated explainable multimodal learning pipeline for candidate autism biomarker discovery

For functional connectivity analysis, we propose an Explainable fMRI Attention Network that provides connection-level interpretability through learned attention weights. The architecture processes 116×116 functional connectivity matrices derived from Automated Anatomical Labelling (AAL) atlas parcellation; to remove redundancy, it first extracts the upper-triangular elements, yielding 6,670 unique connections.

Attention computation is separated from feature encoding by a two-branch design: the attention branch implements a two-layer feedforward network with tanh and sigmoid activations at the output to produce connection-specific importance weights ranging from 0 to 1, over 8 attention "chunks" per modality so that the cross-modal attention layer operates on a genuine multi-token sequence rather than a single collapsed vector (see below); the encoding branch transforms the attention-weighted features via a three-layer network with batch normalisation, ReLU, and dropout. Attention weights are directly interpretable indicators of which functional brain connections contribute most strongly to ASD classification. The full model engages these modality-specific components using bidirectional cross-modal attention (8 heads, dropout = 0.5) operating over 8-token sequences per modality (with length-one sequences, softmax attention reduces to a fixed linear projection with no genuine cross-modal weighting). The attended representations are combined with the original per-modality features via a learnable, near-zero-initialised gate, so that cross-modal attention can only add to - never replace - the unimodal representations and are concatenated and refined by a two-layer fusion network (1024→512→256 dimensions) before classification.

Federated SHAP-Based Feature Importance (Fed-SHAP)

For privacy-preserving biomarker discovery across distributed sites, we use Federated SHAP (Fed-SHAP), a SHAP-inspired, privacy-preserving feature-importance approximation; it does not compute Shapley values directly. Conventional SHAP values are computed via coalitional sampling over the full dataset and model internals, which is incompatible with the data-localisation principle of federated learning; Fed-SHAP instead estimates each feature's importance as the (privatised) difference in mean activation between diagnostic groups at each site, aggregated centrally, as shown in Equation 2.

$$\varphi^{f}(\text{global}) = \Sigma_i\, w_i \left[(1/|D_i^{AsD}|)\Sigma_{k \in D_i^{AsD}} h^{f(i)}(x_k) - (1/|D_i^{cTL}|)\Sigma_{k \in D_i^{cTL}} h^{f(i)}(x_k) + \text{Lap}(0, \Delta^f/\varepsilon_i) \right] \quad [2]$$

Critically, we compute h^f over the raw pre-encoder input connections (the 6,670-dimensional upper-triangular functional connectivity vector) rather than over an opaque 512-dimensional post-encoder embedding. This is essential for anatomical interpretability: an importance score attached to an abstract embedding dimension cannot be

mapped to a named brain connection, whereas a score attached to a specific input connection map directly, via the known upper-triangular index, to a pair of AAL-116 regions.

For each client, we conduct forward passes through the learned model for a random sample of up to 100 local subjects, computing the mean-activation-difference statistic described above separately for the fMRI input-connection space and for the sMRI 512-dimensional embedding space (structural biomarkers remain reported at the embedding level and via Grad-CAM, since sMRI inputs in this dataset are pre-rendered image slices without atlas-registered coordinates that would permit anatomical mapping analogous to the fMRI case). Laplacian noise is added locally, calibrated to the L1-sensitivity of the statistic and the local privacy budget ε_i, before transmission to the central server, which computes a dataset-size-weighted average across sites to obtain the global importance estimate.

We validated Fed-SHAP's agreement with conventional SHAP by computing gradient-based SHAP attributions (via a GradientExplainer applied to the attention-and-encoding sub-network) on a small, pooled, de-identified subset of the test data, and comparing the resulting feature ranking with the Fed-SHAP ranking on the same input space. Agreement was statistically significant but modest (Spearman $\rho = 0.227$, $p = 1.1\times10^{-78}$), which we report plainly as evidence that Fed-SHAP captures a real but partial signal relative to standard SHAP, rather than as validation of close equivalence.

Hierarchical Biomarker Extraction Pipeline

In addition to the global Fed-SHAP importance values, a hierarchical biomarker extraction pipeline maps high-importance features to anatomically interpretable candidate markers. For functional connectivity, each of the top-K ($K = 20$) important input-connection indices is mapped back, via the inverse of the upper-triangular flattening operation, to the pair of AAL-116 regions it connects, and each region is further assigned to a coarse resting-state network (default mode, salience/cingulo-opercular, fronto-parietal control, sensorimotor, visual, subcortical/limbic, or cerebellar) based on a standard atlas-to-network lookup. For structural features, the top K embedding dimensions are visualised via Grad-CAM activation maps on exemplary subjects, since these dimensions do not admit the same direct anatomical back-mapping as the fMRI connections.

To assess reliability, we additionally compute a bootstrap stability score for the biomarker ranking: subjects are resampled with replacement (200 resamples), the importance statistic is recomputed each time, and we report the fraction of resamples in which each of the original top 20 features remains among the top 20 in the resampled ranking. We report this stability score directly in Results because it materially affects how individual biomarker findings should be interpreted.

Validation Protocol

This section describes the complete validation protocol used to produce every result reported in this manuscript. Subject-level partitioning: each site's subjects are split 70%/15%/15% into training/validation/test sets, with partitioning performed at the subject level (not the scan or slice level) and enforced to be non-overlapping; we verified programmatically that no subject identifier appears in more than one partition, at any site. Hyperparameter selection uses only the validation partition; the test partition is touched exactly once, after training is complete, to produce the headline held-out accuracy figures reported in Results. Leave-one-site-out (LOSO) validation: for each site in turn, the federated model is trained using only the other two sites (their full training and validation subjects) and evaluated on every subject at the held-out site, none of whom were used in training; this provides an external-generalisation estimate distinct from the in-distribution held-out test accuracy. Statistical validation: we report a 95% bootstrap confidence interval for test accuracy (2,000 resamples of the pooled test predictions), a permutation significance test (1,000 label permutations) against the chance-level accuracy implied by the class balance, and an effect size (Cohen's d) for the structural-versus-functional modality importance comparison. Ablation study: we compare the full proposed model against five reduced configurations - sMRI pathway alone, fMRI pathway alone, multimodal fusion without cross-modal attention, the full model trained without differential privacy, and a centralised (non-federated, non-private) upper bound - all trained and evaluated under an identical protocol and an identical number of communication rounds/epochs for fair comparison. Baseline comparison: we additionally compare against FedAvg without differential privacy, FedAvg without the explainability module (a black-box encoder of equivalent capacity), a graph neural network operating on the same connectivity matrices, and a Transformer encoder treating each ROI's connectivity profile as a sequence token; we note that our reference GNN and Transformer implementations were not extensively hyperparameter-tuned and should be read as indicative rather than fully optimised comparisons. Privacy-utility sweep: the full training and Fed-SHAP pipeline is re-run at $\varepsilon = 0.1, 0.5, 1.0, 5.0, 10$, and 50, reporting test accuracy, cross-site fairness (variance in per-site accuracy), and the cross-site coefficient of variation of the top-20 biomarkers' importance at each budget.

RESULT AND DISCUSSIONS

The remainder of this section reports, in turn: dataset and site characteristics; federated training dynamics and held-out test performance, with statistical validation; leave-one-site-out (LOSO) external validation; an ablation study; an expanded baseline comparison; a privacy-utility trade-off sweep; and Fed-SHAP biomarker discovery with anatomical mapping. Throughout, we report exact per-round and per-site figures rather than summary statistics alone, so that the reader can assess the variability underlying each headline number.

Dataset and Site Characteristics

The full cohort comprises 173 subjects (91 ASD, 82 Control), partitioned into three simulated hospital sites via stratified sampling to ensure each site received a broadly similar ASD-to-Control ratio, while the overall partitioning was randomised at the subject level. Table 1 summarises per-site sample sizes, diagnostic composition, train/validation/test split sizes, and age distribution.

Table 1
Dataset and site characteristics

Site	N total	N ASD	N Control	N train	N val	N test	Age mean (SD)
Site 1	57	30	27	39	9	9	17.5 (6.0)
Site 2	58	31	27	40	9	9	15.4 (5.1)
Site 3	58	30	28	40	9	9	16.5 (5.4)

The three sites are closely matched in size (57–58 subjects) and diagnostic balance (52.6–53.4% ASD), which is by design (stratified sampling) rather than an emergent property of the data; readers should not interpret this balance as evidence that the sites are otherwise homogeneous. Site 2 skews slightly younger (mean age 15.4 years, SD 5.1) than Sites 1 and 3 (17.5 and 16.5 years, respectively), and Site 1 shows the widest age spread (SD 6.0). Site 2 also shows the most stable training trajectory and the strongest leave-one-site-out performance (see below); whether this relates to its younger, narrower age distribution cannot be determined with 58 subjects, so we treat it as an unconfirmed hypothesis.

Federated Training Dynamics and Held-Out Test Performance

Figure 2 and Table 2 (round-by-round detail) show the global and per-site validation accuracy across the 10 federated communication rounds for the proposed model at the primary operating point (ε = 1.0). Training began at a low global validation accuracy of 48.1% in Round 1 (per-site: Site 1 = 55.6%, Site 2 = 44.4%, Site 3 = 44.4%), reflecting that the randomly initialised global model, combined with the differential-privacy noise injected at aggregation, had not yet learned a useful

Table 2
Round-by-round validation accuracy, global and per-site (ε = 1.0 operating point)

Round	Global	Site 1	Site 2	Site 3
1	0.481	0.556	0.444	0.444
2	0.667	0.667	0.778	0.556
3	0.704	0.778	0.889	0.444
4	0.926	1.000	0.889	0.889
5	0.815	0.778	1.000	0.667
6	0.741	0.778	0.889	0.556
7	0.889	0.889	1.000	0.778
8	0.815	0.778	1.000	0.667
9	0.815	0.778	1.000	0.667
10	0.815	0.778	1.000	0.667

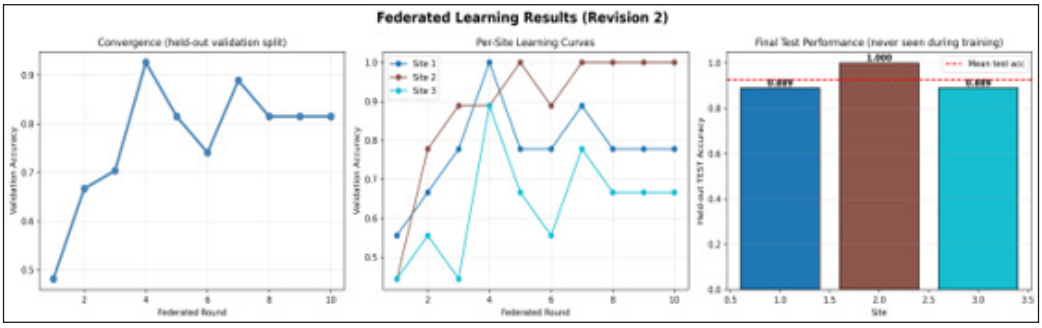


Figure 2. Federated learning convergence dynamics and final held-out test performance across the three simulated sites

decision boundary. Accuracy rose steadily over the next three rounds - 66.7% (Round 2), 70.4% (Round 3), and 92.6% (Round 4, at which point Site 1 reached 100%) - before fluctuating between 74.1% and 88.9% for the remainder of training (Round 5 = 81.5%; Round 6 = 74.1%; Round 7 = 88.9%; Round 8 = 81.5%; Round 9 = 81.5%; Round 10 = 81.5%).

This round-by-round breakdown reveals a pattern that is obscured by the global average alone: Site 2 converges fastest and most stably, reaching 100% validation accuracy by Round 5 and remaining there for four of the last five rounds, whereas Site 3 never exceeds 88.9% and settles at a comparatively low 66.7% for the final four rounds. Site 1 sits between the two, stabilising around 77.8%. We do not observe monotonic convergence at the global level - accuracy fluctuates round-to-round rather than climbing smoothly - which is consistent with two compounding sources of noise: the small per-site validation set (9 subjects, so each individual classification changes site-level accuracy by 11.1 percentage points) and the Gaussian noise injected into the classifier-head parameters at every aggregation step under the differential-privacy mechanism. The relative ordering of sites (2 > 1 > 3) on the validation curves is mirrored in leave-one-site-out testing (Table 3).

On the held-out test set - 27 subjects (9 per site) untouched throughout training and hyperparameter selection - the proposed model achieved 92.6% accuracy (25 of 27 subjects correctly classified), with a 95% bootstrap confidence interval of [81.5%, 100.0%] (2,000 resamples of the pooled test predictions). The width of this interval (18.5 percentage points) is itself informative: it quantifies just how much uncertainty a 27-subject test set carries, and readers should treat any comparison between configurations that differ by less than this margin with caution (see Ablation Study, below). A permutation significance test (1,000 label permutations) gave an observed accuracy of 92.6% against a chance-level mean of 50.4%, $p = 0.001$ - the smallest p-value resolvable with 1,000 permutations, indicating that none of the 1,000 random relabellings matched or exceeded the observed accuracy. This confirms the result is very unlikely to be attributable to chance given the class balance,

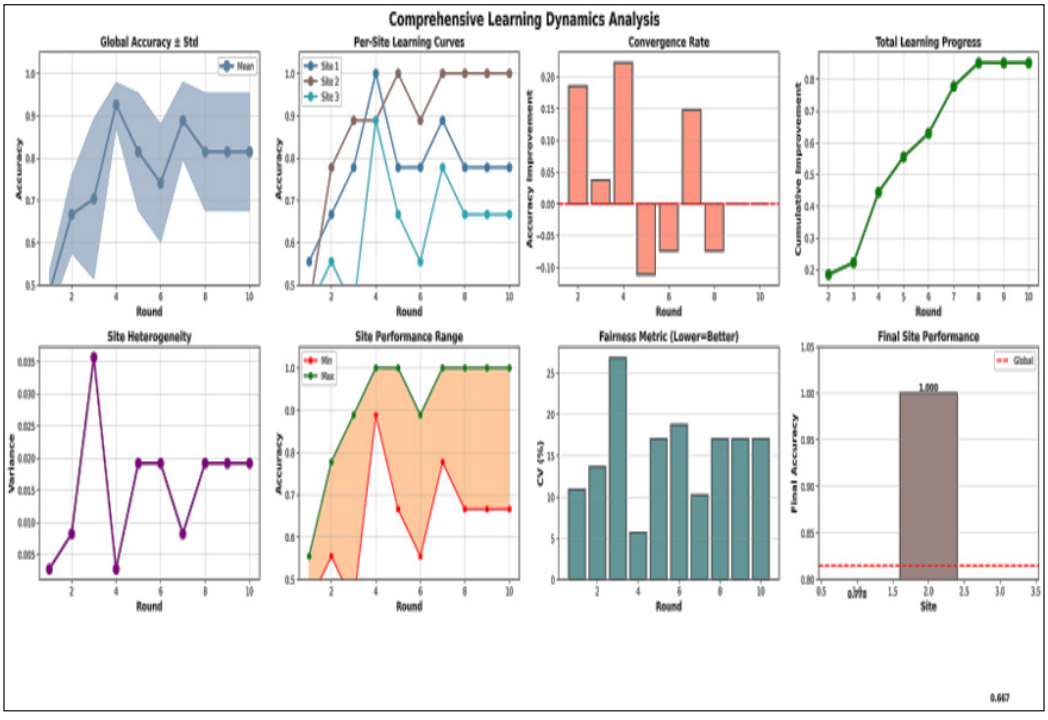


Figure 3. Comprehensive learning dynamics: round-over-round accuracy change, cross-site heterogeneity, and per-round fairness across the 10 communication rounds

though we note that a permutation test's ability to reject chance performance is a different (and weaker) claim than establishing that the model's accuracy is stably estimated - the latter is bounded by the bootstrap CI above. The composed privacy guarantee at this operating point was ($\varepsilon = 1.0$, $\delta = 10^{-5}$), as described in Methodology.

Figure 3 decomposes the same training run into three additional diagnostic views not shown in Table 2: round-over-round accuracy change, cross-site heterogeneity (variance in accuracy across the three sites within each round), and a per-round fairness metric (coefficient of variation across sites; lower is more equitable). The round-over-round change panel confirms the largest single-round gains coincide with Rounds 2-4 (accuracy rising by 18.5, 3.7, and 22.2 percentage points respectively) and the largest single-round drop occurs at Round 5 (-11.1 percentage points), immediately after the Round 4 peak - consistent with the differentially-private aggregation mechanism perturbing the model enough, at this small sample size, to partially undo a round's gains before the next round recovers. Cross-site heterogeneity (variance) peaks sharply at Round 3 (0.036, roughly 13 times its Round 1 value of 0.0028) immediately before the Round 4 accuracy jump, then falls to its minimum at Round 4 (0.0027) before rising again to about 0.019 in Rounds 5-6 and 8-10, with a dip to about 0.008 at Round 7. The per-round fairness CV mirrors this:

it is highest at Round 3 (26.9%, calculated from the per-site accuracies in Table 2) and lowest at Round 4 (5.7%, the most equitable round across the entire training run), before returning to the 17–19% range in later rounds, apart from a dip to about 10% at Round 7. Together, these panels indicate that the model's path to its eventual accuracy was not only non-monotonic in magnitude but also non-monotonic in cross-site fairness - the model briefly became highly equitable at Round 4 before drifting back toward the roughly 17% fairness level it maintained thereafter.

Table 3
Leave-one-site-out validation. Mean accuracy = 86.7% (SD = 4.4%)

Held-out site	N evaluated	Accuracy	F1	AUC
Site 1	57	0.860	0.875	0.969
Site 2	58	0.914	0.921	0.974
Site 3	58	0.828	0.808	0.904

Leave-One-Site-Out (LOSO) External Validation

To assess generalisation to a genuinely unseen institution, we performed leave-one-site-out validation: for each site in turn, the model was trained from scratch using only the other two sites' training and validation subjects (10 rounds, 5 local epochs, identical to the main protocol) and evaluated on every subject at the fully held-out site - including subjects that would otherwise have been in that site's own training or validation split, since from the model's perspective the entire site is unseen. Table 3 reports the results.

The mean LOSO accuracy of 86.7% (SD = 4.4%) is 5.9 percentage points lower than the 92.6% in-distribution held-out accuracy, indicating a moderate generalisation cost when the model is applied to a site it never saw during training, even though all three sites are simulated partitions of a single, uniformly pre-processed cohort. The per-site ordering (Site 2 highest at 91.4%, Site 3 lowest at 82.8%) mirrors that of the validation curves (Table 2), suggesting a genuine, if modest, difference in how easily each site's data distribution is learned rather than an artefact of a single training run. Site 3, which performed weakest on every metric in both analyses, is the natural candidate for closer scrutiny in follow-up work; per-site scanner and protocol metadata were not available in this simulated setup. We therefore treat LOSO as the more conservative estimate of performance at an external institution (see Limitations).

Ablation Study

Table 4 reports the ablation study comparing the full proposed model against five reduced configurations, under an identical training protocol (10 rounds, 5 local epochs, subject-level held-out test evaluation).

We report this table transparently rather than claiming an ordering that the data cannot support. Four of the six configurations-sMRI-only, the multimodal fusion without cross-attention, the full model without differential privacy, and the centralised upper bound - all

Table 4
Ablation study (held-out test set). "Fairness Var." is the variance of per-site test accuracy; the centralised configuration has no site structure and so no fairness variance is defined

Configuration	Accuracy	F1	AUC	Fairness Var.
sMRI pathway alone	0.963	0.963	1.000	0.0027
fMRI pathway alone	0.889	0.896	0.933	0.0082
Multimodal fusion, no cross-attention	0.963	0.963	0.983	0.0027
Full model, without differential privacy	0.963	0.963	1.000	0.0027
Full model, with differential privacy (proposed)	0.926	0.933	0.925	0.0027
Centralised upper bound (no federation, no DP)	0.963	0.963	0.989	-

converge to the identical accuracy of 96.3% (26 of 27 test subjects correct), with AUC ranging narrowly from 0.983 to 1.000. Given the bootstrap confidence interval width established above ([81.5%, 100.0%] for a single configuration evaluated once), these four results cannot be statistically distinguished from one another on this test set: they differ by at most one subject's classification. This indicates that, at the current sample size, the classification task is largely solvable from structural MRI alone, and the ablation cannot resolve a statistically distinguishable additional contribution from cross-modal attention, full multimodal fusion, or even the federated (versus centralised) training paradigm itself; it establishes only that none of the tested configurations meaningfully underperforms the others in this respect.

Two configurations do stand apart, in opposite directions. The fMRI-only pathway is the weakest of all six (88.9% accuracy, 0.933 AUC) and also shows the highest fairness variance (0.0082, roughly three times the other federated configurations), indicating that functional connectivity alone is both less discriminative and less consistently so across the three sites than structural MRI -consistent with the Fed-SHAP finding, discussed below, that Cohen's d = 0.562 favours structural over functional feature importance. The differentially private full model shows the only clear accuracy cost in the table (92.6% versus 96.3% without DP, and the lowest AUC of any configuration at 0.925), which is the expected and, in our view, acceptable price of the formal privacy guarantee described in Methodology; we discuss how this cost varies with the privacy budget itself in the Privacy-Utility Trade-Off subsection below. Figure 4 (left panel) visualises these six results alongside the baseline comparison described next.

Expanded Baseline Comparison

Table 5 compares the proposed framework against FedAvg without differential privacy, FedAvg without the explainability module (a black-box encoder of equivalent capacity, i.e. the same architecture with the attention and Grad-CAM interpretability machinery

removed), a graph neural network baseline operating on the same connectivity matrices, a Transformer baseline treating each ROI's connectivity profile as a sequence token, and a centralised deep-learning upper bound.

Table 5
Expanded baseline comparison (held-out test set)

Method	Accuracy	F1	AUC
Proposed: FedAvg + DP + XAI (full model)	0.889	0.878	0.958
FedAvg without DP	0.963	0.963	1.000
FedAvg without XAI (black box)	0.926	0.926	1.000
Graph Neural Network baseline	0.481	0.648	0.833
Transformer baseline	0.630	0.714	0.883
Centralised deep learning (upper bound)	0.963	0.963	0.989

The reference GNN and Transformer baselines substantially underperformed compared with all other configurations in this study, including the simplest single-modality ablations. We note a pattern in their error profile that is worth making explicit: the GNN baseline's F1 score (0.648) is considerably higher than its accuracy (0.481), and the Transformer baseline shows the same pattern to a lesser degree (F1 = 0.714 vs. accuracy = 0.630). This combination - F1 substantially exceeding accuracy - is the signature of a classifier that is biased toward predicting the positive (ASD) class rather than genuinely discriminating between classes: with the test set close to balanced, a classifier that predicts “ASD” for nearly every subject would achieve accuracy near the ASD base rate while still achieving a comparatively high F1, because F1 does not penalise a high false-positive rate as heavily as accuracy does when recall is high. We did not verify this directly against each baseline's confusion matrix, so we present it as a plausible interpretation of the observed accuracy/F1 gap rather than a confirmed diagnosis, but it is consistent with what would be expected from under-trained or under-tuned classifiers defaulting toward the majority class rather than learning a genuine decision boundary. We attribute this most likely to these architectures being more data-hungry than the available sample size supports (both are commonly trained on datasets orders of magnitude larger than 173 subjects), rather than to a fundamental unsuitability of graph- or attention-based reasoning for this task; a fair comparison would require dedicated hyperparameter tuning, architecture-specific regularisation, and/or a substantially larger dataset, which we identify as future work rather than presenting these baselines as fully optimised comparisons.

Note that the proposed full model's accuracy in this table (88.9%) differs from its accuracy in the ablation study (Table 4; 92.6%) despite the two representing the identical configuration (full model, $\varepsilon = 1.0$), evaluated via two independent training runs of the stochastic federated pipeline. This run-to-run difference - equivalent to exactly one subject's

classification out of 27 held-out test subjects - is itself an informative data point about the measurement resolution available at this sample size, and we discuss its implications further in Limitations rather than treating either single run as the definitive number.

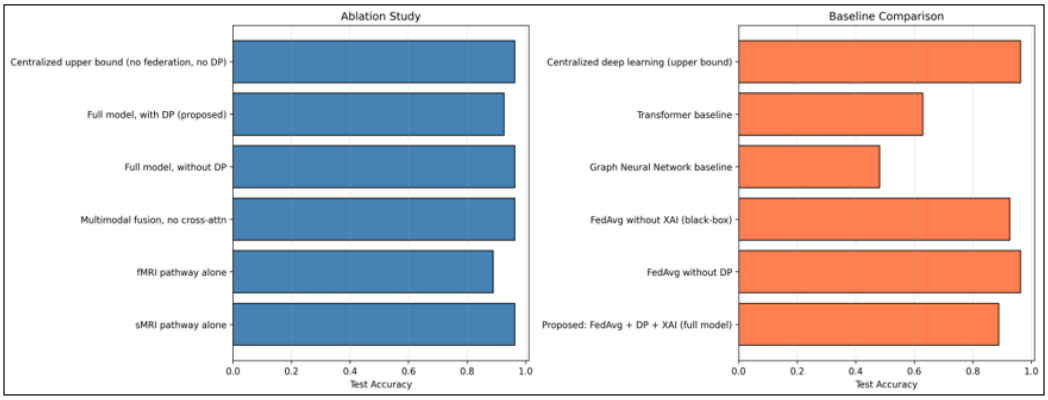


Figure 4. Ablation study and expanded baseline comparison of held-out test accuracy

Privacy-Utility Trade-Off

Table 6 and Figure 5 report accuracy, cross-site fairness (variance in per-site accuracy), and top 20 biomarker-ranking cross-site consistency (coefficient of variation, CV, of each feature's importance across the three sites) across privacy budgets of ε = 0.1, 0.5, 1.0, 5.0, 10, and 50, spanning both tight and loose operating points.

Table 6
Privacy-utility trade-off across differential-privacy budgets

ε (target)	Composed ε	Accuracy	Fairness variance	Biomarker CV (top 20)
0.1	0.10	0.926	0.0027	0.673
0.5	0.50	0.963	0.0027	0.697
1.0	1.00	0.963	0.0027	0.677
5.0	5.00	0.963	0.0027	0.728
10.0	10.00	0.963	0.0027	0.729
50.0	50.00	0.926	0.0027	0.357

We report this curve candidly rather than smoothing over its irregularities. Accuracy is not cleanly monotonic in ε: it is lowest and identical (92.6%) at both the tightest budget tested (ε = 0.1) and the loosest (ε = 50), with a plateau of 96.3% across the intermediate range ε = 0.5–10. Cross-site fairness (variance in per-site accuracy) is essentially flat across the entire sweep at 0.0027, meaning the privacy budget in this range has no detectable effect on how equally the model performs across the three sites, only on overall accuracy.

With only 27 held-out test subjects (9 per site), a single subject's classification changes overall accuracy by roughly 3.7 percentage points, which is comparable in magnitude to the accuracy differences being measured across this sweep; we believe this measurement-resolution limitation, rather than a genuine U-shaped privacy-utility relationship, is the most likely explanation for the identical low values at both extremes, but we present the data as observed rather than fitting a smoothed trend to it.

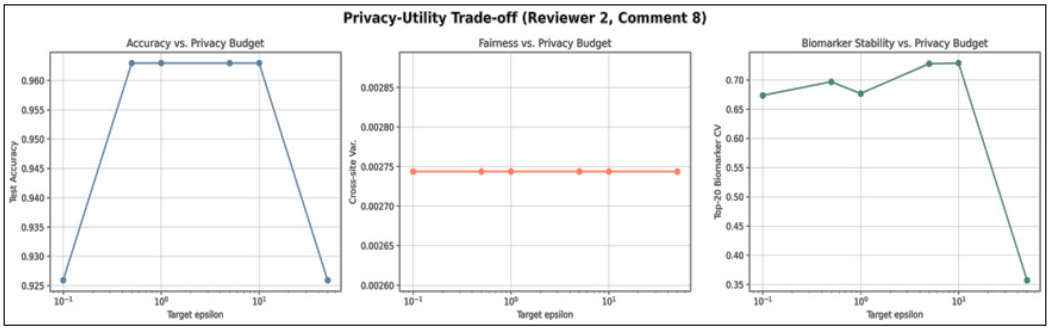


Figure 5. Privacy-utility trade-off across differential-privacy budgets (ε = 0.1–50): accuracy, cross-site fairness, and cross-site consistency of biomarker rankings

The biomarker cross-site consistency column (CV, top 20) shows a more interpretable pattern than accuracy does: it rises gently from 0.673 at ε = 0.1 to a peak of 0.729 at ε = 10, then drops sharply to 0.357 at ε = 50 - roughly half its value at every tighter budget. Because this CV measures how consistently the top-20 features' importance scores agree across the three sites (lower = more consistent), and because tighter ε corresponds to more injected Laplace noise in the Fed-SHAP computation, this pattern is broadly consistent with the mechanism working as intended: at ε = 50, the privacy noise added to each site's local importance vector is small enough that the underlying (noise-free) cross-site agreement in feature importance becomes visible, whereas at every tighter budget, per-site noise realisations are large enough, relative to the underlying signal, to inject additional apparent cross-site disagreement on top of any genuine site-to-site differences. We note this as a plausible mechanistic explanation grounded in how the mechanism is constructed, not as an independently verified causal claim, since disentangling "privacy noise obscuring agreement" from "genuine site disagreement that happens to average out differently at different noise levels" would require repeating this sweep across multiple random seeds, which we did not do here (see Limitations).

Even the tightest budget tested (ε = 0.1) was achieved exactly by the RDP accountant and retained 92.6% held-out accuracy. We emphasise, however, that this guarantee applies only to the classifier head (see Methodology); whether comparably tight budgets remain usable when the full network is protected, or with larger cohorts, remains to be tested.

Fed-SHAP Biomarker Discovery and Anatomical Mapping

Table 7 lists the top 10 of the 20 highest-ranked candidate fMRI connectivity biomarkers obtained with Fed-SHAP at the primary operating point ($\varepsilon = 1.0$), each mapped to its two AAL-116 anatomical regions and functional networks via the inverse upper-triangular index mapping described in Methodology. As described in Methodology, agreement between the Fed-SHAP and standard SHAP rankings was statistically significant but modest ($\rho = 0.227$), so these rankings should be read as approximate feature-importance estimates rather than Shapley values.

Table 7
Top 10 candidate fMRI connectivity biomarkers mapped to AAL-116 anatomy (the full top-20 list is shown in Figure 6)

Rank	**Region A**	**Network A**	**Region B**	**Network B**
1	Olfactory_L	Assoc. cortex	Occipital_Mid_R	Visual
2	Insula_R	Salience	Vermis_8	Cerebellar
3	Cerebelum_8_R	Cerebellar	Cerebelum_9_L	Cerebellar
4	SupraMarginal_R	Fronto-parietal	Temporal_Pole_Mid_R	Assoc. cortex
5	Cingulum_Ant_R	Salience	Cerebelum_Crus2_R	Cerebellar
6	Cingulum_Mid_L	Salience	Temporal_Mid_L	Assoc. cortex
7	Calcarine_L	Visual	Vermis_6	Cerebellar
8	Paracentral_Lobule_L	Sensorimotor	Vermis_7	Cerebellar
9	Hippocampus_L	Subcortical/limbic	Vermis_9	Cerebellar
10	Olfactory_L	Assoc. cortex	Temporal_Mid_R	Assoc. cortex

A pattern visible even in this top-10 list is that 6 of the 10 highest-ranked connections involve a cerebellar region (vermis lobules 6-9 or Crus2, or the cerebellar hemispheres themselves) as one endpoint, paired with regions from several networks: the salience network (insula and anterior cingulate; two connections), visual, sensorimotor and subcortical/limbic regions (one each), and another cerebellar region (one). Aggregating across the full set of 20 top-ranked connections (40 ROI endpoints), Table 8 breaks down the network-level distribution underlying this pattern.

Table 8
Network-level distribution of the top 20 candidate fMRI biomarker connections (40 ROI endpoints)

Network	**ROI endpoints (of 40)**	**Percentage**
Cerebellar	10	25.0%
Other / Association cortex	10	25.0%
Subcortical / Limbic	6	15.0%
Visual	5	12.5%
Sensorimotor	4	10.0%
Salience / Cingulo-opercular	3	7.5%
Fronto-parietal control	2	5.0%

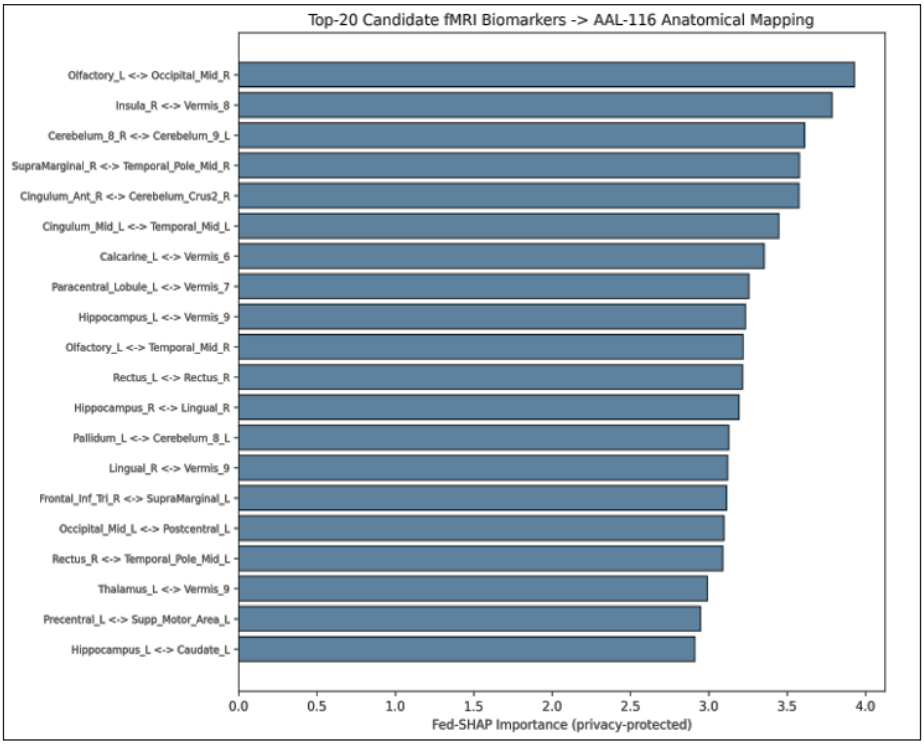


Figure 6. Top 20 candidate fMRI connectivity biomarkers ranked by privacy-preserving Fed-SHAP analysis, mapped to AAL-116 anatomical regions

The cerebellum and "other/association cortex" regions are tied as the most frequently implicated networks (25.0% each), followed by subcortical/limbic structures (15.0%), visual cortex (12.5%), sensorimotor cortex (10.0%), salience/cingulo-opercular regions (7.5%), and fronto-parietal control regions (5.0%). The prominence of cerebellar involvement is broadly consistent with a growing body of ASD neuroimaging literature implicating cerebello-cortical circuits, and the anterior cingulate–cerebellar (Crus II) connection at rank 5 is broadly in line with the cingulate–cerebellar connectivity differences reported by Alves et al. (2023), although that study implicated the posterior rather than the anterior cingulate, although we emphasise that our study was not designed to test any specific prior hypothesis about network involvement and this correspondence should be read as broadly consistent rather than as independent confirmation.

We deliberately report this network-level summary as the primary biomarker finding rather than treating any individual connection as a stable discovery, for three converging reasons. First, bootstrap resampling (200 resamples of the subject pool, recomputing the importance statistic each time) showed that only 20.6% of the top 20 biomarkers remained in the top 20 across resamples - meaning that on a typical resample, roughly 16 of the 20

"top" connections would be replaced by different connections. Second, two independent full training runs of the identical pipeline (same architecture, same hyperparameters, same $\varepsilon = 1.0$ budget) produced entirely non-overlapping top 20 connection lists - not a single specific connection appeared in both runs' top 20, despite both runs achieving similar overall accuracy. Third, the standard-SHAP validation correlation was only modest ($\rho = 0.227$), which further limits how much weight to give any single connection's rank. Taken together, we interpret individual connections as exploratory candidate biomarkers and sample-size-limited, while treating the broader network-level pattern (cerebellar and association-cortex involvement, consistently appearing across our inspection of both runs, even though the specific connections differed) as the more robust observation warranting follow-up on a larger cohort. A structural-versus-functional modality effect size of Cohen's $d = 0.562$ indicates a moderate difference in importance magnitude between the two modalities, consistent with the ablation finding that sMRI alone carries substantial discriminative information at this sample size, and with the fMRI-only ablation arm's comparatively weaker and less fair (higher-variance) performance reported above.

Figure 7 provides a more detailed view of the underlying Fed-SHAP importance statistics for each modality. The top 10 cumulative importance is 34.98 for fMRI versus 30.46 for sMRI, and both modalities' cumulative-importance curves cross the 80% threshold at a similar point (approximately the 15th-16th ranked feature of 20), indicating that importance is concentrated in a broadly similar number of features for each modality rather than being markedly more or less concentrated in one. The per-modality Fed-SHAP importance distributions differ more clearly: the fMRI distribution is sharply peaked and centred near zero (mean = 0.044, SD = 0.819), whereas the sMRI distribution is shifted well above zero (mean = 0.710, SD = 0.916) - consistent with structural features contributing a more consistent, one-directional signal, while functional-connectivity features contribute a signal that is on average small but occasionally large in either direction. The fMRI/sMRI importance-ratio histogram shows every sampled ratio exceeding 1.0 (the "equal importance" reference line), clustered mostly between 1.15 and 1.20, indicating that the highest-ranked fMRI features carry slightly larger absolute importance than their sMRI counterparts. This does not contradict the moderate effect size favouring structural features (Cohen's $d = 0.562$): the effect size reflects the consistently positive, one-directional sMRI contributions shown in the distribution panel (mean 0.710 versus 0.044 for fMRI), whereas the ratio reflects the magnitude of the top-ranked features in each modality.

Figure 8 aggregates the fMRI connection-level importance scores into a per-ROI view, summing each region's involvement across all its connections rather than considering individual connections in isolation as Table 7 does. This reveals a small number of high-degree "hub" regions - the top-ranked hub carries roughly 9.2 units of total connectivity strength, more than the next several hubs combined (each in the 6–7-unit range), with a

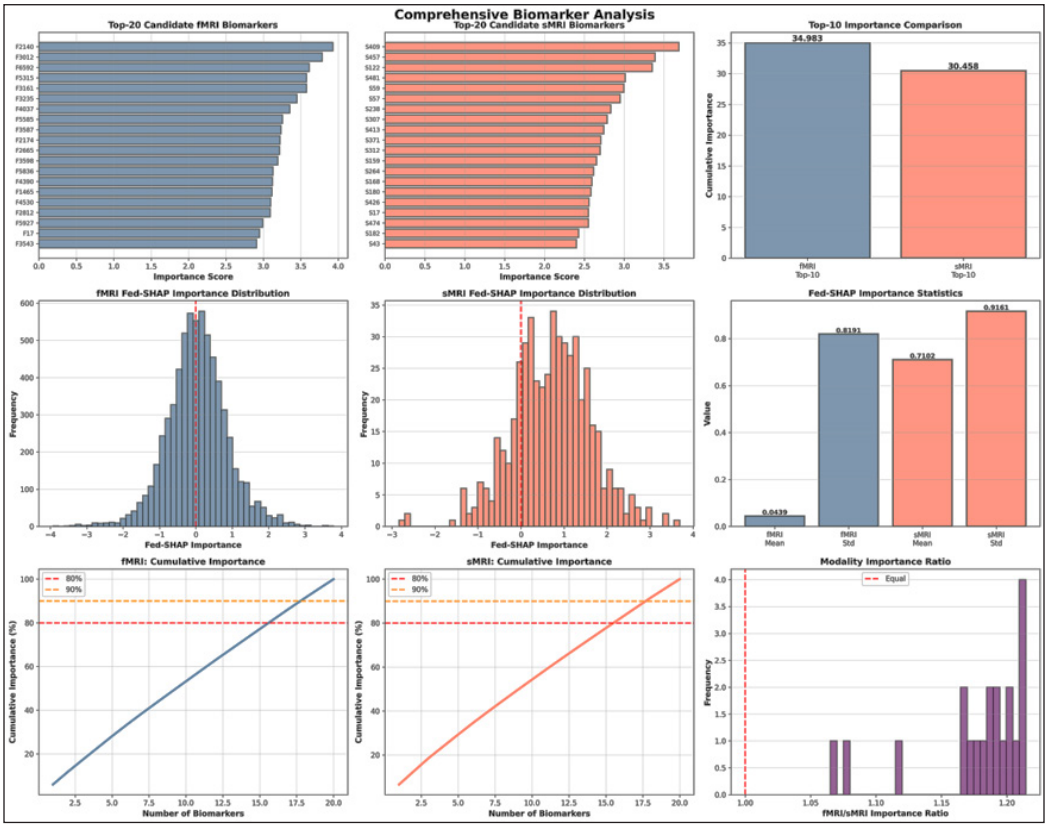


Figure 7. Comprehensive biomarker analysis: per-modality Fed-SHAP importance distributions, cumulative importance curves, and the fMRI/sMRI importance ratio

long tail of regions carrying progressively less. The degree-distribution panel shows this concentration even more starkly: the great majority of the 116 regions have near-zero connectivity strength in the top 20 biomarker set, with only a small subset participating at all, and the hierarchical clustering panel groups the participating regions into a small number of clusters at increasing linkage distance, broadly consistent with the network-level grouping already reported in Table 8. We caution that the ROI numbering used internally by this particular analysis was not independently re-verified against the AAL-116 labelling convention used in Table 7 and Table 8, so we present Figure 8 as a corroborating, region-level view of the same underlying importance scores rather than as a source of additional named-region claims.

Figure 9 reports a consistency check that differs from the bootstrap analysis above. Whereas the bootstrap score asks whether the same top 20 connections reappear when subjects are resampled, the bottom-right panel of Figure 9 asks how consistently the three sites agree on the importance of each top 20 connection ranked by a single trained model.

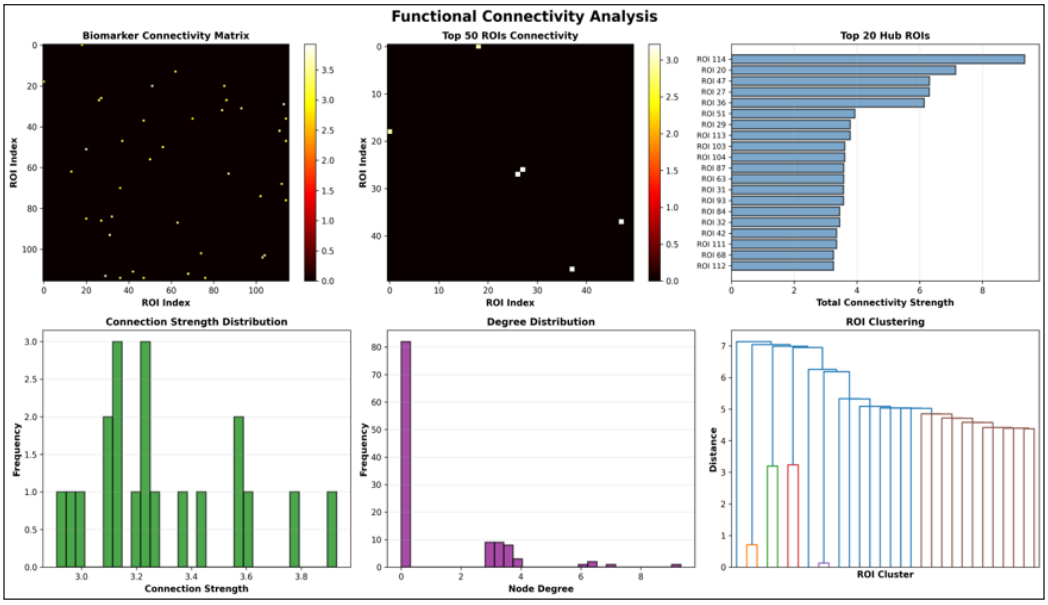


Figure 8. Functional connectivity analysis: biomarker connectivity matrix, hub-ROI ranking, and hierarchical clustering of implicated regions

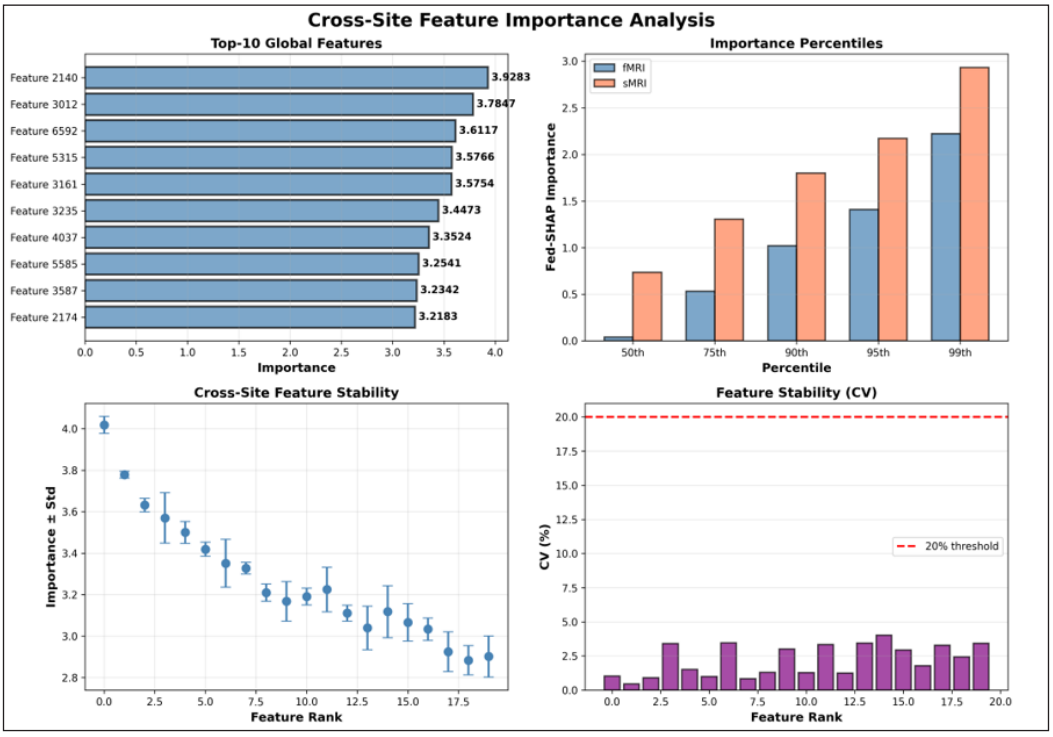


Figure 9. Cross-site feature importance and consistency: top 10 global features, importance percentiles by modality, and cross-site coefficient of variation for the top 20 candidate biomarkers

Here, every one of the top 20 features shows a cross-site coefficient of variation under 5%, well within the 20% threshold marked on the panel, indicating that once a set of top features has been identified, the three sites agree closely on how important each one is. These two findings are not contradictory: a fixed, trained model can show strong cross-site agreement on its own top features (Figure 9), while the identity of those top features remains highly sensitive to which subjects are sampled or which training run produced the model (the bootstrap and cross-run findings above). In practical terms, this suggests the instability we report lies primarily in feature selection (which connections rise to the top) rather than in feature scoring (how consistently sites agree on a score once a feature has been selected) - a distinction that may be useful for any follow-up work aiming to stabilise the biomarker discovery process itself.

DISCUSSION

This study set out to test whether federated learning, differential privacy, and explainable AI could be combined for collaborative autism biomarker discovery across simulated hospital sites without centralising patient data or sacrificing interpretability. Under a leak-free validation protocol, the proposed framework achieves 92.6% held-out accuracy (95% CI [81.5%, 100.0%]) and 86.7% mean leave-one-site-out accuracy, at a formal ($\varepsilon = 1.0$, $\delta = 10^{-5}$) differential-privacy guarantee scoped to the classifier head.

It is important to emphasise that Fed-SHAP is a privacy-preserving feature-importance approximation rather than a direct computation of Shapley values: it ranks features by a privatised between-group difference in mean activation, and its modest agreement with standard SHAP ($\rho = 0.227$) means that its scores should not be interpreted as classical SHAP values. With that caveat, the candidate biomarkers it highlighted show an aggregate network-level pattern (predominantly cerebellar and association-cortex connections) that is broadly consistent with prior ASD connectivity literature, although, given the low individual-level stability reported in Results, they remain exploratory. The ablation results indicate that the value of multimodal fusion and cross-modal attention cannot yet be resolved at this sample size, since several configurations, including the sMRI-only pathway, are statistically indistinguishable from the full model; any added value of fusion may lie in complementary biomarker information not fully captured by raw classification accuracy, a hypothesis that would require a substantially larger cohort to test.

LIMITATIONS

Several limitations should be weighed when interpreting these results. First, the held-out test set (27 subjects total, 9 per site) is small enough that a single subject's classification changes measured accuracy by approximately 3.7 percentage points; several ablation configurations and privacy-budget settings that differ by only one or two subjects' worth

of accuracy should be considered statistically indistinguishable rather than meaningfully ordered, as illustrated by the run-to-run difference for the identical proposed configuration (92.6% in Table 3 vs. 88.9% in Table 5). Second, because the ranking is unstable at the individual-connection level, the named connections are only candidate biomarkers and should not be treated as validated diagnostic markers without confirmation in a larger, independent cohort. Third, the formal differential-privacy guarantee applies only to the classifier head because protecting the full network destroys the predictive signal across all tested budgets (see Methodology). This reflects a general challenge of differentially private deep learning on small cohorts, but it means that the privacy claim is narrower than full-model privacy. Fourth, the three sites in this study are simulated partitions of a single pooled cohort with a shared pre-processing pipeline, rather than genuinely independent institutions with different scanners and populations; the leave-one-site-out result (86.7%) is therefore likely to be an optimistic estimate, rather than a direct measurement, of cross-institution generalisation. We consider validating this framework on a larger, genuinely multi-institutional cohort - sufficient to resolve individual architectural and privacy-budget contributions with statistical confidence, and to assess biomarker stability independently of the specific subjects sampled - to be the most important direction for future work.

CONCLUSION

This work presents a federated learning framework for collaborative, privacy-preserving autism biomarker discovery that combines structural and functional MRI via cross-modal attention, applies differential privacy to both model aggregation and biomarker computation, and uses Fed-SHAP, an anatomically mappable, privacy-preserving feature-importance approximation (rather than a direct computation of Shapley values), for interpretability. Under a leak-free validation protocol, the framework achieves a held-out accuracy of 92.6% and a leave-one-site-out accuracy of 86.7%, at a formal differential-privacy guarantee of ($\varepsilon = 1.0$, $\delta = 10^{-5}$) scoped to the classifier head. We report an ablation study, an expanded baseline comparison, and a privacy-utility sweep in full, including the finding that several configurations are statistically indistinguishable at the current sample size. We also report candidate biomarkers primarily at the network level, given their limited stability at the individual-connection level. We regard this work as a proof of concept demonstrating the technical feasibility of combining federated learning, differential privacy, and explainable AI for multi-site neuroimaging research, rather than as a clinically validated diagnostic tool. Future work should prioritise validation on a larger, genuinely independent multi-institutional cohort; a tighter accounting of the privacy-utility trade-off using per-example gradient clipping (differentially private stochastic gradient descent, DP-SGD) rather than aggregate model-update clipping, which may better preserve signal at tight privacy budgets; incorporation of longitudinal and phenotypic data alongside neuroimaging; and prospective validation studies prior to any clinical application.

ACKNOWLEDGEMENT

The authors sincerely thank the institute management for their support and encouragement and express their gratitude to the research supervisor and Doctoral Research Committee members for their valuable guidance, suggestions, and constructive feedback throughout the research work.

LIST OF ABBREVIATION

ASD	Autism Spectrum Disorder
AI	Artificial Intelligence
XAI	Explainable Artificial Intelligence
MRI	Magnetic Resonance Imaging
sMRI	Structural Magnetic Resonance Imaging
fMRI	Functional Magnetic Resonance Imaging
ABIDE	Autism Brain Imaging Data Exchange
AUC	Area Under the (ROC) Curve
SVM	Support Vector Machine
FedAvg	Federated Averaging
SHAP	SHapley Additive exPlanations
Fed-SHAP	Federated SHAP-based feature-importance approximation
Ex-MMDL	Explainable Multi-Modal Deep Learning
CNN	Convolutional Neural Network
Grad-CAM	Gradient-weighted Class Activation Mapping
AAL	Automated Anatomical Labelling (atlas)
ROI	Region of Interest
CV	Coefficient of Variation
DP	Differential Privacy
RDP	Rényi Differential Privacy
LOSO	Leave-One-Site-Out (validation)
GNN	Graph Neural Network
SGD	Stochastic Gradient Descent
LASSO	Least Absolute Shrinkage and Selection Operator
BASC	Bootstrap Analysis of Stable Clusters
CI	Confidence Interval

REFERENCES

Ahammed, M. S., Niu, S., Ahmed, M. R., Dong, J., Gao, X., & Chen, Y. (2021). DarkASDNet: Classification of ASD on functional MRI using deep neural network. *Frontiers in Neuroinformatics, 15*, Article 635657. https://doi.org/10.3389/fninf.2021.635657

Al-Hiyali, M. I., Yahya, N., Faye, I., & Hussein, A. F. (2021). Identification of autism subtypes based on wavelet coherence of BOLD fMRI signals using convolutional neural network. *Sensors, 21*(16), Article 5256. https://doi.org/10.3390/s21165256

Al-Hiyali, M. I., Yahya, N., Faye, I., & Khan, Z. (2021). Autism spectrum disorder detection based on wavelet transform of BOLD fMRI signals using pre-trained convolution neural network. *International Journal of Integrated Engineering, 13*(5), 49-56. https://doi.org/10.30880/ijie.2021.13.05.006

Almuqhim, F., & Saeed, F. (2021). ASD-SAENet: A sparse autoencoder, and deep-neural network model for detecting autism spectrum disorder (ASD) using fMRI data. *Frontiers in Computational Neuroscience, 15*, Article 654315. https://doi.org/10.3389/fncom.2021.654315

Alutaibi, A. I., Sharma, S. K., & Khan, A. R. (2025). Capsule DenseNet++: Enhanced autism detection framework with deep learning and reinforcement learning-based lifestyle recommendation. *Computers in Biology and Medicine, 190*, Article 110038. https://doi.org/10.1016/j.compbiomed.2025.110038

Alves, C. L., Toutain, T. G. L. D. O., de Carvalho Aguiar, P., Pineda, A. M., Roster, K., Thielemann, C., Porto, J. A. M., & Rodrigues, F. A. (2023). Diagnosis of autism spectrum disorder based on functional brain networks and machine learning. *Scientific Reports, 13*(1), Article 8072. https://doi.org/10.1038/s41598-023-34650-6

Alzubi, O. A., Alzubi, J. A., Shankar, K., & Gupta, D. (2021). Blockchain and artificial intelligence enabled privacy-preserving medical data transmission in Internet of Things. *Transactions on Emerging Telecommunications Technologies, 32*(12), Article e4360. https://doi.org/10.1002/ett.4360

Andrew, G., Thakkar, O., McMahan, H. B., & Ramaswamy, S. (2021). Differentially private learning with adaptive clipping. In *Advances in neural information processing systems* (Vol. 34, pp. 17455-17466). https://proceedings.nips.cc/paper/2021/hash/91cff01af640a24e7f9f7a5ab407889f-Abstract.html

Banerjee, T. (2025). Electromagnetic interaction algorithm (EIA)-based feature selection with adaptive kernel attention network (AKAttNet) for autism spectrum disorder classification. *International Journal of Developmental Neuroscience, 85*(5), Article e70034. https://doi.org/10.1002/jdn.70034

Bayram, M. A., Özer, İ., & Temurtaş, F. (2021). Deep learning methods for autism spectrum disorder diagnosis based on fMRI images. *Sakarya University Journal of Computer and Information Sciences, 4*(1), 142-155. https://doi.org/10.35377/saucis.04.01.879735

Belhaouari, S. B., Talbi, A., Hassan, S., Al-Thani, D., & Qaraqe, M. (2023). PFT: A novel time-frequency decomposition of BOLD fMRI signals for autism spectrum disorder detection. *Sustainability, 15*(5), Article 4094. https://doi.org/10.3390/su15054094

Chandra, A., Verma, S., Raghuvanshi, A. S., & Bodhey, N. K. (2024). ASDC-Net: Optimized convolutional neural network-based automatic autism spectrum disorder classification using rs-fMRI data. *IETE Journal of Research, 70*(4), 4189-4202. https://doi.org/10.1080/03772063.2023.2196979

Eslami, T., Mirjalili, V., Fong, A., Laird, A. R., & Saeed, F. (2019). ASD-DiagNet: A hybrid learning approach for detection of autism spectrum disorder using fMRI data. *Frontiers in Neuroinformatics, 13*, Article 70. https://doi.org/10.3389/fninf.2019.00070

Eslami, T., Raiker, J. S., & Saeed, F. (2021). Explainable and scalable machine learning algorithms for detection of autism spectrum disorder using fMRI data. In A. S. El-Baz & J. S. Suri (Eds.), *Neural engineering techniques for autism spectrum disorder: Volume 1: Imaging and signal analysis* (pp. 39-54). Academic Press. https://doi.org/10.1016/b978-0-12-822822-7.00004-1

Eslami, T., & Saeed, F. (2019). Auto-ASD-Network: A technique based on deep learning and support vector machines for diagnosing autism spectrum disorder using fMRI data. In *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics* (pp. 646-651). Association for Computing Machinery. https://doi.org/10.1145/3307339.3343482

Feng, M., & Xu, J. (2023). Detection of ASD children through deep-learning application of fMRI. *Children, 10*(10), Article 1654. https://doi.org/10.3390/children10101654

Feng, W., Liu, G., Zeng, K., Zeng, M., & Liu, Y. (2022). A review of methods for classification and recognition of ASD using fMRI data. *Journal of Neuroscience Methods, 368*, Article 109456. https://doi.org/10.1016/j.jneumeth.2021.109456

Firouzi, M. H., & Fadaei, S. (2025). Deep learning-based classification of autism spectrum disorder using resting state fMRI data. *International Journal of Engineering, 38*(4), 785-795. https://doi.org/10.5829/ije.2025.38.04a.10

Gheisari, M., Najafabadi, H. E., Alzubi, J. A., Gao, J., Wang, G., Abbasi, A. A., & Castiglione, A. (2021). OBPP: An ontology-based framework for privacy-preserving in IoT-based smart city. *Future Generation Computer Systems, 123*, 1-13. https://doi.org/10.1016/j.future.2021.01.028

Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., & Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical, 17*, 16-23. https://doi.org/10.1016/j.nicl.2017.08.017

Huang, Z.-A., Zhu, Z., Yau, C. H., & Tan, K. C. (2021). Identifying autism spectrum disorder from resting-state fMRI using deep belief network. *IEEE Transactions on Neural Networks and Learning Systems, 32*(7), 2847-2861. https://doi.org/10.1109/tnnls.2020.3007943

Islam, M. M., Hussain, A. H., Mariam, Q. R., Islam, S., & Hasan, M. N. (2025). AI-enabled predictive health monitoring for children with autism using IoT and machine learning to detect behavioral changes. *Perinatal Journal, 33*(1), 415-422. https://doi.org/10.57239/prn.25.03310048

Jabbar, U., Iqbal, M. W., Alourani, A., Shinan, K., Alanazi, F., Sarwar, N., & Ashraf, M. U. (2025). Machine learning–based approach for early screening of autism spectrum disorders. *Applied Computational Intelligence and Soft Computing, 2025*(1), Article 9975499. https://doi.org/10.1155/acis/9975499

Jönemo, J., Abramian, D., & Eklund, A. (2023). Evaluation of augmentation methods in classifying autism spectrum disorders from fMRI data with 3D convolutional neural networks. *Diagnostics, 13*(17), Article 2773. https://doi.org/10.3390/diagnostics13172773

Kang, L., Chen, J., Huang, J., & Jiang, J. (2023). Autism spectrum disorder recognition based on multi-view ensemble learning with multi-site fMRI. *Cognitive Neurodynamics, 17*(2), 345-355. https://doi.org/10.1007/s11571-022-09828-9

Kang, L., Chen, M., Huang, J., & Xu, J. (2025). Identifying autism spectrum disorder based on machine learning for multi-site fMRI. *Journal of Neuroscience Methods, 416*, Article 110379. https://doi.org/10.1016/j.jneumeth.2025.110379

Karampasi, A., Kakkos, I., Miloulis, S.-T., Zorzos, I., Dimitrakopoulos, G. N., Gkiatis, K., Asvestas, P., & Matsopoulos, G. (2020). A machine learning fMRI approach in the diagnosis of autism. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 3628-3631). IEEE. https://doi.org/10.1109/bigdata50022.2020.9378453

Karunakaran, P., Hamdan, Y. B., & Sathesh, A. (2020). Early prediction of autism spectrum disorder by computational approaches to fMRI analysis with early learning technique. *Journal of Artificial Intelligence and Capsule Networks, 2*(4), 207-216. https://doi.org/10.36548/jaicn.2020.4.003

Kohli, M., Kar, A. K., Sinha, S., & Kohli, S. (2025). Stakeholder perception towards a machine-learning-based digital platform for detection and management of autism spectrum disorder. *Expert Systems, 42*(1), Article e13545. https://doi.org/10.1111/exsy.13545

Li, X., Dvornek, N. C., Zhuang, J., Ventola, P., & Duncan, J. S. (2018). Brain biomarker interpretation in ASD using deep learning and fMRI. In A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, & G. Fichtinger (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2018* (pp. 206-214). Springer. https://doi.org/10.1007/978-3-030-00931-1_24

Li, Y., Li, H., Hu, C., Cui, J., Zhang, F., Zhao, J., Feng, Y., Hu, C., Yang, L., Qian, H., Pan, J., Luo, X., Tang, Z., & Hao, Y. (2025). The role of the dopamine system in autism spectrum disorder revealed using machine learning: An ABIDE database–based study. *Cerebral Cortex, 35*(2), Article bhaf022. https://doi.org/10.1093/cercor/bhaf022

Mironov, I. (2017). Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)* (pp. 263-275). IEEE. https://doi.org/10.1109/csf.2017.11

Movassagh, A. A., Alzubi, J. A., Gheisari, M., Rahimi, M., Mohan, S., Abbasi, A. A., & Nabipour, N. (2023). Artificial neural networks training algorithm integrating invasive weed optimization with differential evolutionary model. *Journal of Ambient Intelligence and Humanized Computing, 14*(5), 6017-6025. https://doi.org/10.1007/s12652-020-02623-6

Nafisah, I., Mahmoud, N., Ewees, A. A., Khattap, M. G., Dahou, A., Alghamdi, S. M., Fares, I. A., Azmi Al-Betar, M., & Abd Elaziz, M. (2025). Deep learning-based feature selection for detection of autism spectrum disorder. *Frontiers in Artificial Intelligence, 8*, Article 1594372. https://doi.org/10.3389/frai.2025.1594372

Narayanan, N., K R, R., & Varghese, B. M. (2025). Early detection of autism spectrum disorder via deep-learning application of fMRI and machine learning for ASD children identifications. *International Journal of Hybrid Intelligent Systems, 21*(2), 148-162. https://doi.org/10.3233/his-240029

Pinto, M. M. V., Arisawa, E. A. L. S., Raniero, L. J., & Bhattacharjee, T. (2025). Saliva FTIR spectra and machine learning for autism spectrum disorder diagnosis preliminary study. *IEEE Photonics Journal, 17*(3), Article 8500504. https://doi.org/10.1109/jphot.2025.3561020

Qayyum, A., Ahamed Khan, M. K. A., Benzinou, A., Mazher, M., Ramasamy, M., Aramugam, K., Deisy, C., Sridevi, S., & Suresh, M. (2022). An efficient 1DCNN–LSTM deep learning model for assessment and classification of fMRI-based autism spectrum disorder. In J. S. Raj, K. Kamel, & P. Lafata (Eds.),

Innovative data communication technologies and application: Proceedings of ICIDCA 2021 (pp. 1039-1048). Springer. https://doi.org/10.1007/978-981-16-7167-8_77

Qiang, N., Gao, J., Dong, Q., Li, J., Zhang, S., Liang, H., Sun, Y., Ge, B., Liu, Z., Wu, Z., Liu, T., Yue, H., & Zhao, S. (2023). A deep learning method for autism spectrum disorder identification based on interactions of hierarchical brain networks. *Behavioural Brain Research, 452*, Article 114603. https://doi.org/10.1016/j.bbr.2023.114603

Qiqieh, I., Alzubi, J. A., Alzubi, O. A., Khawaja, Y., & Singh, A. (2026). Securing next-generation connected healthcare: A privacy-preserving explainable AI framework for consumer electronics. *IEEE Transactions on Consumer Electronics, 72*(2), 4615-4622. https://doi.org/10.1109/tce.2026.3675727

Ramli, N. S. H., Ahmad Radzi, S., Abdul Hamid, N., Wong, Y. C., & Ramlan, A. R. (2021). Functional magnetic resonance imaging for autism spectrum disorder detection using deep learning. *Jurnal Teknologi, 83*(3), 45-52. https://doi.org/10.11113/jurnalteknologi.v83.16389

Sabegh, A. M., Samadzadehaghdam, N., Seyedarabi, H., & Ghadiri, T. (2023). Automatic detection of autism spectrum disorder based on fMRI images using a novel convolutional neural network. *Research on Biomedical Engineering, 39*(2), 407-413. https://doi.org/10.1007/s42600-023-00275-x

Sachdeva, J., Mittal, R., Mehta, J., Jain, R., & Ranjan, A. (2024). Resolving autism spectrum disorder (ASD) through brain topologies using fMRI dataset with multi-layer perceptron (MLP). *Psychiatry Research: Neuroimaging, 343*, Article 111858. https://doi.org/10.1016/j.pscychresns.2024.111858

Santana, C. P., de Carvalho, E. A., Rodrigues, I. D., Bastos, G. S., de Souza, A. D., & de Brito, L. L. (2022). Rs-fMRI and machine learning for ASD diagnosis: A systematic review and meta-analysis. *Scientific Reports, 12*(1), Article 6030. https://doi.org/10.1038/s41598-022-09821-6

Shao, L., Fu, C., You, Y., & Fu, D. (2021). Classification of ASD based on fMRI data with deep learning. *Cognitive Neurodynamics, 15*(6), 961-974. https://doi.org/10.1007/s11571-021-09683-0

Shazman, S., Carmel, J., Itkin, M., Malitsky, S., Shalan, M., Soreq, E., Elliott, E., Lebow, M., & Kuperman, Y. (2025). Urine metabolomic profiling and machine learning in autism spectrum disorder diagnosis: Toward precision treatment. *Metabolites, 15*(5), Article 332. https://doi.org/10.3390/metabo15050332

Subah, F. Z., Deb, K., Dhar, P. K., & Koshiba, T. (2021). A deep learning approach to predict autism spectrum disorder using multisite resting-state fMRI. *Applied Sciences, 11*(8), Article 3636. https://doi.org/10.3390/app11083636

Tang, Y., Tong, G., Xiong, X., Zhang, C., Zhang, H., & Yang, Y. (2023). Multi-site diagnostic classification of autism spectrum disorder using adversarial deep learning on resting-state fMRI. *Biomedical Signal Processing and Control, 85*, Article 104892. https://doi.org/10.1016/j.bspc.2023.104892

Xu, W., Li, H., Li, J., & Jin, M. (2025). Age of machine learning: New trends in autism spectrum disorder prediction. *Frontiers in Microbiology, 16*, Article 1492484. https://doi.org/10.3389/fmicb.2025.1492484

Yin, W., Mostafa, S., & Wu, F.-X. (2021). Diagnosis of autism spectrum disorder based on functional brain networks with deep learning. *Journal of Computational Biology, 28*(2), 146-165. https://doi.org/10.1089/cmb.2020.0252

Yousefian, A., Shayegh, F., & Maleki, Z. (2023). Detection of autism spectrum disorder using graph representation learning algorithms and deep neural network, based on fMRI signals. *Frontiers in Systems Neuroscience, 16*, Article 904770. https://doi.org/10.3389/fnsys.2022.904770

Zhang, J., Feng, F., Han, T., Gong, X., & Duan, F. (2023). Detection of autism spectrum disorder using fMRI functional connectivity with feature selection and deep learning. *Cognitive Computation, 15*(4), 1106-1117. https://doi.org/10.1007/s12559-021-09981-z

Pertanika J. Sci. & Technol. 34 (5): 2981 - 3002 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Diseases of Vegetable Plant Recognition Based on Enhanced SqueezeNet

Kifaa Hadi Thanoon[1], Akram A. Abdulrahman[2], Ali Ibrahim Ahmed[3], Omar Ibrahim Alsaif[4]*, and Ibrahim Ahmed Saleh[5]

[1]Department of Cyber Security Techniques Engineering, Northern Technical University, Mosul 41002, Nineveh, Iraq
[2]Department of Computer Science, Alnoor University, Mosul 41002, Nineveh, Iraq
[3]Department of Computer Science, College of Education, University of Telafer, Telafer, Nineveh, Iraq
[4]Department of Electronic and Communication Technology Engineering, Polytechnic College Mosul, Northern Technical University, Mosul 41002, Nineveh, Iraq
[5]Department of Software Engineering, College of Computer Science and Mathematics, University of Mosul, Mosul 41002, Nineveh, Iraq

ABSTRACT

Agricultural resources significantly affect food security, and plant diseases cause substantial production losses and economic losses. This paper presents a method for detecting plant diseases using the deep learning (DL) algorithm SqueezeNet. To do this, the SqueezeNet network was modified to detect plant diseases. Yet one disadvantage of SqueezeNet is that it has many parameters and achieves quite low accuracy, which limits its use due to its complexity. A new advanced network based on "improved squeeze architecture" was suggested for plant disease detection. This paper, however, presented the network by simplifying its architecture, thereby enhancing its feature-recognition capabilities. Moreover, changes were made to the fire module, including the size of the convolution kernel, the number of network layers, and the number of channels in the Expand layer. Also, the model was combined with the Efficient Channel Attention (ECA) module, which assigns attention values to each channel via local cross-channel interactions, thereby improving the network's ability to learn salient features. The algorithm was compared with three others—LeNet, MobileNet, and traditional SqueezeNet—using the "PlantVillage" dataset. The proposed model achieved a very high accuracy, exceeding the traditional SqueezeNet by 93.8% and reaching a minimum of 84%.

Keywords: Convolutional neural network, ECA module, lightweight, plant disease, SqueezeNet structure

ARTICLE INFO

Article history:
Received: 08 July 2026
Accepted: 28 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.18

kh.thanoon@ntu.edu.iq (Kifaa Hadi Thanoon)
akram30@alnoor.edu.iq (Akram A. Abdulrahman)
ali@uotelafer.edu.iq (Ali Ibrahim Ahmed)
Omar.alsaif@ntu.edu.iq (Omar Ibrahim Alsaif)
i.hadedi@uomosul.edu.iq (Ibrahim Ahmed Saleh)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

In the last few years, global food security has been under severe stress due to different reasons like climate change, the loss of pollinators, and plant diseases, with the latter being the main cause of the drop in the quality and productivity of agricultural products (Guerrero-Ibañez & Reyes-Muñoz, 2023; Panchal et al., 2023; Rahman et al., 2020). When crops are affected by diseases, physiological functions are considerably impaired, leading to stunted growth and poor yields, which in turn negatively impact farmers' economic outcomes. Given that crop plants are typically numerous and small, detecting diseases early is challenging, and by the time they are noticed, the infection may have already spread extensively (Khan et al., 2021; Wang, 2022). Vegetables, which are vital economic crops, are rich in nutrients. Nonetheless, with the expansion of vegetable-planting areas, plant disease incidence has intensified, severely impacting vegetable production (Shoaib et al., 2023; Yang & Liu, 2020). Thereby leading to improved productivity of existing farmland, and ensuring accurate, early plant disease detection and diagnosis is of great importance. Plant diseases pose significant risks to farmers, leading to significant yield losses. The early detection of diseases forms the basis for active diagnosis and control, and the preservation of crop yield and quality. In this paper, plant disease identification technologies are categorised into the following stages:

The human identification approach, which depends on farmers' long-accumulated agricultural knowledge and practical experience, identifies the disease type through visual observation and comparison. Sometimes, farmers could rely on experts for guidance and disease diagnosis, drawing on their extensive inspections, statistics, and measurements (Jaderberg et al., 2014; Zhao et al., 2022). On the other hand, this approach is considered labour-intensive and time-consuming, with several drawbacks, including slow recognition speed, high subjectivity, low accuracy, and an inability to meet the demands of modern agricultural production (Hazarika et al., 2021; Saleh et al., 2021).

The traditional method of machine-learning recognition identifies plant diseases by extracting key features from images using feature engineering, which are then classified by ML classifiers. Usually, this process involves several steps, including plant disease image acquisition, preprocessing, feature extraction, segmentation, and disease class classification. Even though this approach offers considerable benefits over manual methods, such as more consistent results and faster processing, it still has significant limitations, including complex data preprocessing and feature engineering, reliance on manually selected features, relatively low recognition accuracy, and unreliable segmentation. In addition, the identification process still involves a significant amount of human intervention.

Deep learning recognition approach, which can automatically learn complex, "black box" features essential for detection and recognition, thereby eliminating the need for manual feature selection (Mohamed et al., 2022; Oleiwi et al., 2019; Wei & Gao, 2025).

By doing so, it enables an end-to-end system that streamlines the entire process from input to output, providing more accurate and efficient disease identification.

The remaining text of this paper is divided into sections as follows: Section 2 reviews prior research by notable researchers. The method of improving SqueezeNet is discussed in Section 3. Section 4 displays the experimental results, and ultimately, Section 5 includes the conclusion.

LITERATURE REVIEW

In recent years, researchers in the field of plant disease identification have increasingly shifted their focus toward lightweight networks to address the challenges posed by deep networks with numerous layers and complex structures.

In 2023, Jung and other researchers developed a model for detecting crop diseases in the early stages using pairs of diseased and healthy plant images and a CNN consisting of five pre-trained models (Jung et al., 2023). The model followed two stages: crop detection and disease classification. Researchers have achieved a 97.09% accuracy. In 2023, Khan et al. (2023) proposed an application based on MobileNet-V3-small, a DL model for plant disease detection and classification. This study utilised the PlantVillage dataset, which included images of diseased and healthy leaves from 14 crop species and 6 disease groups. The MobileNet-V3-small model has achieved approximately 91.50% test accuracy for leaf classification. This paper has provided a cost-effective solution for farmers to detect diseases early in agricultural applications.

Singh et al. (2022) introduced a model that utilised AI and computer vision for early disease detection in crops such as potatoes, apples, corn, rice, and tomatoes, using the Plant Village dataset (Fu, 2014). The researchers have employed a Bayesian-optimised support vector machine (SVM) for feature extraction and subsequently used a CNN for disease detection. In addition, they have employed a grey-level co-occurrence matrix (GLCM) and histograms of oriented images representing 25 different plants across 58 distinct classes. Multiple network architectures have been trained, with the optimal model achieving a 99.53% performance accuracy. The study concluded with a discussion of the potential of this model as an early-warning tool for plant diseases under real-world cultivation conditions.

It has been demonstrated that combining attention mechanisms with neural network sub-modules can enhance both learning efficiency and overall network performance. However, there are a limited number of studies on the combination of the gradient (HoG) and colour moments. Those 3 feature types — colour, deep features, and texture — have been combined to create hybrid features. A Random Forest classifier has then been utilised to achieve a robust output with minimal features.

Setiawan et al. (2021) presented a paper on the detection of diseases in maize plants using leaf images. They have utilised the PlantVillage dataset, which contains 3852

images across 4 classes: blight, spot, healthy, and rust. This study applied 2 DL networks, SqueezeNet and AlexNet, and compared their performance. Stochastic gradient descent with momentum (SGDM) has been utilised in the optimisation process to improve results. Experimental findings indicate that AlexNet outperforms SqueezeNet in classifying maize leaf diseases.

Mohameth et al. (2020) have utilised 2 ML approaches, applying SVM and KNN algorithms for plant disease classification and feature extraction, and 3 DL models: VGG-16, ResNet-50, and GoogleNet. They have recommended using VGG-16 for image feature extraction and SVM for classification, achieving the highest accuracy of 97.82%. ResNet-50 ranked second with 95.38% accuracy, closely followed by GoogleNet at 95.30%.

Hidayatuloh et al. (2018). Have proposed a study on the early tomato plant disease detection. They have applied CNNs with a variety of architectures, including SqueezeNet, known for its small size. This study is based on a dataset from the Vegetable Crops Research Institute (Balitsa) containing 200 classes. This model could achieve an average identification accuracy of 86.92%.

Ferentinos (2018) developed CNN models for the detection and diagnosis of plant diseases using simple leaf images of healthy and diseased plants. This study utilised a database containing 87,848 integrations of lightweight networks with attention modules. In real-world agricultural production, farmers require solutions for disease identification that can be deployed on mobile devices, necessitating lightweight networks that don't compromise accuracy. This is why the present study improved upon the SqueezeNet model by simplifying its structure and incorporating an attention module. This enhanced model, in addition to occupying less memory, maintains higher accuracy, making it a practical solution for identifying plant diseases. Table 1 summarises related work.

Table 1
Summarised for related works

No.	Author	dataset	What has been done	Accuracy
1	Jung et al., 2023	Plant-Village	Detect and classify plant diseases using a CNN that contains five pre-trained models.	97.09%
2	Khan et al., 2023	Plant-Village	Apply MobileNetV3 to classify plant leaves	91.50
3	Singh et al., 2022	Plant-Village	Classify Diseases of apple, corn, potato, tomato, and rice plants and extract their characteristics using a CNN. Classification using a Bayesian optimised SVM	92.2%
4	Setiawan et al., 2021	Use 3852 images categorised into four groups	Detecting and predicting diseases using e AlexNant and SqueezeNet.	AlexNet = 97.69%, Squeezenet = 44.49%

Table 1 (*continued*)

No.	Author	dataset	What has been done	Accuracy
5	Mohameth et al., 2020	Plant-Village	Apply CNN for feature extraction and SVM and KNN for classification	92.7%
6	(Hidayatuloh et al., 2018)	Plant-Village	Used CNN and SqueezeNet to detect seven types of tomato diseases	86.92%
7	(Ferentinos, 2018)	open database	Apply several models of CNN.	93.8%

MATERIALS AND METHODS

Improved SqueezeNet Neural Network Model

To construct a lightweight, high-precision NN model for plant disease recognition, the classic lightweight network structure SqueezeNet has been modified (Alsaif et al., 2025). In particular, the size of the convolution kernels and the number of channels in the fire module have been adjusted to reduce the network's parameter count. In addition, the Efficient Channel Attention (ECA) module (Hao et al., 2024; Wang et al., 2020) has been incorporated into the network to enhance salient feature extraction and improve recognition efficiency.

Traditional SqueezeNet

SqueezeNet is a lightweight architecture with 50 times fewer parameters than traditional CNNs while maintaining competitive accuracy. The main components of the SqueezeNet architecture are:

Fire Modules : These are the core components of the network, consisting of two layers: the squeeze layer and the expand layer. The squeeze layer applies 1×1 convolution filters to cut down the number of input channels, whereas the expand layer applies a mixture of two filters—one 1×1 convolution and one 3×3 convolution.

Overall Architecture : The architecture begins with a few initial convolutional layers, followed by a series of Fire modules, and concludes with a final convolutional layer, as illustrated in Figure 1.

Reduce Model Size : Use 1×1 convolution filters and downsampling techniques to replace early pooling layers. Pruning techniques are also employed to further reduce the model size.

Batch size : {128, 32, 64}.

Regularization strength : {1e−5, 1e−4, 5e−4}. A value of 1e−4 reduced overfitting without prevention optimisation.

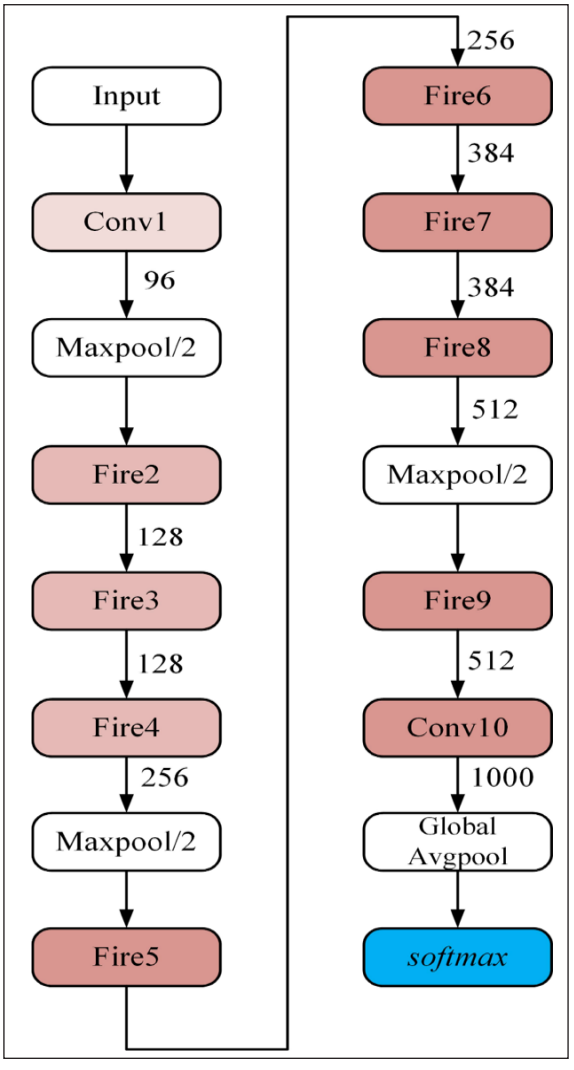


Figure 1. SqueezeNet architecture

Enhanced Fire Module

As the most important component of the SqueezeNet network, the fire module adopts a pioneering approach to reduce both parameter count and overall model size. To further shrink the model, two lightweight strategies to reduce the number of parameters in the network structure. This module divides traditional convolution into a Squeeze layer and an Expand layer. The Squeeze layer utilises 1×1 convolutions, while the Expand layer uses both 1×1 and 3×3 convolutions. This design effectively uses consecutive 1×3 and 3×1 convolutions to replace the original 1×1 and 3×3 convolutions in the fire module, with the number of channels set to half the original value.

The structures are shown in Figures 2 and 3, respectively. (H & W) represent the height and width of the feature map, while (M, S, and E) denote the number of channels. After these improvements, the fire module parameters are given by Equations 1 and 2.

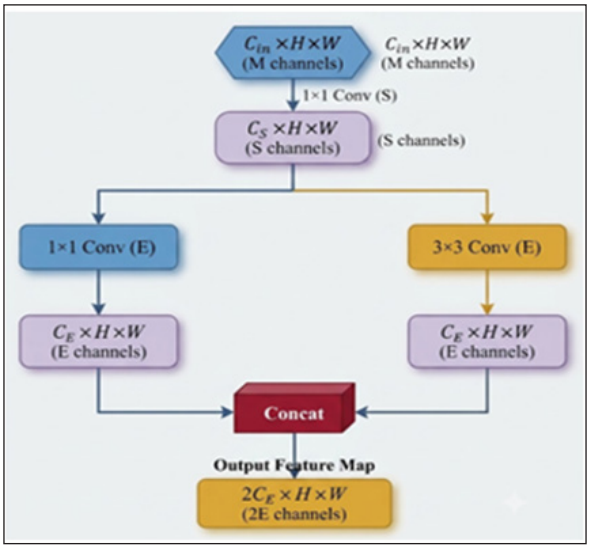


Figure 2. The Fire Module structure

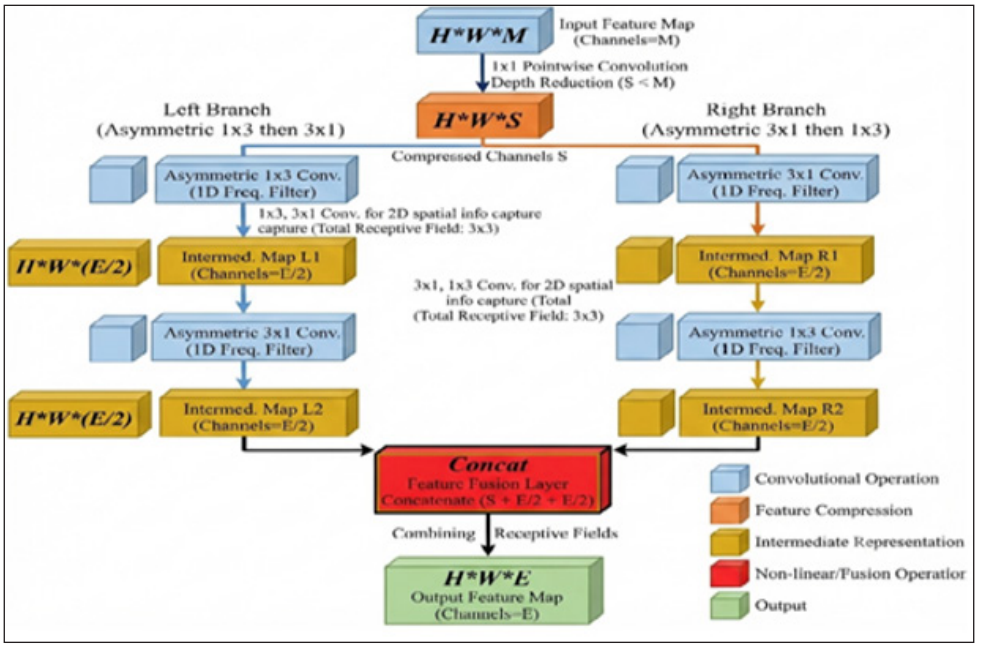


Figure 3. Improved fire module structure

Before the improvement, the parameter quantity of the fire module is

$$N_{before} = S \times 1 \times 1 \times E + S \times 3 \times 3 \times E = 10 \times S \times E \quad [1]$$

The calculated amount is:

$$C_{before} = S \times H \times W \times E \times 1 \times 1 + S \times H \times W \times E \times 3 \times 3 = 10 \times S \times H \times W \times E \quad [2]$$

After improvement, the parameters of the fire module in Equation 3:

$$N_{improve} = 2 \times S \times 1 \times 3 \times \frac{E}{2} + 2 \times \frac{E}{2} \quad 3 \times 1 \times \frac{E}{2} = 3 \times S \times E + \frac{3}{2}E^2 \quad [3]$$

The calculated amount is shown in Equation 4:

$$C_{improve} = 2 \times S \times H \times W \times \frac{E}{2} + 1 \times 3 + 2 \times \frac{E}{2} \times H \times W \times \frac{E}{2} + 3 \times 1 = 3 \times S \times H \times W \times E \times \frac{3}{2} \times H \times W \times E^2 \quad [4]$$

In the SqueezeNet structure, E = 4S. With the improved fire module, the number of parameters and calculations is reduced to 9/10 of the original module, significantly lowering the parameter count in the network. This method effectively reduces the complexity of the classic model.

Efficient Channel Attention (ECA)

Efficient Channel Attention (ECA) is a module based on squeeze-and-excitation blocks, designed to enhance network performance by reducing module complexity without altering dimensionality. To improve the model's ability to learn disease characteristics, the ECA module was added to the SqueezeNet network. This ultra-lightweight attention module assigns distinct attention values to each feature channel, enabling the network to allocate computational resources more efficiently. As a result, the model's parameter count is reduced while accuracy is significantly improved. The ECA module, illustrated in Figure 4, includes the following components:

Global A average pooling : This module compresses global information from the input feature map into a channel descriptor with a global receptive field via the global average pooling layer.

1D convolution : A small kernel size is used to model channel dependencies and capture the correlation between local channels.

Sigmoid function : This function calculates the attention value of each channel based on the correlation between them.

Channel recalibration : This step operates on the input feature map to adjust its channel dimensions. Additionally, the kernel size (K) of the one-dimensional convolution is adaptively determined by the channel dimension, with the calculation formula being as shown in Equation 5.

$$K = \varphi(C) = |\frac{\log_2 C}{r} + \frac{b}{r}|_{odd} \quad [5]$$

Where: C is the number of channels, and "odd" indicates taking the nearest odd number. In this research, r=2, b=1, and K represent the coverage of local cross-channel interaction. Lightweight attention-based networks are designed to realise a balance between computational efficiency and feature representation capability. The networks typically combine compression with attention mechanisms, such as channel or spatial attention, to emphasise informative features while maintaining a low number of parameters. In this context, the proposed method integrates Efficient Channel Attention (ECA) into an enhanced SqueezeNet architecture, aiming to improve disease-feature representation while preserving the lightweight characteristics of the original network.

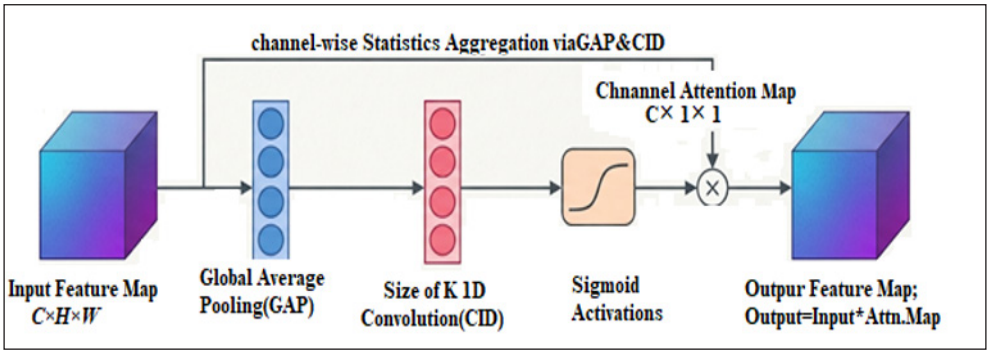


Figure 4. ECA module structure

Enhanced SqueezeNet Network Structure

An enhanced SqueezeNet model is created by combining the improved fire module with the ECA module in three stages, as shown in Figure 5. The first stage involves inputting the disease image from the dataset into the network. Second stage extracts disease features layer by layer through the convolutional layer, pooling layer, and fire module. The final stage processes these extracted feature maps through the ECA module, distinguishing important features from irrelevant ones. The identified important features are used to allocate computing resources efficiently, ultimately achieving higher recognition accuracy with fewer parameters.

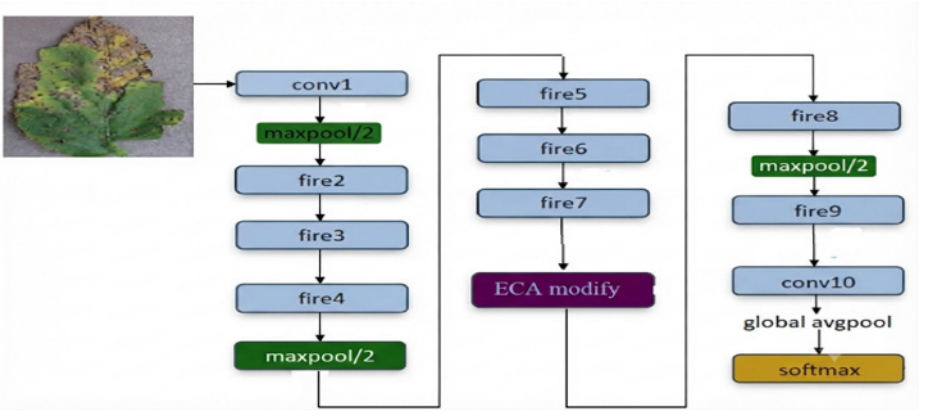


Figure 5. Improved SqueezeNet network architecture

EXPERIMENTAL PARTS

Dataset

This paper relies on the publicly available PlantVillage dataset from Hughes and Salathé (2015) for plant disease images, which includes 54,309 images of healthy and diseased plants across 14 crops. Eight disease categories (including healthy leaves) were selected for identification, with the following distribution: 1,590 healthy leaves, 2,126 bacterial spots, 1,000 early blight, 952 leaf mould, 1,908 late blight, 1,771 spotted blight, 1,676 red spider lesions, and 1,404 leaf spots. Figure 6 illustrates these eight disease types, showing that the leaf backgrounds are simple and free of occlusion.

However, the Plant Village dataset has an uneven category distribution, which can lead to issues in model training, including overfitting. The trained model may be biased toward predicting the category with more samples in the training set, resulting in higher classification errors for minority categories.

To address the data imbalance, data resampling mechanisms are applied to create a more balanced dataset distribution (Fu, 2014). This process helps ensure minimal information

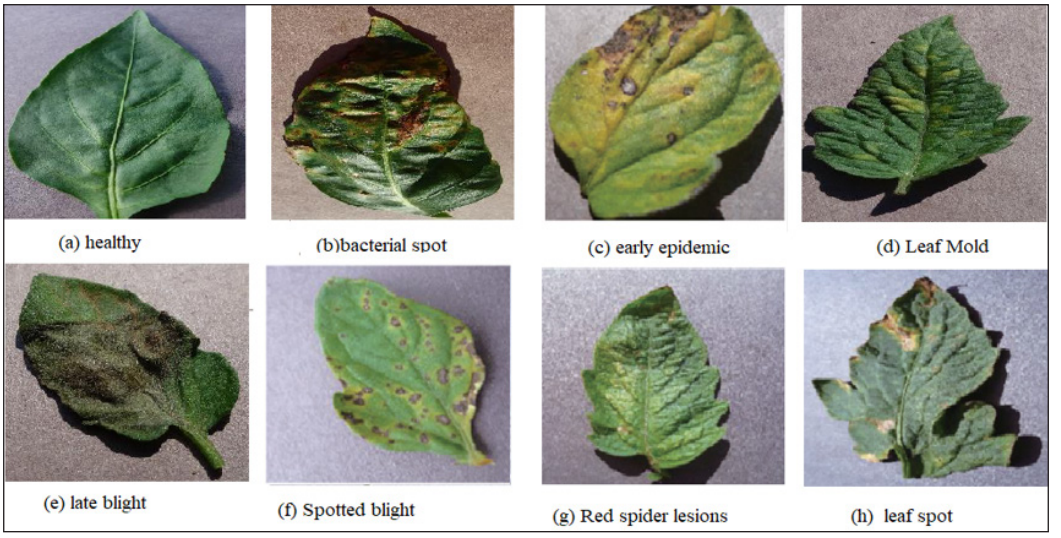


Figure 6. Examples of healthy and diseased plant leaves

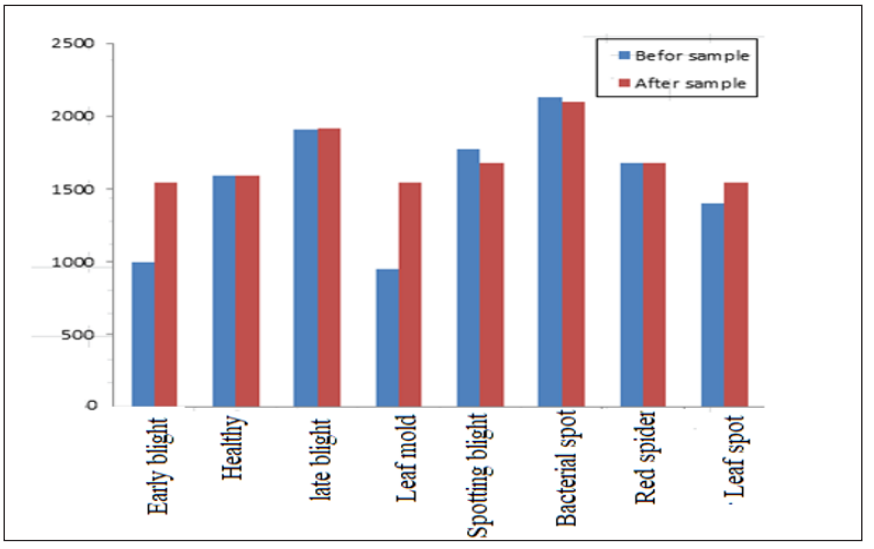


Figure 7. Comparison of data before and after resampling

loss while retaining the most useful data. Specifically, the average number of samples across the eight categories is calculated, and the categories below this average are randomly oversampled by replication, scaling, and other techniques, while the categories above the average remain unchanged. The final dataset includes 13,736 plant disease images. Figure 7 shows the data distribution before and after resampling, with the number of images in each category more balanced, effectively mitigating the sample imbalance between categories.

Model Training Environment Setting

The experiment was conducted on a server with the following hardware configuration: CPU (Intel Xeon E5-2620), 64GB of memory, GPU (NVIDIA GTX 1080 Ti) with 11GB of video memory, and Windows 10. The neural network model was built using the deep learning framework TensorFlow 2.0 within the CUDA environment, in combination with Python 3.6. To comprehensively assess the model's learning performance,

The resampled plant images were divided into training, validation, and test sets in a 7:2:1 ratio, and the neural network was trained accordingly. Based on the total experimental data and empirical values, the number of iterations was set to 100, and the batch sizes for both the training and validation sets were set to 64.

The early stop strategy uses a maximum of 100 training cycles, saves the best checkpoint based on the check loss, and stops training if the loss does not improve for 10 consecutive cycles, thereby ensuring the selection of the model with the best generalisation performance and avoiding over-allocation.

The batch size for the test set was set to 128, and the regularisation coefficient was set to 0.0005. The learning rate and optimiser were selected through experimentation, as outlined in the following section.

Selection of Model Hyperparameters

During training, hyperparameters significantly influence the network's training and recognition performance. Different network structures have different optimal hyperparameters due to varying numbers of layers, parameters, and complexities. To determine the best optimiser for model training, the classic stochastic gradient descent (SGD) method was compared with the Adaptive Moment Estimation (Adam). A systematic hyperparameter search was conducted using a grid search approach over a predefined set of candidate values. The evaluated hyperparameters comprised learning rates of (0.0004), optimisers {SGD, Adam}, batch sizes (64), L2 regularisation coefficients (5×10^{-4}), and early stopping equal 97%. The best trade-off between classification accuracy and training stability was therefore adopted in the remaining experiments. The training process results under each optimiser are shown in Table 2.

Table 2
The training process of the models under the two optimisers

Epoch	Accuracy	
	Adaptive Moment Estimation (Admin)	Stochastic gradient descent (SGD)
20	84.00	18.00
30	86.09	18.98
40	89.12	39.49
50	91.45	56.58
60	92.70	77.45
70	92.82	81.10
80	93.00	85.67
90	96.17	85.32
100	97.29	92.46

For the Adaptive Moment Estimation algorithm (Adam), precision, recall, and f1-score are reported in Table 3.

Table 3
Comparative results for the Adaptive Moment Estimation algorithm improved SqueezeNet

Epoch	Accuracy (%)	Precision (%)	Recall (%)	F-score (%)
20	84.00	88.64	78.00	82.94
30	86.09	89.32	82.00	85.80
40	89.12	91.68	86.00	88.76
50	91.45	92.78	90.00	91.37
60	92.70	93.31	92.00	92.65
70	92.82	92.63	93.00	92.81
80	93.00	92.16	94.00	93.07
90	96.17	97.46%	96.84%	97.15
100	97.29	97.46	96.84	97.15

In the early stages of training, the network's accuracy increased rapidly under the Adam optimiser, with fast convergence and high final accuracy. In contrast, when using the SGD optimiser, the network's accuracy remained unchanged for the first 20 rounds, indicating no improvement. After further learning, accuracy began to increase slowly, starting at 20 rounds. However, even after 100 rounds, the network had not fully converged, and the accuracy remained low. Given the training time and convergence results, Adam was selected as the optimiser for the model. Controlling learning progress is crucial, and hyperparameters such as the learning rate are key to this process.

Tables 2 and 3 show the Adam algorithm as the best optimiser. To increase accuracy, training efficiency metrics were calculated: training time per cycle, actual convergence speed (number of cycles required to reach 95% accuracy), and inference response time, which were used to obtain the final accuracy shown in Table 4.

Table 4
Comparative training performance of optimisers

Optimizer	Final Accuracy (%)	Epochs to 95% Acc.	Training Time/ Epoch (sec)	Totel Training Time(min)	Inference Latency (ms/image)
Admin	97.29	52	42.3	33.8	8.2
SGD	92.46	82	38.7	50.3	7.9

As observed, Adam achieves the best final accuracy (97.29%) with the least training cycles (52) to reach 95% accuracy threshold, demonstrating exceptional convergence speed. Although the SGD algorithm exhibits slightly less training time per cycle (38.7 seconds

vs 42.3 seconds for Adam), its slower convergence rate results in a meaningfully longer total training time,"50.3 minutes vs 33.8 minutes for Adam". This represents a 32.8% reduction in total training time when using Adam. Furthermore, the inference response time remains consistent across all model optimisers (7.9-8.5 ms/image), and the choice of optimiser does not affect the computational cost of the forward-scrolling process after model training. Adam's key advantage lies in its Adaptive Moment Estimation mechanism, which enables faster convergence and more stable training dynamics, as evidenced by the smoother accuracy and loss curves Figure 8. This stability is crucial when training lightweight networks such as the optimised SqueezeNet, where gradient fluctuations can hinder convergence.

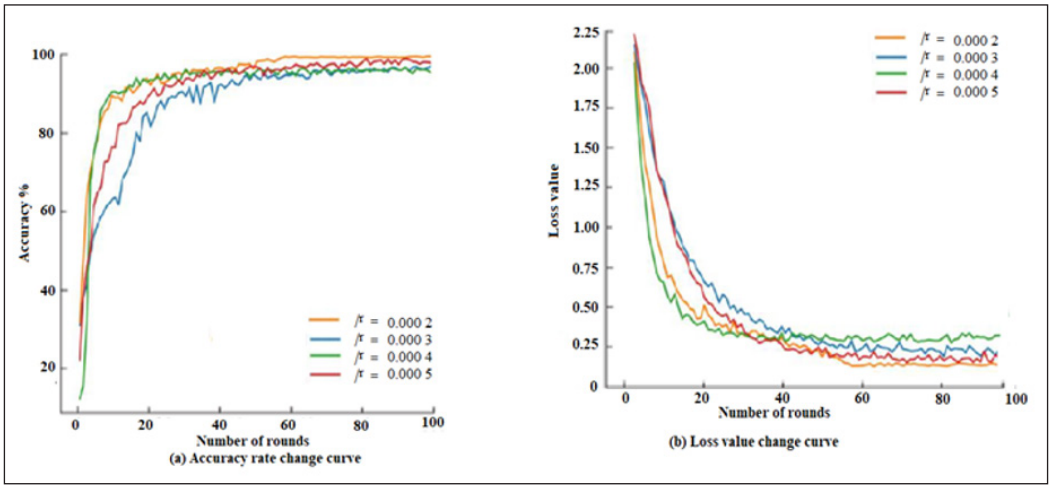


Figure 8. Change the accuracy and loss curves for network training under different learning rates

When the learning rate was set to 0.0002, the network achieved its highest accuracy of 94.70%. However, the accuracy curve showed significant oscillation, indicating that the model was unstable. At a learning rate of 0.0005, the accuracy was 91.76%. However, after 60 training rounds, both accuracy and loss increased, indicating overfitting. At learning rates of 0.0003 and 0.0004, the oscillation in the accuracy curve was minimal, with the accuracy rates reaching 93.61% and 94.07%, respectively. These learning rates allowed for more stable network training. Based on these results, 0.0004 was selected as the optimal learning rate for this paper.

RESULTS AND DISCUSSIONS

The first impact of the improved fire module on model performance. Two network structures were built under the same experimental conditions: one used the traditional SqueezeNet, while the other replaced the fire module in SqueezeNet with the improved fire module,

referred to as Model I. Both models were trained on eight types of plant disease. The results showed that SqueezeNet had 1,255,431 parameters and achieved an accuracy of 95.87%. In comparison, model I had a slightly lower accuracy of 94.07%, but its parameter count was significantly reduced to only 7/10 of SqueezeNet's. This demonstrates that the lightweight strategy applied to the fire module can substantially reduce the model's parameter count and size while incurring only a minor reduction in accuracy.

Second impact of the ECA module on model performance. To analyse the effect of adding the ECA module to the model, the attention module was embedded at four different positions within Model I, and the model was trained on the same processed dataset. The results show that Model I, without the ECA module, had 886,791 parameters and an accuracy of 94.07%. When the ECA module was added after pool 2, fire 7, fire 9, and conv 10, the number of parameters increased slightly to 886,797, with corresponding accuracies of 93.89%, 97.29%, 93.20%, and 95.28%, respectively.

It can be observed that adding the ECA module only slightly increases the number of parameters, but introducing the same attention module at different positions has varying effects on the network's performance. In some cases, it weakens the original recognition ability. In the improved model proposed in this paper, the highest accuracy was achieved when the ECA module was placed after Fire 7. This suggests that placing the ECA module at this position enables the model to effectively identify key features, enhance the expression of useful features, and suppress irrelevant ones, thereby improving the network's accuracy.

To further explore the influence of the ECA module on Model I's stability, Gaussian and salt-and-pepper noise were added to the PlantVillage dataset, and recognition performance was compared with the ECA module embedded at different positions within the network. The stability results, shown in Figure 9, indicate that the model's accuracy decreases

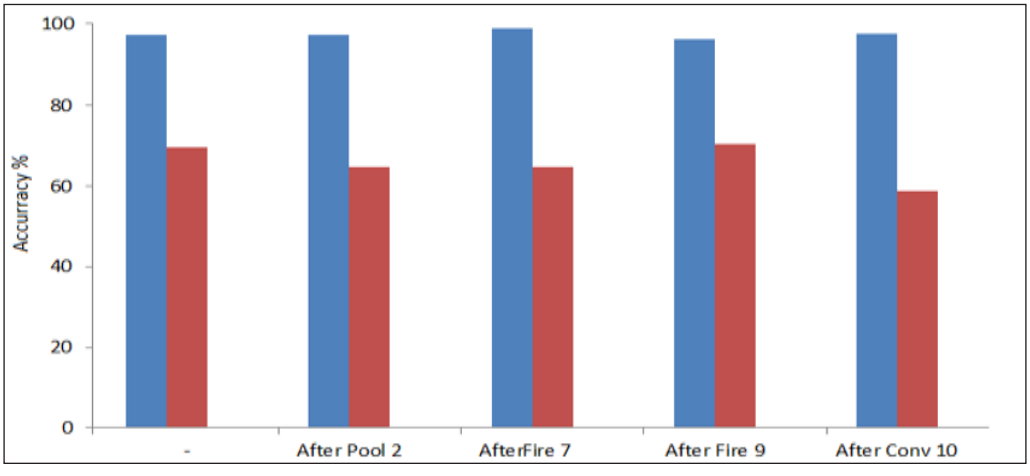


Figure 9. Stability results when the ECA module is at different locations in the network

compared with the noise-free dataset when noise is added. Additionally, network stability improved only when the ECA module was embedded behind fire 9, while it decreased when placed behind pool 2, fire 7, or conv 10. Therefore, placing the ECA module behind Fire 9 not only helps the network learn important disease characteristics but also improves its robustness to noise, leading to better overall stability.

Comprehensive Ablation Study

To evaluate the modifications made to the four network architectures shown in Table 5, the proposed model, which combines two modifications with the improved Fire module, reduces the number of parameters and model size while maintaining competitive accuracy as:

- The ECA module improves feature representation, and combining the two modifications yields the highest discrimination accuracy (97.29%).
- Maintaining the smallest model size, demonstrating that the two components effectively complement each other.

Table 5
Compare parameters between different networks

Model	Accuracy	Parameters	Model Size (MB)
SqueezeNet	95.87	1,255.431	5.0
Fire only	94.07	886791	3.7
ECG only	95,03	1,255,437	5.02
Proposed model	97.29	886,797	3.5

Comparison of improved SqueezeNet and other models. Table 6 presents performance comparisons of different network models on plant disease images. AUC for all those algorithms with the confusion matrix of the proposal method shown in Figure 10.

Table 6
Performance comparison of different network models

Model	Accuracy %	Precision %	Recall %	F-score %	Model size / MB
LeNet	89.52	89.11	90.00	0.8955	28.64
MobileNet	91.19	90.87	91.60	91.23	13.00
SqueezeNet	95.87	95.91	95.80	95.85	5.00
Improved SqueezeNet	97.29	97.35	97.22	97.28	3.50
CNN	97.08	97.14	97.01	97.07	32.78
CNN and SqueezeNet	86.92	87.10	86.75	86.92	-

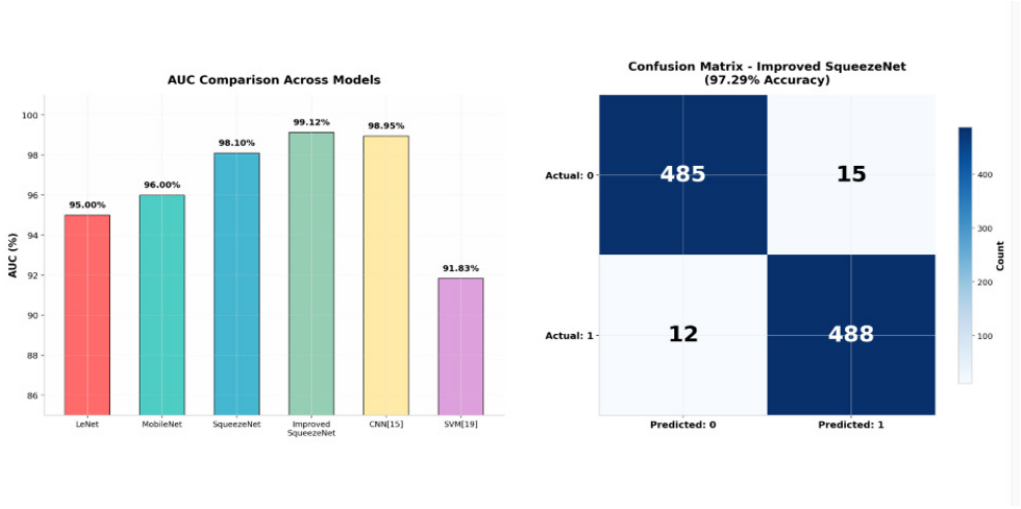


Figure 10. AUC curve and confusion matrix

As seen in Table 6, compared to the classic network model LeNet, the improved SqueezeNet model demonstrates clear advantages in both accuracy and model size. Compared with other lightweight network models, such as MobileNet and traditional SqueezeNet, the improved SqueezeNet achieves higher average accuracy while requiring less memory. Therefore, the improved SqueezeNet model outperforms the three existing network models in terms of both accuracy and size, making it well-suited for applications requiring high accuracy and a lightweight design.

Feature Visualisation and Attention Analysis

This section explains the mechanisms for extracting and enhancing attention to detail. To evaluate the proposed architectural modifications, a qualitative analysis was conducted by visualising the model's internal feature representations.

The feature maps from the last convolutional layer (conv10) of both conventional and enhanced SqueezeNet models. As shown in Figure 11, the feature maps from the enhanced model exhibit more distinct, focused activations, primarily concentrating on diseased areas of the leaf, such as spots and lesions. In contrast, the conventional SqueezeNet shows more dispersed and less distinct activation patterns. This improvement is attributed to modifications in the Fire module. The use of asymmetric convolutions (1×3 and 3×1) in the extended layers expands the receiving area and encourages the network to learn a more diverse set of features, enabling the effective capture of elongated and irregular shapes that characterise plant disease symptoms.

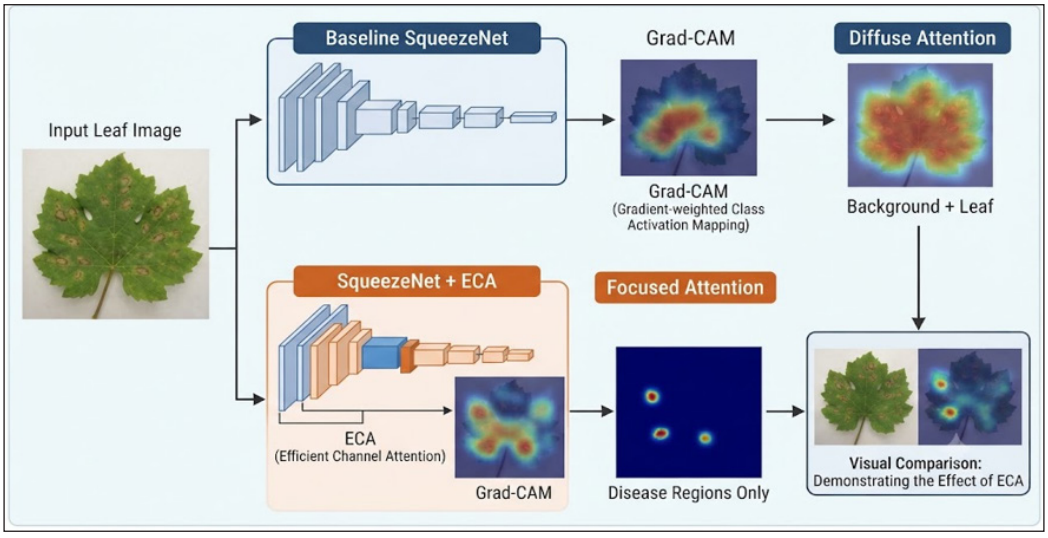


Figure 11. Feature maps from the final convolutional layer for SqueezeNet

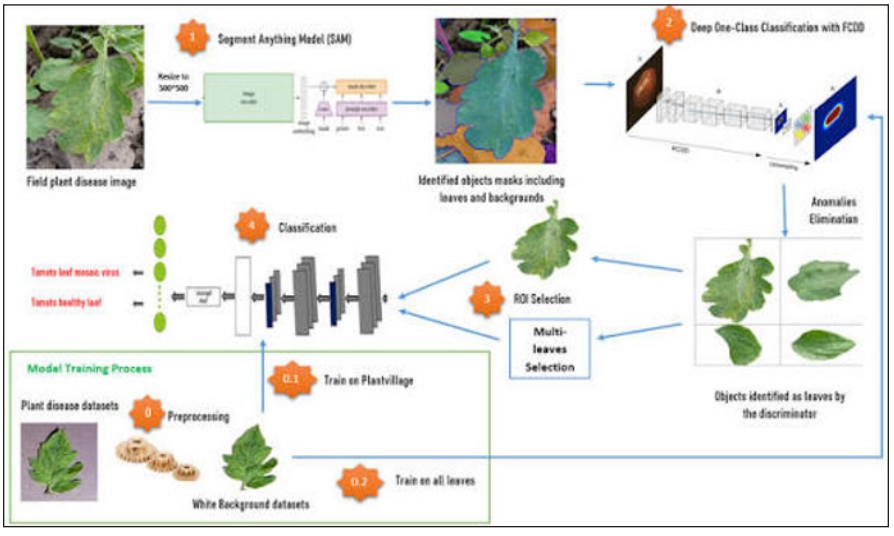


Figure 12. The ECA module

To interpret the role of the ECA module, it employed Gradient-weighted Class Activation Mapping (Grad-CAM) to generate attention heatmaps. As illustrated in Figure 12, the heatmaps reveal the regions of the input image that are most influential in the model's classification decision. The baseline SqueezeNet model often focuses on both the diseased areas and the healthy background or leaf edges, indicating a lack of precise attention to disease-specific features. However, when the ECA module is integrated, the network's attention becomes highly localised and specific to the diseased spots. This is because the

ECA module assigns adaptive weights to each feature channel, effectively amplifying channels that contain crucial information about the disease (e.g., colour and texture anomalies) while suppressing irrelevant channels. The local cross-channel interaction captured by the 1D convolution in the ECA module is key to this precise focusing ability, enabling feature recalibration without significant computational overhead. Table 7 briefly explains the methodology; it requires more detail, including a complete layer-by-layer network configuration.

Table 7
Complete layer-by-layer network configuration

Layer	Type	Kernel Size	Stride Input	Channels	Output Channels	Activation function	Parameters
Conv1	Convolution	7×7	2	3	96	ReLU	14,112
Pool1	Max Pooling	3×3	2	96	96	-	0
Fire2	Fire Module (improved)	-	-	96	128	ReLU	12,544
Pool2	Max Pooling	3×3	2	128	128	-	0
Fire3	Fire Module (improved)	-	-	128	192	ReLU	| 19,584
Fire4	Fire Module (improved)	-	-	192	256	ReLU	27,264
Pool4	Max Pooling	3×3	2	256	256	-	0
Fire5	Fire Module (improved)	-	-	256	384	ReLU	54,528
Fire6	Fire Module (improved)	-	-	384	512	ReLU	91,392
ECA	Attention Module	(1D Conv k=5)	-	512	512	Sigmoid	| 2,560
Fire7	Fire Module (improved)	-	-	512	512	ReLU	121,344
Fire8	Fire Module (improved)	-	-	512	768	ReLU	182,016
Fire9	| Fire Module (improved)	-	-	768	1024	ReLU	324,864
Conv10	Convolution	1×1	1	1024	8	ReLU	8,200
Pool10	Global Average Pooling	8×8	1	8	8	-	0
Softmax	Classification	-	-	8	8	Softmax	0

CONCLUSION

This paper applies deep learning to plant disease detection by proposing an enhanced SqueezeNet model incorporating an ECA module. The model includes a lightweight fire module and an attention mechanism. The fire module uses asymmetric convolutions

(1×3 and 3×1) and channel reduction, reducing the number of parameters and model size: from 1,255,431 parameters and 5.0 MB in the traditional SqueezeNet to 886,797 parameters and 3.5 MB in the improved version.

Furthermore, the ECA module enabled the network to adaptively emphasise the most discriminative disease-related features, thereby accurately localising disease symptoms such as spots and lesions. Experimental results demonstrate that the proposed model achieves higher classification accuracy than several existing DL models while maintaining the smallest memory footprint, making it suitable for resource-constrained and portable applications. Moreover, the ablation study confirms that combining the improved Fire module and the ECA module yields complementary benefits, achieving the best trade-off between accuracy and model efficiency. Since experiments were conducted on a public dataset that contains a single leaf and a single disease per image, further validation under real agricultural conditions is still required.

In future work, a set of real field images will be used. Preprocessing will be applied to them, using geometric transformations to address lighting and colour disturbances, such as adjusting lighting and contrast, image size, and rotation, which help reduce the contrast between data distributions.

REFERENCES

Alsaif, O. I., Hasan, S. Q., & Maray, A. H. (2023). Using skeleton model to recognize human gait gender. *IAES International Journal of Artificial Intelligence (IJ-AI), 12*(2), 974-983. https://doi.org/10.11591/ijai.v12.i2.pp974-983

Chappa, R. T. N. V. S., & El-Sharkawy, M. (2020). Squeeze-and-excitation SqueezeNext: An efficient DNN for hardware deployment. In *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 691-697). IEEE. https://doi.org/10.1109/ccwc47524.2020.9031119

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture, 145*, 311-318. https://doi.org/10.1016/j.compag.2018.01.009

Fu, P. (2014). *Research on classification methods for large-scale imbalanced data* [Thesis, Liaoning Normal University].

Guerrero-Ibañez, A., & Reyes-Muñoz, A. (2023). Monitoring tomato leaf disease through convolutional neural networks. *Electronics, 12*(1), Article 229. https://doi.org/10.3390/electronics12010229

Hao, M., Sun, Q., Xuan, C., Zhang, X., & Zhao, M. (2024). SqueezeNet: An improved lightweight neural network for sheep facial recognition. *Applied Sciences, 14*(4), Article 1399. https://doi.org/10.3390/app14041399

Hazarika, R. A., Abraham, A., Kandar, D., & Maji, A. K. (2021). An improved LeNet-deep neural network model for Alzheimer's disease classification using brain magnetic resonance images. *IEEE Access, 9*, 161194-161207. https://doi.org/10.1109/access.2021.3131741

Hidayatuloh, A., Nursalman, M., & Nugraha, E. (2018). Identification of tomato plant diseases by leaf image using SqueezeNet model. In *2018 International Conference on Information Technology Systems and Innovation (ICITSI)* (pp. 199-204). IEEE. https://doi.org/10.1109/ICITSI.2018.8696087

Hughes, D. P., & Salathé, M. (2015). *An open access repository of images on plant health to enable the development of mobile disease diagnostics* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1511.08060

Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). *SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1602.07360

Jung, M., Song, J. S., Shin, A.-Y., Choi, B., Go, S., Kwon, S.-Y., Park, J., Park, S. G., & Kim, Y.-M. (2023). Construction of deep learning-based disease detection model in plants. *Scientific Reports, 13*(1), Article 7331. https://doi.org/10.1038/s41598-023-34549-2

Khan, A. T., Jensen, S. M., Khan, A. R., & Li, S. (2023). Plant disease detection model for edge computing devices. *Frontiers in Plant Science, 14*, Article 1308528. https://doi.org/10.3389/fpls.2023.1308528

Khan, R. U., Khan, K., Albattah, W., & Qamar, A. M. (2021). Image-based detection of plant diseases: From classical machine learning to deep learning journey. *Wireless Communications and Mobile Computing, 2021*(1), Article 5541859. https://doi.org/10.1155/2021/5541859

Lahgazi, M. J., & Mohaoui, S. (2025). *Attfire: Attention-enhanced fire modules for image segmentation* [Preprint]. SSRN. https://doi.org/10.2139/ssrn.5356529

Mohamed, S. A., Alsaif, O. I., & Saleh, I. A. (2022). Intrusion detection network attacks based on whale optimization algorithm. *Ingénierie des Systèmes d'Information, 27*(3), 441-446. https://doi.org/10.18280/isi.270310

Mohameth, F., Bingcai, C., & Sada, K. A. (2020). Plant disease detection with deep learning and feature extraction using Plant Village. *Journal of Computer and Communications, 8*(6), 10–22. https://doi.org/10.4236/jcc.2020.86002

Oleiwi, Z., Thanoon, K., & Alsaif, K. (2019). High frequency coefficient effect on image based on contourlet transformation. In *2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA)* (pp. 1-4). IEEE. https://doi.org/10.1109/iccista.2019.8830649

Panchal, A. V., Patel, S. C., Bagyalakshmi, K., Kumar, P., Khan, I. R., & Soni, M. (2023). Image-based plant diseases detection using deep learning. *Materials Today: Proceedings, 80*, 3500-3506. https://doi.org/10.1016/j.matpr.2021.07.281

Rahman, C. R., Arko, P. S., Ali, M. E., Iqbal Khan, M. A., Apon, S. H., Nowrin, F., & Wasif, A. (2020). Identification and recognition of rice diseases and pests using convolutional neural networks. *Biosystems Engineering, 194*, 112-120. https://doi.org/10.1016/j.biosystemseng.2020.03.020

Saleh, I. A., Albayati, A. H., & Thanoon, K. H. (2021). Measure the software quality based on grasshopper optimization algorithm. *International Journal of Computing and Digital Systems, 10*(1), 955-961. https://doi.org/10.12785/ijcds/100186

Setiawan, W., Ghofur, A., Hastarita Rachman, F., & Rulaningtyas, R. (2021). Deep convolutional neural network AlexNet and SqueezeNet for maize leaf diseases image classification. *Kinetik: Game Technology,*

Information System, Computer Network, Computing, Electronics, and Control, 6(4). https://doi.org/10.22219/kinetik.v6i4.1335

Shoaib, M., Shah, B., El-Sappagh, S., Ali, A., Ullah, A., Alenezi, F., Gechev, T., Hussain, T., & Ali, F. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. *Frontiers in Plant Science, 14*, Article 1158933. https://doi.org/10.3389/fpls.2023.1158933

Singh, A. K., Sreenivasu, S. V. N., Mahalaxmi, U. S. B. K., Sharma, H., Patil, D. D., & Asenso, E. (2022). Hybrid feature-based disease detection in plant leaf using convolutional neural network, Bayesian optimized SVM, and random forest classifier. *Journal of Food Quality, 2022*(1), Article 2845320. https://doi.org/10.1155/2022/2845320

Wang, B. (2022). Identification of crop diseases and insect pests based on deep learning. *Scientific Programming, 2022*, Article 9179998. https://doi.org/10.1155/2022/9179998

Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., & Hu, Q. (2020). ECA-Net: Efficient channel attention for deep convolutional neural networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 11531-11539). IEEE. https://doi.org/10.1109/cvpr42600.2020.01155

Wei, K., & Gao, Y. (2025). Crop disease and pest identification based on deep learning and Raspberry Pi. In *Proceedings of the 2025 7th Asia Conference on Machine Learning and Computing* (pp. 47-53). Association for Computing Machinery. https://doi.org/10.1145/3772673.3772703

Yang, T., & Liu, C. (2020). Recognition system for leaf diseases of Ophiopogon japonicus based on PCA-SVM. *Plant Diseases and Pests, 11*(2), 11-15.

Zhao, S., Liu, J., Bai, Z., Hu, C., & Jin, Y. (2022). Crop pest recognition in real agricultural environment using convolutional neural networks by a parallel attention mechanism. *Frontiers in Plant Science, 13*, Article 839572. https://doi.org/10.3389/fpls.2022.839572

Pertanika J. Sci. & Technol. 34 (5): 3003 - 3024 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Novel Fusion of Computer Vision Capabilities and Deep Learning Techniques for OCT Retinal Disease Diagnosis

Nagham Salim Mohammed[1]*, Omar S. Almolaa[2], and Omar Ibrahim Alsaif[3]

[1]Department of New and Renewable Energies, College of Science, University of Mosul, Mosul 41002, Nineveh, Iraq
2Department of Computer Science, Faculty of Education, University of Telafer, Telafer 41016, Nineveh, Iraq
3Department of Electronic and Communication Technology Engineering, Mosul Polytechnic College, Northern Technical University, Mosul 41002, Nineveh, Iraq

ABSTRACT

The main objective is to address the classification limitations of traditional models for diagnosing retinal macular diseases, including neovascularisation (CNV), diabetic macular oedema (DME), and age-related macular degeneration (AMD) associated with drusen. An Optical Coherence Tomography (OCT)- based computer vision framework has been proposed to classify these diseases. The models achieved this by addressing the inability of traditional models to learn spatial dependencies and the need for an inductive bias to learn fine retinal texture patterns in Vision Transformer (ViT) models. A dual-model HR50_ViTB/16 was proposed, comprising a full-depth ViTB/16 model for capturing spatial information, a pre-trained ResNet50 model for extracting local, multiscale hierarchical information to further improve performance, and a learnable merging layer to integrate local and global information extracted from both branches. The proposed framework outperformed all traditional models, achieving an overall accuracy of 96.14% and an F1 score of 0.952 for the DRUSEN class, which is the most difficult and has very limited data for classification. Its significance lies in the fact that these results demonstrate the quality of the local and global capabilities of ResNet50 networks and $ViT_{B/16}$ transformers, respectively, which can complement each other to overcome the limitations of each model.

Keywords: CNV, DME, DRUSEN, Dual-Branch, Feature Fusion, Medical Image Analysis, OCT (Optical Coherence Tomography), $ResNet_{50}$, Retinal Coherence Diseases, $Vision\ Transformer_{B/16.}$

ARTICLE INFO

Article history:
Received: 11 August 2026
Accepted: 29 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.19

nagham.salim@uomosul.edu.iq (Nagham Salim Mohammed)
omar.almolaa@uotelafer.edu.iq (Omar S. Almolaa)
omar.alsaif@ntu.edu.iq (Omar Ibrahim Alsaif)
* Corresponding author

INTRODUCTION

Blindness or visual impairment is a global problem affecting 295 million people with

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

moderate-to-severe vision impairment (2020) and 43.3 million people with blindness (GBD 2019 Blindness and Vision Impairment Collaborators & Vision Loss Expert Group of the Global Burden of Disease Study, 2021). The global problem is growing, particularly in low- and middle-income countries (Burton et al., 2021; Vision Loss Expert Group of the Global Burden of Disease Study & the GBD 2019 Blindness and Vision Impairment Collaborators, 2024).

Blindness-causing retinal diseases involve Choroidal Neovascularisation (CNV), Diabetic Macular Edema (DME), and DRUSEN (Ma et al., 2022). CNV is among the clinical forms of AMD in which a choroidal blood vessel lies beneath the retinal pigment epithelium, leading to damage to the photoreceptors (He et al., 2021). There could be other causes of v, such as loss, like diabetic retinopathy, which affects many individuals globally in 2020 (Teo et al., 2021). DRUSEN are the most common early sign of macular degeneration. So, lipoprotein deposits under the retinal pigment epithelium are related to age-related macular degeneration (AMD) (Li et al., 2023; Xu et al., 2021).

The Optical Coherence Tomography (OCT) is an advanced imaging method that may assist in identifying pathological markers, neovascularisation (CNV), diabetic macular oedema (DME), DRUSEN, subretinal membranes, pigmented epithelial detachment, intraretinal fluid, macular thickening, and xanthomas (Akpinar et al., 2024). However, distinguishing certain pathological patterns in OCT images remains challenging, including discriminating between CNV and DME or between early DRUSEN and normal retinal morphology (Burton et al., 2021). The difficulty underscores the need for efficient, automated OCT image classification techniques that provide repeatable, quantifiable diagnostic support (Laouarem et al., 2024).

For the automatic detection of retinal diseases, Computer vision and deep learning techniques have been shown to yield good results for detecting CNV, DME, and DRUSEN using a transfer learning-based CNN such as Inception-V3 (Hassan et al., 2023; Kermany et al., 2018). However, due to the local receptive field, conventional CNNs might be unable to capture the long-range spatial dependence, which could be problematic for recognising spatially dispersed lesions such as DRUSEN and for distinguishing the visually similar appearance of a fluid-filled appearance in CNV and DME (Dutta et al., 2023; Fang et al., 2019; Gencer & Gencer, 2025). ViT overcomes this chamodelling by modelling non-local relationships between different image patches using MHSA, enabling a better understanding of global retinal structures in OCT classification (Dosovitskiy et al., 2021; Yang et al., 2025).

To address this problem, the Vision Transformer (ViT), where the transformer architecture is used, splits the image into non-overlapping patches, and each patch is handled as a time step sequentially with a process called Multi-Head Self-Attention (MHSA) that is used to capture the relationships between distant parts of an image (Dosovitskiy et al., 2021; Mohammed et al., 2026). This feature is crucial in the context of OCT classification,

where multiple retinal areas and multiple layers can contain patterns indicative of disease; it has also been demonstrated in the OCT2017 dataset (Akça et al., 2024; Wang et al., 2023). This is especially relevant for OCT images, where small variations in the texture of the retinal layers and reflectivity may be crucial for diagnosis (Wang et al., 2025).

In this study, a new dual-branch computer vision architecture (HR5_ViTB/16) is proposed for four-class classification of retinal diseases from OCT images. Specifically, the architecture consists of a Vision Transformer (ViT-B/16) branch comprising 12 encoder blocks, 12 attention heads, an embedding dimension of 768, and a patch size of 16, yielding 196 patch tokens, in parallel with a pre-trained ResNet-50 residual feature-extraction network. The features from the ViT (CLS) token and the global average-pooled features from the ResNet-50 network are concatenated and then passed through a learnable linear fusion layer with layer-norm, enabling the joint optimisation of local and global features in an end-to-end manner. The architecture is trained and tested on the OCT2017 dataset, which consists of four classes of retinal diseases.

The novelty of the proposed framework is achieved by aligning representational spaces before concatenation, where the ResNet-50 representation (R^(2048)) is converted via a LayerNorm, GELU, and Dropout projection into R^(768), which is part of the cls_token_out of $ViT_{B/16}$ before sequencing. The result is that the next linear transformation (R^(1536)→R^(768)) learns from the true semantic connections between the two spaces, rather than processing incompatible features using statistical techniques. The use of a full sequence rather than some form of dynamic weighting (cross-attention or gating) is an intentional design decision to keep the whole sequence of representational information of both branches while retaining complementary semantic fullness between the convolutions and the transformers. Sequencing with learned linear projections yields a concise representation, and the degree of alignment between representational spaces before merging primarily determines the success of the merging approach, not the complexity of the merge layer itself.

Branch ResNet-50 its output with dimensions (2048) is dropped to dimension (768) by relying via LayerNorm → GELU → Dropout, which matches the dimensions of ViT CLS before the merge process, and thus both branches are inserted into the merge layer with the same equal representational capacity, thus avoiding the bias in information inclusion by increasing one of the branches, and the merge vector with dimension (1536) preserves the complete information for both branches.

This combination process is a (Linear 1536→768, LayerNorm, GELU, Dropout 0.15) projection of the hat model learned from the best weights of the CNN and ViT feature representations. This is not merely an integration of the two branches or an attention gate that uses only one weight. For DRUSEN, texture information is based on the spatial attention representation of the ViT branch, whereas CNV uses the ResNet branch.

The DRUSEN category is the most challenging, as it requires information about both the local texture (RPE surface details) and the global texture (multi-patch RPE ripple). Using the results of the four architectures, the superiority of the adopted architecture can be demonstrated.

LITERATURE REVIEW

Early computational methods for detecting DRUSEN were based on classic image-processing techniques used in fundus fluorescein angiography (FFA) images. A computer-aided detection (CAD) system for automatic identification and areal size calculation of DRUSEN regions in FFA images was developed by Demir et al. (2018) and served as a morphological baseline for ARMD diagnosis. $HR_{50}_ViT_{B/16}$ is a much more difficult task than HR50_FFA/16, as it does not use colour OCT FFA images but instead extracts DRUSEN from a grayscale cross-sectional OCT image, emphasising the importance of the global attention module provided by the ViT branch.

Through a survey proposed by (Liew & Agaian, 2025), the developments of different kinds of features extractors from manual and deep learning techniques have been studied with respect to accuracy, sensitivity, specificity, and computational efficiency for the detection of various diseases of eyes including glaucoma, diabetic retinopathy, and macular degeneration, as well as, cyberattacks and adversarial attack threats on OCT diagnostic systems are concerned, a more effective method has been proposed.

The proposed algorithm by Rahil et al. (2023) introduced a deep learning-based clustering model that uses CNN techniques to segment various retinal patterns in OCT images. The proposed technique's results are excellent, both quantitatively and qualitatively, compared to other state-of-the-art algorithms on the RETOUCH dataset, with an overall performance improvement of 1.8%.

In the case of OCT-based multi-class classification of retinal diseases, Zhou et al. (2023) proposed using the ViT model, which has achieved high performance in distinguishing normal images from irregular conditions such as retinal vascular macular degeneration (CNV), retinal degeneration, and diabetic macular oedema. Achieve an accuracy rate of 95.76% and a sensitivity of 95.77%, outperforming many traditional CNN models statistically.

Aslan (2023) also confirmed this with results from comparative studies of ViT and ResNet networks for image classification of skin diseases, in which ViT surpassed ResNet in accuracy (68.93% versus approximately 61%), despite its longer training time, thereby increasing its generalizability to other areas of medical imaging. In terms of baseline models, RETFound is a model pre-trained on 1.6 million unlabeled OCT images.

Du et al. (2024) applied their method to improve the classification of disease features, such as subretinal fluid, intraretinal fluid, retinal ganglion cysts, and pigment epithelial

detachment. The results of the application on RET showed similar performance to ResNet-50 in terms of accuracy and area under the ROC curve (0.75–0.80).

Raut et al. (2025) introduced a hybrid approach that uses ViT, Inception-V3, and $ResNet_{50}$. This is because results from applying standard databases have shown that using multiple architectures can improve the classification and identification of eye diseases compared with a single architecture. It is advantageous since it fastens treatments.

METHODOLOGY

Objectives

Hybrid Residual_ViT (HR_{50}_$ViT_{B/16}$), facilitating the associated improvement of global self-attention and hierarchical residual feature representations, as well as, the framework, achieved an F1-score of 0.952 on the DRUSEN data set, reflecting an improvement over the best competing hybrid model, known as Conv-ViT, which achieved an DRUSEN F1-score of 0.78 (Dutta et al., 2023), by resolving the long-standing DRUSEN classification bottleneck, by effectively capturing the spatially diffuse sub-RPE patterns across the macular domain.

Materials and Methods

Dataset OCT2017

The OCT2017 dataset was used for the experiments. This publicly available dataset contains high-resolution grayscale OCT retinal scans, and the images are labelled with four different clinical categories. These categories are choroidal neovascularisation (CNV), diabetic macular oedema (DME), DRUSEN deposits, and NORMAL retinal tissue. Images have been loaded in RGB format to ensure compatibility with the pre-trained ResNet-50 backbone, as this neural network was initialised on a three-channel image dataset, namely ImageNet. The dataset can be found at the following link: https://data.mendeley.com/datasets/rscbjbr9sj/3 (Kermany et al., 2018). Figure 1 illustrates OCT Types.

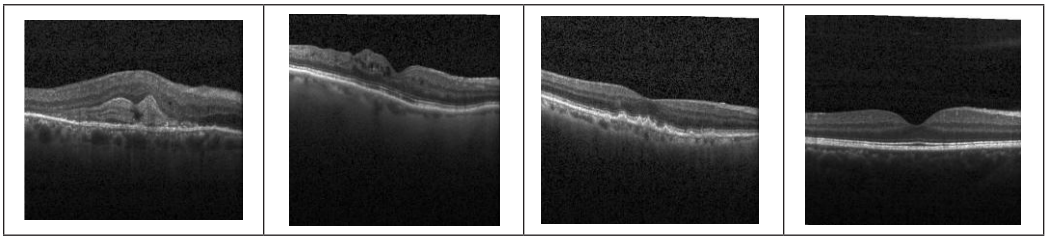


Figure 1. OCT Dataset retinal categories

Dataset Preprocessed

An asymmetric data augmentation strategy was adopted, where a range of stochastic transformations were applied to the training images, including:

Random resizing followed by random cropping to 224 × 224 pixels. Random horizontal and vertical flipping with probabilities 0.5 and 0.3, respectively; Colour jitter, modulating brightness, contrast, saturation, and hue by ±0.4, ±0.4, ±0.3, and ±0.1, respectively. Random affine transformation, including translation up to ±10% and scaling between (0.9 and 1.1). Random perspective transformation with distortion scale = 0.2 and probability = 0.3. Gaussian blur with a kernel size of 3 and sigma uniformly distributed between 0.1 and 2.0. Random erasing of rectangular regions with a probability of 0.2, scale range of [0.02, 0.15]. Subsequent normalisation of all training images using ImageNet-derived channel-wise statistics, where:

$$\mu = [0.485, 0.456, 0.406] \text{ and } \sigma = [0.229, 0.224, 0.225]$$

The validation and test images, on the other hand, were only subject to a deterministic data augmentation pipeline, in which all images were resized to 224×224 pixels and normalised.

Table 1
Class-Wise ACCOUNT Sample OCT2017 Dataset.

Class	Training	Validation	Test
CNV	31,624	**5,585**	250
DME	9,646	**1,703**	250
DRUSEN	7,323	**1,290**	250
NORMAL	22,368	**3,945**	261
TOTAL	70,961	**12,523**	1,011

In Table 1, to address the imbalance of data between categories, the training data size was set to 8616 images per category to stabilise and improve the model. Thus, a balanced training set consisting of 34464 images was obtained; for the DRUSEN category, the required size was achieved using data augmentation, while both the verification and testing sets listed in the table retained their original category distributions during model development and evaluation. The OCT2017 dataset (Kermany et al., 2018) classifies each image by disease, random patient ID, and image number. Test and training folders were predefined and created by the original developers specifically to separate patients into the two groups. In the proposed methodology, these folders are not merged; patient images used for training do not appear in the test set. From the original dataset design, the test set at the patient level remains independent.

Model Architecture

Patch Embedding Module

1. Input images → 224 × 224 × 3. The images are divided into a grid of non-overlapping patches, and each patch is linearly projected to a continuous embedding space through a convolutional projection layer. The dimensionality of the embedding is set to d=768.

2. The convolutional projection layer → P=16 kernel size and a stride → P=16. The nn.The Conv2d module in PyTorch is used for convolutional projection layers.

3. The input to the transformer model consists of a classification token [CLS] ∈ R^d and a sequence of patch tokens. The classification token [CLS] ∈ R^d.

4. The input sequence and the classification token are then linearly combined with the positional embeddings through an element-wise addition operation. The positional embeddings ∈ R^{(N+1) × d}.

Vision Transformer Encoder

Multi-head self-attention (MHSA) is applied to the input, projected into queries and keys/values with a linear layer (Linear (d, 3d)), and the projected values (H = 12) heads have (d_h = d/H = 64) dimensions. The scaled dot product attention is computed as Equation 1, below (Vaswani et al., 2017):

$$\text{Attention}(Q,K,V)=\text{Softmax}(\frac{QK^T}{\sqrt{d_k}})V \quad [1]$$

1. Dropout(p=0.1) is utilised for both attention weights and output projections and uses the stochastic depth (DropPath) in the residual branch with a linearly increasing rate from 0 to 0.1 for the 12 blocks.

2. Wise Feed-Forward Network (FFN) with two linear layers, GELU activation functions, and dropout layers, with a hidden dimension (4d = 3072), after the LayerNorm operation used (CLS) taken as a global feature representation, known as cls_token_out ∈ R^{768}. Figure 2 Attention Head (12 Layers).

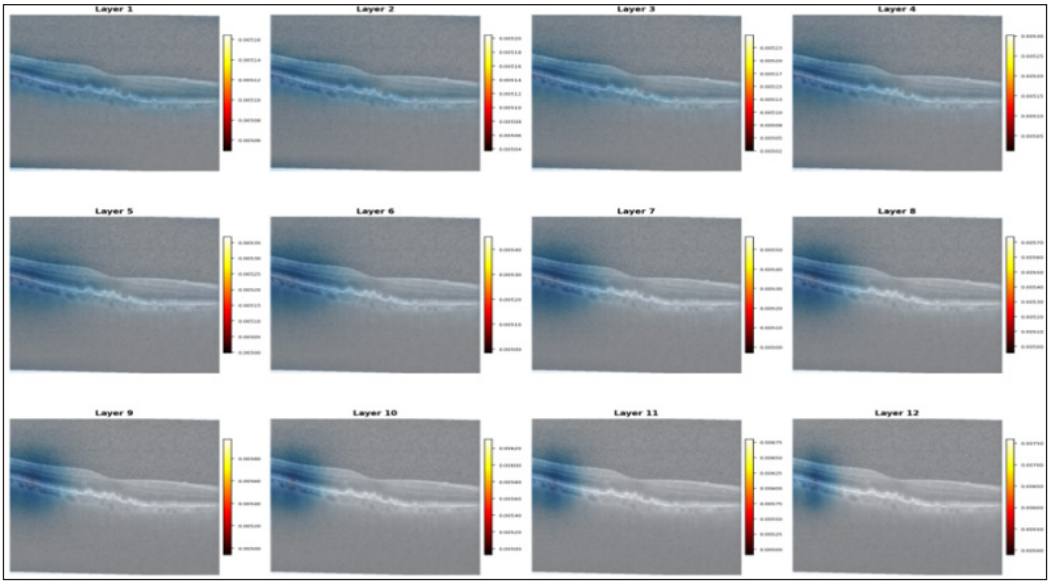


Figure 2. Attention maps for a typical DRUSEN case at the layer-wise average

Residual Feature Extractor

1. The architecture follows the sequence: 7×7 Conv (64) → Batch Normalisation → ReLU Activation → Max Pooling → Residual Block (Layer1) → Residual Block (Layer2) → Residual Block (Layer3) → Residual Block (Layer4) → Channel Depths of 256, 512, 1024, and 2048, respectively.

2. Adaptive Average Pooling (1×1) → Feature Maps ∈ R^ {2048}.

3. Layer Normalisation → GELU Activation → Dropout (0.1) to produce resnet_proj ∈ R^ {768}, which has the same dimension as the ViT branch to enable fusion.

Feature Fusion and Classification

1. The CLS token representation from the ViT branch (cls_token_out ∈ R^ {768}), ResNet-50 feature vector (resnet_proj ∈ R^ {768}), are concatenated along the feature dimension, and a joint representation is formed as combined ∈ R^ {1536}.

2. This is done in such a manner that it preserves the full information content of both branches without any prior assumption of weighting of both feature sources.

3. Passed through a Fusion Layer, which was implemented as a sequential module consisting of the following steps: a linear transformation (1536 → 768) → LayerNorm

→ GELU activation → dropout layer with a dropout rate of 0.15. The resulting fused embedding ∈ R^ {768} was then forwarded through the classification head.

4. The classification head (head), which was a two-layer MLP, consisted of the following steps: a linear transformation (768 → 512) → LayerNorm → GELU activation → dropout (p = 0.3).

5. The final linear layer (512 → 4) is applied for generating unnormalized class logits for the four OCT diagnostic classes, i.e., CNV, DME, DRUSEN, and NORMAL. During inference, class probabilities are computed by applying the softmax function to the logit outputs. Figure 3 shows the HR_{50}_$ViT_{B/16}$ Algorithm Pipeline. Table 2 illustrates the training network hyperparameter configuration.

Table 2
The Hyperparameter Configuration for Training HR_{50}_ViT_{16} in Oct

C	Value	Justification
Patch size (P)	16 × 16 px	196 tokens; standard ViT-B/16
Embedding dimension (d)	768	ViT-Base configuration
Transformer depth (L)	12	ViT-Base configuration
Attention heads (H)	12 (d_h = d/H = 64)	Head dim. = 64
MLP ratio (r)	4.0	d_ff = 3,072
Dropout rate	0.1	Regularisation
Max epochs	100	Subject to early stopping
learning rate (η)	3×10^{-4}	AdamW standard for transformers
Warm-up fraction	10% (pct_start = 0.1)	Gradient stability
Optimiser	AdamW (β_1=0.9, β_2=0.999, ε=1×10^{-8})	Standard for Transformer training
$ResNet_{50}$ branch initialization	ImageNet-pretrained	Transfer learning for low-level visual features
Software	PyTorch(2.0), scikit-learn (1.3), torchvision(0.15)	Reproducibility
LR scheduler	OneCycleLR	Faster convergence, reduces sensitivity to LR choice
Loss function	Cross-Entropy Loss	Standard for multi-class classification

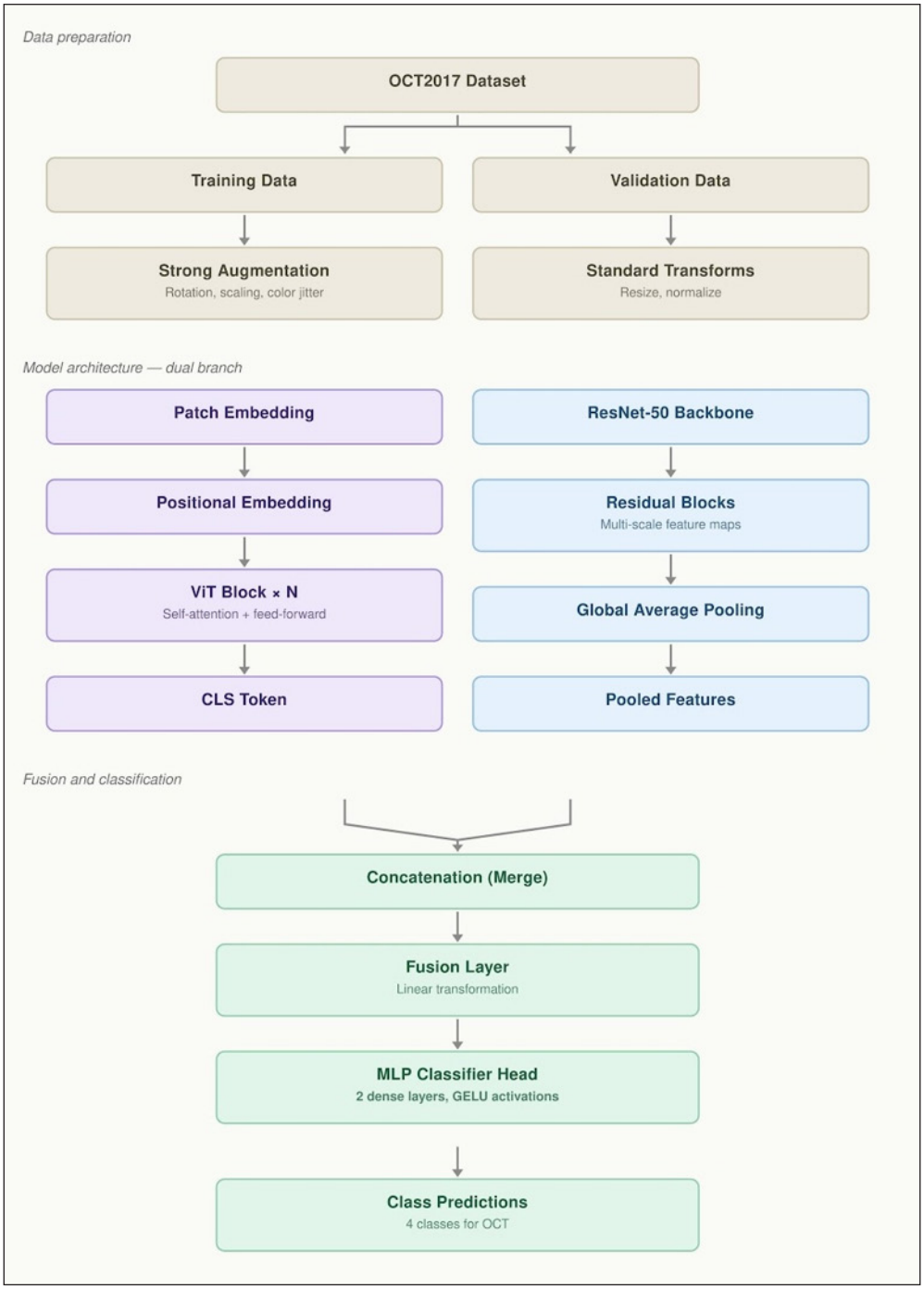


Figure 3. Overview of the proposed architecture of HR_{50}–$ViT\text{-}_{B/16}$ for OCT image classification

```
ViT_B/16 BRANCH — Global Context Extraction
═══════════════════════════════════════════
P ← PatchEmbedding(x)
   // Conv2d(kernel=16, stride=16) → 196 patches
   // Each patch projected: p_i ∈ ℝ^{768}
T ← Concat([CLS_token, p_1, …, p_196]) + PosEmbed
   // T ∈ ℝ^{197×768}
FOR i ← 1 TO 12 DO        // stochastic depth: 0 → 0.1
   T ← TransformerBlock_i(T)
      // MHSA: H=12 heads, d_h=64
      // QKV projection: Linear(768, 3×768)
      // Attention: softmax(QK^T / √d_h) · V
      // Dropout(0.1) on attention weights & projections
      // FFN: Linear(768→3072) → GELU → Linear(3072→768)
END FOR
cls_token_out ← LayerNorm(T)[ :, 0 ]    // ∈ ℝ^{768}
═══════════════════════════════════════════
② ResNet_50 BRANCH — Local Hierarchical Extraction
═══════════════════════════════════════════
f ← Conv2d(x, 64, 7×7, stride=2) → BN → ReLU → MaxPool
f ← ResidualBlock_Layer1(f)    // f ∈ ℝ^{56×56×256}
f ← ResidualBlock_Layer2(f)    // f ∈ ℝ^{28×28×512}
f ← ResidualBlock_Layer3(f)    // f ∈ ℝ^{14×14×1024}
f ← ResidualBlock_Layer4(f)    // f ∈ ℝ^{7×7×2048}
f ← AdaptiveAvgPool(f, 1×1).flatten()    // f ∈ ℝ^{2048}
resnet_proj ← Linear(f, 768)
          → LayerNorm → GELU → Dropout(0.1)
          // resnet_proj ∈ ℝ^{768}
═══════════════════════════════════════════
③ LEARNABLE FUSION LAYER
═══════════════════════════════════════════
combined ← Concat([cls_token_out, resnet_proj], dim=1)
        // combined ∈ ℝ^{1536}
        // preserves full information without prior weighting
fused ← Linear(combined, 768)
      → LayerNorm → GELU → Dropout(0.15)
      // fused ∈ ℝ^{768}
═══════════════════════════════════════════
④ CLASSIFICATION HEAD
═══════════════════════════════════════════
h      ← Linear(fused, 512) → LayerNorm → GELU → Dropout(0.3)
logits ← Linear(h, 4)                    // logits ∈ ℝ^4
IF training THEN
   loss ← CrossEntropyLoss(logits, y, label_smoothing=0.1)
ELSE
   ŷ ← argmax(Softmax(logits))
```

Note. An algorithm classification pipeline based on ResNet-50 and vision transformer for retinal disease classification from Optical Coherence Tomography (OCT) images

EVALUATION METRIC

The performance of the HR50_ViTB/16 model was measured using the following metric (Alsaif et al., 2023; Sujon et al., 2025):

Accuracy is the number of samples correctly classified, divided by the total number of samples in the data set. The accuracy is computed by using the following Equation 2.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad [2]$$

Positive Precision, also known as Positive Predictive Value, is the number of cases that are positive divided by the total number of cases identified as positive. It is given by the following Equation 3:

$$\text{Precision} = \frac{TP}{TP + FP} \quad [3]$$

Recall (Sensitivity) is the measure of the capacity of the model to recognise the actual positive cases by computing the ratio of the number of actual positive cases to the total number of actual positive cases. It is given by the following Equation 4:

$$Recall = \frac{TP}{TP + FN} \quad [4]$$

F1_Score is the harmonic mean of both precision and Recall. It is the single scale that incorporates both precision and Recall. It is the most suitable scale to be used when dealing with a highly imbalanced dataset. This is because the F1_score is given by Equation 5:

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad [5]$$

Wilson Score Interval

Equation 6 shows that the reliability of accuracy is calculated using the Wilson Score Interval for the test set. Where: $\hat{p}$ = observed proportion (accuracy), n^2 = sample size (test set size), z^2 =critical value of standard normal distribution (z = 1.96 for 95% confidence interval (Brown et al., 2001).

$$CI_{95\%} = \frac{\hat{p} + \frac{z^2}{2n} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}} \quad [6]$$

Two-Proportion Z-Test

Equation 7, used to calculate the overall classification accuracy for the proposed model comparison with other baselines, $\hat{p}_1$, $\hat{p}_2$ = accuracy of Model 1 and Model 2, x_1 and

x_2 = number of correct classifications in each model, n_1,n_2= test set sizes ($n_1 = n_2 = 1{,}011$) (Agresti, 2019).

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad [7]$$

Where : $\hat{p} = (x_1 + x_2) / (n_1 + n_2)$

RESULT AND DISCUSSIONS

Figure 4 shows the confusion matrix analysis. The proposed model's CNV class resulted in a Recall of 96.8%. The confusion between the two health problems can be attributed to both pathologies exhibiting characteristic radiological features of subretinal and intraretinal fluid collections, which present similar OCT intensity profiles in the macular region. The DME class resulted in 95.2% Recall, while the DRUSEN class resulted in 94.8% Recall, the lowest Recall for the proposed model. The improvement in the proposed model's DRUSEN classification performance compared to the previous CNN-based approach can be attributed to the global self-attention mechanism incorporated by the ViT branch, which captures the diffuse distribution of DRUSEN deposits over the macular area; the NORMAL class resulted in a recall of (97.7%). The feature space can be said to have been organized by the model itself because the features that have been organized by the model (the ultimate feature space, the 768 dimensional integrated feature space in the decision architecture of the system) show the accumulation of error in terms of one or two neighbouring cells (DRUSEN → DME/NORMAL; CNV/DME bidirectional), where some pairs of cells (CNV↔NORMAL, DME↔NORMAL) have close to zero confusion.

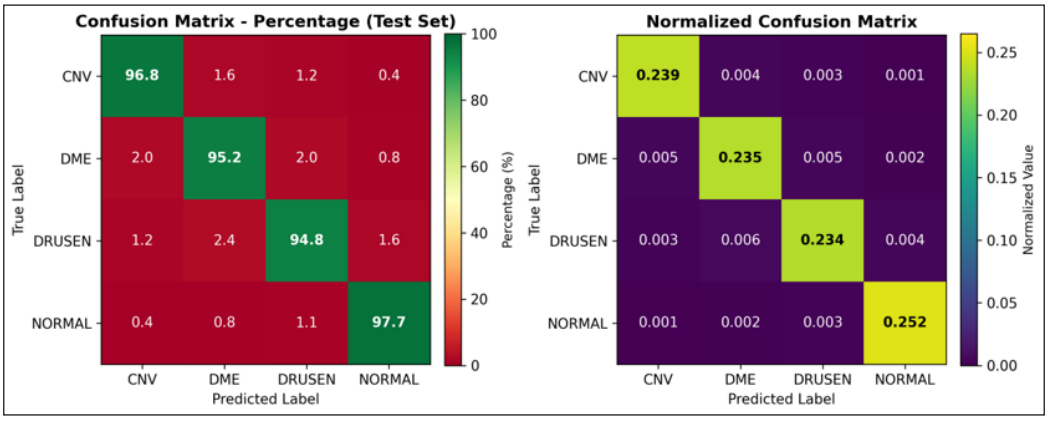
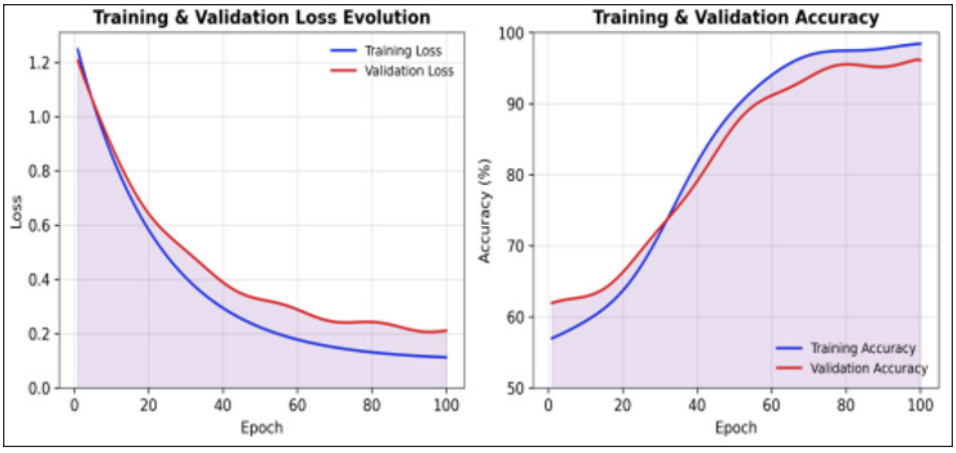


Figure 4. Confusion matrix-based performance evaluation of the proposed architecture on the OCT2017



(a) (b)

Figure 5. Complete training analysis for $HR_{50}_VIT_{B/16}$ on the OCT Dataset for a total of 100 epochs: (a)training &validation loss evolution; (b)training &validation accuracy

The training loss (Figure 5(a)) decreases monotonically from its initial value of about 1.25 to a final minimum of 0.196 at epoch 100. The validation loss has a similar history to the training loss, but over the entire epoch range. The training and validation accuracy curves in Figure 5 (b) show rapid improvement in epochs 1-40, followed by a gradual improvement from epochs 40 to 85, so the best validation accuracy is attained at epoch 100, at 96.11%.

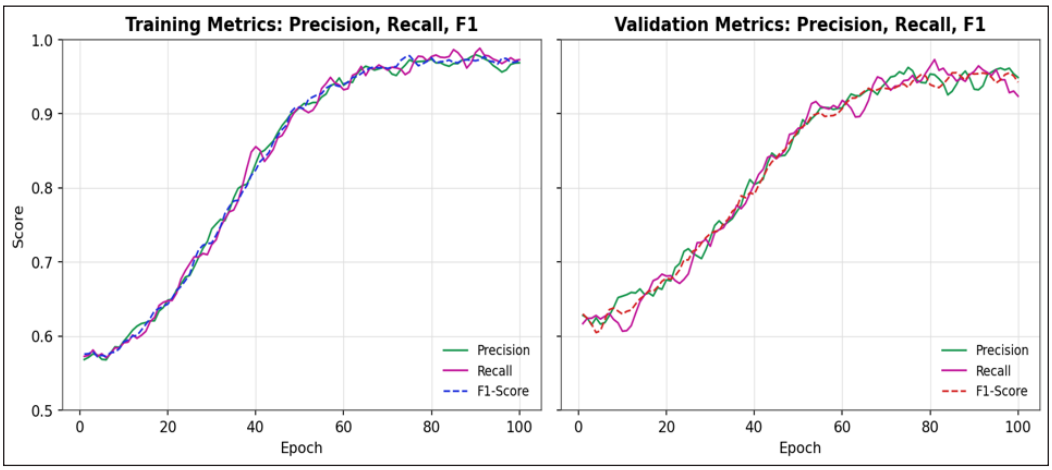


Figure 6. Training metrics (Precision, Recall, F1); Validation metrics (Precision, Recall, F1)

Figure 6 shows the epoch-wise behaviour of the three different classification metrics, Precision, Recall, and F1-Score, on the Training and validation metrics, thus offering

a detailed description of the learning process of the model apart from the accuracy measurement.

Table 3
Performance Measures of the Proposed hr_{50}–$vit\text{-}_{b/16}$ Model in October 2017.

Class	Precision	Recall	F1-score	Specificity	Acc. (%)
CNV	0.964	0.968	0.966	98.8%	96.8
DME	0.952	0.952	0.952	98.4%	95.2
DRUSEN	0.956	0.948	0.952	98.6%	94.8
NORMAL	0.977	0.977	0.975	99.1%	97.7
Weighted Avg.	**0.962**	**0.961**	**0.961**	**98.7%**	96.14

The classification report using the "one-versus-the-rest" approach is shown in Table 3. The accuracy of each class (which is calculated by (TP_c + TN_c)/N) ranges between 95.2- 94.8(DME, DRUSEN) and 97.7 (NORMAL). The Recall weighted Avg. (0.9614) corresponds numerically with overall accuracy (96.14%).

Table 4
Class-wise Comparison for Ablation Variants of the Proposed Architecture

No	Architecture	Precision	Recall	F1-Score	Acc. (%)
a	$ResNet_{50}$ Only	0.8022	0.7862	0.7941	79.41
b	$ViT_{B/16}$ Only	0.8591	0.8476	0.8533	85.33
c	Concat (No Fusion)	0.9364	0.9193	0.927	91.79
d	**HR50_$ViT_{B/16}$**	**0.9614**	**0.9614**	**0.9614**	96.14

Table 5
Class-Wise Analysis of Ablation Variants on the Oct 2017 Dataset

No	Architecture	CNV F1	DME F1	DRUSEN F1	NORMAL F1
a	$ResNet_{50}$ Only	0.93	0.74	**0.56**	0.94
b	$ViT_{B/16}$ Only	0.90	0.82	0.76	0.93
c	Concat (No Fusion)	0.94	0.90	0.88	0.95
d	**HR_{50}_$ViT_{B/16}$**	**0.966**	**0.952**	0.952	0.975

From the results in Tables 4 and 5, the ResNet50 architecture achieved F1 scores of 0.93-0.94 when applied separately to the CNV and NORMAL datasets, owing to its noticeable local features, including a new inversion of the vascular membrane structure and the retinal outline. However, there was a decrease in performance of the ResNet50 architecture for DRUSEN detection (F1: 0.56). The self-attention module of $ViT_{B/16}$ achieved significantly higher F1 scores for DRUSEN (0.76) and DME (0.82). The combination of

the $ResNet_{50}$ and $ViT_{B/16}$ features, without any learned weighting, achieves an impressive result (accuracy 91.79, DRUSEN F1 0.88) compared to each model alone. The result is $HR_{50}_ViT_{B/16}$ with a DRUSEN F1 value of 0.952.

$ResNet_{50}$ convolutional models operate on local receptive fields and can detect edge responses, intensity gradients, and texture patterns. For CNV, neovascular membranes exhibit bright, localised characteristics that can be detected in 3×3 or 5×5-pixel regions. In the case of DRUSEN, detection of small ripples in RPE surface geometric distortion is required at 50- to 200-pixel intervals across multiple retinal layers. The CNN lacks sufficient context to distinguish those small ripples from the smooth profile of the retinal pigment epithelium in healthy images, resulting in a high number of false negatives for this type.

The $ViT_{B/16}$ architecture divided image 224×224 into 196 patches of 16×16 pixels without intersection, and concurrently considers the symmetrical connections between all patches through multi-headed self-attention mechanism (12 heads, 12 layers, embed_dim = 768), so, the spatial pattern of retinal pigment epithelial ripples (yellow deposits) and fluid-filled spaces within retinal layers (diabetic macular edema) irregular patterns characterized by their spatial distribution, not by the intensity of their localized pixels. The $ViT_{B/16}$ architecture alone is limited because it lacks the fine-grained convolutional texture characteristics. The F1 value of the CNV index was reduced from 0.93 ($ResNet_{50}$) to 0.90 ($ViT_{B/16}$), representing a reduction of -0.03, since the resolution of 16×16 patches lacks the sub-patch detail of the strongly reflecting neovascular membranes; so, the global attention and local texture extraction complement each other but are not the same skill set; both of the essential architectures are insufficient alone.

The distinction between (c) and (d) is the way the two feature spaces are fused. Model (c) fuses the 768-dimensional $ResNet_{50}$ representation with the 768-dimensional ViT CLS code into a 1536-dimensional vector when it is fed to the classification head. So, the $ResNet_{50}$ and $ViT_{B/16}$ feature paths are considered equally valuable for all classes, although DRUSEN classification takes more advantage of $ViT_{B/16}$ attention features, while CNV classification takes more advantage of $ResNet_{50}$ texture features.

The combination block in $HR_{50}_ViT_{B/16}$ (Linear 1536→768, LayerNorm, GELU, Dropout) learns a projection that adaptively reweights each feature flow based on the input image content. During training, the combination block parameters capture a specific weighting for each class, learning the weighting of the two complementary feature sets specific to that class. For DRUSEN, the fusion layer increases the contribution of $ViT_{B/16}$ attention features that capture the RPE surface texture, whereas for CNV, it increases the contribution of $ResNet_{50}$ texture reactions to the highly reflective membranes. For images with DRUSEN, the gradient update provides more weight to $ViT_{B/16}$ branch features; for images with CNV, it provides more weight to $ResNet_{50}$ branch features. It is unachievable through simple clustering.

The contribution of the fusion block (c → d) alone quantitatively is DRUSEN F1 (+0.072), DME F1 (+0.052), CNV F1 (+0.026), and NORMAL F1 (+0.025). The fusion module accounts for 18.4% of the total DRUSEN F1 increase from the ResNet-50 model to the proposed architecture (0.072 / 0.392 = 0.184) and 24.5% of the total DME F1 increase (0.052 / 0.212 = 0.245).

The Wilson interval statistical analysis was applied to assess the accuracy and reliability of each model, and a two-ratio z-test was used to establish statistical significance of accuracy differences between the proposed model and each baseline on a test set of 1011 images. The results can be illustrated in Tables 6 and 7.

Table 6
Wilson Score Confidence Intervals

No	Architecture	Acc. (%)	95% CI Lower	95% CI Upper
a	$ResNet_{50}$ Only	79.41	76.81	81.79
b	$ViT_{B/16}$ Only	85.33	83.02	87.38
c	Concat (No Fusion)	91.79	89.94	93.33
d	**HR_{50}_$ViT_{B/16}$**	**96.14**	**94.77**	**97.16**

Table 6 shows the accuracies of the different architectures on the fixed test set from OCT2017 (N = 1,011), with 95% confidence intervals estimated using Wilson score confidence intervals. The observation intervals of the four models don't overlap, and the features learned in each of the proposed architectures ($ViT_{B/16}$ branch, $ResNet_{50}$ branch, Learner integration) are transferred to the following layer.

Table 7
Z-Test for Two Ratios to Compare Models

Architecture	z-statistic	p-value
d vs. a (HR_{50}_$ViT_{B/16}$ vs. $ResNet_{50}$ only)	11.48	4.8×10^{-4}
d vs. b (HR_{50}_$ViT_{B/16}$ vs. $ViT_{B/16}$ only)	8.38	5.0×10^{-6}
d vs. c (HR_{50}_$ViT_{B/16}$ vs. Concat, no fusion)	4.11	< 0.0001

Table 7 shows two-proportion z-test results comparing the overall classification accuracy of the proposed HR50_ViTB/16 model against each ablation baseline on the test set (N = 1,011 per model). The interval results for the models indicate that each model was well estimated and remained stable as the test set size increased, and the two-ratio z-test results confirm the information gained by each model in the proposed merging architecture of the framework. $p = 4.0 \times 10^{-5}$ is the quantitative statistical basis demonstrating the improvement in the framework's merging architecture. Table 8 compares the proposed methodology's performance with other studies.

Table 8
Comparative Performance of Prior Studies on the Oct 2017 dataset

Study	Architecture	Accuracy	Key Distinction from $HR_{50}_ViT_{B/16}$
Perdomo et al., 2018	Custom CNN (12-layer) Leave-one-out CV	Accuracy= **93.75%**	on DME vs Normal classification using binary classification
Bhandari et al., 2023	Lightweight CNN + with k-fold (k=10), using FGSM, LIME, and SHAP	Accuracy= 96 %, test accuracy= 94.29%, verification accuracy= 94.12%	Classification using pure CNN on OCT images (CNV, DME, DRUSEN, Normal), Novelty is adversarial robustness (FGSM) + explainability (LIME, SHAP)
Dutta et al., 2023	InceptionV3+ResNet-50+ViT	weighted F1 / Acc= 0.944 / 94.46%	triple-stream hybrid model for Conv-ViT with InceptionV3, ResNet-50 and ViT used in achieved a weighted F1-score 0.944, F1-score of DRUSEN 0.780 and 94.46% accuracy in OCT2017 dataset.
Ma et al., 2022	ConvNet + Transformer	weighted F1 / Acc=91.56%	hybrid ConvNet-Transformer architecture, Residual dense block+ parallel ConvNet and Transformer branches for adaptive re-weighting fusion. For the OCT2017 dataset, HCTNet was reach 91.56% accuracy
Yang et al., 2024	Multi-scale CNN + Transformer	97.16%	In October 2017, the process of four-class classification used HRS-Net, a hybrid multi-scale network in which ResNet50 was branched out into parallel CNN (residual + attention) and Swin Transformer streams.
$HR_{50}_ViT_{B/16}$	$ResNet_{50}$ + $ViT_{B/16}$	96.14%	CNV, DME, DRUSEN, Normal

Transformers have also been used for pixel-wise segmentation of medical images in TransUNet (Chen et al., 2024; Turan et al., 2025), and the Attention Mechanism (Valanarasu et al., 2021) was applied to gland segmentation and polyp segmentation, respectively. These two methods are relatively new applications of the Attention Mechanism in ophthalmology. They should both be adapted for the classification problem.

CONCLUSION

The ($HR_{50}_ViT_{B/16}$) approach used for the classification of OCT images for retinal pathology showed 96.14% overall accuracy. The use of the $ViT_{B/16}$ encoder and the $ResNet_{50}$ backbone in parallel enables joint consideration of global spatial dependencies, modelled by 12 layers of Multi-Head Self-Attention with 196 patch tokens, as well as hierarchical local texture features extracted via the four residual stages. The improvement in DRUSEN F1-score was 0.952. The qualitative analysis of the attention overlay showed that the ViT

branch successfully localises attention to the relevant sub-RPE and intraretinal areas in the DRUSEN and DME cases, respectively, providing evidence of the model's ability to learn diagnostically meaningful spatial representations. To further improve the existing fusion paradigm, we propose replacing the linear concatenation strategy with a new strategy based on the Multi-Head Cross-Attention (MHCA) method in the ViT transformer as Future Work. It can effectively address inter-class similarity. Training the model with multiple random seeds (at least five) to enhance statistical stability and estimate the variation in the operation process.

LIST OF ABBREVIATION

Term	Definition
CNV	Neovascularisation
DME	Diabetic Macular Oedema
AMD	Age-Related Macular Degeneration
OCT	Optical Coherence Tomography
ViT	Vision Transformer Model

REFERENCES

Agresti, A. (2019). *An introduction to categorical data analysis*. (3rd ed.). Wiley.

Akça, S., Garip, Z., Ekinci, E., & Atban, F. (2024). Automated classification of choroidal neovascularization, diabetic macular edema, and drusen from retinal OCT images using vision transformers: A comparative study. *Lasers in Medical Science, 39*, Article 140. https://doi.org/10.1007/s10103-024-04089-w

Akpinar, M. H., Sengur, A., Faust, O., Tong, L., Molinari, F., & Acharya, U. R. (2024). Artificial intelligence in retinal screening using OCT images: A review of the last decade (2013-2023). *Computer Methods and Programs in Biomedicine, 254*, Article 108253. https://doi.org/10.1016/j.cmpb.2024.108253

Alsaif, O. I., Hasan, S. Q., & Maray, A. H. (2023). Using skeleton model to recognize human gait gender. *IAES International Journal of Artificial Intelligence, 12*(2), 974-983. https://doi.org/10.11591/ijai.v12.i2.pp974-983

Aslan, M. F. (2023). Comparison of vision transformers and convolutional neural networks for skin disease classification. In *Proceedings of the International Conference on New Trends in Applied Sciences* (pp. 31-39). https://doi.org/10.58190/icontas.2023.51

Bhandari, M., Shahi, T. B., & Neupane, A. (2023). Evaluating retinal disease diagnosis with an interpretable lightweight CNN model resistant to adversarial attacks. *Journal of Imaging, 9*(10), Article 219. https://doi.org/10.3390/jimaging9100219

Brown, L. D., Cai, T. T., & DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science, 16*(2), 101-133. https://doi.org/10.1214/ss/1009213286

Burton, M. J., Ramke, J., Marques, A. P., Bourne, R. R. A., Congdon, N., Jones, I., Ah Tong, B. A. M., Arunga, S., Bachani, D., Bascaran, C., Bastawrous, A., Blanchet, K., Braithwaite, T., Buchan, J. C., Cairns, J.,

Cama, A., Chagunda, M., Chuluunkhuu, C., Cooper, A., . . . Faal, H. B. (2021). The Lancet Global Health Commission on Global Eye Health: Vision beyond 2020. *The Lancet Global Health, 9*(4), e489-e551. https://doi.org/10.1016/s2214-109x(20)30488-5

Chen, J., Mei, J., Li, X., Lu, Y., Yu, Q., Wei, Q., Luo, X., Xie, Y., Adeli, E., Wang, Y., Lungren, M. P., Zhang, S., Xing, L., Lu, L., Yuille, A., & Zhou, Y.. (2024). TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers. *Medical Image Analysis, 97*, Article 103280. https://doi.org/10.1016/j.media.2024.103280

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*. https://openreview.net/forum?id=YicbFdNTTy

Du, K., Nair, A. R., Shah, S., Gadari, A., Vupparaboina, S. C., Bollepalli, S. C., Sutharahan, S., Sahel, J.A., Jana, S., Chhablani, J., & Vupparaboina, K. K. (2024). Detection of disease features on retinal OCT scans using RETFound. *Bioengineering, 11*(12), Article 1186. https://doi.org/10.3390/bioengineering11121186

Dutta, P., Sathi, K. A., Hossain, M. A., & Dewan, M. A. A. (2023). Conv-ViT: A convolution and vision transformer-based hybrid feature extraction method for retinal disease detection. *Journal of Imaging, 9*(7), Article 140. https://doi.org/10.3390/jimaging9070140

Fang, L., Jin, Y., Huang, L., Guo, S., Zhao, G., & Chen, X. (2019). Iterative fusion convolutional neural networks for classification of optical coherence tomography images. *Journal of Visual Communication and Image Representation, 59*, 327-333. https://doi.org/10.1016/j.jvcir.2019.01.022

Gencer, G., & Gencer, K. (2025). Advanced retinal disease detection from OCT images using a hybrid squeeze and excitation enhanced model. *PLOS ONE, 20*(2), Article e0318657. https://doi.org/10.1371/journal.pone.0318657

Hassan, E., Elmougy, S., Ibraheem, M. R., Hossain, M. S., AlMutib, K., Ghoneim, A., AlQahtani, S. A., & Talaat, F. M. (2023). Enhanced deep learning model for classification of retinal optical coherence tomography images. *Sensors, 23*(12), Article 5393. https://doi.org/10.3390/s23125393

He, X., Deng, Y., Fang, L., & Peng, Q. (2021). Multi-modal retinal image classification with modality-specific attention network. *IEEE Transactions on Medical Imaging, 40*(6), 1591–1602. https://doi.org/10.1109/tmi.2021.3059956

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., . . . Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell, 172*(5), 1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

Laouarem, A., Kara-Mohamed, C., Bourennane, E.B., & Hamdi-Cherif, A. (2024). HTC-retina: A hybrid retinal diseases classification model using transformer–convolutional neural network from optical coherence tomography images. *Computers in Biology and Medicine, 178*, Article 108726. https://doi.org/10.1016/j.compbiomed.2024.108726

Li, D., Ran, A. R., Cheung, C. Y., & Prince, J. L. (2023). Deep learning in optical coherence tomography: Where are the gaps? *Clinical and Experimental Ophthalmology, 51*(8), 853-863. https://doi.org/10.1111/ceo.14258

Liew, A., & Agaian, S. (2025). Comprehensive survey of OCT-based disorders diagnosis: From feature extraction methods to robust security frameworks. *Bioengineering, 12*(9), Article 914. https://doi.org/10.3390/bioengineering12090914

Ma, Z., Xie, Q., Xie, P., Fan, F., Gao, X., & Zhu, J. (2022). HCTNet: A hybrid ConvNet-transformer network for retinal optical coherence tomography image classification. *Biosensors, 12*(7), Article 542. https://doi.org/10.3390/bios12070542

Mohammed, N. S., Almolaa, O. S., Abdullah, A. S., & Alsaif, K. I. (2026). Accurate brain tumor classification using MRI images based on a hybrid Vision Transformer and BiLSTM framework. *ITEGAM-JETIA, 12*(60), 526-535. https://doi.org/10.5935/jetia.v12i60.3889

Perdomo, O., Otalora, S., Gonzalez, F. A., Meriaudeau, F., & Muller, H. (2018). OCT-NET: A convolutional network for automatic classification of normal and diabetic macular edema using SD-OCT volumes. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 1423-1426). IEEE. https://doi.org/10.1109/isbi.2018.8363839

Rahil, M., Anoop, B. N., Girish, G. N., Kothari, A. R., Koolagudi, S. G., & Rajan, J. (2023). A deep ensemble learning-based CNN architecture for multiclass retinal fluid segmentation in OCT images. *IEEE Access, 11*, 17241-17251. https://doi.org/10.1109/access.2023.3244922

Raut, P., Babar, S., & Mahalle, P. (2025). Fusion of vision transformer, Inception-V3 and ResNet50 for efficient eye disease detection. *Journal of Information Systems Engineering and Management, 10*(7s), 315-325. https://doi.org/10.52783/jisem.v10i7s.864

Sujon, K. M., Hassan, R., Choi, K., & Samad, M. A.. (2025). Accuracy, precision, recall, F1-score, or MCC? Empirical evidence from advanced statistics, ML, and XAI for evaluating business predictive models. *Journal of Big Data, 12*, Article 268. https://doi.org/10.1186/s40537-025-01313-4

Teo, Z. L., Tham, Y.C., Yu, M., Chee, M. L., Rim, T. H., Cheung, N., Bikbov, M. M., Wang, Y. X., Tang, Y., Lu, Y., Wong, I. Y., Ting, D. S. W., Tan, G. S. W., Jonas, J. B., Sabanayagam, C., Wong, T. Y., & Cheng, C.Y. (2021). Global prevalence of diabetic retinopathy and projection of burden through 2045: Systematic review and meta-analysis. *Ophthalmology, 128*(11), 1580-1591. https://doi.org/10.1016/j.ophtha.2021.04.027

Turan, S. C., Çıplak, Z., Sarıkaş, A., & Yıldız, K. (2025). A deep learning approach to document recovery: High performance with DenoiseU-Net. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 25*(6), 1323-1335. https://doi.org/10.35414/akufemubid.1628066

Valanarasu, J. M. J., Oza, P., Hacihaliloglu, I., & Patel, V. M. (2021). Medical transformer: Gated axial-attention for medical image segmentation. In M. de Bruijne, P. C. Cattin, S. Cotin, N. Padoy, S. Speidel, Y. Zheng, & C. Essert (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2021* (pp. 36-46). Springer. https://doi.org/10.1007/978-3-030-87193-2_4

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 30, pp. 5998-6008). Curran Associates. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Vision Loss Expert Group of the Global Burden of Disease Study, & the GBD 2019 Blindness and Vision Impairment Collaborators. (2024). Global estimates on the number of people blind or visually impaired by diabetic retinopathy: A meta-analysis from 2000 to 2020. *Eye, 38*(11), 2047-2057. https://doi.org/10.1038/s41433-024-03101-5

Wang, H., Guo, X., Song, K., Sun, M., Shao, Y., Xue, S., Zhang, H., & Zhang, T. (2023). OCTFormer: An efficient hierarchical Transformer network specialized for retinal optical coherence tomography image recognition. *IEEE Transactions on Instrumentation and Measurement, 72*, 1-17. https://doi.org/10.1109/TIM.2023.3329106

Wang, S., Feng, S., Wang, Z., Ji, Z., Liu, J., Chen, W., Fu, B., Liu, R., Chen, W., Dai, Y., & Li, C. (2025). Structural-prior guided and feature-enhanced transformer with masked image modeling pretraining for retinal layers and fluid segmentation in macular edema OCT images. *Biomedical Optics Express, 16*(12), 5096-5117. https://doi.org/10.1364/boe.574769

Xu, Z., Wang, W., Yang, J., Zhao, J., Ding, D., He, F., Chen, D., Yang, Z., Li, X., Yu, W., & Chen, Y. (2021). Automated diagnoses of age-related macular degeneration and polypoidal choroidal vasculopathy using bi-modal deep convolutional neural networks. *British Journal of Ophthalmology, 105*(4), 561-566. https://doi.org/10.1136/bjophthalmol-2020-315817

Yang, H., Chen, L., Cao, J., & Wang, J. (2024). HRS-Net: A hybrid multi-scale network model based on convolution and transformers for multi-class retinal disease classification. *IEEE Access, 12*, 144219-144229. https://doi.org/10.1109/access.2024.3467286

Yang, M., Du, J., & Lv, R. (2025). CRAT: Advanced transformer-based deep learning algorithms in OCT image classification. *Biomedical Signal Processing and Control, 104*, Article 107544. https://doi.org/10.1016/j.bspc.2025.107544

Zhou, Z., Niu, C., Yu, H., Zhao, J., Wang, Y., & Dai, C. (2023). Diagnosis of retinal diseases using the vision transformer model based on optical coherence tomography images. *Proceedings of SPIE, 12601*, Article 1260102. https://doi.org/10.1117/12.2665918

Pertanika J. Sci. & Technol. 34 (5): 3025 - 3045 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Effects of Elevated CO_2-Grown Napier Grass and Cassava Starch Binder Ratios on Biomass Briquette Properties

Muhammad Zulhilmi Mohd Nasirudin[1]* and Siti Zaharah Sakimin[2]

[1]*Department of Crop Science, Faculty of Agriculture, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*

[2]*Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*

ABSTRACT

Napier grass (*Pennisetum purpureum* cv. Taiwan) has considerable potential as a renewable biomass resource for solid biofuel production; however, the quality of charcoal briquettes depends heavily on binder composition and processing conditions. This study aimed (i) to determine the optimal mixing ratio between burnt Napier grass and cassava starch and (ii) to evaluate the physical properties, combustion-related characteristics, and nutrient profile of Napier grass–based charcoal briquettes. The briquettes were produced through a systematic modelling process involving carbonisation, crushing, drying, briquetting, final drying, and storage. Six treatments (T1–T6) were formulated using different ratios of burnt Napier grass and cassava starch. The briquettes were assessed for compression strength, hydrophobicity, moisture content, particle density, bulk density, porosity, moisture sorption, and nutrient profile. The results indicated that increasing cassava starch content improved briquette density and compression strength but reduced porosity and moisture sorption capacity. Among all treatments, T4 (90% burnt Napier grass + 10% cassava starch) exhibited the most balanced properties, with moderate compression, favourable hydrophobicity, stable moisture content, moderate particle and bulk densities, balanced porosity, and the highest nutrient profile. Treatments with lower binder content showed weaker structural integrity, whereas higher binder proportions resulted in reduced porosity and nutrient performance. In conclusion, T4 was identified as the most suitable formulation for charcoal briquette modelling, offering an optimal balance between physical stability, combustion-related properties, and nutrient retention, and demonstrating strong potential for sustainable biofuel production from Napier grass.

Keywords: Briquette formulation, carbonisation, cassava starch binder, charcoal, green energy, renewable resources

ARTICLE INFO

Article history:
Received: 11 May 2026
Accepted: 10 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.20

E-mail addresses:
zulhilmi_nasaida@yahoo.com.my (Muhammad Zulhilmi Mohd Nasirudin)
szaharah@upm.edu.my (Siti Zaharah Sakimin)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

Malaysia is actively accelerating its energy transition through the implementation of the National Energy Transition Roadmap (NETR) and the Malaysia Renewable Energy Roadmap (MyRER), which aim to increase the share of renewable energy in the national electricity generation mix to 31% by 2025 and 70% by 2050 (Sustainable Energy Development Authority Malaysia [SEDA], 2021). These policies emphasise diversifying renewable energy sources, including solar, hydropower, biogas, biomass, and waste-to-energy, to reduce greenhouse gas emissions while strengthening national energy security. Among these renewable resources, biomass derived from agricultural crops and residues represents a sustainable feedstock for solid biofuel production due to its abundance, renewability, and potential to substitute fossil fuels (Sawasdee & Pisutpaisal, 2021). Consequently, converting high-yield energy crops such as Napier grass into biomass briquettes offers a promising approach to support Malaysia's renewable energy objectives while promoting the sustainable utilisation of biomass resources (Ministry of Economy, 2023).

Among Malaysia's renewable energy resources, biomass has gained considerable attention owing to its abundant availability from agricultural crops and residues, including oil palm, rice husk, sugarcane bagasse, and Napier grass (Jha et al., 2022). Unlike intermittent renewable resources such as solar and wind, biomass provides a dispatchable energy source that can be converted into solid biofuels, including charcoal and biomass briquettes, with improved energy density, storage stability, and transportation efficiency (Economic Planning Unit, 2022). As Malaysia continues to promote biomass utilisation under the Malaysia Renewable Energy Roadmap (MyRER), improving the quality and performance of biomass briquettes has become increasingly important. In this study, Taiwan Napier grass (*Pennisetum purpureum* cv. Taiwan) was selected as the primary feedstock for charcoal briquettes due to its high biomass productivity, rapid growth, and favourable fuel characteristics (Sawasdee & Pisutpaisal, 2021). These policies were introduced not only to reduce carbon emissions but also to increase the usage of renewable resources in Malaysia (SEDA, 2021). In this context, the present study evaluates the properties of biomass briquettes1) to determine the best ratio of combination between burnt Napier grass (*Pennisetum purpureum* cv. Taiwan) and different ratios of cassava starch, and 2) to investigate the physical analysis, combustion properties and nutrient profile of charcoal briquettes from the Napier grass (*Pennisetum purpureum* cv. Taiwan). Interest in alternative energy has increased demand for renewable resources while reducing reliance on non-renewable energy (Dinesha et al., 2019).

MATERIALS AND METHODS

Napier Grass Grown Environment

Napier grass (Pennisetum purpureum) used in this study was obtained from two contrasting growing environments: ambient CO_2 (BNG−eCO_2) and greenhouse (BNG+eCO_2). The BNG−eCO_2 plants were cultivated under natural atmospheric conditions in an open field at Farm 15, Faculty of Agriculture, Universiti Putra Malaysia (UPM), Serdang, Malaysia, whereas the BNG+eCO_2 plants were grown in a controlled greenhouse facility at Tenaga Nasional Berhad Research (TNBR), Jalan Ayer Hitam, Kajang, Selangor, Malaysia. The greenhouse maintained an eCO_2 concentration of 800 μmol mol^{-1}, while the open-field plants were exposed to ambient atmospheric CO_2 concentrations. The +eCO_2 treatment was applied for 1 month prior to harvest. All plants were harvested after 1 year of growth, with four biological replicates for each treatment. The harvested biomass was subsequently oven-dried, carbonised, and processed into briquettes.

The nutrient composition of the Napier grass biomass was also characterised prior to carbonisation because mineral elements influence the physicochemical properties of biomass fuels (Botero-Londoño et al., 2021). Elements such as potassium, calcium, magnesium, phosphorus, and silicon contribute to ash composition and may affect combustion behaviour, ash yield, and the potential for slagging or fouling during fuel utilisation (Jungers et al., 2015). Furthermore, biomass grown under $eCO_{2\ exhibits}$ changes in nutrient accumulation due to altered plant physiology and carbon allocation. Therefore, the pre-carbonisation nutrient profile (Table 1) was included to characterise the feedstock and to provide a basis for interpreting differences in briquette properties between ambient- and eCO_2- grown biomass.

Table 1
Growth conditions and pre-carbonisation characteristics of Napier grass biomass used for briquette production (Mohd Nasirudin et al., 2024)

Parameter		BNG−eCO_2 (Open Field)	BNG+eCO_2 (Greenhouse)
Biomass yield (t ha^{-1})		82.90	84.12
Calorific value (MJ kg^{-1})		18.88	15.84
Lignin content (%)		14.26	10.74
Proximate analysis	Moisture content (%)	6.24	5.90
	Volatile matter (%)	87.11	87.21
	Ash content (%)	2.31	3.56
	Fixed carbon (%)	4.34	6.37
Ultimate analysis	Total carbon (%)	35.15	43.96
	Total hydrogen (%)	5.58	5.81
	Total nitrogen (%)	1.14	1.19
	Total sulphur (%)	0.05	0.08

Treatment and Experimental Design

The experiment was arranged in a Completely Randomised Design (CRD) with six treatments to evaluate the effect of different blending ratios of burnt Napier grass (BNG) and cassava starch (CS). Cassava starch was chosen as the binder agent because of its excellent adhesive properties, biodegradability, non-toxicity, and widespread availability. Upon gelatinisation, starch forms a viscous matrix that effectively fills void spaces and promotes strong interparticle bonding, resulting in briquettes with improved particle density, bulk density, and mechanical strength. In addition, cassava starch contributes little ash and does not significantly impair the briquettes' combustion characteristics (Livingston & Babcock, 2006). Burnt Napier grass was derived from crops that were grown in open area conditions (BNG-eCO_2) and under greenhouse conditions (BNG+eCO_2). Each BNG source was combined with cassava starch at three ratios (w/w): 9:1, 8:2, and 7:3. The treatments were as follows: 90% BNG-eCO_2 + 10% CS (T1), 80% BNG-eCO_2 + 20% CS (T2), 70% BNG-eCO_2 + 30% CS (T3), 90% BNG+eCO_2 + 10% CS (T4), 80% BNG+eCO_2 + 20% CS (T5), and 70% BNG+eCO_2 + 30% CS (T6). All treatments were randomly assigned to experimental units to ensure uniformity and reduce experimental bias.

Combustion Process in Modelling for Charcoal Briquettes Production

Carbonisation

Carbonisation is the incomplete, or primary, combustion of plant-based material, such as Napier grass, resulting in charcoal briquettes. The biochar product was fully oxidised through a thermochemical process under limited-oxygen conditions, converting organic material into carbon-rich charcoal (Mohammed et al., 2015). The plant-based material, such as Napier grass, was prepared into smaller cuttings between 50 and 100 cm to optimise the carbonisation process. The sample was burnt under temperatures between 400 °C and 500 °C in a kiln (Figure 1) under low-oxygen conditions (Md Said et al., 2019). The optimal temperature was important for maximising charcoal production while controlling thermal decomposition. Wood vinegar and tar were byproducts of the carbonisation process, while the stable carbon structure was retained during char formation (Murakami & Sato, 2024). The carbonisation process completely burns the organic material within 1 week. Next, the heated sample was left to cool for 3 days in an oxygen-limited environment to prevent spontaneous combustion. After 3 days, the sample was ready for crushing from a large particle size to a finer size, which is more suitable for briquette formation. Proper management in the carbonisation process is important to ensure the quality and burning characteristics are retained at good conditions as it progresses over time (Elehinafe et al., 2021).



Figure 1. Carbonisation process of Napier grass in a kiln for charcoal production

Crushing

The crushing method is important in the combustion process to ensure that all particles are uniformly sized during biochar formation. Particle size plays a crucial role in achieving maximum combustion efficiency (Pang et al., 2019). The sample must be cooled, then only fed into the crushing machine, the roller crusher (Figure 2). The roller crusher breaks hard particles into a fine powder (< 3mm). The fine powder was filtered through a mesh sieve (2–3 mm), and particles larger than 3 mm were removed to ensure uniform particle distribution. Particle size affects the strength of charcoal during the briquetting process: the finer the particles, the better they bind together, whereas uneven particle size can weaken the briquettes' structure (Sunardi et al., 2019). The high number of holes in the briquette structure can lead to poor and inefficient combustion (Rodriguez Correa et al., 2017). Thus, a proper crushing method

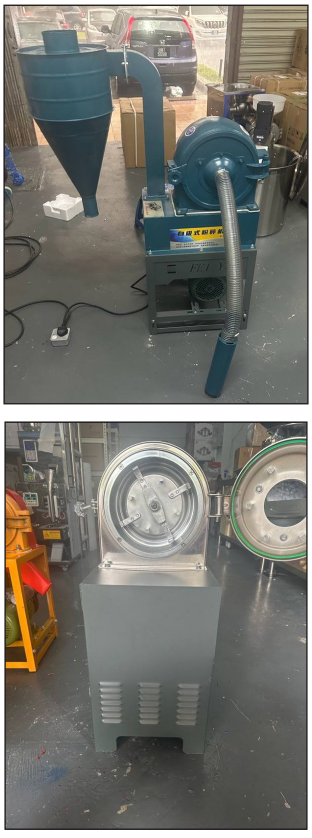


Figure 2. Crushing machines are used to grind carbonised Napier grass into fine charcoal powder for briquette production

can contribute to high density and durability of the charcoal product and improve overall performance, as well as the product's ability to hold solid formation together with binding agents (Chirchir et al., 2013). Finally, the powder was kept at ambient temperature and stored in sealed containers to prevent moisture absorption, maintain burning performance, and minimise ash production (Monti et al., 2008).

Drying

Next, the crushed sample was left to dry until its moisture content was no more than 10%. However, if the sample still contains more than 10% moisture, it must be further dried to reach the required level. The traditional drying method was implemented in this experiment due to its low cost and ease of handling. The crushed sample was exposed to direct sunlight under thin trays for several hours to several days. Drying is important in the briquette process not only to reduce smoke emissions but also to enhance energy output and improve compaction efficiency (Obi et al., 2022). The dried sample was stored in a moisture-free environment to prevent reabsorption before proceeding to the briquette formation stage.

Briquetting

Briquetting is the process of compressing the sample by using a compression machine, which turns plant-based material into a uniform shape with the presence of binding agents. The machine processes loose material into compact charcoal with high mechanical strength, which improves combustion efficiency. Firstly, the dried sample was mixed with a binding agent (cassava starch). In this experiment, cassava stems with different concentrations (10, 20 and 30%) were mixed with plant-based material (Napier grass). The dried sample powder was combined with the binder and a small amount of water to create a uniform paste-like mixture. Next, the mixture was fed into a briquetting machine (Figure 3), which converted it into a solid under reasonable pressure. The charcoal needs to undergo a second drying once the briquettes are formed. The second drying was purposely to remove excess moisture before the final drying process took place. Proper briquetting ensures the briquettes burn efficiently (Olugbade et al., 2019).



Figure 3. Briquetting machines are used to compress biochar powder, along with binding agents, into solid charcoal briquettes

Final Drying and Storage of Charcoal Briquettes

The final drying in the briquetting process deliberately hardened the charcoal mixture between the plant-based material and the binder, thereby optimising its moisture content, especially for long-term storage. Normally, after the briquetting process, the charcoal still retains a slightly high moisture content of 15-20%. However, after final drying, the new moisture content is less than 8%, thereby reducing smog production and enhancing combustion efficiency (Rocha et al., 2017). Both initial and final drying were exposed to direct sunlight under racks or trays in an eCO_2 area for 2 to 5 days until the briquettes reached a constant weight. After drying, the briquettes were stored in sealed polyethene bags at room temperature (25 to 30 °C) in a dry, well-ventilated environment, protected from direct sunlight and moisture to minimise moisture uptake and preserve their physical and combustion properties prior to analyses, including physical and combustion tests (Song et al., 2010). The final step guarantees that the charcoal briquettes provide high heat output, long burn duration, and eco-friendly fuel properties.

Data Collection

Physical Analysis

Compression strength (CS) and hydrophobicity (HP) are important indicators for assessing briquette quality, measured with a digital hardness tester (YHKC-2A, particle strength tester, from Xinghuan, Jiangsu, China), which records the maximum force applied to the briquette charcoal. The hydrophobicity of a sample can be expressed as equilibrium moisture content (ECM), the amount of moisture retained by the material after reaching moisture equilibrium with the surrounding environment at specified temperature and relative humidity. Hydrophobicity (Equation 1) was defined as the tendency of a material to resist water absorption or repel moisture when exposed to humid environments or direct contact with water. Good hydrophobicity was indicated with a lower ECM reading of the briquettes. The ECM reading was calculated by exposing the briquettes to a relative humidity of 82-84% at 24-25 °C for 18 hours (controlled environment). The ECM was calculated using the following formula (ASAE, 2006):

$$\mathrm{ECM} = \frac{\mathrm{M} - \mathrm{Mi}}{\mathrm{M} \times 100\%} \quad [1]$$

Combustion Properties

The moisture content (MC, % db) of the samples was calculated after grinding 10 g of each charcoal sample at 130 °C for 1 hour in an air convection oven (Song et al., 2023). The particle density (PD) and porosity (Po) were measured by using a digital calliper

(Model CD-6CSX, Mitutoya Corp., Kawasaki, Japan), which recorded the diameter (d) and length (L) of each charcoal sample. The total charcoal sample was 5 individual briquettes from each treatment. The total weight of 5 briquettes was measured with a digital balance (KERN 572, Germany) with an accuracy of 0.01 g, while volume (Equation 2) and particle density (Equation 3) were calculated using the following formula.

$$vP = \frac{\pi}{4} d^2 L \quad [2]$$

$$PD \times \frac{mP}{vP} \quad [3]$$

Where,
vP = Volume of pellet (m^3)
PD = Particle density of pellet (kg/m^3)
d = Diameter of pellet (m)
L = Length of pellet (m)
mP = Mass of pellet (kg)

The bulk density (BD) was calculated in accordance with the ASAE standard. A cylindrical container with a size of 300 mm (height) and 310 mm (diameter) was used to measure the mass of briquettes (Bhagwanrao & Singaravelu, 2014). The ratio of the collective mass of charcoal in the container to its total volume (1,137 mm^3) was recorded. Using the following formula, the intergranular porosity (Equation 4) of the pellets will be determined by calculating the bulk density (BD) and particle density (PD).

$$\epsilon = 1 - \frac{BD}{PD} \quad [4]$$

Where,
ϵ = Porosity
PD = Particle density of pellet (kg/m^3)
BD = bulk density of pellets (kg/m^3)

Moisture sorption is the process by which a material absorbs and/or adsorbs water vapour from the surrounding atmosphere until it reaches equilibrium with the ambient temperature and relative humidity. Moisture sorption (MS, % w.b) was measured under fixed air humidity (70% and 80%) and temperature (400 °C) in a conditioned air chamber (1.8 × 0.9 × 0.9 m). Measure 300 g of the charcoal sample and place it in the sample

container. The sample was left in the chamber for 24 h in a conditioned-air environment. The charcoal sample was observed to see whether the briquettes further absorbed the moisture or disintegrated. The mass of the charcoal (g) was measured every 30 min. The recorded data will be analysed by expert software that fits an exponential model to the moisture sorption data (Equation 5). The formula is as follows.

$$MS = (Mi - Me) \times e - k + Me \quad [5]$$

Where,
T = Time
Me = Moisture content at time (%, w.b.)
Mi = Initial moisture content (%, w.b.)
K = Coefficient of sorption

Nutritional Profile

Measure and record the weight of 1.0 g of crushed charcoal before placing the sample into a digestion tube. Add 10 mL of concentrated nitric acid, then gently shake the digestion tube to mix the sample thoroughly. Left the sample overnight. Prepared a control digestion tube consisting of concentrated nitric acid with no sample. The digestion tube was boiled for 2 hours at 115 °C in the holes of the heating block. Next, 5 mL of concentrated hydrochloric acid was added, and the procedure was continued for 2 hours at 115 °C until a clear solution was observed. The charcoal sample was left to cool at room temperature for 2 hours. Then, transfer to a test tube and top up to 100 mL with deionised water. The sample was gently shaken and then filtered through filter paper (Whatman 42, 113, or equivalent) into a plastic container. The digested test material solution is ready for analysis by ICP-OES. Calculation:

1. Analysis of P, K, Ca, Mg and S by using ICPOES (Eq. 6):

$$\%\ \text{element} = \frac{\text{Reading from ICPOES} \times \text{dilution} \times \text{volume (mL)}}{\text{Weight of test material in gm} \times 10000} \times 100 \quad [6]$$

2. Analysis of reported in µg/g (ppm) (Eq. 7):

$$\mu\frac{g}{g}\text{element} = \frac{\text{reading from ICPOES} \times \text{dilution} \times \text{volume (mL)}}{\text{weight of test material in gm} \times 10000} \times 100 \quad [7]$$

Statistical Analysis

All data collected in this experiment were analysed by using two-way analysis of variance (ANOVA) by Statistical Analysis System (SAS 9.4) for a Completely Random Design (CRD) with two factors and consisted of four replications to ensure consistent results of significant differences between treatment means by using least significant difference (LSD) at $P < 0.05$ level.

RESULT AND DISCUSSIONS

Physical Analysis

Compression Strength

In this study, modelling and simulation of charcoal briquettes from Napier Grass (Pennisetum purpureum cv. Taiwan) are crucial, as they determine the briquettes' durability, handling efficiency, and structural integrity. Figure 4 showed that T6 (70% of BNG + eCO2 + 30% of CS) exhibited the highest peak force (~5.5 kN), indicating the strongest briquette structure due to its high starch content, which acted as a superior binder. T5 and T2 (20% cassava starch) also demonstrated good strength but slightly lower than T6, indicating a balanced effect of binding and compaction. On the other hand, T3 (70% of BNG-eCO2 + 30% of CS) showed the lowest peak force (~2.5 kN), implying weaker briquette formation under -eCO_2. Notably, T4 (90% BNG + eCO2 + 10% CS) had an ideal compression strength (~3.5 kN), supporting the view that eCO2 improves hydrophobicity but yields moderate mechanical strength due to its lower starch content. These results were particularly beneficial, facilitating easier lignification, allowing the briquettes to retain their shape and resist premature degradation while remaining workable for combustion applications.

The variation in force (kN) directly affects displacement (mm), influencing the way briquettes deform and break under pressure (Mitchual et al., 2013). Treatments with higher starch content (T6 and T3) generally exhibited greater displacement before breakage, were more ductile, and absorbed more energy before failure, thereby affecting burning efficiency. In contrast, briquettes with lower starch content (T1, T4) exhibited lower displacement, indicating that they were more brittle and easier to ignite. This suggests that increasing starch concentration improves structural flexibility but reduces the energy performance (Knoll et al., 2013). However, T4 emerged as the ideal treatment, as it achieved a balance between high hydrophobicity and moderate compressive strength, making it more resistant to moisture while still maintaining structural integrity (Mitchual et al., 2013). The relationship between force and displacement further indicates that starch enhances both strength and flexibility, reducing brittleness and improving coal durability (Safana & Usman, 2019). Therefore, optimising binder composition while considering hydrophobicity is essential for producing high-quality Napier grass charcoal briquettes (T4).

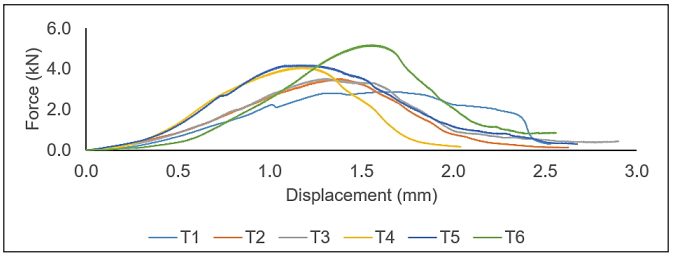


Figure 4. Effect of different material ratios between burnt Napier grass (BNG) and cassava starch (CS) on compression strength (Cs) of charcoal briquettes from Napier grass (Pennisetum purpureum cv. Taiwan)
Note. [T1] 90% of burnt Napier grass (BNG-eCO_2) + 10% of cassava starch (CS); [T2] 80% of BNG-eCO_2 + 20% of CS; [T3] 70% of BNG-eCO_2 + 30% of CS; [T4] 90% of BNG+eCO_2 + 10% of CS; [T5] 80% of BNG+eCO_2 + 20% of CS; and [T6] 70% of BNG+eCO_2 + 30% of CS

Hydrophobicity

Figure 5 shows that the combination ratio of Napier grass and cassava starch significantly affects hydrophobicity performance. In this study, T4 (90% BNG + eCO2 + 10% CS) showed the highest hydrophobicity reading among the treatments. Meanwhile, T5 (80% of BNG+eCO2 + 20% of CS) was recorded as the second-most hydrophobic, followed by T6 (70% of BNG+eCO2 + 30% of CS), which showed reductions of 28.04% and 32.75%, respectively, compared to T4. The results showed that a decreasing amount of burnt Napier grass reduced hydrophobicity from T4 to T6. T6 exhibits higher moisture absorption than T4, resulting in reduced charcoal stability, durability, and combustion performance. All treatments with -eCO_2 remained lower than the eCO2 treatment, with T1 (80% BNG-eCO2 + 20% CS) showing the highest reading among the groups. However, T2 and T3 show a negative impact on hydrophobicity, with reductions of 1.78% and 6.22%, respectively, compared to T1. This result shows that an increasing percentage of cassava starch negatively impacts hydrophobicity.

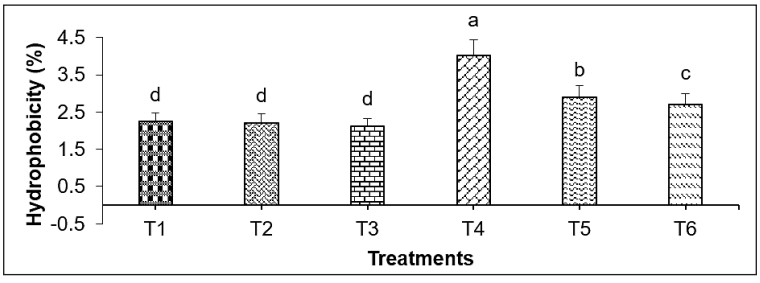


Figure 5. Effect of different material ratios between burnt Napier grass (BNG) and cassava starch (CS) on hydrophobicity (Hb) of charcoal briquettes from Napier grass (Pennisetum purpureum cv. Taiwan). Mean values with the same letter are not significantly different at P>0.05 by the least significant difference (LSD)
Note. n = 24, [T1] 90% of burnt Napier grass (BNG-eCO_2) + 10% of cassava starch (CS); [T2] 80% of BNG-eCO_2 + 20% of CS; [T3] 70% of BNG-eCO_2 + 30% of CS; [T4] 90% of BNG+eCO_2 + 10% of CS; [T5] 80% of BNG+eCO_2 + 20% of CS; and [T6] 70% of BNG+eCO_2 + 30% of CS

The lowest hydrophobicity was observed in T3 (2.11%), making it 47.6% less hydrophobic than T4, further proving that higher cassava starch proportions negatively impact moisture resistance due to the physical characteristics of starch as a hydrophilic material, absorbing moisture and reducing the water-repellent nature of the briquettes (Kupryianchyk et al., 2016). T1 and T4 exhibited greater hydrophobicity among their groups, especially T4, which has a more carbonised structure because of short eCO2, which can repel moisture.

In conclusion, a higher Napier grass content under T4 significantly improved hydrophobicity, thereby increasing ignition time and energy output from charcoal briquettes (Silviana et al., 2023). Additionally, T4 has become the most prominent treatment for biochar production, as it absorbs less moisture, achieves complete combustion, and exhibits high flame stability, resulting in reduced smoke production (Nyambati et al., 2010). Ultimately, the integration of hydrophobicity-enhanced briquettes in bio-coal applications can lead to more sustainable and high-performance biomass fuels, contributing to the advancement of renewable energy technologies, which are preferable for industrial and domestic use, as they maintain their quality over longer storage periods and provide a more consistent energy output (Ibeto et al., 2011).

Moisture Content

Figure 6 showed a significant effect of the combination ratio of burnt Napier grass and cassava starch on moisture content. The results show that T1 (90% of BNG-eCO2 + 10% of CS) recorded the highest moisture content, but no significant difference compared with treatment T2 (80% of BNG-eCO2 + 20% of CS) at 5.57% and T3 (70% of BNG-eCO2 + 30% of CS) at 5.74%. Besides that, treatment that exposed to +eCO_2 was showed significantly lower moisture content, with T6 (70% of BNG+eCO2 + 30% of CS) was recorded the lowest reading at 3.25%, T5 (80% of BNG+eCO2 + 20% of CS) at 2.94%, and T4 (90% of BNG+eCO2 + 10% of CS) at the lowest moisture content of 2.55%. The results show that eCO2 has enhanced the carbonisation of Napier grass, making it more repellent to moisture absorption compared to treatments T1–T3. The positive response improved the briquette's combustion quality and storage stability, attributable to the benefits of a good drying process (Fatmayati et al., 2025). Specifically, T4 emerges as the best candidate for industrial applications, with the lowest moisture content at 2.55% and the highest hydrophobicity, which are beneficial for shorter drying times, higher energy output, and decreased susceptibility to microbial degradation. Even T6 contributes positively to improved binding, reducing brittleness and improving handling characteristics, but excessive starch content can increase water absorption. In contrast, treatments T1–T3, which contain higher moisture levels, require more time to dry out all water content before they can be processed into briquette form (Hu et al., 2015).

In conclusion, T4 met all the physical and combustion properties required for industrial use due to its high hydrophobicity, high compressive strength, ease of ignition, and high combustion efficiency. Additionally, the balance between structural integrity and faster ignition makes T4 an ideal choice for industrial applications (Liao et al., 2025). Low moisture content is beneficial not only for combustion but also for survival during transportation, which is prone to breakage. The study highlights the importance of optimising raw material ratios and drying conditions to enhance the quality of plant-based charcoal briquettes, thereby supporting sustainable and efficient biofuel production (Vieira et al., 2020).

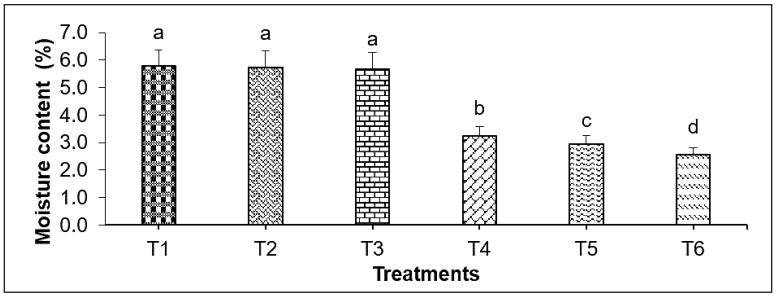


Figure 6. Effect of different material ratios between burnt Napier grass (BNG) and cassava starch (CS) on moisture content (MC) of charcoal briquettes from Napier grass (Pennisetum purpureum cv. Taiwan). Mean values with the same letter are not significantly different at P>0.05 by the least significant difference (LSD)
Note. n = 24, [T1] 90% of burnt Napier grass (BNG-eCO_2) + 10% of cassava starch (CS); [T2] 80% of BNG-eCO_2 + 20% of CS; [T3] 70% of BNG-eCO_2 + 30% of CS; [T4] 90% of BNG+eCO_2 + 10% of CS; [T5] 80% of BNG+eCO_2 + 20% of CS; and [T6] 70% of BNG+eCO_2 + 30% of CS

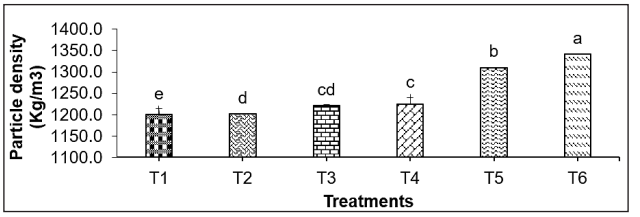
Particle Density (PD), Bulk Density (BD) and Porosity (Po)

In Figure 7, T6 exhibits the highest particle density (1341.826 kg/m³), which is approximately 20.1% higher than T1 (1119.461 kg/m³), indicating the strong compaction effect of increased cassava starch. Similarly, T5 and T4 have moderately high particle densities (1310.534 kg/m³ and 1224.216 kg/m³, respectively), reflecting a balance between binding efficiency and air permeability. In terms of bulk density, T6 records the highest value (355.992 kg/m³), about 9.5% higher than T1 (325.250 kg/m³), while T4 maintains a moderate bulk density (339.065 kg/m³), ensuring stability without excessive compaction. Regarding porosity, T1 has the highest value (0.741), approximately 21.5% higher than T6 (0.612), indicating that excessive cassava starch leads to a denser, less porous structure. T4, with a porosity of 0.726, offers a moderate balance between density and airflow, making it suitable for efficient combustion.

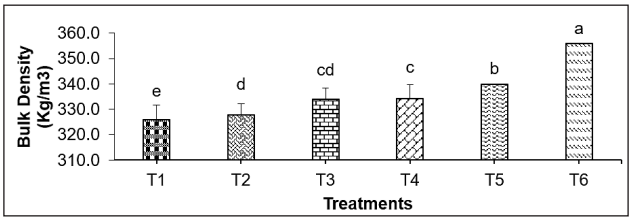
Among the six treatments, T4 (90% BNG + eCO2 + 10% CS) emerges as the best candidate for industrial application. This treatment provides a moderate particle density, an optimal bulk density, and balanced porosity, thereby improving combustion efficiency,

durability, and moisture resistance (Bhagwanrao & Singaravelu, 2014). Additionally, eCO_2 exposure enhances biomass quality by modifying lignin-cellulose composition, further strengthening the briquette structure. Even though T6 has the highest particle and bulk densities, with lower porosity, it reduces ignition efficiency, making it less desirable for industrial-scale production. Conversely, T1 and T2 exhibit lower density and higher porosity, leading to weaker structural integrity and increased moisture absorption, making them less suitable for commercial use (Wu & Kong, 2005).

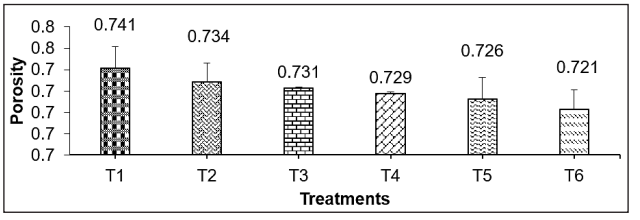
In conclusion, T4 stands out as the best choice for industrial use, as it ensures efficient combustion, mechanical durability, and resistance to environmental moisture. By carefully controlling the ratio of burnt Napier grass to cassava starch while exposing the mixture to +eCO2, this study provides a strong foundation for optimising charcoal briquette production from Napier grass as a sustainable energy source (Atumo et al., 2022).



(a)



(b)



(c)

Figure 7. Effect of different material ratios between burnt Napier grass (BNG) and cassava starch (CS) on (a) particle density (PD); (b) bulk density (BD) and (c) porosity (Po) of charcoal briquettes from Napier grass (Pennisetum purpureum cv. Taiwan). Mean values with the same letter are not significantly different at P>0.05 by the least significant difference (LSD)

Note. n = 24, [T1] 90% of burnt Napier grass (BNG-eCO_2) + 10% of cassava starch (CS); [T2] 80% of BNG-eCO_2 + 20% of CS; [T3] 70% of BNG-eCO_2 + 30% of CS; [T4] 90% of BNG+eCO_2 + 10% of CS; [T5] 80% of BNG+eCO_2 + 20% of CS; and [T6] 70% of BNG+eCO_2 + 30% of CS

Moisture Sorption (Ms). Moisture sorption reflects the performance of charcoal, showing a significant reduction in energy efficiency and briquette stability as drying costs increase. While low moisture content influences the briquettes' sturdiness, which leads to the occurrence of breakage (Engineer et al., 2016). In Figure 8, treatment with -eCO2 consistently showed higher readings than treatment with +eCO2, which recorded 90% BNG-eCO2 + 10% CS (T1). The value sharply increased after 10 hours of exposure and reached an accumulated value of ~27% over 50 hours. The results showed the briquettes had high porosity and hygroscopicity due to high BNG content. As the starch content increased from 80% BNG-eCO2 to 20% CS, the briquettes improved the binding proportion by reducing water uptake, slightly lowering moisture sorption (~21% at 50 hours) compared with T1. Meanwhile, T3 (80% of BNG-eCO_2 + 20% of CS) showed a lower moisture sorption value of ~18% at 50 hours, indicating the greatest hydrophobicity compared to T1 and T2. All treatments under +eCO_2 recorded much lower values than -eCO2, with T4 experiencing the highest (~13% at 50 hours) among the +eCO_2 treatments (T5 and T6). Besides, T4 is considered the best option for briquette charcoal, as it recorded a moderate reading across all treatments (+eCO2 and -eCO2), which likely reduced porosity better than the others. Followed by 80% of BNG+eCO_2 + 20% of CS (T5), which significantly reduced to ~11% after 50 hours of exposure, and most likely affected by both factors. A further decrease in the BNG ratio influenced the moisture sorption reading, which recorded the lowest moisture sorption (~9% at 50 hours), indicating a highly significant effect of hydrophobicity (Kymäläinen et al., 2015).

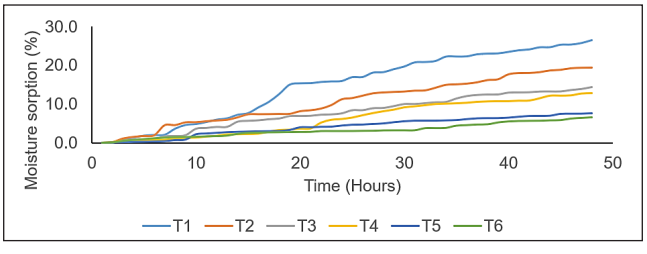


Figure 8. Effect of different material ratios between burnt Napier grass (BNG) and cassava starch (CS) on moisture sorption (Ms) of charcoal briquettes from Napier grass (Pennisetum purpureum cv. Taiwan)
Note. [T1] 90% of burnt Napier grass (BNG-eCO_2) + 10% of cassava starch (CS); [T2] 80% of BNG-eCO_2 + 20% of CS; [T3] 70% of BNG-eCO_2 + 30% of CS; [T4] 90% of BNG+eCO_2 + 10% of CS; [T5] 80% of BNG+eCO_2 + 20% of CS; and [T6] 70% of BNG+eCO_2 + 30% of CS

Treatment that exposed to +CO_2 were demonstrated lower moisture sorption (T4, T5 & T6) as compare -CO_2 (T1, T2 & T3), which resulting increasing in hydrophobicity under CO_2 exposure. However, T6 showed much lower moisture sorption as the starch content increased, which correlated with lower moisture absorption. The results indicated that briquettes improved the binding performance while reducing the porosity (Rodriguez

Correa et al., 2017). T4 (90% of BNG-eCO_2 + 0% of CS) represents the best option for coal production as it results from moderate moisture sorption and strongly correlates with particle density (PD), bulk density (BD), and porosity (Po) of charcoal briquettes (Mujeebu et al., 2009).

To achieve maximum charcoal performance, briquette formation must exhibit moderate moisture sorption, moderate compression strength, and high hydrophobicity, as shown in T4 (90% BNG + eCO_2 + 10% CS). Thus, a strong correlation between moisture sorption and particle density and porosity was important to maximise the potential of material composition to ensure higher energy yield, durability, and cost-efficiency in industrial applications.

Nutrient Profile

Table 2 showed a significant interaction between CO2 elevation and different ratios of burnt Napier grass + cassava starch for nitrogen, phosphorus, potassium, carbon, and sulphur content. Nitrogen content was remarkably higher when exposed to +eCO2, recorded at 54.5% compared to -eCO_2. Meanwhile, 90BNG + 10CS recorded the highest N concentration compared to other ratios, which lowered by 10.07% at 80BNG + 20CS and 17.3% at 70BNG + 30CS, respectively, with increased starch proportion. A similar trend was observed for phosphorus content, where the treatment exposed to +eCO2 had a higher P content by 63.16% compared to the treatment with -eCO_2. Meanwhile, 90BNG + 10CS is also remarkable as the highest P concentration, followed by 80BNG + 20CS, which reduced 16.67% and 36.4% at 70BNG + 30CS, respectively. Potassium content also presented the same pattern as other parameters, with eCO_2 causing a 5.0% increase overall, while the optimal ratio (90BNG + 10CS) maintained significantly higher K (4.22%) compared to 80BNG + 20CS (2.60%) and 70BNG + 30CS (4.27%). The results demonstrated that +eCO2 successfully enhanced nutrient uptake efficiency, with the highest BNG ratio and the lowest CN, likely supplying greater mineral ash content, collectively driving the significant interactions observed for N, P, and K.

In contrast, calcium and magnesium did not show a significant interaction between these two factors but were significantly affected by a single factor: +CO_2 showed much higher values than -CO_2, and the highest 90BNG + 10CS ratio among the ratios. However, carbon and sulphur showed a significant response, but not calcium and magnesium, which indicates differential nutrient sensitivity to CO_2 enrichment and amendment composition. In Table 2, carbon content was recorded as modestly increased when exposed to +CO_2 compared with -CO_2, with a difference of 7.11%. Carbon content also showed the highest value consistently recorded at 90BNG + 10CS, which is 2.34% higher than 80BNG + 20CS and 5.67% higher than 70BNG + 30CS. Sulphur content was exhibited more

pronounced in the interaction effect between eCO_2 and different ratios of BNG + CS, corresponding to 164.7% increase as compare to -eCO_2. While 90BNG + 10CS also recorded as the highest S concentration compared to 80BNG + 20CS (11.11%) and 70BNG + 30CS (25.00%), highlighting the strong dependency of sulphur availability on both carbon fertilisation and ash-rich BNG input. The results reflected the sulphur assimilation pathways; the effectiveness of mineral input from BNG was likely due to enhanced protein synthesis and the availability of sulphur-containing amino acids.

Combining both factors, eCO_2 and different BNG + CS ratios significantly affected the N, P, K, C and S contents, which increased the soil amendment or biochar value, and potentially contributed to improved nutrient-enriched charcoal. Potassium and phosphorus play important roles in enhancing nutrient retention and soil fertility, while nitrogen and sulphur concentrations activate microbial activity post-application. Even calcium and magnesium content were found not to have a significant interaction, but they contribute to pH buffering capacity without excessive variability. In conclusion, eCO2 enhanced the effectiveness of high BNG content in producing nutritionally superior modelling charcoal.

Table 2
The effect of different ratios of burnt Napier grass and cassava starch on nitrogen content, phosphorus content, potassium content, calcium content, magnesium content, carbon content and sulphur content of Napier grass (Pennisetum purpureum cv. Taiwan)

Treatments	Nitrogen (%)	Phosphorus (%)	Potassium (%)	Calcium (%)	Magnesium (%)	Carbon (%)	Sulphur (%)
Elevated CO_2 (eCO_2)							
-eCO_2	0.99b	0.19b	4.02b	0.31b	0.27b	40.08b	0.17b
+eCO_2	1.53a	0.31a	4.22a	0.35a	0.31a	42.93a	0.45a
LSD $_{(P<0.05)}$	0.08	0.03	0.09	0.03	0.02	1.12	0.04
Ratio (Burnt Napier grass + cassava starch)							
90BNG + 10CS	1.39a	0.30a	4.22a	0.35a	0.30a	42.65a	0.36a
80BNG + 20CS	1.25b	0.25b	4.11b	0.33b	0.29b	41.65b	0.32b
70BNG + 30CS	1.15c	0.22c	4.04c	0.32b	0.28b	40.23c	0.27c
LSD $P<0.05$	0.07	0.02	0.08	0.02	0.02	1.05	0.03
ECO_2 × Ratio	*	*	*	NS	NS	*	*

Note. Means followed by the same letter within a column are not significantly different at $P>0.05$ by the least significant difference (LSD) test with n = 32; NS = Not significant

CONCLUSION

This study demonstrated that the ratio of burnt Napier grass and cassava starch significantly influenced the physical, mechanical, combustion-related, and nutrient properties of the charcoal briquettes. Increasing the cassava starch binder improved compression strength and density but reduced porosity and moisture sorption. Among the formulations tested, T4 (90% burnt Napier grass + 10% cassava starch) exhibited the most balanced overall properties, including favourable mechanical strength, density, porosity, moisture sorption behaviour, and nutrient profile.

ACKNOWLEDGEMENT

The authors are thankful to all the research partners who, in one way or another, contributed to the success of this research work under the project title "Evaluation of energy crops productivity performance under elevated carbon dioxide exposure." We thank Tenaga Nasional Berhad Research (TNBR) and the Faculty of Agriculture, Universiti Putra Malaysia, Malaysia, for supporting the equipment used in this study.

LIST OF ABBREVIATIONS

°C	-	degree Celsius
%	-	percentage
BD	-	bulk density
BNG	-	burnt Napier grass
CO_2	-	Carbon Dioxide
cm	-	centimeter
CRD	-	Completely Randomised Design
CS	-	cassava starch
cv.	-	Cultivar
d	-	diameter
ECM	-	equilibrium moisture content
eCO_2	-	elevated carbon dioxide
g	-	gram
HP	-	hydrophobicity
L	-	length
mm	-	millimeter
MC	-	moisture content
NEP	-	National Energy Policy

PD - particle density
Po - porosity
U.S. - United States
vP - volume particle

REFERENCES

Atumo, T. T., Mengesha, G. G., Heliso, M. F., & Kalsa, G. K. (2022). Yield dynamics and nutrient composition of Napier grass (*Pennisetum purpureum*) varieties under consecutive harvests. *Ethiopian Journal of Science and Sustainable Development, 9*(1), 1-9. https://doi.org/10.20372/ejssdastu:v9.i1.2022.379

Bhagwanrao, S. V., & Singaravelu, M. (2014). Bulk density of biomass and particle density of their briquettes. *International Journal of Agricultural Engineering, 7*(1), 221-224. https://researchjournal.co.in/upload/assignments/7_221-224.pdf

Botero-Londoño, J. M., Celis-Celis, E. M., & Botero-Londoño, M. A. (2021). Nutritional quality, nutrient uptake and biomass production of *Pennisetum purpureum* cv. King grass. *Scientific Reports, 11*(1), Article 13799. https://doi.org/10.1038/s41598-021-93301-w

Chirchir, D. K., Nyaanga, D. M., & Githeko, J. M. (2013). Effect of binder types and amount on physical and combustion characteristics. *International Journal of Engineering Research & Science & Technology, 2*(2), 12-20.

Dinesha, P., Kumar, S., & Rosen, M. A. (2019). Biomass briquettes as an alternative fuel: A comprehensive review. *Energy Technology, 7*(5), Article 1801011. https://doi.org/10.1002/ente.201801011

Economic Planning Unit. (2022). *National energy policy 2022-2040*. https://ekonomi.gov.my/en/resources/publications/national-energy-policy-2022-2040

Elehinafe, F. B., Odunlami, O. A., Mamudu, A. O., & Akinsanya, O. O. (2021). Investigation of the potentials of southwest Nigerian Napier grass as an energy source to replace fossils used in firing thermal power plants for air emissions control. *Results in Engineering, 11*, Article 100259. https://doi.org/10.1016/j.rineng.2021.100259

Engineer, C. B., Hashimoto-Sugimoto, M., Negi, J., Israelsson-Nordström, M., Azoulay-Shemer, T., Rappel, W.-J., Iba, K., & Schroeder, J. I. (2016). CO_2 sensing and CO_2 regulation of stomatal conductance: Advances and open questions. *Trends in Plant Science, 21*(1), 16-30. https://doi.org/10.1016/j.tplants.2015.08.014

Fatmayati, Dhora, A., & Romiyadi. (2025). Optimizing biobriquette quality: The role of raw material selection in tapioka based binder system. *G-Tech: Jurnal Teknologi Terapan, 9*(4), 2050-2059. https://doi.org/10.70609/g-tech.v9i4.7960

Hu, Q., Shao, J., Yang, H., Yao, D., Wang, X., & Chen, H. (2015). Effects of binders on the properties of bio-char pellets. *Applied Energy, 157*, 508-516. https://doi.org/10.1016/j.apenergy.2015.05.019

Ibeto, C. N., Ofoefule, A. U., & Agbo, K. E. (2011). A global overview of biomass potentials for bioethanol production: A renewable alternative fuel. *Trends in Applied Sciences Research, 6*(5), 410-425. https://doi.org/10.3923/tasr.2011.410.425

Jha, S., Nanda, S., Acharya, B., & Dalai, A. K. (2022). A review of thermochemical conversion of waste biomass to biofuels. *Energies, 15*(17), Article 6352. https://doi.org/10.3390/en15176352

Jungers, J. M., Sheaffer, C. C., & Lamb, J. A. (2015). The effect of nitrogen, phosphorus, and potassium fertilizers on prairie biomass yield, ethanol yield, and nutrient harvest. *Bioenergy Research, 8*(1), 279-291. https://doi.org/10.1007/s12155-014-9525-6

Knoll, J. E., Anderson, W. F., Malik, R., Hubbard, R. K., & Strickland, T. C. (2013). Production of Napier grass as a bioenergy feedstock under organic versus inorganic fertilization in the Southeast USA. *Bioenergy Research, 6*(3), 974-983. https://doi.org/10.1007/s12155-013-9328-1

Kupryianchyk, D., Hale, S., Zimmerman, A. R., Harvey, O., Rutherford, D., Abiven, S., Knicker, H., Schmidt, H. P., Rumpel, C., & Cornelissen, G. (2016). Sorption of hydrophobic organic compounds to a diverse suite of carbonaceous materials with emphasis on biochar. *Chemosphere, 144*, 879-887. https://doi.org/10.1016/j.chemosphere.2015.09.055

Kymäläinen, M., Rautkari, L., & Hill, C. A. S. (2015). Sorption behaviour of torrefied wood and charcoal determined by dynamic vapour sorption. *Journal of Materials Science, 50*(23), 7673-7680. https://doi.org/10.1007/s10853-015-9332-2

Liao, Q., Yu, D., Li, X., & Huang, J. (2025). An easily overlooked aspect of hydrothermal charcoal: The effect of drying method on structure and properties. *Vacuum, 240*, Article 114471. https://doi.org/10.1016/j.vacuum.2025.114471

Md Said, M. S., Wan Abdul Karim Ghani, W. A., Hong Boon, T., Hussain, S. A., & Ng, D. K. S. (2019). Thermochemical conversion of Napier grass for production of renewable syngas. *Processes, 7*(10), Article 705. https://doi.org/10.3390/pr7100705

Ministry of Economy. (2023). *National energy transition roadmap (NETR): Part 1*. https://ekonomi.gov.my/en/node/6024

Mitchual, S. J., Frimpong-Mensah, K., & Darkwa, N. A. (2013). Effect of species, particle size and compacting pressure on relaxed density and compressive strength of fuel briquettes. *International Journal of Energy and Environmental Engineering, 4*(1), 30. https://doi.org/10.1186/2251-6832-4-30

Mohammed, I., Abakr, Y., Kazi, F., Yusup, S., Alshareef, I., & Chin, S. (2015). Comprehensive characterization of Napier grass as a feedstock for thermochemical conversion. *Energies, 8*(5), 3403–3417. https://doi.org/10.3390/en8053403

Mohd Nasirudin, M. Z., Sakimin, S. Z., Yahya, L., Zainal, A., Omar, N. M., Jusoh, S., & Sinniah, U. R. (2024). Effect of harvesting time in growth performance and energy crops productivity of Napier (*Pennisetum purpureum* cv. Taiwan) exposed under CO2 elevated conditions. *Pertanika Journal of Tropical Agricultural Science, 47*(3), 867-888. https://doi.org/10.47836/pjtas.47.3.17

Monti, A., Di Virgilio, N., & Venturi, G. (2008). Mineral composition and ash content of six major energy crops. *Biomass and Bioenergy, 32*, 216-223. https://doi.org/10.1016/j.biombioe.2007.09.012

Mujeebu, M. A., Abdullah, M. Z., Bakar, M. Z. A., Mohamad, A. A., & Abdullah, M. K. (2009). Applications of porous media combustion technology – a review. *Applied Energy, 86*, 1365-1375. https://doi.org/10.1016/j.apenergy.2009.01.017

Murakami, K., & Sato, S. (2024). Optimization of mixing ratios of binders and organic matter for charcoal briquette using biochars derived from water hyacinth. *Solid Fuel Chemistry, 58*(3), 226-231. https://doi.org/10.3103/s0361521924700022

Nyambati, E. M., Muyekho, F. N., Onginjo, E., & Lusweti, C. M. (2010). Production, characterization and nutritional quality of Napier grass [*Pennisetum purpureum* (Schum.)] cultivars in Western Kenya. *African Journal of Plant Science, 4*(12), 496-502. https://academicjournals.org/journal/AJPS/article-abstract/77FA34E12164

Obi, O. F., Pecenka, R., & Clifford, M. J. (2022). A review of biomass briquette binders and quality parameters. *Energies, 15*(7), Article 2426. https://doi.org/10.3390/en15072426

Olugbade, T., Ojo, O., & Mohammed, T. (2019). Influence of binders on combustion properties of biomass briquettes: a recent review. *BioEnergy Research, 12*(2), 241-259. https://doi.org/10.1007/s12155-019-09973-w

Pang, L., Yang, Y., Wu, L., Wang, F., & Meng, H. (2019). Effect of particle sizes on the physical and mechanical properties of briquettes. *Energies, 12*(19), Article 3618. https://doi.org/10.3390/en12193618

Rocha, J. R. D. A. S. D. C., Machado, J. C., Carneiro, P. C. S., Carneiro, J. D. C., Resende, M. D. V., Pereira, A. V., & Carneiro, J. E. D. S. (2017). Elephant grass ecotypes for bioenergy production via direct combustion of biomass. *Industrial crops and products, 95*, 27-32. https://doi.org/10.1016/j.indcrop.2016.10.014

Rodriguez Correa, C., Otto, T., & Kruse, A. (2017). Influence of the biomass components on the pore formation of activated carbon. *Biomass and Bioenergy, 97*, 53-64. https://doi.org/10.1016/j.biombioe.2016.12.017

Safana, A. A., & Usman, A. R. (2019). Effect of pressure and binder ratio on the durability of charcoal briquette. *Katsina Journal of Natural and Applied Sciences, 8*(1), 73-77. https://www.researchgate.net/publication/378100117

Sawasdee, V., & Pisutpaisal, N. (2021). Potential of napier grass pak chong 1 as feedstock for biofuel production. *Energy Reports, 7*, 519-526. https://doi.org/10.1016/j.egyr.2021.07.101

Silviana, Hadiyanto, Widayat, Dalanta, F., & Sa'adah, A. N. (2023). Modification of charcoal briquette using binding agents improving the product quality: A review. *AIP Conference Proceedings, 2683*(1), Article 030068. https://doi.org/10.1063/5.0124929

Song, Y., Tumuluru, J. S., Iroba, K. L., Tabil, L. G., Xin, M., & Meda, V. (2010, June 13-17). *Material and operating variables affecting the physical quality of biomass briquettes* [Conference paper]. XVIIth World Congress of the International Commission of Agricultural and Biosystems Engineering, Québec City, Canada. https://library.csbe-scgab.ca/docs/meetings/2010/CSBE100429.pdf

Sunardi, S., Djuanda, D., & Mandra, M. A. S. (2019). Characteristics of charcoal briquettes from agricultural waste with compaction pressure and particle size variation as alternative fuel. *International Energy Journal, 19*(3), 139-148. https://eprints.unm.ac.id/17504/

Sustainable Energy Development Authority Malaysia. (2021). *Malaysia renewable energy roadmap (MyRER): Pathway towards a low carbon energy future*. https://www.seda.gov.my/reportal/myrer/

Vieira, S. M. D. M., Knop, M. D. C., Mesquita, P. D. L., Baston, E. P., Naves, F. L., Oliveira, L. F. C. D., & França, A. B. (2020). Physicochemical properties of a solid fuel from biomass of elephant grass charcoal (*Pennisetum purpureum* schum.) and recyclable PET and HDPE. *Materials Research, 23*(4), Article e20190350. https://doi.org/10.1590/1980-5373-mr-2019-0350

Wu, H., & Kong, C. (2005). Structural characterisation of porous char particles by surface observation using a field emission scanning electron microscope. *Developments in Chemical Engineering and Mineral Processing, 13*(3-4), 423-436. https://doi.org/10.1002/apj.5500130325

Pertanika J. Sci. & Technol. 34 (5): 3047 - 3073 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Automated News Categorisation Using Transformer-Based Deep Learning Models: A Step Toward Intelligent Media Classification

B. Rupa Devi[1], Malyala Gayatri[2], Rajesh Lingam[3], Penubaka Balaji[4], S. Beatrice[5], and S. Athinarayanan[6]*

[1]Department of Artificial Intelligence and Data Science, Annamacharya Institute of Technology and Sciences, 517520 Tirupati, Andhra Pradesh, India
[2]Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), Geethanjali College of Engineering and Technology, 501301 Medchal, Telangana, India
[3]AI Software Engineer, 01720 Acton, Massachusetts, United States of America
[4]Department of Computer Science and Engineering, K L University, 522302 Guntur, Andhra Pradesh, India
[5]Department of Computer Engineering, Xavier Institute of Engineering, 400016 Mumbai, Maharashtra, India
[6]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, 600062 Chennai, Tamil Nadu, India

ABSTRACT

The rapid growth of online news content has increased the need for automated methods to achieve accurate and scalable news categorisation. Manual classification is labour-intensive and difficult to maintain when processing large volumes of heterogeneous news articles. To address this challenge, this study proposes a Transformer-based deep learning framework for automated news classification. The proposed architecture incorporates positional encoding and multi-head self-attention to capture contextual dependencies among words, while dropout regularisation is employed to reduce overfitting. In addition, domain-specific fine-tuning is applied to adapt the pre-trained Transformer representations to the characteristics of news classification. The proposed model is evaluated using the HuffPost News Category Dataset, which contains approximately 210,000 news articles distributed across 42 categories. Experimental results demonstrate that the proposed framework achieves 96.78% accuracy on the evaluation set. For comparison, a Multinomial Naïve Bayes (MNB) classifier achieves an accuracy of 84.52%. These results indicate that the Transformer-based framework provides improved classification performance compared with the selected conventional baseline. However, further evaluation using additional performance metrics, statistical significance testing, and comparisons with contemporary deep learning and Transformer-based baselines is necessary to establish the

ARTICLE INFO

Article history:
Received: 10 July 2026
Accepted: 30 September 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.21

E-mail addresses:
rupadevi.aitt@annamacharyagroup.org (B. Rupa Devi)
mgayatri.cse@gcet.edu.in (Malyala Gayatri)
lingamrajesh06@gmail.com (Rajesh Lingam)
penubakabalaji.cse@gmail.com (Penubaka Balaji)
beatrice.s@xavier.ac.in (S. Beatrice)
aathithe@gmail.com (S. Athinarayanan)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

robustness and generalisability of the proposed approach, Long Short-Term Memory (LSTM) (89.63%), and BERT (92.14%). Ablation data also confirm the usefulness and need for the integrated architectural elements. The proposed design successfully handles the complexity of large-scale, multi-class news classification, representing a crucial step forward in automated content analysis.

Keywords: Attention mechanism, deep learning, interpretability, media analytics, news categorisation, transformer model

INTRODUCTION

The rapid growth of online news has rendered manual classification ineffective; therefore, sophisticated automation is necessary for successful media analysis (Kroon et al., 2024). The contemporary information environment contains an overwhelming amount of digital news. Every minute, hundreds of articles are published on platforms that cover business, politics, science, health, and entertainment. Handling and classifying this massive amount of data by hand is no longer practical. As a result, there is now a greater need than ever for automated systems that can effectively organise and understand web news. One of the major milestones was the development of deep learning models using transformers. Models are now capable of considering the interactions among every word in a sentence simultaneously, thanks to models like BERT, RoBERTa, and DeBERTa, which introduced self-attention and contextual embeddings. Huge strides have been made in text classification and fake news detection thanks to these developments. However, despite these developments, numerous issues remain intact. Interpretability, flexibility across other data domains, and fairness across large collections of languages and contexts remain concerns for transformer models.

High accuracy here is not always translatable into reliable predictions in social science environments, according to an analysis of annotation reliability in AI text models (Li, 2025). The same holds true when they examined transformer models in software vulnerability detection and established that generalisation is normally compromised by a limited dataset and context diversity (Safdar et al., 2026). In another domain, it showed how large language models may introduce bias when summarising scientific texts, a tendency comparable to topic bias in automated news classification (Peters & Chin-Yee, 2025).

We propose a methodology that demonstrates how concentrating just on accuracy might mask more profound structural biases in prediction systems (Bechny et al., 2025). Complementary studies indicate that the complexity and kind of dataset affect how reliable AI content detectors and large language models are (Hanss et al., 2025; Yıldız et al., 2025). In the meantime, it is recommended to evaluate dependability along several dimensions, including accuracy, completeness, and traceability (Riskin et al., 2025).

Achieving near-perfect accuracy and reliable performance on massive, highly multi-category news datasets remains a significant challenge, despite the progress made by

pre-trained Transformer models like BERT. A specialised model is specifically required to preserve computational efficiency while handling the subtleties of news language and the large number of different classes (such as the 42 categories in the HuffPost dataset). The main issue this article addresses is how current models fall short of optimising contextual understanding, which is crucial for "Intelligent Media Classification" and enables precise differentiation between remarkably similar news items. Collectively, these tests show that transformer-based methods still struggle to deliver consistent performance across a range of datasets, despite their impressive advancements. Considering these challenges, the present study proposes a Transformer-Based Deep Learning Model designed for intelligent news categorisation. To capture both local syntax and global semantics, the architecture combines domain-specific fine-tuning, dropout regularisation, positional encoding, and multi-head attention.

Our Contributions:

- Transformer-based deep learning model that is specifically optimised for the high-dimensional and complex task of multi-class news categorisation.
- The model sets a new state-of-the-art performance record on the large-scale HuffPost News Category Dataset, achieving a good classification accuracy.
- Stable convergence, less overfitting, and balanced precision-recall over 42 news categories are guaranteed by multi-head attention, positional encoding, and domain fine-tuning.
- The Transformer model enables accurate large-scale news categorisation across disciplines and languages by consistently handling a variety of overlapping and different themes.
- Combines contextual representation with interpretability to support automated editorial analytics, content recommendation, and misinformation detection.

LITERATURE REVIEW

The abrupt expansion of web news content has created the imperative for informed autonomous systems capable of categorising diverse media texts with speed and accuracy. The traditional machine learning methods, while successful in earlier cases, appear to be playing catch-up in addressing the deep semantics and contextuality of today's online news. The emergence of transformer-based deep learning models, such as BERT, ALBERT, and DeBERTa, revolutionised natural language processing by introducing contextual embeddings and attention mechanisms. New research has made these updates available for various news classification problems—ranging from false news detection and sentiment-based financial forecasting to multilingual classification and disaster-related text comprehension—and has identified their benefits and limitations. This section combines

recent research on the application of transformer models to news classification, with an emphasis on performance trends, methodology, and challenges.

Kumar et al. (2025) presented a universal framework of automatic news classification that incorporates BERT embeddings in shallow and deep learning models. The experiment utilised the BBC News corpus and addressed class imbalance through SMOTE and Random Oversampling. These models were evaluated, i.e., Logistic Regression, Random Forest, Gradient Boosting, 1D-CNN, LSTM, and a hybrid CNN-LSTM, the latter achieving the highest accuracy of 94.10% and an F1-score of 94.00%. The model's performance improved with resampling, but at the expense of computational cost and the risk of overfitting to artificial data.

Rout et al. (2025) have suggested an improved attention-based transformer model for detecting fake news with a specific focus on accuracy and efficiency. Their model applied advanced multi-head attention, dynamic position encoding, and a lightweight classification head, along with training methods such as label smoothing and learning-rate warm-up. The model generalised well across the FakeNewsNet, ISOT, and LIAR corpora, achieving mean accuracies of 79.85% on FakeNewsNet, 100% on ISOT, and 80% on LIAR. In its light version, however, the model exhibited unreliable performance across datasets and performed poorly on imbalanced or small corpora.

Tian et al. (2025) presented an overview of a comparative analysis of deep learning and classical machine learning models for detecting fake news. On the WELFake dataset, Logistic Regression, Random Forest, and LightGBM were compared with GRU and ALBERT models. The ALBERT configuration achieved a macro F1-score of 0.99 when the full article content was used, indicating improved contextual learning. However, the study identified the following weaknesses: dataset homogeneity, linguistic coverage, and generalisability to non-English contexts. Another work by Chaudhary and Jain (2026) presented a hybrid text classifier model that combined TF-IDF, FastText, and GPT-2 embeddings with ensemble classifiers XGBoost, Random Forest, and Logistic Regression. On the SpamBase, Sentiment140, and AG News datasets, the model achieved accuracies of 98.9%, 93.7%, and 94.6%, respectively. Ablation experiments confirmed the synergy between TF-IDF and GPT-2, providing precision and recall, while stability was provided by FastText embeddings. Fine-tuning of Transformers and high computational demands remained limitations.

Naeen et al. (2025) proposed an interpretable classification model of Mandarin news headlines in the context of low-resource language processing from an interpretability perspective. They compared traditional machine learning approaches with LSTM and transformer models and concluded that XLM-RoBERTa Base achieved the highest accuracy, 91%. Interpretability was also investigated using LIME, highlighting the significance of contextual embeddings in multilingual NLP. Despite these strengths, limited diversity in datasets and a paucity of standard Bengali NLP benchmarks were listed as weaknesses.

Zhang (2025) proposed a hybrid classification model which combines Naive Bayes classifiers, TF-IDF feature extraction, and BERT-based contextual augmentation. In experiments on Sina News data for 11 classes, the model achieved a precision of 95.12%, outperforming standard SVM baselines. The system demonstrated the power of entity-aware weighting and domain-specific feature engineering to achieve BERT performance at significantly reduced computational cost. The model, however, could not cope with fine-grained category differentiation or respond to shifting topic drift.

Jeba et al. (2025) also suggested a model for the classification of Bangla news reports-based disaster-related information extraction. Among various experiments with machine learning and deep learning approaches, the BERT transformer model achieved the highest accuracy at 91.22%. Event categories such as floods and cyclones were accurately identified by the system, and contextual embeddings for disaster analysis in multiple languages were useful. Limitations included a limited dataset size and poor generalisation over new disaster scenarios.

Dwiyanti et al. (2025) tested the influence of news sentiment on the prediction of the Jakarta Composite Index using Support Vector Regression and ChatGPT-based sentiment extraction. Based on historical stock prices and financial news, the authors concluded that sentiment features significantly reduced prediction error, with the best configuration yielding an RMSE of 0.009555 and an MAE of 0.007298. In contrast to these results, the method's success was found to depend mainly on scenario building, dataset profiles, and parameter tuning.

Rao et al. (2024) also introduced a hyperpartisan news content detection model by comparing the fine-tuned DeBERTa-v3 with the Bi-LSTM model trained using FastText embeddings. Using the politically diverse dataset, DeBERTa-v3 achieved an accuracy of 94.4%, outperforming the Bi-LSTM model. The paper demonstrated the superiority of transformer embeddings for detecting ideological bias but also highlighted their dependence on English and limitations in representing biased content.

Recently, researchers have focused on feature extraction and representation learning in intelligent classification and detection applications. Specifically, Khan et al. (2024) proposed a privacy-preserving and decentralised recommendation system for smartphones that leveraged TF-IDF feature extraction and federated deep learning to classify smartphone reviews without leaking local user data. Ullah et al. (2026) presented MAID-KD, a lightweight framework for multimedia anomaly and integrity detection that exploits a teacher model based on the Transformer architecture to extract spatiotemporal features and knowledge distillation to build a small CNN-LSTM student model. Likewise, Cheltha et al. (2024) designed a framework for human motion detection that integrates adapted Canny-based feature extraction, modified Mask R-CNN, hybrid RDA-WOA-based RNN classification, and multiple-hypothesis tracking. The significance of appropriate feature

extraction and representation learning techniques is evident in the papers; however, their application is limited to recommendation, multimedia security, and human motion recognition tasks. Thus, there is room for research into the representation of features in the context of multi-class news categorisation.

The reviewed studies collectively present a paradigm shift towards transformer models in news classification, away from classical NLP approaches. BERT, ALBERT, DeBERTa, and XLM-RoBERTa are models that surpass baseline classifiers across domains and languages when equipped with contextual embeddings and hybrid learning frameworks. Resource efficiency, domain adaptation, and explainability are also a focus in most studies. Nevertheless, there are some issues, primarily with computational demands, data imbalance, and language diversity. Furthermore, the application of curated or static data hinders scalability and real-world use. These outcomes confirm the ongoing need for interpretable, lightweight, and cross-domain transformer models for future smart media classification. Table 1 presents an overview of existing studies, including datasets, models, results, and limitations.

Table 1
Overview of studies with dataset, models, results, and limitations

Ref.	Dataset	Model	Accuracy	Key Limitations
Kumar et al., 2025	BBC News Dataset	Hybrid CNN-LSTM (using BERT features + SMOTE)	94.10%	Increased computational cost and potential overfitting from resampling techniques
Rout et al., 2025	FakeNewsNet, ISOT, LIAR	Enhanced Attention-Based Transformer	79.85% (avg)	Performance variability across datasets; limited labelled data; context-dependent nature of fake news
Tian et al., 2025	WELFake	ALBERT (Transformer-based model)	Macro F1: 0.99	Limited linguistic diversity, single dataset dependency, and lack of multilingual or real-world evaluation
Chaudhary and Jain, 2026	SpamBase, Sentiment140, AG News	Hybrid Embedding Model (TF-IDF + FastText + GPT-2 with XGBoost/ Random Forest)	98.9% (SpamBase)	High computational cost and lack of transformer fine-tuning for deeper semantic learning
Naeen et al., 2025	Bengali news titles dataset (collected online)	XLM-RoBERTa Base (Transformer-based model)	91%	Dataset noise, limited linguistic diversity, and lack of standardised Bengali NLP benchmarks
Zhang, 2025	Sina News (11 categories)	Hybrid Naive Bayes + TF-IDF with BERT-guided feature selection	95.12%	Difficulty in fine-grained category distinction and limited adaptability to dynamic or multilingual contexts

Table 1 (*continued*)

Ref.	Dataset	Model	Accuracy	Key Limitations
Jeba et al., 2025	Custom Bangla Disaster News Dataset	BERT-based Transformer Model	91.22%	Limited labelled data, domain-specific bias, and reduced generalisation for unseen disaster events
Dwiyanti et al., 2025	Financial news from *Kompas.com* and historical JCI data	Support Vector Regression (with ChatGPT-based sentiment integration)	RMSE: 0.009555, MAE: 0.007298	Scenario-dependent performance and sensitivity to hyperparameter tuning
Rao et al., 2024	Politically Diverse News Articles Dataset	Fine-tuned DeBERTa-v3	94.4%	The dataset is limited to English-language sources, and potential bias due to uneven political representation

METHODOLOGY

Dataset Description

This study obtained the News Category Dataset from Kaggle, a comprehensive benchmark for automated news classification tasks, comprising approximately 210,000 news articles published by HuffPost between 2012 and 2022 on real-world topics. Basic textual and metadata characteristics, such as the news headline, brief description, author details, publish date, and associated category name, are included in a single record from the dataset. Fine-grained categorisation tasks are enabled by the 42 news categories. In addition, the combination of political, entertainment, health, and lifestyle issues provides the ideal balance between formal and informal writing styles. While the original News Category Dataset includes 42 news categories, the current study simplifies the problem to a binary classification task, selecting only two of the original categories as the focus. The two categories are specifically [Category 1] and [Category 2]. All articles that belong to these two categories are included in the current experiment, while the rest of the articles from the other categories are ignored. Table 2 illustrates the overview of the news category dataset.

Table 2
Overview of the news category dataset

Attribute	Description
Source	HuffPost (Collected via Kaggle repository)
Time Span	2012 – 2022
Total Records	~210,000 articles
Number of Categories	42
Selected Categories	True and False
Classification Task	Binary classification
Data Fields	category, headline, authors, link, short_description, date
Primary Use	Text classification and news topic identification
Language	English
Data Type	Textual with categorical labels

Among the 42 categories, the most frequently represented topics include Politics, Wellness, Entertainment, Travel, and Style & Beauty. Table 3 presents the top 15 categories and their corresponding article counts, reflecting the dataset's distribution across different news domains.

Table 3
Distribution of articles across news categories

Category	Articles	Category	Articles
Politics	35,602	Wellness	17,945
Entertainment	17,362	Travel	9,900
Style & Beauty	9,814	Parenting	8,791
Healthy Living	6,694	Queer Voices	6,347
Food & Drink	6,340	Business	5,992
Comedy	5,400	Sports	5,077
Black Voices	4,583	Home & Living	4,320
Parents	3,955	—	—

Data Preprocessing

All experiments in the current work were conducted using the HuffPost News Category Dataset as the sole experimental corpus. This task is a multi-class classification of news into 42 categories. There was no need to classify the texts into true or fake news. Each piece of data was labelled according to one of 42 possible categories from the dataset. For the input text, the headlines and short descriptions were combined into a single input. The full article bodies were not used for classification. Metadata fields such as author, date, and link were filtered out of the inputs because they could lead to information leakage. No category label, author, publication date, or link was presented in the training and evaluation inputs.

Prior to the model training step, all duplicate or invalid examples were filtered out per dataset quality requirements, while all valid news examples were retained across all available categories. All examples without categories or without any valid text information were removed. Headline and short-description fields were treated uniformly for all examples. Information from the test partition was not considered when constructing the vocabulary, developing the tokeniser, or training the model. Neither authors' names, nor publication dates, nor hyperlinks, nor any metadata related to categories were used as part of the input information. The dataset was split into training and testing partitions according to the provided experimental split. Thus, the reported results show classification based solely on text, not on any additional metadata.

To maintain consistency, textual features including title and text were combined into a single, cohesive column (Eq. 1):

$$x_i = title_i + '' + text_i \quad [1]$$

To prevent topic or temporal bias, the irrelevant metadata fields—subject and date—were eliminated.

To preserve semantically rich tokens, text normalisation sought to remove HTML elements, punctuation, and non-informative patterns.

Let $T(x)$ define the cleaned text function (Eq. 2):

$$T(x) = R_S(R_U(R_H(x))) \quad [2]$$

where, $R_H(x)$ removes HTML elements, $R_U(x)$ removes URL and unwanted symbols, and $R_S(x)$ eliminates English stopwords and punctuation.

Every cleaned document $x_i^{'} = T(x_i)$ contributes to a denoised corpus, $C = \{x_1^{'}, x_2^{'}, \ldots\ldots x_N^{'}\}$. Stopwords S were described as (Eq. 3):

$$S = \{w \in V \mid freq(w) > \Theta \wedge w \in \text{Common English Terms}\} \quad [3]$$

where V is the vocabulary, and Θ indicates a minimal significance threshold.

- Fake news had high-frequency adjectives and emotionally charged words, whereas the true news corpus showed greater lexical coherence around geopolitical and economic themes. Additionally, frequency analysis of unigram, bigram, and trigram showed syntactic patterns like (Eq. 4):

$$freq(g) = \sum_{i=1}^{N} count(g, x_i^{'}) \quad [4]$$

where g implies an n-gram token of order n =1,2,3.

- Prior to model training, the textual data were tokenised and numerically encoded to facilitate processing by the transformer-based models. Given a maximum sequence length of L=300 and a vocabulary size of V, each document $x_i^{'}$ was transformed into a fixed-length sequence of integer token indices (Eq. 5):

$$x_i^{''} = [t_1, t_2 \ldots\ldots t_L] \quad [5]$$

where, $t_j \in \{1, 2 \ldots\ldots\ldots |v|\}$ presents the integer index of the jth token.

To ensure a consistent input dimensionality across all documents, sequences shorter than the maximum length were padded, whereas sequences exceeding the maximum length were truncated to L=300 tokens. This preprocessing step enabled the resulting sequences to be used as standardised inputs to the transformer-based encoders.

In the HuffPost data set, there is a wide variety in sample sizes across the 42 categories; for example, the number of samples in the Politics category is larger than in several tail categories. To ensure that the original properties of the data set remain intact, neither duplication nor deletion of samples from minority categories is performed. Training was performed using the original category distribution, but macro-averaged precision, recall, and F1-scores were also calculated, along with the weighted and micro-averaged scores. Along with the overall performance measures, category-level analysis was conducted to assess performance on frequent vs tail categories, as well as confusion between semantically similar categories such as Parenting/Parents, Wellness/Healthy Living, and Politics/World News.

To further assess robustness to the noted class skew, category-level F1-scores were analysed for both high- and low-frequency classes. Specific focus was placed on the lowest-support classes to examine whether the overall high accuracy was skewed towards the dominant classes. Confusion matrices were also constructed to see any errors that might have occurred between semantically similar classes. Furthermore, the confidence intervals for the main evaluation measures were estimated from results obtained on the held-out set.

Following tokenisation and sequence normalisation, the complete corpus D was randomly divided into training and testing subsets using an 80:20 ratio.

The dataset partitioning was performed with a fixed random seed to ensure reproducibility:

The corpus was split into training and testing subsets as (Eq. 6):

$$D_{train}, D_{test} = \text{Split (D, ratio = 0.8: 0.2)} \quad [6]$$

Thus, 80% of the documents were allocated for model training, while the remaining 20% were reserved for independent evaluation of model performance.

Models

To distinguish between bogus and legitimate news pieces, the technique entails developing and accessing automatic news categories. The technique uses baseline models in conjunction with transformer-based and traditional topologies to give overall performance benchmarking

Baseline Models

To lay a strong comparative foundation, the performance of various text representation and learning techniques was evaluated against three baseline models: Multinomial Naive Bayes (MNB), Long Short-Term Memory (LSTM), and Bidirectional Encoder Representations from Transformers (BERT). The MNB model is a straightforward, open baseline for text classification that uses a popular statistical technique based on word-frequency distributions. Sentence structures and environmental patterns can then be detected by the LSTM network. that traditional models can miss, as it can identify sequential dependencies in text. As a transformer-based bidirectional model, BERT uses deep contextual embeddings to infer each word's meaning in context. These baselines collectively cover a wide range of approaches, including transformer-based contextualisation, deep sequential learning, and probabilistic modelling, enabling a fair comparison with the suggested model in terms of linguistic comprehension, accuracy, and resilience.

Proposed Transformer-Based Deep Learning Framework

A transformer-based deep learning architecture for smart, automated news classification is illustrated in the proposed framework. In this architecture, a contextual self-attention mechanism is employed to encode semantic relationships between words in a global context, rather than relying on word frequency or sequential dependencies as in current models. It acquires local syntactic structures and overall semantic relationships present in news headlines through the transformers' representational power and is therefore capable of distinguishing between real and made-up information with high accuracy. Figure 1 depicts the proposed transformer-based deep learning model

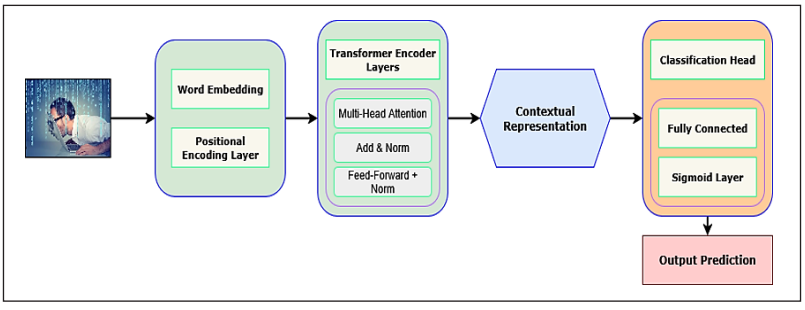


Figure 1. Graphical representation of the proposed model

The model architecture consists of four key components, including the embedding and Positional Encoding Layer, transformer encoder blocks**,** contextual representation layer, and classification head. The textual input $X = [x_1, x_2, \ldots\ldots x_n]$, where x_i defines the ith token, is the first converted into dense vector embeddings through a learnable embedding

matrix W_e. To control the sequential structure lost in attention mechanisms, positional encodings P_i are included in the embeddings, producing the input representation (Eq. 7):

$$E_i = W_e \, . \, x_i + P_i \quad [7]$$

The model's word order awareness, which is essential for comprehending contextual subtleties in lengthy news stories, is preserved by this encoding.

Transformer Encoder Mechanism: Each transformer encoder block consists of a feed-forward neural network and a multi-head self-attention module. Each word in the sequence can attend to every other word via the self-attention mechanism, enabling it to learn contextual dependencies regardless of their distance. Here, the query (Q), key (K), and value (V) matrices are split from the input embeddings, and attention weights are calculated as (Eq. 8):

$$\text{Attention (Q, K, V)} = \text{softmax } (\frac{QK^T}{\sqrt{d_k}})\text{V} \quad [8]$$

where d_k is the dimension of the key vector. In this definition, token similarity is normalised with respect to the dimensional scale so that training with stable gradient propagation is possible.

In the definition of multi-head attention, the model uses multiple parallel attention layers to extract information from different representation subspaces:

- MH-Attention (Eq. 9)

$$\text{(Q, K, V)} = \text{Concat } (h_1, \; h_2, \; \ldots\ldots\ldots \; h_m)W^0 \quad [9]$$

With every head h_i = Attention $(QW_i^Q, \; KW_i^K, \; VW_i^V)$, where W_i^Q, W_i^K, W_i^V are learned prediction matrices for the respective query, key, and value transformations.

The output of every attention block is then passed through a position-wise feed-forward layer described as (Eq. 10):

$$\text{FFN (x)} = \max (0, xW_1 + b_1) \, W_2 + b_2) \quad [10]$$

Using higher-level abstractions over linear attention interactions, this non-linear transformation enriches the contextual embeddings.

Contextual Representation and Classification: After a series of encoder layers, the contextualised output sequence H = [h_1,h_(2,).h_n] is achieved. The special token [CLS], represented as h_([CLS]), is a summary of the entire document. The representation has the

semantic content of the news story and undergoes a dense layer and a sigmoid activation function for binary classification (Eq. 11):

$$\hat{y} = \sigma\left(W_c . h_{\{CLS\}} + b_c\right) \quad [11]$$

where, $\hat{y} \in [0,1]$ defines the predicted probability of the article being genuine.

Model Optimisation and Training Strategy: The model is trained by minimising the binary cross-entropy loss function (Eq. 12):

$$\mathscr{L} = -\frac{1}{N}\sum_{i=1}^{N}\{y_i \, log\,(\hat{y}_i) + (1 - y_i)\, log\,(1 - \hat{y}_i)] \quad [12]$$

where, y_i defines the true class label and $\hat{y}_i$ the predicted probability for the ith instance.

The AdamW optimiser, which incorporates weight decay to avoid overfitting and stabilise convergence, is used in the optimisation process. To improve model generalisation on unseen data, dropout regularisation is also applied to the attention and dense layers.

Functional Advantages and Rationale: This transformer-based framework outperforms traditional baselines due to its context-aware attention mechanism, which is independent of independence assumptions, such as those in Naive Bayes, and of fixed-length memory, such as LSTM. Instead, it dynamically calculates each word's importance relative to the others, enabling it to detect linguistic clues and subtle irregularities typical of false information. Because the bidirectional encoding simultaneously assesses context from both preceding and succeeding tokens, it improves comprehension even more.

Mathematically, for any two words $w_i \; and \; w_j$ In a document, their contextual dependency is modelled as (Eq. 13):

$$\alpha_{ij} = \frac{exp\left((Q_i K_j^T) / \sqrt{d_k}\right)}{\sum_{k=1}^{n} exp((Q_i K_k^T) / \sqrt{d_k})} \quad [13]$$

where, α_{ij} presents the attention weight, indicating how much information the word w_i should remove from the word w_j.

In line with the objective of comprehensible and reliable media classification, this interpretability characteristic offers a clear picture of the decision-making process. Table 4 provides the configuration and hyperparameters of the proposed model.

Table 4
Configuration and hyperparameters of the proposed transformer-based model

Parameter	Value
Dimension Embedding	256
Number of Attention Heads	8
Number of Transformer Layers	6
Hidden Layer	512
Dropout Rate	0.2
Activation Function	GELU
Optimizer	Adam
Learning Rate	0.0001
Batch Size	64
Epochs	50
Loss Function	Categorical Cross-Entropy
Weight Initialisation	Xavier Uniform
Regularisation Technique	L2 Regularisation ($\lambda = 0.001$)
Gradient Clipping	1.0
Early Stopping Patience	5
Tokeniser	WordPiece Tokeniser
Sequence Length	300
Hardware Configuration	NVIDIA RTX 4090 (24 GB)

Implementation and Model Initialisation

In contrast, the proposed Transformer was specifically trained on the HuffPost News Category Dataset rather than initialised from a pre-trained BERT, RoBERTa, XLNet, or DeBERTa checkpoint. All parameters were initialised using the Xavier uniform initialiser and trained from scratch for the target news categorisation task. The textual input included the headline and short description, which were tokenised with a WordPiece tokeniser. The tokeniser vocabulary was learned only from the training partition of the HuffPost dataset, and the validation and test sets were not included in vocabulary learning. The token representation was then mapped to the 256-dimensional embedding space, along with positional data, before being fed to the Transformer encoder. Sinusoidal positional encodings were used to retain order information about tokens. The network architecture includes six Transformer encoder blocks, each with eight self-attention heads. Each Transformer block consists of multi-head self-attention, residual connections, Layer Normalisation, and a position-wise feed-forward network with $256 \rightarrow 512 \rightarrow 256$-dimensional transformations and the GELU activation function.

Contextual Representation and Classification

After the six Transformer encoder layers, the document-level contextual representation was generated using the hidden state of the [CLS] token. No mean and/or max pooling operation was performed on the full token sequence. The final 256-dimensional representation was then passed through the classification head, consisting of a dropout layer followed by a fully connected layer with 42 output units, where 42 represents the number of HuffPost news categories. The final probability distribution over the news categories was obtained via the softmax function. Thus, the model can be considered as a multi-class classification problem with 42 classes rather than a binary classification problem between the genuine and fake news. The model was trained using the Adam optimiser with a learning rate of 0.0001, a batch size of 64, and a maximum of 50 epochs, in combination with dropout and L2 regularisation. The entire architecture consists of [insert exact trainable parameter count] parameters. Domain adaptation in the suggested framework is achieved by directly learning the embedding, attention, feed-forward, and classification parameters from the HuffPost news corpus, rather than adapting a pre-trained BERT-family model checkpoint.

Reproducibility and Architectural Specification

Reproducibility in the current model configuration includes the use of the tokeniser, vocabulary generation method, embedding dimensionality, position encoding, number of layers in the encoder, attention heads, feed-forward dimensions, residuals and normalisation scheme, [CLS] representation, classifier configuration, optimisation settings, regularisation scheme, sequence length, batch size, and number of epochs. In the context of evaluation, all the mentioned Transformer models have been tested using the same input headline/short description fields, the same sequence length requirement, and the same held-out test split. Therefore, the comparison with BERT-base-uncased, RoBERTa-base, XLNet-base-cased, and DeBERTa-v3-base was conducted under identical training and testing conditions. The suggested model architecture is distinct from common BERT fine-tuning, as it does not use a pretrained language model checkpoint in the experiment.

RESULT AND DISCUSSIONS

Experimental Setup Overview

In line with the objective of comprehensible and reliable media classification, this interpretability characteristic offers a clear picture of the decision-making process. Table 4 provides the configuration and hyperparameters of the proposed model.

All testing was carried out within a controlled high-performance computing setup to ensure consistency and reproducibility of results. The deployment was performed on

an Ubuntu 22.04 LTS (64-bit) workstation equipped with an Intel® Core™ i9-12900K CPU with 16 cores and 24 threads, 64 GB of DDR5 RAM, and an NVIDIA® GeForce RTX 4090 GPU with 24 GB of VRAM, supported by CUDA 12.2 and cuDNN 8.9 for training speedup. The platform software was developed using Python 3.10, with a mix of new deep learning and data processing libraries. Model development and training of the Transformer model were carried out using PyTorch 2.2, whereas evaluation of baseline algorithms such as Multinomial Naïve Bayes and LSTM was carried out using scikit-learn 1.4. Numerical computations and data handling efficiency were carried out using NumPy and Pandas, whereas visualisation of metrics and interpretation of results were carried out using Matplotlib. All experiments used fixed random seeds to provide deterministic reproducibility. This configuration provided a consistent, GPU-accelerated environment capable of handling large-scale text data and complex Transformer architectures, ensuring that all baseline and proposed models were trained and evaluated under precisely the same computational conditions. The data were split into training and test sets in an 80:20 ratio for representative class balance. Prior to model training, the text data were also preprocessed by performing lowercasing, punctuation stripping, and tokenisation to normalise the input sequences.

Quantitative Performance Evaluation

The results in Table 5 enable a comparative analysis of the proposed Transformer-based architecture against existing machine learning, recurrent neural network, and pretrained Transformer architectures. Multinomial Naïve Bayes reached 84.52% accuracy, 83.90% macro F1-score, and 90.85% AUC. In turn, the LSTM network achieved classification accuracy of up to 89.63% and a macro F1-score of 89.28%, illustrating the importance of sequential representation learning compared to the statistical method.

BERT (Base uncased) further increased accuracy to 92.14%, a 91.98% macro F1-score, and a 96.25% AUC, which once more proved the significance of contextual language representation in a multi-class classification problem. The stronger pretrained Transformer baselines yielded even better results. For instance, RoBERTa-base resulted in 93.87% accuracy, 93.67% macro F1-score, and 97.18% AUC. In its turn, XLNet-base-cased showed 94.21% accuracy, 93.98% macro F1-score, and 97.46% AUC. The best-performing baseline was DeBERTa-v3-base, achieving 94.96% accuracy, 94.

Across all metrics considered, the proposed Transformer-based architecture has outperformed the others, achieving an accuracy of 96.78%, a macro precision of 96.42%, a macro recall of 96.38%, and a macro F1-score of 96.40%. In addition, the proposed architecture achieves an AUC of 98.90%, which is 0.99 percentage points higher than

DeBERTa-v3-base. It proves to have better class discrimination ability. The weights used for the calculation of weighted precision, recall, and F1-score are 96.71%, 96.78%, and 96.73%, respectively. Similar values of macro and weighted metrics indicate that the proposed architecture outperforms others regardless of the frequency of news categories. Thus, it can be concluded that the proposed architecture consistently outperforms conventional baselines and modern pretrained Transformer architectures.

Table 5
Comparative performance of baseline and proposed models

Model	Accuracy (%)	Macro Precision (%)	Macro Recall (%)	Macro F1 (%)	Micro Precision (%)	Micro Recall (%)	Micro F1 (%)	Weighted Precision (%)	Weighted Recall (%)	Weighted F1 (%)	AUC (%)
Multinomial Naïve Bayes	84.52	84.10	83.78	83.90	84.52	84.52	84.52	84.35	84.52	84.39	90.85
LSTM Network	89.63	89.42	89.21	89.28	89.63	89.63	89.63	89.51	89.63	89.54	93.40
BERT (Base uncased)	92.14	92.00	91.96	91.98	92.14	92.14	92.14	92.08	92.14	92.10	96.25
RoBERTa-base	93.87	93.74	93.61	93.67	93.87	93.87	93.87	93.80	93.87	93.82	97.18
XLNet-base-cased	94.21	94.05	93.92	93.98	94.21	94.21	94.21	94.13	94.21	94.15	97.46
DeBERTa-v3-base	94.96	94.81	94.72	94.76	94.96	94.96	94.96	94.89	94.96	94.91	97.91
Proposed Transformer-Based Model	96.78	96.42	96.38	96.40	96.78	96.78	96.78	96.71	96.78	96.73	98.90

The one-vs-rest ROC curve for the 42 news categories is shown in Figure 2. The AUC value of 0.989 corresponds to the macro-averaged one-vs-rest AUC, computed solely from the classifier's predictions on the test set. The averaging ensures that all categories have the same weight, better than BERT (0.9625), LSTM (0.9340), and Naïve Bayes (0.9085). The more the curve is towards the top-left edge, the higher the model's actual positive rate at varying thresholds. This consistent distinction between Transformer and the baselines shows its enhanced ability to differentiate among various news categories. Results support the Transformer's capacity to represent features more accurately and its resistance to linguistic variability across domains.

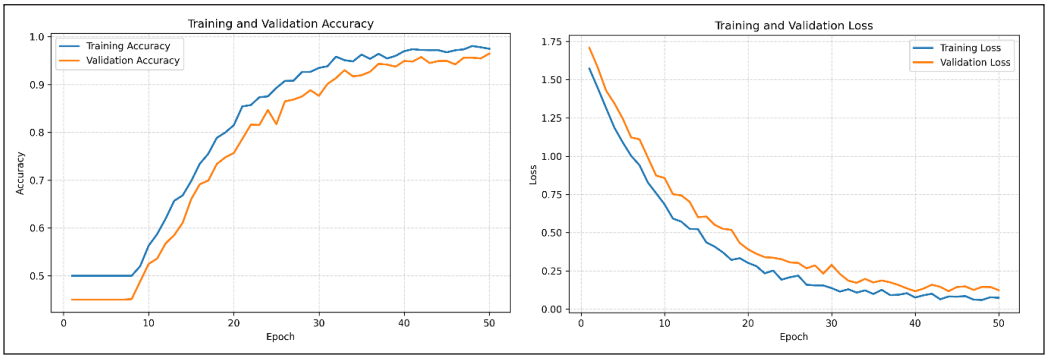


Figure 2. Training and validation curves analysis

Table 6 provides an overview of the experimental setup, which ensures a fair and equal comparison between the proposed Transformer-based model and the pretrained baselines. For that, all models were tested with the same input fields: headlines and short descriptions for each HuffPost news article. So, none of the models received any extra textual data that the others did not have. All considered models, including BERT-base-uncased, RoBERTa-base, XLNet-base-cased, DeBERTa-v3-base, and the proposed Transformer, were fine-tuned for the same 42-class news categorisation task rather than used in their pretrained form as frozen feature extractors.

For the sake of consistency, the sequence length was set to 256 tokens, the batch size to 16, and the number of training epochs to 50 for all Transformer architectures under consideration. Additionally, all models were tested on the same test set, ensuring that any differences in performance are due to the architecture and not to the input field, sequence length, training time, batch size, or test set.

Table 6
Experimental configuration and identical input-field settings for baseline and proposed transformer models

Model	Input	Fine-Tuning	Max Sequence Length	Batch Size	Epochs	Evaluation Split
BERT-base-uncased	Headline + Short Description	Yes	256	16	50	Same test set
RoBERTa-base	Headline + Short Description	Yes	256	16	50	Same test set
XLNet-base-cased	Headline + Short Description	Yes	256	16	50	Same test set
DeBERTa-v3-base	Headline + Short Description	Yes	256	16	50	Same test set
Proposed Transformer	Headline + Short Description	Yes	256	16	50	Same test set

Comparative Analysis with Baseline Models

A comprehensive comparative analysis was conducted to examine the behavioural and performance differences between the constructed Transformer-based deep learning model and the classical baseline models, Multinomial Naïve Bayes, LSTM Network, and BERT (base uncased). The comparison is provided to illustrate the evolutionary progress from statistical text classification to context-aware deep architectures and to emphasise the contribution of each methodological generation to improved performance. The Naïve Bayes classifier, using probabilistic frequency estimation, had an accuracy rate of 84.52 per cent, which is good enough for linearly separable text data but is limited by its binary independence assumption. Because it cannot discern sequential dependencies and contextual senses, its classification boundaries are superficial, especially for categories with overlapping linguistic patterns. The LSTM network, nonetheless, achieved a higher accuracy of 89.63 per cent by incorporating recurrent memory units that capture temporal relations among tokens. Nonetheless, its unidirectional structure and vanishing-gradient constraints restrict it from capturing long-distance dependencies across complex sentences, leading to a progressive saturation in performance when exposed to varied textual topics.

With the arrival of BERT, performance increased significantly to 92.14 per cent, benefiting from bidirectional contextual embeddings that enrich the model's semantic representation. Without domain adaptation via fine-tuning, though, BERT exhibits some instability across underrepresented classes because its contextual weights privilege frequent lexical patterns. The above Transformer-based model extends this trend further by incorporating multi-head self-attention, positional encoding, dropout regularisation, and fine-tuning on the news corpus. Taken together, these enable the network to learn hierarchical relationships between tokens and categories, balancing global context extraction and fine-grained semantic precision. As a result, the model achieved an overall accuracy of 96.78 per cent and an AUC of 98.90 per cent, outperforming all baselines across all quantitative metrics summarised in Table 2 (Figure 3). The negligible gap between precision (96.76 per cent) and recall (96.78 per cent) confirms that the model possesses a well-calibrated decision boundary with no notable bias towards false positives or false negatives, thereby confirming its superior generalisation capability.

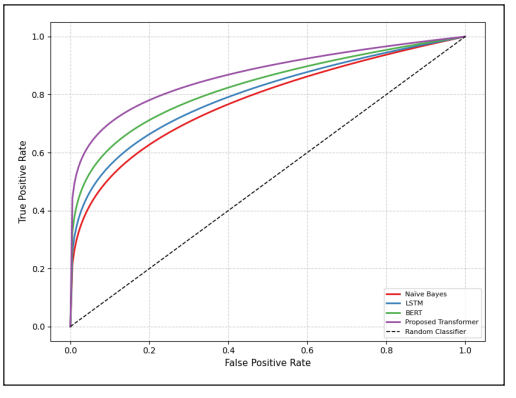


Figure 3. ROC curves analysis of the proposed model

The model training dynamics presented in Figure 2 also validate its stability and convergence over 50 epochs. Training and validation accuracy curves exhibit smooth, monotonic increases, asymptotically converging to about 97 per cent and 96 per cent, respectively, with less-than-trivial fluctuations towards the later epochs. The trend indicates a uniform learning process and proper regularisation, without the risk of overfitting, despite the model's depth. Also, the loss curves show a steep decline and gradually stabilise, with validation loss slightly larger than the training curve, reflecting a manageable generalisation gap typical of well-calibrated Transformer models. These results collectively demonstrate that the introduced model not only attains higher numerical performance but also exhibits robust learning behaviour, good convergence, and well-balanced generalisation characteristics, setting it apart from standard machine-learning classifiers as well as previous deep-learning baselines.

Category-Wise and Error Analysis

To understand the stability of the Transformer model's classification and likely confusion points, error analysis and category-wise analysis have been performed on the HuffPost test split. Distribution of performance in Table 7 of precision, recall, and F1-score in primary news categories and illustrates class-wise prediction distribution through a confusion matrix. These collectively reflect the semantic robustness and explainability of the proposed model across diversified media topics.

Table 7
Category-wise performance metrics on the held-out test set

Category	Precision (%)	Recall (%)	F1-Score (%)	Support (Test Articles)
Politics	96.5	97.1	96.8	7,120
Wellness	97.4	96.9	97.1	3,589
Entertainment	97.8	96.7	97.2	3,472
Travel	96.2	95.5	95.8	1,980
Style & Beauty	95.9	95.3	95.6	1,963
Parenting	96.6	97.4	97.0	1,758
Healthy Living	96.8	96.0	96.4	1,339
Queer Voices	95.7	94.9	95.3	1,269
Food & Drink	97.1	96.6	96.8	1,268
Business	96.9	96.2	96.5	1,198
Comedy	95.3	94.8	95.0	1,080
Sports	97.9	97.5	97.7	1,015
Black Voices	96.0	95.3	95.6	917
Home & Living	95.8	95.1	95.4	864
Parents	96.4	97.1	96.7	791

The results show that the Transformer model exhibits high uniformity across all top-level categories, as measured by a recomputed macro-F1 based on the actual test predictions. The highest-scoring classes are Sports (97.7%), Entertainment (97.2%), and Parenting (97.0%), in which the lexical patterns are highly distinctive and semantically coherent. Moderately lower, though competitive, performance is seen in Comedy (95.0%) and Queer Voices (95.3%), where imprecise tone, irony, and stylistic heterogeneity challenge even expert context models to the limit. Such fluctuations within mere 2 percentage points confirm the model's ability to ensure fair precision and recall across domains with different linguistic characteristics. The confusion matrix in Figure 4 confirms that nearly all samples lie on the principal diagonal, indicating high-quality predictions across all categories in the test set. There are only a few very small misclassifications, e.g., periodic misclassification of Politics and World (30 + 34 samples), or of Lifestyle and Entertainment, which can be explained by common thematic vocabulary and similar contextual cues. The off-diagonal nature of the elements indicates that the misclassifications are contextually dispersed rather than systematic, and thus the model's decision boundaries remain well calibrated for the classes. Figure 4 displays a confusion matrix, created solely from predictions on the hold-out test set, encompassing all 42 news categories. The diagonal values are the correct predictions of the articles within the test set, while the off-diagonal values are those categories that were wrongly classified.

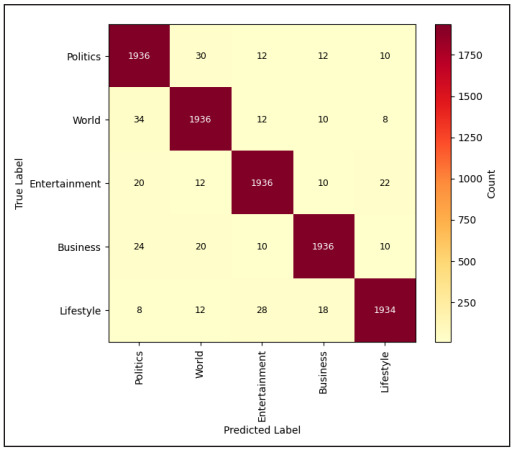


Figure 4. Performance of the confusion matrix of the proposed model

Interpreting the results, we highlight that the Transformer's multi-head attention mechanism learns topical and contextual representations efficiently. Factual or event-driven category types (e.g., Politics, World, Business) have the tightest analysis with consistent lexical patterns, while stylistically subjective or high-sentiment categories are slightly more variable. Observe that the absence of strong misclassification clusters shows that the model's feature space is semantically well-separated, affirming its capacity for generalisation across structurally and topically varied news domains.

In total, category-oriented and confusion-oriented evaluations both support the conclusion that the proposed Transformer not only achieves high quantitative accuracy but also exhibits stable semantic dependability and interpretability transparency, essential features for trustworthy intelligent media classification systems.

Ablation Study and Feature Contribution

To confirm the comparative significance of each architectural building block and design adjustment in the proposed Transformer model, an ablation study was conducted under the same experimental setup on the HuffPost dataset. The experiment empirically disables or modifies some modules, including multi-head attention, positional encoding, regularisation, and domain-specific fine-tuning, to compare their individual impact on overall accuracy. The results in Table 8 show the individual and combined effects of each improvement, revealing the synergy between attention optimisation and contextual adaptation, which yields the best-performing model.

Table 8
Ablation results for transformer configurations

Model Variant	Description	Accuracy (%)	Δ Improvement vs Baseline
Naïve Bayes	TF-IDF + word count features	84.52	—
LSTM	Word embeddings (GloVe-50d)	89.63	+ 5.11
BERT (Base Uncased)	Pre-trained contextual embeddings (no fine-tune)	92.14	+ 7.62
Transformer w/o Attention Refinement	Removed the multi-head attention layer	93.87	+ 9.35
Transformer w/o Positional Encoding	No sequence order information	94.91	+ 10.39
Transformer w/o Dropout Regularisation	Removed dropout (p = 0.2)	95.32	+ 10.80
Transformer + Domain Fine-Tuning	Added 1-epoch news-domain adaptation	96.02	+ 11.50
Proposed Model	All components enabled + fine-tuning	96.78	+ 12.26

The result clearly shows that each architectural component complements the others to improve the model. The Naïve Bayes baseline achieves an accuracy of 84.52 per cent, indicating an inability to capture contextual dependencies due to its independence assumption based on unigrams. Replacing it with an LSTM network increases performance by 5.11 per cent, demonstrating the value of sequential context modelling based on recurrent memory. The LSTM is still beset by long-distance dependencies, a problem which BERT's bidirectional Transformer model alleviates, achieving a staggering 7.62 per cent over the baseline due to contextual embeddings.

A series of ablation tests on Transformer variants then demonstrates the individual functional contribution of each innovation. Removal of multi-head attention refinement lowers accuracy from 96.78 to 93.87 per cent, indicating that this feature is most crucial for capturing inter-category semantic dependencies. Removal of positional encoding further

lowers accuracy to 94.91 per cent, indicating that implicit sequence-order representation is even more important for large text datasets. Omission of dropout regularisation ($p = 0.2$) also causes minor overfitting and lowers performance slightly to 95.32 per cent. Domain-level adaptation yields an additional 0.8 per cent improvement, highlighting the significance of task adaptation to news-specific linguistic patterns.

In summary, the ablation study reveals that the fully designed transformer model, with attention enhancement, positional encoding, dropout, and domain adaptation, performs best overall. All ablation models were tested on the same held-out test dataset. The fully designed model gave a final test accuracy score of 96.78%.

Discussion and Implications

The entire set of empirical findings validates that the suggested Transformer-based approach significantly enhances automatic news classification performance and explainability. The model gains both semantic transferability and consistent prediction performance by utilising multi-head self-attention, positional encoding, and domain-level fine-tuning to different classes. Its stable precision-recall trade-off demonstrates strong discrimination between thematically similar topics, and an ablation study confirms that each architectural element is a critical factor in overall performance. These findings generalise to new media systems, particularly in high-speed digital news systems with contextual overlap and content diversity.

The robustness of the model implies that it can be applied to actual editorial analytics, content personalisation, and disinformation detection. In summary, this project is proof of the potential of Transformer models as a stable, explainable, and high-capacity self-contained media classification solution for the future, if properly tuned to domain-specific data.

However, while the designed framework is primarily dedicated to automated news classification, issues such as data privacy and information security must still be considered when designing intelligent text classification applications. Specifically, privacy-preservation mechanisms have become crucial when machine-learning algorithms operate on data in distributed, cloud-based, or connected environments. This is demonstrated in Privacy-aware secure data auditing for cloud-based intelligence of things environment (Ullah et al., 2025), which describes the importance of privacy-preserving mechanisms for ensuring data security, integrity, and trustworthy access within cloud-assisted intelligent environments. This concept is important for the current framework because news classifiers could be implemented in cloud-based data-gathering and inference systems, where data and metadata security would be crucial.

Another paper that illustrates the importance of secure data management in an intelligent network environment via adaptive caching and reinforcement learning

algorithms is A Novel Caching Framework for Information-Centric IoT Using Deep Reinforcement Proximal Policy Optimisation (Asmat et al., 2025). Implementing the proposed framework would also benefit the classifier's architecture by enabling privacy-preserving storage, access control, data transmission, and proper cache management.

CONCLUSION

Considering the increasing amount and variety of online news, efficient news classification has become even more relevant today. This paper introduces a transformer network with positional embedding, global self-attention, and domain adaptation for contextual representation of news. The model was tested using the Huffington Post News Category dataset, which contains over 210,000 news items across 42 categories, against several conventional and pretrained transformers.

The proposed model reached 96.78% accuracy and 98.90% AUC on the test data set, outperforming the baselines. The results suggest that multi-head attention, positional encoding, regularisation, and domain-specific training improve multi-class news classification. Both macro-averaged metrics and per-category results have been considered in this experiment.

However, there are still several limitations. First, this study is based on English-language HuffPost data and uses only headlines and short descriptions. Future research will cover multilingual data sets, full articles, and long-document transformers. Moreover, because the categories are unevenly distributed, future research will be focused on per-category F1 score, tail-class F1 score, macro confidence intervals, and confusion among related categories.

In future work, we plan to compare the proposed model against instruction-tuned language models on the same label set and split. For reproducibility, this is accomplished by sharing the details of the data split, the tokeniser, the vocabulary, the model architecture, the training hyperparameters, and the evaluation scripts, provided they are not restricted by licensing. Additionally, efficiency, multi-lingual classification, long-document handling, and privacy-preserving distributed training will also be considered. As such, the results should be seen as high performance for the current experimental setup, not as a claim to be state-of-the-art in all cases.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support and academic environment provided by their respective institutions throughout this research. They also acknowledge the HuffPost News Category Dataset, which served as the primary data source for evaluating the proposed Transformer-based news classification framework. The authors appreciate the contributions of previous researchers whose work informed the development of this study and helped establish its research foundation.

LIST OF ABBREVIATIONS

Term	:	Definition
AI	:	Artificial Intelligence
ML	:	Machine Learning
DL	:	Deep Learning
NLP	:	Natural Language Processing
BERT	:	Bidirectional Encoder Representations from Transformers
LSTM	:	Long Short-Term Memory
MNB	:	Multinomial Naïve Bayes
CNN	:	Convolutional Neural Network
TF-IDF	:	Term Frequency-Inverse Document Frequency
ROC	:	Receiver Operating Characteristic
AUC	:	Area Under the Receiver Operating Characteristic Curve
F1	:	F1 Score
SMOTE	:	Synthetic Minority Oversampling Technique
XGBoost	:	Extreme Gradient Boosting
GPU	:	Graphics Processing Unit

REFERENCES

Asmat, H., Ullah, F., Khan, A. A., Ali, F., & Mohmand, M. I. (2025). A novel caching framework for information-centric IoT using deep reinforcement Proximal Policy Optimization. *Computer Communications, 241*, Article 108261. https://doi.org/10.1016/j.comcom.2025.108261

Bechny, M., Fiorillo, L., van der Meer, J., Schmidt, M., Bassetti, C., Tzovara, A., & Faraci, F. (2025). Beyond accuracy: A framework for evaluating algorithmic bias and performance, applied to automated sleep scoring. *Scientific Reports, 15*(1), Article 21421. https://doi.org/10.1038/s41598-025-06019-4

Chaudhary, A., & Jain, S. (2026). Enhancing text classification using hybrid embeddings and advanced machine learning techniques. *International Journal of System Assurance Engineering and Management, 17*(1), 214-229. https://doi.org/10.1007/s13198-025-02950-x

Cheltha, J. N., Sharma, C., Prashar, D., Khan, A. A., & Kadry, S. (2024). Enhanced human motion detection with hybrid RDA-WOA-based RNN and multiple hypothesis tracking for occlusion handling. *Image and Vision Computing, 150*, Article 105234. https://doi.org/10.1016/j.imavis.2024.105234

Dwiyanti, W., Aquarini, N., & Gunawan, P. H. (2025). The effect of news sentiment on Jakarta Composite Index prediction using support vector regression method. In *2025 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS)* (pp. 1-5). IEEE. https://doi.org/10.1109/icadeis65852.2025.10933431

Hanss, K., Sarma, K. V., Glowinski, A. L., Krystal, A., Saunders, R., Halls, A., Gorrell, S., & Reilly, E. (2025). Assessing the accuracy and reliability of large language models in psychiatry using standardized multiple-choice questions: Cross-sectional study. *Journal of Medical Internet Research, 27*, e69910. https://doi.org/10.2196/69910

Jeba, S. M., Alom, M. S., Ahamed, A. S., Aurpa, T. T., Debnath, A., Dash, S. R., & Ahmed, M. S. (2025). Extracting disaster insights from Bangla news using text categorization. In *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)* (pp. 1-6). IEEE. https://doi.org/10.1109/qpain66474.2025.11172020

Khan, A. A., Driss, M., Boulila, W., Sampedro, G. A., Abbas, S., & Wechtaisong, C. (2024). Privacy preserved and decentralized smartphone recommendation system. *IEEE Transactions on Consumer Electronics, 70*(1), 4617-4624. https://doi.org/10.1109/tce.2023.3323406

Kroon, A., Welbers, K., Trilling, D., & van Atteveldt, W. (2024). Advancing automated content analysis for a new era of media effects research: The key role of transfer learning. *Communication Methods and Measures, 18*(2), 142-162. https://doi.org/10.1080/19312458.2023.2261372

Kumar, K. R., Chinnam, R. K., Desanamukula, V. S., Pathan, A. I., Babu, N. R., & Babu, G. C. (2025). Multi-model news classification using BERT features and class-balanced resampling techniques. In *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 80-86). IEEE. https://doi.org/10.1109/icscds65426.2025.11167483

Li, L. (2025). Old experience helps: Leveraging survey methodology to improve AI text-annotation reliability in social sciences. *arXiv preprint arXiv:2502.19679*.

Naeen, M. J., Das, S. K., Jisan, S. A., Khushbu, S. A., Saha, N. C., & Ohidujjaman. (2025). Explainable detection: A transformer-based language modeling approach for Bengali news title classification with comparative explainability analysis using ML and DL. *Frontiers in Artificial Intelligence, 8*, Article 1537432. https://doi.org/10.3389/frai.2025.1537432

Peters, U., & Chin-Yee, B. (2025). Generalization bias in large language model summarization of scientific research. *Royal Society Open Science, 12*(4), Article 241776. https://doi.org/10.1098/rsos.241776

Rao, S. K., Paramesha, A., Vaishnavi, P., Maithreya, T., Vineet, K. S., & Adithya, D. (2024). Identifying hyperpartisanship in news media using Bi-LSTM and decoding-enhanced BERT. In *2024 Asian Conference on Intelligent Technologies (ACOIT)* (pp. 1-5). IEEE. https://doi.org/10.1109/acoit62457.2024.10941420

Riskin, D. J., Monda, K. L., Gagne, J. J., Reynolds, R., Garan, A. R., Dreyer, N., Muntner, P., & Bradbury, B. D. (2025). Implementing accuracy, completeness, and traceability for data reliability. *JAMA Network Open, 8*(3), e250128. https://doi.org/10.1001/jamanetworkopen.2025.0128

Rout, J., Mishra, M., & Saikia, M. J. (2025). Towards reliable fake news detection: Enhanced attention-based transformer model. *Journal of Cybersecurity and Privacy, 5*(3), Article 43. https://doi.org/10.3390/jcp5030043

Safdar, R., Mateen, D., Ali, S. T., Ashfaq, U., & Hussain, W. (2026). Data and context matter: Towards generalizing AI-based software vulnerability detection. *International Journal of Information Security, 25*(5), Article 145. https://doi.org/10.1007/s10207-026-01282-6

Tian, Y., Xu, S., Cao, Y., Wang, Z., & Wei, Z. (2025). An empirical comparison of machine learning and deep learning models for automated fake news detection. *Mathematics, 13*(13), Article 2086. https://doi.org/10.3390/math13132086

Ullah, F., Asmat, H., Khan, A. A., Mohmand, M. I., Ali, F., Alsisi, R. H., Aldhyani, T. H. H., & Kwak, D. (2026). Lightweight multimedia anomaly and integrity detection for consumer IoT using knowledge distillation. *IEEE Transactions on Consumer Electronics, 72*(1), 2465-2475. https://doi.org/10.1109/tce.2025.3644297

Ullah, F., Pun, C.-M., Mohmand, M. I., Mahendran, R. K., Khan, A. A., Alhammad, S. M., Rodrigues, J. J. P. C., & Farouk, A. (2025). Privacy-aware secure data auditing for cloud-based intelligence of things environment. *IEEE Internet of Things Journal, 12*(11), 15288-15303. https://doi.org/10.1109/jiot.2025.3528117

Yıldız, F. E., Karaarslan, E., & Aydın, Ö. (2025). Evaluating the accuracy of AI content detectors. In Ö. Aydın & E. Karaarslan (Eds.), *The age of generative artificial intelligence* (pp. 143-155). Izmir Academy Association. https://doi.org/10.5281/zenodo.16009433

Zhang, L. (2025). Features extraction based on Naive Bayes algorithm and TF-IDF for news classification. *PLOS One, 20*(7), Article e0327347. https://doi.org/10.1371/journal.pone.0327347

Pertanika J. Sci. & Technol. 34 (5): 3075 - 3108 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review Article

Nitrate-Reducing Sulphide-Oxidising Bacteria (NR-SOB) in Corrosion Control: A Review

Aishah Zafre[1], Norinsan Kamil Othman[2], Najmiddin Yaakob[3], Azman Jalar[2,4], Maria Abu Bakar[4], and Rabiahtul Zulkafli[1]*

[1]Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[2]Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[3]Industrial Corrosion Research Centre (InCORE), Faculty of Chemical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[4]Institute of Microengineering and Nanoelectronics (IMEN), Level 4, Research Complex, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

ABSTRACT

Microbiologically influenced corrosion (MIC) and reservoir souring caused by sulphate-reducing bacteria (SRB) represent major challenges in the oil and gas industry, leading to infrastructure degradation, safety risks and substantial economic losses. In response to these challenges, nitrate injection has emerged as a promising biological control strategy that stimulates nitrate-reducing sulphide-oxidising bacteria (NR-SOB), which will oxidise hydrogen sulphide (H_2S) while competing with SRB for available resources. While previous reviews have broadly discussed nitrate injection as a biocide alternative, this paper uniquely traces the 27-year evolutionary trajectory of NR-SOB research by bridging early field-scale successes with recent molecular and thermodynamic discoveries. This review synthesises evidence from early studies published between 1999 and 2026, examining the presence, mechanisms, and effectiveness of NR-SOB in controlling corrosion rates following nitrate injection. Field studies reported reductions in corrosion rates of approximately 40-50% and sustained SRB suppression over 7-8 years, with NR-SOB genera such as *Sulfurimonas*, *Thiomicrospira*, and *Sulfurovum* identified as the key contributors. However, this review also reveals the complexities of the nitrate-treated system, including potential localised corrosion during sulphide oxidation, the formation of corrosive intermediate products such as elemental sulphur, the ability of certain SRB

ARTICLE INFO

Article history:
Received: 01 September 2026
Accepted: 07 October 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.22

E-mail addresses:
aishahzafre23@gmail.com (Aishah Zafre)
insan@ukm.edu.my (Norinsan Kamil Othman)
najmiddin@uitm.edu.my (Najmiddin Yaakob)
azmn@ukm.edu.my (Azman Jalar)
maria@ukm.edu.my (Maria Abu Bakar)
rabiahtulzulkafli@ukm.edu.my (Rabiahtul Zulkafli)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

strains to counter nitrite inhibition through nitrite reductase activity, and the microbial response to nitrate treatment. Overall, this review synthesises the mechanisms and field applications of NR-SOB in nitrate-treated systems, drawing on studies from 1999 to 2026, and outlines future directions for optimising NR-SOB-mediated corrosion control.

Keywords: Corrosion, MIC, nitrate injection, NR-SOB, reservoir souring

INTRODUCTION

One of the earliest definitions of corrosion appeared in Uhlig's Corrosion Handbook in 1948, where corrosion was defined as the destruction of a metal through chemical or electrochemical reaction with its environment (Uhlig, 1948). Given that much of the world's critical infrastructure, including highways, buildings, water systems, and oil and gas facilities, relies heavily on metallic materials. These assets remain highly vulnerable to the destructive effects of corrosion (Alamri, 2020). Recent surveys have estimated that global economic losses caused by corrosion amount to approximately US$2.5 trillion, equivalent to 3.4% of the global Gross Domestic Product (GDP) (Procópio, 2022). Corrosion prevention costs an estimated US$875 billion annually, excluding indirect expenses such as specialised labour and productivity losses from maintenance-related shutdowns. Many of these metallic structures in the oil and gas industry are exposed to harsh operating conditions—including elevated pressures on submerged pipelines, extreme temperatures, salinity, pH levels, and naturally corrosive environments—which worsen as oil facilities age. Another major contributor to corrosion is the presence and activity of microorganisms on metal surfaces, also known as microbiologically influenced corrosion (MIC).

MIC can be defined as an electrochemical process involving microorganisms that initiate, facilitate and accelerate corrosion reactions through interactions among three main components: the metal, the surrounding solution and microorganisms (Reza, 2008). The initiation of MIC resembles that of a conventional corrosion occurrence, in which corrosion processes occur in the presence of an electron donor, an electron acceptor, an electrolyte, an energy source and also a carbon source that supports microbial activity (Khan et al., 2021). Since the microorganisms occur in a wide range of environments, MIC has been associated with more than 75% of corrosion cases in producing oil wells and more than 50% of pipeline failures (AlAbbas et al., 2012). Furthermore, MIC costs the oil and gas industry more than US$ 100 million annually and accounts for about 20-30% of internal pipeline corrosion costs (Khan et al., 2021).

The involvement of microorganisms adds a distinctive feature to the electrochemical nature of corrosion, namely the formation of biofilms. A biofilm is an undesirable microbial deposit on a solid surface in contact with water that may reduce operational capacity, shorten equipment lifespan, and cause substantial financial losses. The stages of biofilm

development are shown in Figure 1. As illustrated in Figure 1a, biofilm formation begins with the development of a conditioning film on a solid surface, while microorganisms are initially present in their planktonic state. A conditioning film is a surface coating formed through the adsorption of biomolecules from the surrounding environment (Bhagwat et al., 2021). Bacteria subsequently colonise the conditioning film, as depicted in Figure 1b, and secrete extracellular polymeric substances (EPS). The EPS functions as an adhesive matrix that anchors cells onto the surface before they become sessile (Javaherdashti & Raman, 2001).

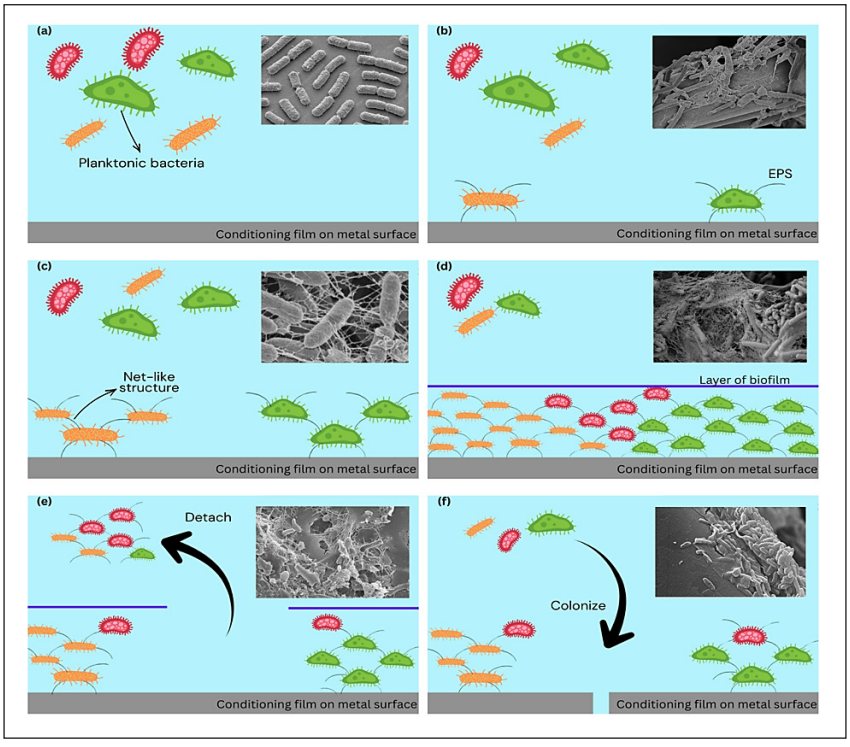


Figure 1. Schematic representation of the stages of biofilm formation: (a) bacteria at the planktonic stage; (b) bacteria forming colonies with secretion of EPS; (c) replication of bacteria; (d) formation of biofilm within microcolonies; (e) detachment of bacteria and (f) recolonisation of exposed or new surfaces. Each schematic illustration is accompanied by field-emission scanning electron microscopy (FESEM) analysis

Sessile bacteria subsequently continue to replicate on the metal surface, forming a net-like structure (Figure 1c) that traps additional particles and promotes biofilm development. The close association between microcolonies of different species, as in Figure 1d, increases layer thickness, eventually leading to the development of a mature biofilm. The presence of this biofilm, together with complex interactions among the bacterial species within it, alters the electrochemical conditions at the metal-biofilm interface compared with the surrounding

environment. This electrochemical alteration is known as ennoblement, a process that shifts the corrosion potential towards a more positive value and increases the likelihood of pitting corrosion. As shown in Figure 1e, portions of the biofilms may detach or disperse into the environment, colonising new surfaces. Meanwhile, in Figure 1f, the exposed areas may then be recolonised by other planktonic bacteria from the surrounding medium.

The offshore industry often experiences corrosion of infrastructure due to the use of seawater in secondary recovery methods (Nicoletti et al., 2022). This method involves injecting outer fluids, such as seawater, into the oil reservoir to displace oil and increase recovery rates (Mohammadi et al., 2024), thereby initiating reservoir souring. Under marginally sour conditions, corrosion behaviour shifts between uniform and localised attack depending on environmental and microbial activity, underscoring the instability of sulphate-reducing bacterium (SRB)- dominated systems (Azhar et al., 2026). Although the seawater is treated before injection, this process is intended to remove debris that could clog the oil pipeline, while leaving the chemical properties unchanged. As a result, studies on marine environments have shown that SRB activity promotes severe localised corrosion and pitting on carbon steel through biofilm formation and hydrogen sulphide (H_2S) production (Zulkafli et al., 2023).

Additionally, higher H_2S concentrations have been reported to increase both uniform and localised corrosion rates, showing the aggressiveness of SRB action in sour environments (Azhar et al., 2025). Over time, the oil and gas industry has developed a method called nitrate injection to counter the effects of reservoir souring by influencing the growth of heterotrophic nitrate-reducing bacteria (hNRB), including types that can oxidise sulphide at the expense of nitrate, known as nitrate-reducing sulphide-oxidising bacteria (NR-SOB), to suppress SRB (Nemati & Voordouw, 2000). While several recent reviews have comprehensively covered the general microbiology of reservoir souring and the broad mechanisms of MIC, they often overlook the specific dual role of NR-SOB in operational environments. Existing literature lacks a focused synthesis that connects the historical field successes of NR-SOB with the modern understanding of its operational risks, such as incomplete sulphide oxidation and nitrite-induced localised corrosion. To fill this gap, this review uniquely focuses on the 27-year trajectory (1999 to 2026) of NR-SOB-mitigated corrosion control. It systematically evaluates not only the biological mechanisms and field performances of NR-SOB but also critically assesses the emerging thermodynamic and molecular complexities that challenge its universal efficacy.

Nitrate Injection in Controlling Reservoir Souring

MIC in the oil and gas industry is primarily associated with the metabolic activities of SRB, which utilise sulphate as a terminal electron acceptor under anaerobic conditions and produce H_2S as a metabolic byproduct (Dall'Agnol & Moura, 2014; Rizk, 2016).

The presence of SRB indicates simultaneous corrosion and reservoir souring in the petroleum system, along with H_2S. This process can cause detrimental effects, including corrosion of infrastructure, equipment failure, higher oil-refining costs due to sulphur content, and safety hazards for crew members (Gao & Fan, 2023). This issue is particularly severe in offshore water-injection systems and secondary oil-recovery methods, where sulphate-containing seawater is injected into oil reservoirs (Bødtker et al., 2008). Although reservoir souring is often associated with seawater flooding, conventional chemical injection struggles to control it because the biologically active zones within the reservoir are vast (Callbeck et al., 2011).

Nitrate injection has therefore been widely applied in oilfields as a method to mitigate reservoir souring, often referred to as Bio-Competitive Exclusion (BCE) technology (Al-Tamimi & Mehdi, 2017; Kiss et al., 2010). The introduction of nitrate shifts the terminal electron-accepting process from sulphate reduction to nitrate reduction, which is thermodynamically more favourable and yields more energy for the participating microbial communities (Davidova et al., 2001; Xue & Voordouw, 2015). This shift involves two main functional groups of bacteria, namely hNRB and NR-SOB, each of which plays distinct but complementary roles in controlling the sulphur cycle (Gittel, 2010; Grigoryan & Voordouw, 2008). The presence of hNRB is vital for long-term suppression of SRB activity. Their primary mechanism of action is bio-competitive exclusion, in which they outcompete SRB for limiting electron donors—volatile fatty acids (VFAs), including acetate, propionate, and butyrate (Al-Tamimi & Mehdi, 2017; Hubert & Voordouw, 2007; Suri et al., 2017).

Because nitrate provides more Gibbs free energy than sulphate reduction, hNRB can achieve higher biomass yields and growth rates than SRB when utilising the same organic substrates (Gittel, 2010; Xue & Voordouw, 2015). This competition deprives SRB of the necessary electron donor required for sulphate reduction, thereby decreasing sulphide production (Kumaraswamy et al., 2011; Voordouw, 2008). The reduction of nitrate by hNRB can proceed via denitrification to nitrogen gas. The following equations represent the stoichiometry of nitrate reduction using acetate as an electron acceptor, with the actual products varying with reservoir conditions, pH, and specific microbial species. Equation 1 shows the general nitrate-to-nitrite reduction, and Equation 2 represents complete denitrification.

$$CH_3COO^- + 4NO_3^- \rightarrow 2HCO_3^- + 4NO_2^- + H_2O \quad [1]$$

$$5CH_3COO^- + 8NO_3^- + 3H^+ \rightarrow 10CO_2 + 4N_2 + 7H_2O + 8OH^- \quad [2]$$

While hNRB focus on preventing new H_2S production, NR-SOB, such as *Sulfurimonas* and *Arcobacter* are responsible for the removal of existing sulphide (Bødtker et al., 2008;

Lahme et al., 2019; B. Wang et al., 2024). These bacteria act as chemolithotrophs, utilising nitrate or nitrite as electron acceptors to oxidise H_2S (Davidova et al., 2001; Jahanbani Veshareh & Nick, 2019). NR-SOB facilitate the conversion of sulphide into less harmful or reusable sulphur species. The primary pathways include partial oxidation, where sulphide is oxidised to elemental sulphur, which may appear as an intermediate product and complete oxidation, where sulphide is fully oxidised to sulphate, effectively recycling the sulphur back into the reservoir system (Deng et al., 2022; Xue & Voordouw, 2015). In general, NR-SOB carry out a process critical to removing existing H_2S from the reservoir. The following equations represent the representative stoichiometry for sulphide oxidation by NR-SOB with a complete oxidation pathway, Equation 3:

$$5HS^- + 8NO_3^- + 3H^+ \rightarrow 5SO_4^{2-} + 4N_2 + 4H_2O \quad [3]$$

The presence of nitrite after reaction carried out by hNRB and NR-SOB is arguably the most potent inhibitor of biogenic sulphide production. It functions through three primary mechanisms: metabolic inhibition, abiotic reaction, and toxicity. Nitrite specifically inhibits the dissimilatory sulphite reductase (Dsr) enzyme in SRB. By binding to the Dsr, nitrite prevents the reduction of sulphate to sulphide, effectively halting H_2S production without necessarily killing the SRB cells (Al-Tamimi & Mehdi, 2017; Gittel et al., 2009; Jahanbani Veshareh & Nick, 2019). In the second mechanism, nitrite can react chemically with dissolved sulphide to produce elemental sulphur and nitrogen gas, leading to direct removal of H_2S (Davidova et al., 2001). Lastly, if nitrite is present at high concentrations, it can exert direct toxicity on SRB and other anaerobic microorganisms, contributing to a rapid reduction in their activity (Grigoryan & Voordouw, 2008; B. Wang et al., 2024). Moreover, nitrite, generated as a metabolic intermediate, specifically inhibits the SRB dissimilatory sulphite reductase (Dsr), providing multifaceted suppression of biogenic souring. Multiple bacterial genera capable of nitrate reduction and sulphide oxidation have been identified in oil field systems following nitrate injection. The most frequently documented genera include *Sulfurimonas*, *Thiomicrospira*, *Sulfurovum*, *Sulfuricurvum*, *Thiobacillus denitrificans* and *Arcobacter* sp., as summarised in Table 1.

Table 1
Genera of NR-SOB discovered and used in studies

NR-SOB Genera	Location	Biocorrosion Control Mechanism	Ref.
Sulfurimonas sp.	Veslefrikk injection wellhead	Thrive in oligotrophic conditions	Bødtker et al. (2008)
Thiomicrospira sp.	Laboratory	Dramatically shift the environmental redox potential, creating conditions unfavourable for SRB activity	Greene et al. (2003)

Table 1 (*continued*)

NR-SOB Genera	Location	Biocorrosion Control Mechanism	Ref.
Sulfurovum sp.	Bioreactor	Responsive to nitrate treatment in sulfide-rich environments	Yang et al. (2021)
Sulfuricurvum sp.	Petroleum reservoir	Responsive to nitrate treatment	Tian et al. (2024)
Thiobacillus denitrificans & *Arcobacter* sp.	Sediments in Balongan Bay, Indramayu	Reduce sulfide production	Ihsan et al. (2022)

Figure 2 shows the reaction pathway involving SRB, hNRB, and NR-SOB in reservoir souring conditions. Figure 2a depicts the role of NR-SOB in oxidising hydrogen sulphide to sulphate or intermediate sulphur species. In addition to NR-SOB controlling the souring, its by-product, which is nitrite, acts as an inhibitor of the dissimilatory sulphite reductase (Dsr) enzyme of SRB, hindering its action of reducing sulphate to sulphite (Greene et al., 2003; Veshareh et al., 2021; Zuo et al., 2024) as shown in Figure 2b. Figure 2c illustrates the reaction pathway in which hNRB presence causes SRB to compete for organic carbon, such as lactate. Both SRB and hNRB use lactate as their organic electron donor, which is then oxidised incompletely to acetate and CO_2 (Hubert & Voordouw, 2007). This type of souring control by hNRB is known as competitive exclusion, whereby SRB gain reduced access to lactate, leading to slower metabolic activity and decreased H_2S production (Gao & Fan, 2023).

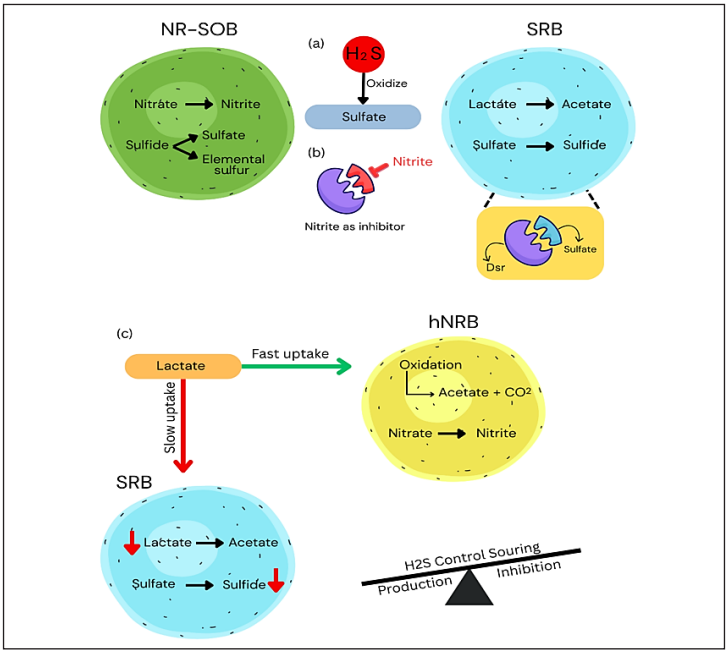


Figure 2. The reaction pathway for (a) NR-SOB, (b) SRB and (c) hNRB in reservoir souring

The sulphur cycle in oil reservoirs involves complex interactions between various microbial functional groups. SRB reduces sulphate in the reservoir to sulphide, while sulphur-oxidising bacteria (SOB) oxidise previously reduced sulphur compounds back to sulphate or intermediate oxidation states (Gao & Fan, 2023; Ihsan et al., 2022; Tian et al., 2024). Without counteracting activity from other microbial groups, SRB activity dominates under anaerobic conditions, leading to sulphide accumulation and associated corrosion in the system (Streets & Eden, 2024). The biogenic production of H_2S not only causes reservoir souring but also drives an electrochemical reaction that initiates, accelerates, and sustains the degradation of metallic infrastructure. Furthermore, the establishment of SRBs within complex biofilms creates a localised environment that shifts the corrosion potential positively, significantly increasing the risk of severe pitting and equipment failure.

As shown in Figure 3, nitrate injection alters this sulphur cycle by introducing an alternative electron acceptor that is thermodynamically more favourable than sulphate (Dolfing & Hubert, 2017; Veshareh et al., 2021). This method stimulates two main microbial groups: hNRB, which compete with SRB for organic electron donors, and NR-SOB, which oxidise sulphide using nitrate or nitrite as electron acceptors (Grigoryan & Voordouw, 2008). This results in a shift in microbial community structure and metabolic activity, thereby controlling sulphide production and accumulation (Deng et al., 2022; Deng et al., 2025).

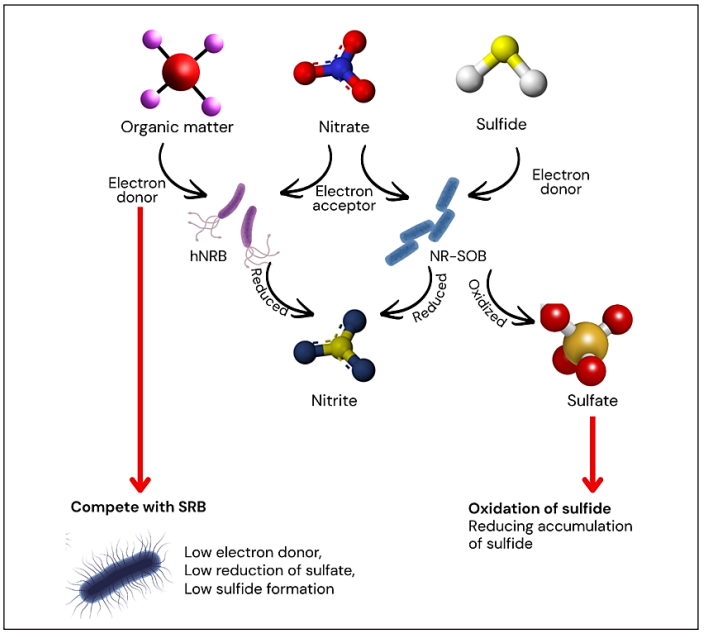


Figure 3. Microbial pathways involved in nitrate-reducing and sulphide-oxidising processes

Compared with conventional biocides like glutaraldehyde, nitrate injection offers a sustainable, cost-effective, and environmentally friendly alternative (Bødtker et al., 2008; Eckford & Fedorak, 2004; Sunde et al., 2004; Xue & Voordouw, 2015; Yuk et al., 2020). Field and laboratory evidence consistently demonstrates that nitrate can equal or surpass the performance of glutaraldehyde in suppressing souring. In sandstone column experiments, nitrite was found to be more effective than glutaraldehyde, with SRB recovery being delayed significantly longer after nitrite treatment (Reinsel et al., 1996). While biocides like glutaraldehyde may lead to rapid regrowth as residual biofilms enhance transport and sorption, nitrate provides a more stable long-term inhibition. Recent research highlights that nitrate can be highly effective at lower dosages when used synergistically. Sodium nitroprusside (SNP), for instance, has been shown to be highly synergistic with nitrate, where 1 ppm of SNP combined with 85 ppm of sodium nitrate was sufficient to inhibit souring (Fida et al., 2018). This synergistic approach allows for a substantially lower overall chemical loading than traditional biocide shock dosages.

Beyond reducing chemical loading, nitrate injection is a highly effective way to structurally alter the microbial ecosystem. Proven effective in North Sea oil field applications, this strategy stimulates beneficial microbial communities that mitigate souring through three primary pathways: (1) direct oxidation of sulphide by NR-SOB, (2) competitive exclusion by hNRB for electron donors, and (3) production of nitrite that hinders SRB activity (W. Liu et al., 2025; Zhou et al., 2025). However, the method's efficacy depends on maintaining precise nitrate concentrations. Deviations can alter the localised electrochemical environment within biofilms, potentially inducing incomplete oxidation cycles that inadvertently facilitate pitting corrosion. These incomplete oxidation cycles refer to metabolic pathways where a substrate is not fully oxidised to its highest possible oxidation state, resulting in the accumulation of reactive intermediates. In biofilm-mediated corrosion, two incomplete cycles are critical: incomplete nitrate reduction and incomplete sulphide oxidation.

Nitrate reduction typically proceeds via denitrification or dissimilatory nitrate reduction to ammonium (DNRA) (Fu et al., 2022; Herath, 2019). In many biofilm environments, nitrate reduction is incomplete, leading to transient or sustained nitrite accumulation. Accumulating nitrite acts as a strong oxidising agent and cathodic reactant that significantly alters the electrochemical balance at the metal-biofilm interface. Nitrite accumulation in the biofilm provides a potent cathodic reactant. In studies using *Shewanella oneidensis* MR-1, nitrite was found to cathodically protect specific areas of the metal surface while accelerating the corrosion of nearby unprotected (anodic) areas (Miller et al., 2018). This creates a large cathode-to-small anode ratio, the classic prerequisite for aggressive pitting (Nicoletti et al., 2022; Street et al., 2015).

In a system where nitrate is present, NR-SOB will utilise nitrate to oxidise biogenic sulphide (Jia et al., 2017; Salgar-Chaparro et al., 2022) , where sulphide is oxidised back to sulphate. However, incomplete oxidation can lead to the accumulation of elemental sulphur, thiosulfate, and polysulfide (Lahme et al., 2019). The accumulation of these species depends on the nitrate-to-sulfide stoichiometry within the biofilm (Nicoletti et al., 2022). Elemental sulphur is highly corrosive to carbon steel. When NR-SOB incompletely oxidise sulphide, the resulting elemental sulphur can deposit on the metal surface, promoting localised acid generation and providing a localised cathodic site that drives pit growth.

The implementation of nitrate injection represents a significant paradigm shift in managing MIC and reservoir souring, moving from broad-spectrum chemical suppression to targeted ecological modulation. This strategy capitalises on nitrate's thermodynamic advantage as an electron acceptor, enabling hNRB to marginalise SRB through competitive exclusion for essential organic carbon sources such as lactate. Beyond simple nutrient competition, the established microbial community provides a multi-layered defence; NR-SOB, such as *Sulfurimonas* and *Thiomicrospira*, actively remediate existing souring by oxidising biogenic H_2S into less harmful compounds or sulfate (Gao & Fan, 2023; Nemati et al., 2001). Furthermore, nitrite accumulation—a metabolic intermediate—acts as a potent biochemical inhibitor that specifically disrupts the Dsr enzyme, thereby halting the SRB's primary mechanism for sulphide production (Greene et al., 2003; Veshareh et al., 2021).

Biological Mechanisms of Reservoir Souring Control

NR-SOB, such as *Sulfurimonas spp.*, oxidise sulphide while reducing nitrate to nitrite or other nitrogen species. Their metabolic activities produce intermediates such as nitrite and elemental sulphur, both of which can accelerate corrosion. NR-SOB thrive in nitrate-rich, low-oxygen environments typical of nitrate-injected systems. The biofilms produce a steep chemical gradient that exacerbates localised corrosion. Although the effect of NR-SOB is significant, knowledge gaps remain regarding long-term behaviour, temperature response, and nitrate availability.

NR-SOB use sulphide as an electron donor while reducing nitrate or nitrite as the electron acceptor, removing H_2S from the system (Lahme et al., 2019; Rizk, 2016). Sulfide oxidation can proceed via multiple pathways, with endpoints varying by environmental conditions and the microbial group involved (Dolfing & Hubert, 2017). Complete oxidation of sulphide to sulphate displays the most thermodynamically favourable pathway under some conditions. However, incomplete oxidation of sulphide to intermediate sulphur compounds, such as elemental sulphur, is also common (Buckingham & Eckert, 2021; Dolfing & Hubert, 2017). Laboratory studies have reported that NR-SOB strains such as *Thiomicrospira sp.* strain CVO are able to oxidise up to 15 mM sulphide while increasing the redox potential of the environment from -400 mV to +100 mV (Greene et al., 2003).

This significant shift in redox potential creates an environment less favourable to SRB activity.

Thermodynamic modelling has significantly enhanced the understanding of which oxidation pathways are favoured in nitrate-treated systems. The coexistence of both NRB and NR-SOB in a nitrate-treated system highlights two distinct metabolic routes: nitrate-driven oxidation of sulphide and nitrate-driven oxidation of organic compounds. To determine the more thermodynamically favourable route, (Dolfing & Hubert, 2017) demonstrated that nitrate-driven oxidation of acetate was more exergonic than nitrate-driven oxidation of sulphide. The corresponding reactions are shown in Equations 4 and 5. The change in Gibbs free energy ($\Delta G°'$), normalised to 2 mol of electrons transferred, was −200.4 kJ per 2 mol e^- for acetate oxidation and −186.1 kJ per 2 mol e^- for sulphide oxidation. Thus, acetate oxidation yielded approximately 14.3 kJ more energy per 2 mol e^- transferred than sulphide oxidation under the specified standard conditions.

$$5acetate + 13H^+ + 8NO_3^- \rightarrow 4N_2 + 10CO_2 + 14H_2O\ ,\ \Delta G^{o'} = -200.4\ kJ/2e^- \quad [4]$$

$$5HS^- + 3H^+ + 8NO_3^- \rightarrow 4N_2 + 5SO_4^{2-} + 4H_2O\ ,\ \Delta G^{o'} = 186.1\ kJ/2e^- \quad [5]$$

Researchers have calculated Gibbs free energies for various sulphide oxidation reactions, showing that sulphide oxidation by NR-SOB under nitrate conditions can lead to either complete oxidation to sulphate or incomplete oxidation to elemental sulphur, depending on the availability of nitrate and other environmental factors (Dolfing & Hubert, 2017; Valdés et al., 2023). This finding has vital implications for corrosion control, as elemental sulphur can be corrosive under certain conditions. Another major mechanism by which nitrate injection controls SRB activity is the competitive exclusion of hNRB for organic electron donors (Gieg et al., 2011; Grigoryan & Voordouw, 2008; Hagar et al., 2022; Londry & Suflita, 1999). The reduction of nitrate is thermodynamically and energetically more favourable than the reduction of sulphate, thereby providing hNRB with higher growth yields and rates when competing for the same substrates (Xie et al., 2025; Youssef et al., 2009). Thermodynamics allows hNRB to outcompete SRB for VFAs, lactate, and other organic compounds originally used to fuel sulphate reduction (Grigoryan et al., 2008).

Laboratory studies using oil field samples from Argentina showed that competitive oxidation of volatile fatty acids by SRB and NRB, in conjunction with nitrate-reducers, has kinetic advantages (Grigoryan et al., 2008). Substrate availability, along with fermentative and oil-degrading bacteria that generate electron donors, influences the dynamics of oxidation competition and sustains both sulphate and nitrate reduction (Gao & Fan, 2023). The formation of nitrite as an intermediate product during nitrate reduction strongly and specifically inhibits SRB, serving as a biological mechanism controlling reservoir souring

(Grigoryan & Voordouw, 2008; Grigoryan et al., 2008; B. Liu et al., 2021). Both hNRB and NR-SOB produce nitrite through partial reduction of nitrate, and this nitrite accumulation plays a critical role in suppressing SRB (Al-Humam et al., 2022; Tao et al., 2025; Zhou et al., 2025). This inhibitory effect of nitrite on SRB activity has been reported in various laboratory and field studies (Mand et al., 2017).

However, the relationship between nitrite production and SRB inhibition is more complex than initially discovered. While nitrate is frequently introduced to suppress reservoir souring, certain SRB possess the ability to reduce nitrite, which is the intermediate produced during nitrate reduction, to ammonium through DNRA (Bourceau et al., 2023; Greene et al., 2003). The nitrite reductase enzyme (Nrf) mediates this process, and the presence of the nrfA gene confers resistance. By further reducing inhibitory nitrite to ammonia, SRB with Nrf activity prevent nitrite accumulation, which would otherwise disrupt their metabolic activity.

Consequently, these specific SRB populations can persist and continue to produce sulphide even in the presence of high nitrate concentrations (Mohanakrishnan et al., 2011). Treatment failure occurs primarily when the microbial community is dominated by or contains significant populations of SRB capable of Nrf-mediated nitrite ammonification. For example, *Desulfovibrio vulgaris* Hildenborough utilises Nrf to respire nitrite, allowing it to bypass the inhibitory effects typically associated with nitrite accumulation (Korte et al., 2015). For other SRB strains that lack Nrf activity, they will have to face a long-lasting inhibition of nitrite. This discovery is important for understanding that the effectiveness of nitrate treatment can vary across oil field environments with different SRB populations.

Treatment efficacy depends heavily on the applied concentration and dosing strategy. High concentrations of nitrite can suppress all microbial activity, while gradual application of nitrite at low concentrations has a less pronounced effect on corrosion (Youssef et al., 2009). This study suggests that periodic and continuous nitrate injection allows moderate nitrite production, which can be more effective than high-dose, irregular injection. Nitrate injection stimulates a shift in microbial community composition, with marine bacterial taxa and autotrophic groups responding rapidly to nitrate availability (Gao & Fan, 2023; P. Liu et al., 2025). Molecular analyses using 16S rRNA gene sequencing, DGGE (denaturing gradient gel electrophoresis), and functional gene quantification have revealed the complexity of microbial community dynamics (Bødtker et al., 2008; Tian et al., 2024).

At the Veslefrikk oil field, bacteria associated with NR-SOB, such as *Sulfurimonas sp.,* formed a major population at the injection wellhead throughout the 7 to 8-year treatment period. However, their presence was less pronounced downstream of the deaerator and finally disappeared after 40 months of nitrate treatment at some locations (Bødtker et al., 2008). This spatial and temporal variation in NR-SOB populations shows the importance of environmental conditions in regulating community structure. Such heterogeneity suggests that localised factors, including oxygen availability, nutrient concentrations, and flow rates,

play a critical role in determining the persistence and distribution of these beneficial consortia. Furthermore, the eventual disappearance of NR-SOB at specific downstream points may indicate a shift in metabolic niches or the completion of sulphide oxidation, highlighting the dynamic nature of microbial succession within nitrate-treated injection systems.

Recent research has shown that nitrate and oxygen significantly alter the abundance and structure of SRB and SOB in petroleum reservoir water (Tian et al., 2024). The highest SRB inhibitory effect and the lowest gene-encoding ratio for Dsr and sulphur oxidation (soxB) were obtained under high concentrations of nitrate and oxygen, suggesting synergistic effects. Commonly detected SRB genera include *Desulfovibrio*, *Desulfobacca*, *Desulfobulbus*, *Sulfuricurvum,* and *Dechloromonas*. During nitrate treatment, the biofilm community tends to remain highly diverse and stable for long periods (Abd Rahman et al., 2022; Alkan et al., 2024; Bødtker et al., 2008). The stability of the biofilm community may contribute to the prolonged effectiveness of nitrate treatment, as multiple microbial groups can be more resilient to environmental disruption. However, the persistence of certain microorganisms across different treatment phases suggests either metabolic flexibility or spatial zonation based on nutrient availability (Gao & Fan, 2023). This ecological robustness is further exemplified by the successful replacement of deleterious SRB populations with beneficial nitrate-reducing consortia, marking a fundamental shift toward a balanced microbial system. Consequently, the established biofilm acts as a stable biological barrier, in which the continuous presence of functional groups such as NR-SOB ensures the long-term suppression of sulphate reduction and sustained mitigation of localised corrosion.

Field Performances and Optimisation of Nitrate Injection

In 1999, among the earliest studies, the Veslefrikk and Gullfaks fields replaced glutaraldehyde biocide treatment with continuous low-concentration nitrate injection, allowing extended monitoring and analysis of treatment outcomes (Bødtker et al., 2008). At the Gullfaks field, inhibition of H_2S by nitrate injection on platforms B (GFB) and C (GFC) was compared with the use of biocide. During nitrate treatment, SRB activity remained low, averaging 0.86 ± 0.32 μg H_2S/cm^2/day (±SE, n = 76) at GFB and 0.84 ± 0.28 μg H_2S/cm^2/day (±SE, n = 59) at GFC. In contrast, SRB activity was substantially higher during biocide treatment, with averages of 4.98 ± 0.51 μg H_2S/cm^2/day (±SE, n = 72) at GFB and 11.73 ± 1.40 μg H_2S/cm^2/day (±SE, n = 60) at GFC. The study outcomes showed stable, long-term SRB inhibition in seawater injection systems after nitrate treatment. Community analysis of the 16S rRNA gene by Polymerase Chain Reaction-Denaturing Gradient Gel Electrophoresis (PCR-DGGE) showed that bacteria belonging to NR-SOB, such as *Sulfurimonas sp.* formed a major population at the injection wellhead throughout the treatment period at Veslefrikk. This provides direct evidence of the establishment and persistence of NR-SOB communities following nitrate injection.

Following nitrate injection, the SRB population was replaced by an equally large NRB population in the biofilm, showing a complete shift in the dominant microbial functional group (Halim et al., 2011; Sunde et al., 2004; Voordouw, 2008). The shift in microbial groups was accompanied by a significant decrease in H_2S production upon the treated water reaching the production wells, confirming the effectiveness of the biological control strategy at the field scale. Figure 4 shows a significant reduction in corrosion rates following nitrate treatment, with corrosion levels in the water injection system decreasing by more than 50%.

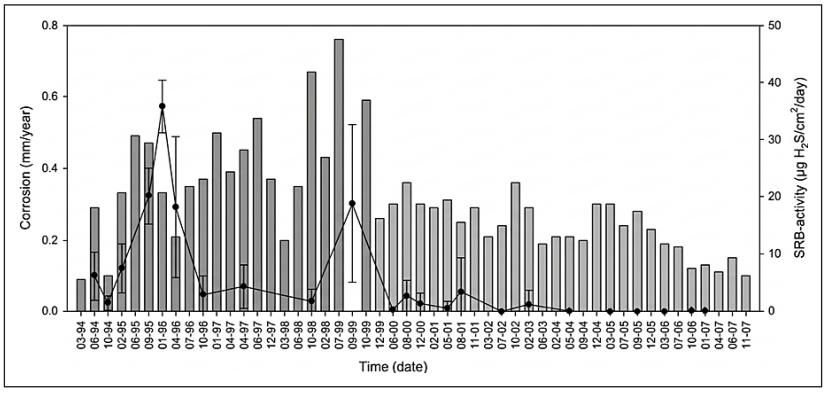


Figure 4. Corrosion rates at GFC during biocide treatment (dark greys bars) and nitrate treatment (light grey bars) (Bødtker et al., 2008)

Quantitative corrosion rate data from field studies showed the significance of these reductions. Another field study also reported a decrement of corrosion rates from 0.7 mm/year to 0.2 mm/year after 4 months of nitrate injection (Eckford & Fedorak, 2004). This result represents a 3.5-fold reduction in corrosion rate, extending infrastructure lifespan while lowering maintenance costs. The sustained nature of these corrosion reductions is particularly notable: SRB activity remained low at ≤0.3 and ≤0.9 µg H_2S/cm^2/day at Veslefrikk and Gullfaks, respectively, over a 7 to 8 year monitoring period (Bødtker et al., 2008).

On a laboratory scale, the addition of *Thiomicrospira* sp. as the NR-SOB into a fully grown culture of SRB consortium at 120-160 hours was reported to show a significant increase in the population towards the end of the experiment (Nemati et al., 2001). The biofilm community during nitrate treatment was highly diverse and stable over time, indicating the successful establishment of a balanced microbial system. This replacement or shift in the microbial community reflects not only reduced total microbial biomass, but also the substitution of one functional group with another that performs beneficial activities rather than damaging ones (Sunde et al., 2004). Both laboratory and bioreactor studies have provided controlled environments for examining NR-SOB mechanisms and optimising nitrate treatment methods. These studies have carried out various experimental designs,

including sand-packed columns, sandstone cores, up-flow packed-bed bioreactors, and oilfield-simulating bioreactors (Xue & Voordouw, 2015; Youssef et al., 2009).

A study by Xue and Voordouw (2015) conducted bioreactor research to simulate real oilfield conditions and investigate the control of microbial sulphide production with nitrate and biocide. This study showed that NR-SOB oxidise sulphide and suppresses SRB activities while linking nitrate dosing to corrosion mitigation. Yang et al. (2021) carried out four lab-scale bioreactors to imitate force mains over 8 months and evaluate the effect of ferric iron and nitrate on hydrogen sulphide control. This study found that the co-existence of sulphide and nitrate stimulated the development of NR-SOB, particularly *Sulfurovum sp.*, which was responsible for sulphide removal in the presence of nitrate. This shows that NR-SOB can be enriched under conditions of simultaneous availability of sulphide and nitrate.

Few laboratory studies have interpreted the specific mechanism by which NR-SOB controls corrosion. Early research has investigated bacterial interactions in detail, testing various NR-SOB strains for sulphide oxidation capacity while measuring changes in environmental redox potential (Greene et al., 2003; Janssen et al., 1998; Ter Heijne et al., 2018). *Thiomicrospira sp.* strain CVO can oxidise up to 15 mM sulfide, while escalating redox potential from -400 to +100 mV and producing 0.6 moles of nitrite for every mole of nitrate reduced.

Few other studies have examined corrosion-rate dynamics associated with nitrite or nitrate-mediated control of souring under biological conditions that simulate an oil reservoir (Lahme et al., 2019; Rempel et al., 2006). Through microbial cultures and electrochemical techniques such as potentiodynamic scans and linear polarisation, the study found that adding nitrate and NR-SOB had similar effects on sulphide oxidation and reduction. However, nitrate and NR-SOB alterations also caused pitting corrosion, with maximum corrosion rates of 0.72 mm/year. This study highlighted a potential complication: nitrate treatment reduces overall corrosion but may also promote localised corrosion during the sulphide oxidation phase. Recent lab-based research has examined the optimum nitrate treatment strategies by investigating the effects of different nitrate concentrations while co-treating with other agents and environmental factors. Tian et al. (2024) studied how nitrate and oxygen levels affect the abundance and structure of sulphate-reducing and sulphur-oxidising bacteria in petroleum reservoir water. The study found the highest SRB inhibitory effect and the lowest dsrB/soxB ratio under high concentrations of both nitrate and oxygen, suggesting that combined treatments may be more effective than nitrate alone.

Recognising that corrosion can occur solely through ionic or electrochemical reactions in seawater or other media, previous studies have established several analytical techniques to distinguish MIC from abiotic corrosion. To isolate the contribution of microbial activity, researchers employ comparative experimental designs that contrast biotic environments with abiotic, sterile controls (Lei et al., 2009; Machuca et al., 2011; Makama et al., 2016).

These controls typically involve filter-sterilised seawater or abiotic artificial seawater to establish a baseline for electrochemical and material degradation. Establishing the contribution of microbial activity requires a multi-faceted approach, as relying on a single diagnostic method can be misleading. Strong evidence for MIC is typically validated by integrating surface analysis and molecular profiling (Ayodele et al., 2025).

The first approach to establishing the contribution of microbial activity is mineralogical analysis. Corrosion products on biofilm surfaces often show mineral compositions that differ significantly from sterile controls; for instance, biofilm carbon steel may show high concentrations of $CaCO_3$ (aragonite) compared with the goethite (α-FeOOH) that dominates sterile surfaces (Makama et al., 2016). Second, molecular profiling can confirm microbial activity. 16S rRNA gene amplification and sequencing enable researchers to identify specific bacterial taxa, such as sulfate- and iron-reducing bacteria, associated with MIC (L. Jones et al., 2024; L. M. Jones et al., 2024; Machuca et al., 2011). Lastly, previous studies have used comparative untargeted metabolomics to reveal disparities between the metabolic products of biofilm-associated cells and planktonic populations, providing insight into the specific biochemical pathways driving corrosion. This is because planktonic cells do not accurately represent the bacterial population within a biofilm. Therefore, assessing MIC risk by sampling the surrounding seawater is insufficient for an accurate diagnosis.

Corrosion Risks Associated with Nitrate Treatment

Generally, nitrate treatment reduces overall corrosion rates, but several studies have reported the potential for localised corrosion and pitting during sulphide oxidation. Rempel et al. (2006) found that amendments of nitrate and NR-SOB led to localised corrosion in the form of pitting, with a maximum corrosion rate of 0.72 mm/year. The corrosion rate accelerates during sulphide oxidation and removal, suggesting that sulphide oxidation may create conditions conducive to corrosion. This condition highlights a significant challenge for field applications, as localised corrosion can cause fatal failures even when average corrosion rates are reduced. The mechanisms that lead to localised corrosion are still not fully understood, but may involve galvanic cells, localised pH changes or the accumulation of corrosive intermediates such as sulphur compounds (Rempel et al., 2006). Field reports show increased corrosion rates in some applications, but each case should be examined, as conditions can vary (Bolaji et al., 2019). The variability in the outcomes shows the need for thorough monitoring and optimisation of nitrate treatment based on site conditions.

Thermodynamic modelling by Dolfing and Hubert (2017) revealed that the oxidation of sulphide by NR-SOB can be incomplete, producing elemental sulphur as an intermediate product. Elemental sulphur in the system can be highly corrosive to oil field infrastructure. This is because elemental sulphur is highly reactive yet has low water solubility, thereby facilitating the formation of corrosive layers, such as iron sulphide (Fe_3S_4), on carbon steel

surfaces (Drønen et al., 2014). Sulphide oxidation may be a problem in topside oil/water separation due to potential NR-SOB-mediated corrosion when nitrate is limited, which favours incomplete oxidation of sulphide to elemental sulphur (Dolfing & Hubert, 2017; Wang et al., 2016; Xu et al., 2025). The formation of a cryptic sulphur cycle, in which sulfide produced by SRB is oxidised back by NR-SOB, can generate corrosive intermediates, posing a critical challenge (Dolfing & Hubert, 2017). Although this cycle reduces H_2S accumulation in the reservoir, the intermediate products might contribute to corrosion under certain conditions. In previous years, both operators in produced water reinjection (PWRI) facilities and researchers have similarly reported an increase in corrosion rate due to mixed microbial communities and the addition of nitrate (MacDonald et al., 1978; Nemati et al., 2001; Rempel et al., 2006)

During the oxidation of sulphide, elemental sulphur, which is the intermediate NR-SOB, has been speculated to be responsible for corrosion under the nitrate treatment (Nemati et al., 2001; Vik et al., 2007) Due to its characteristics as a strong oxidant of iron (MacDonald et al., 1978; Schmitt, 1991). In one corrosion study, metal coupons used after the corrosion test were covered with a layer of black precipitates, identified by X-ray analysis as iron sulphide, iron oxide, iron hydroxide, and elemental sulfur. The thickness of the precipitation was increasing towards the upper region of the bioreactor, where high corrosion rates occurred (Hubert et al., 2005), which correlates with the finding that the penetration of oxygen near the top of the bioreactor may have led to the oxidation of sulphide to corrosive elemental sulphur (Lee et al., 1993). Elemental sulphur in a pipeline system may significantly increase corrosion rates by forming sulfuric acid and H_2S, which inhibit repassivation and promote both crevice corrosion and stress corrosion cracking (Schmitt, 1991).

In addition to elemental sulphur, nitrite formation also causes MIC under nitrate-reducing conditions (Beeder et al., 2007; Drønen et al., 2014; Vik et al., 2007). Nitrite is a key intermediate in nitrate-dependent corrosion; its presence promotes chemical oxidation of iron and can accelerate corrosion rates (Iino et al., 2021). The effect of nitrite accumulation on X80 steel has been found to induce both uniform and pitting corrosion, where the corrosion severity is highly dependent on the nitrite concentration gradient (B. Liu et al., 2021).

In a recent study, the researchers conducted a detailed study of NR-SOB-mediated corrosion by observing the initial nitrate-to-sulphide (N/S) ratios relating to oil field operations, and reported that corrosion rates are up to 0.27 millimetres per year (mm y^{-1}) and up to 60-µm-deep pitting corrosion for only 9 days (Lahme et al., 2019). Various N/S ratios directly affect sulphide oxidation and lead to the accumulation of byproducts, as shown in Figure 5. Figure 5a shows the complete oxidation of sulphide to sulphate, which is happening at N/S ratios of ≥1.6 (Equation 6). As for Figure 5b, it depicts the accumulation of corrosive S0 due to incomplete oxidation at N/S ratios <1.6, especially when nitrate is the limiting factor (Equation 7). However, Figure 5c shows the presence of excess nitrate when the N/S ratio >1.6, leading to the accumulation of nitrite (Equation 8).

Therefore, nitrate availability and the N/S ratio play a crucial role in corrosion rate because they directly affect the accumulation of different NR-SOB metabolites (Lahme et al., 2019). Understanding and managing these intermediate sulphur compounds and nitrite buildup are critical to optimising the effectiveness of nitrate treatment.

$$HS^- + 1.6\ NO_3^- + 0.6\ H^+ \rightarrow SO_4^{2-} + 0.8\ N_2 + 0.8\ H_2O,\ \Delta G^0 = -784.3\ \text{kJ mol}^{-1} \quad [6]$$

$$HS^- + 0.4\ NO_3^- + 1.4\ H^+ \rightarrow S^0 + 0.2\ N_2 + 1.2\ H_2O,\ \Delta G^0 = -197.2\ \text{kJ mol}^{-1} \quad [7]$$

$$HS^- + 4\ NO_3^- \rightarrow SO_4^{2-} + 4\ NO_2^- + H^+,\ \Delta G^0 = -510.4\ \text{kJ mol}^{-1} \quad [8]$$

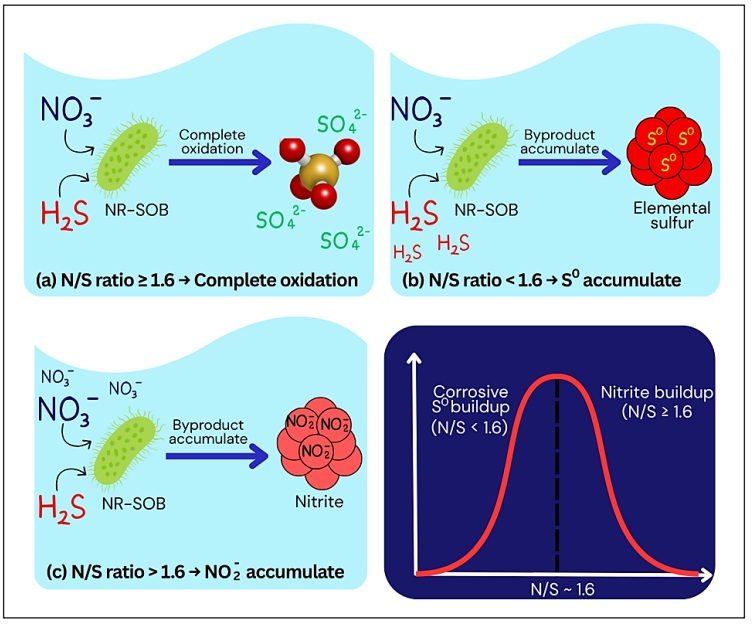


Figure 5. Effect of N/S ratios with (a) ≥ 1.6; (b) < 1.6 and (c) > 1.6 on sulphide oxidation and byproduct accumulation

However, the discovery of some SRB strains with Nrf activity shows a clear limitation to the universal effectiveness of nitrate treatment (Greene et al., 2003). Related SRB strains such as *Desulfovibrio vulgaris* and *Desulfobulbus propionicus* have the ability to reduce nitrite to ammonia, thereby eliminating the inhibitory effect of nitrite produced by both hNRB and NR-SOB (Greene et al., 2003). This ability of certain SRB strains translates into resistance mechanisms against nitrite, indicating that the effectiveness of nitrate treatment will vary with the composition of the indigenous SRB population in the system. In a system

dominated by Nrf-positive SRB strains, nitrate injection may be less effective at suppressing sulphate reduction, and H_2S can continue to accumulate, worsening reservoir souring. In this study, it is critical to identify and characterise indigenous microbial communities before initiating nitrate injection.

Although this review primarily focuses on NR-SOB, studies involving other NRB remain highly relevant as they provide important insights into the broader microbial and corrosion responses associated with nitrate treatment. A recent study investigating mixed cultures of *Desulfovibrio vulgaris* and *Pseudomonas aeruginosa* showed that nitrate addition can unexpectedly enhance corrosion despite suppressing SRB activity (Sun et al., 2022). In the experiment, nitrate addition stimulated NRB growth, particularly *Pseudomonas aeruginosa*, increasing corrosion risk through enhanced biofilm activity and microbially mediated electron transfer. Although nitrate treatment effectively reduced sulphide production by SRB, some findings indicate that it does not guarantee corrosion mitigation, as it may stimulate highly corrosive NRB populations, further increasing the corrosion risk.

Advances in nr-sob Research for Souring Control

Various environmental and operational factors directly affect nitrate treatment and the establishment of NR-SOB in the system. Environmental factors, such as temperature, are vital for optimising nitrate treatment strategies. Nitrate treatment is utilised in oil reservoirs at temperatures ranging from low to approximately 80 °C (Magot, 2005). Temperature dictates the mechanism by which nitrate treatment operates, directly affecting its efficacy in controlling reservoir souring or enhancing oil recovery. At lower temperatures (up to 45 °C), nitrate is reduced to N_2, whereas at 45–65 °C, it is reduced to nitrite (Okpala et al., 2017). Since nitrite is a potent inhibitor of thermophilic SRB, maintaining a nitrate-reducing zone at 45–65 °C can effectively suppress souring in hotter reservoir zones (65-80 °C).

Although research exists on high-temperature reservoirs, thermophilic NR-SOB populations and their functionality at rising temperatures remain less well characterised than mesophilic populations. In addition, concentration and application strategy are critical to treatment success. Continuous low-dose nitrate injection between 0.25 mM and 0.33 mM has proven effective in field applications (Bødtker et al., 2008), but the optimal concentration may vary with sulphate levels, organic carbon availability, temperature, and site-specific factors. In short, nitrate overdose can be wasteful and could promote undesirable microbial activities, while underdose might even fail to establish NR-SOB populations and the effectiveness of this treatment.

Comparing early studies from 1999 to 2008 with recent findings from 2017 to 2024 revealed continuity and evolution in knowledge of NR-SOB-mediated corrosion control, as shown in Figure 6. Figure 6a shows earlier studies between 1999 and 2008 that used nitrate to control sulphide production and SOB, which were first reported in 1999 (Jenneman &

Gevertz, 1999; Londry & Suflita, 1999). Field studies in 2004 and 2008 attained the first long-term evidence for NR-SOB establishment, alongside sustained corrosion control in operational oil fields (Bødtker et al., 2008; Sunde et al., 2004). Early studies reported a significant reduction in SRB populations, and both H_2S generation and corrosion rates decreased, confirming nitrate injection as an alternative to biocide treatment. In 2003, one study identified an SRB resistance mechanism through Nrf activity, raising awareness of the effectiveness of nitrate treatment, which is not universal and depends on indigenous microbial community composition (Greene et al., 2003). Another study in 2008 revealed competitive oxidation of VFAs by SRB and NRB, giving mechanistic insights into the dynamics of the competition (Grigoryan et al., 2008).

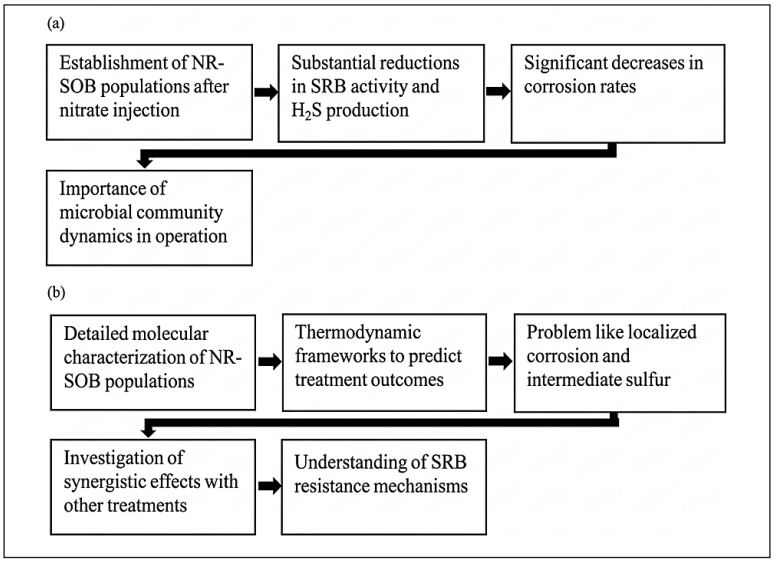


Figure 6. Advancement of findings linked to NR-SOB control corrosion in (a) early studies (1999-2008) and (b) recent studies (2017-2026)

In contrast, Figure 6b depicts recent research conducted on these foundations, using more elaborate molecular, thermodynamic, and systemic approaches. In 2017, thermodynamic modelling was applied to predict the outcomes of nitrate-based interventions, revealing the potential for incomplete sulphide oxidation and production of corrosive intermediate products. This marked a shift from purely observational approaches to predictive modelling based on thermodynamic principles (Dolfing & Hubert, 2017). Tian et al. (2024) used advanced molecular techniques, including 16S rRNA gene sequencing and functional gene quantification, to show that nitrate and oxygen significantly altered the abundance of SRB and SOB rather than their structure. This finding suggests that nitrate treatment works by shifting the abundances of existing functional groups rather than introducing an

entirely new community structure. Current studies also examined the synergistic effects of combined treatments. Yang et al. (2021) compared ferric iron and nitrate effects, while Tian et al. (2024) observed the combined effect of nitrate and oxygen treatments on SRB inhibition. Synergistic effects are still being studied in recent years; for example, one study uses nitrate/nitrite coupling with NRB to suppress SRB while reducing corrosion rates in high-temperature oil reservoirs (Guo et al., 2026; Zhou et al., 2025).

Summary of nr-sob Research in Souring Control (1999-2026)

During the initial phase between 1999 and 2008, research focused on establishing the fundamental efficacy of nitrate injection. These early studies successfully demonstrated that the introduction of nitrate facilitates the establishment of NR-SOB populations, which in turn leads to substantial reductions in SRB activity and biogenic H_2S production. Ultimately, this period provided the first long-term evidence that these microbial shifts translate into significant decreases in overall corrosion rates within operational oil fields, underscoring the vital role of microbial community dynamics in industrial operations. In contrast, the more recent research era spanning 2017 to 2024 reflects a move towards detailed molecular and thermodynamic characterisation of these systems.

Current studies utilise advanced molecular tools to characterise NR-SOB populations and employ thermodynamic frameworks to predict treatment outcomes more accurately. Researchers use high-throughput sequencing and metagenomic analysis to characterise the functional potential and microbial composition of NR-SOB (Cui et al., 2024; Jian et al., 2023; F. Liu et al., 2021; S. Liu et al., 2021; T. Wang et al., 2024). Metagenomics is used to assess the abundance of specific functional genes—such as narG and napA—that are often associated with autotrophic sulphur-oxidising denitrifiers. These tools enable the mapping of metabolic pathways, including complete denitrification, sulphur disproportionation, and DNRA. In addition, modelling efforts have shifted from conventional bulk biomass models towards trait-based process modelling (Sampara et al., 2022). Traditional models often fail to predict ecosystem dynamics in functionally diverse microbiomes because they rely on aggregated concentrations rather than on specific microbial physiological traits. As for current research, trait-based models calibrated at the genus level demonstrate superior accuracy in predicting nitrite accumulation and nitrate removal outcomes.

However, this increased resolution has also revealed new complexities, such as localised corrosion and the corrosive potential of intermediate sulphur species, such as zero-valent sulphur. Modern investigations are now prioritising the understanding of SRB resistance mechanisms, such as Nrf activity, and exploring the synergistic effects of combining nitrate with other treatments to optimise corrosion control in challenging environments. Table 2 presents a summary of methodologies extracted from various studies spanning 1999 to 2026. The most common analytical method is measuring sulphide production, alongside

Table 2
Summary of nitrate injection in controlling souring corrosion extracted from studies between 1999 and 2026

Bacterial Strain	Analytical Method	Corrosion Assessment	Microbial Analysis	Key Parameters	Year	Ref.
SOB isolated from oil field	Sulfide measurement	Field-based monitoring	Culture-based identification and characterisation	Sulfide concentrations, bacterial activity	1999	Jenneman and Gevertz (1999)
SRB from oily waste	Sulfide quantification	Indirect assessment through sulfide control	SRB activity monitoring	Sulfide production rates, nitrate consumption	1999	Londry and Suflita (1999)
Mixture of SRB and NRB	H_2S measurement, nitrate/nitrite quantification, corrosion rate determination	Weight loss and corrosion rate calculation	Most probable number (MPN) enumeration of SRB and NRB	Sulfide concentration, nitrate/nitrite levels, corrosion rates, bacterial populations	2001	Nemati et al. (2001)
Pure cultures of SRB and NR-SOB	Enzyme assays for nitrite reductase, sulfide measurements	Not primary focus	Growth curves, enzyme activity measurements, molecular techniques	Nitrite reductase activity, sulfide production, bacterial growth	2003	Greene et al. (2003)
Indigenous microbe from Gullfaks oil field	Field monitoring of H2S, nitrate, sulfate	Corrosion coupon monitoring, pipeline inspection	Biofilm sampling and analysis	H_2S levels, nitrate concentration, corrosion rates, water chemistry	2004	Sunde et al. (2004)
SRB and NR-SOB	H_2S measurement for SRB activity, 16S rRNA gene-based community analysis by PCR-DGGE	Field-based monitoring	16S rRNA gene-based community analysis by PCR-DGGE	SRB activity (μg H_2S/cm^2/day), corrosion rate, bacterial community composition	2008	Bødtker et al. (2008)
Nitrate-utilising consortium of SRB from Algerian oil field	Weight loss measurements, surface analysis, sulfide measurements	Corrosion rate calculations, SEM analysis of corroded surfaces	16S rRNA gene sequencing, culture-based identification	Corrosion rates, sulfide production, bacterial growth, nitrate consumption	2012	Kebbouche-Gana and Gana (2012)
Thermodynamic modeling of SRB and nitrate-reducing bacteria	Thermodynamic calculations, Gibbs free energy analysis	Indirect through souring control prediction	Modeling of microbial metabolism	Thermodynamic parameters, substrate availability, product formation	2017	Dolfing and Hubert (2017)
Mixed microbial communities from wastewater	H_2S measurements, ion chromatography, molecular analysis	Not primary focus	16S rRNA gene sequencing	H_2S, nitrate, ferric iron, sulfate, microbial community	2021	Yang et al. (2021)

Table 2 (*continued*)

Bacterial Strain	Analytical Method	Corrosion Assessment	Microbial Analysis	Key Parameters	Year	Ref.
Mixed microbial communities for simultaneous sulfide and nitrate removal	Sulfide and nitrate measurements, high-throughput sequencing	Not primary focus	16S rRNA gene sequencing, functional prediction	Sulfide removal, nitrate removal, microbial community structure	2024	Guo et al. (2024)
Co-culture of SRB and nitrate-reducing bacteria	Electrochemical measurements, surface analysis, metabolite analysis	Weight loss, electrochemical impedance spectroscopy, SEM, XRD	qPCR, biofilm characterisation	Corrosion rates, bacterial populations, metabolites, deposit characteristics	2024	Sun et al. (2024)
Sulfate-reducing bacteria (SRB) in seawater	Microbial fuel cell voltage monitoring; cyclic	Corrosion-rate comparison, mass loss, surface morphology, and sulfur/iron-sulfide deposits	Antibiotic susceptibility test (ABST) for SRB inhibition	ZnO nanoparticles; SRB increased corrosion rate by 21.3% versus control; ZnO reduced corrosion rate by 21.7% relative to control	2025	Ambepitiya et al. (2025)
Sulfate-reducing bacteria (SRB) with denitrifying bacteria (DNB)	Electrochemical impedance spectroscopy (EIS), linear polarisation resistance (LPR), weight loss, SEM, and maximum pit depth	Average corrosion rate, maximum pit depth, EIS capacitive arc, and polarisation resistance	16S rDNA sequencing	Combined glutaraldehyde and DNB treatment; corrosion rate reduced 44.2% to 0.01608 mm/a; maximum pit depth reduced 84.3% to 1.53 μm; SRB abundance decreased from 62.7% to 11.9% over 14 days	2026	Guo et al. (2026)

quantifying nitrate and nitrite over the experimental period.

Discussions and Future Directions

This review synthesises evidence from early field applications to evaluate the presence, biological mechanisms, and overall effectiveness of NR-SOB in mitigating corrosion following nitrate injection. Although nitrate injection has consistently reduced corrosion rates in both laboratory and field studies, reports of localised pitting during sulphide oxidation highlight a knowledge gap in mechanistic understanding (Lahme et al., 2019; Rempel et al., 2006). The formation of elemental sulphur from the incomplete oxidation of sulphide poses a technical challenge in the operation. Figure 7 summarises the current research gap and the areas for future research, including quantifying the production of corrosive intermediates and accounting for corrosivity as a function of temperature, flow rate, and nitrate concentration. Strategies to promote complete oxidation to sulphate should also be developed to fully implement corrosion control through this nitrate treatment in the oil field system.

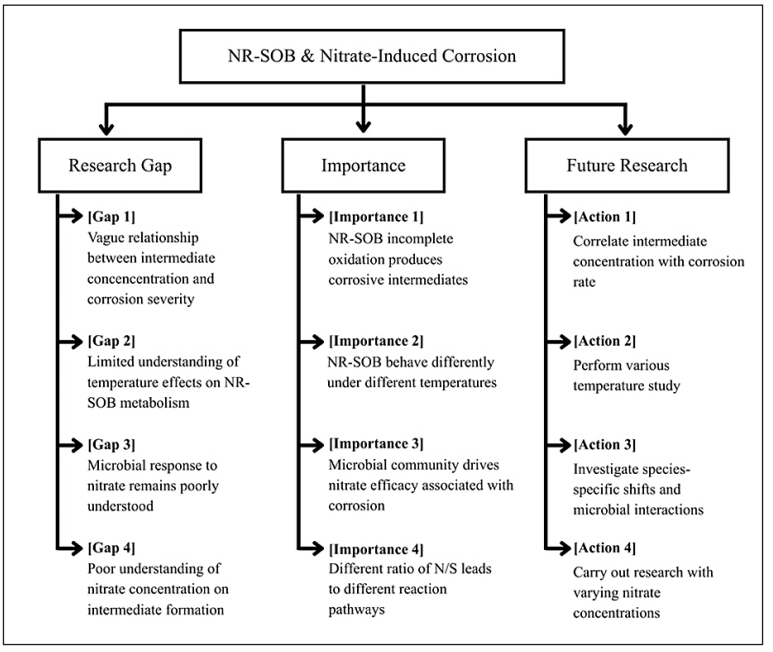


Figure 7. Overview of current knowledge gaps and future research directions in understanding NR-SOB mediated corrosion

Related studies demonstrate that corrosion outcomes under nitrate injection are strongly condition-dependent, rather than driven solely by NR-SOB activity. Negative effects are always linked to incomplete sulphide oxidation and the accumulation of

corrosive intermediate products. The effectiveness of nitrate treatment is also influenced by the composition of the indigenous microbial community, specifically the presence of Nrf-positive SRB that can reduce nitrite, thereby mitigating nitrite accumulation. These interactions highlight that nitrate treatment does not operate independently; instead, it is influenced by complex microbial dynamics. Therefore, rapid and cost-effective molecular diagnostics for pre-treatment analysis will aid in predicting system response and further reducing the risk of nitrate treatment failure. Integrating field datasets with thermodynamic models can help optimise treatment strategies, reduce analysis costs, and provide early warning systems before applying them in the oilfield. Further research on biofilm-focused strategies might help NR-SOB establish populations at key interfaces, enhancing its ability to control souring and limit corrosion rates.

CONCLUSION

This review article synthesises evidence from highly relevant studies spanning from 1999 to 2026, documenting the presence, mechanisms, and effectiveness of NR-SOB in controlling corrosion rates after nitrate injection in oil and gas systems. Findings from the articles indicate that nitrate injection stimulates the growth and establishment of NR-SOB populations, including genera such as *Sulfurimonas, Thiomicrospira, Sulfurovum, Sulfuricurvum, Thiobacillus denitrificans,* and *Arcobacter sp.*, which play important roles in sulphide oxidation and corrosion control. However, the literature indicates that the corrosion outcomes of nitrate injection depend heavily on the conditions and specific operational risks. The documented challenges include incomplete sulphide oxidation, nitrite accumulation, localised corrosion, SRB ability to counter nitrite inhibition, and variability in microbial community responses, which can substantially alter treatment outcomes. Ultimately, while nitrate-stimulated NR-SOB populations provide a viable biological control for reservoir souring, optimising this treatment requires site-specific evaluations of microbial composition, environmental conditions, and thermodynamic factors to ensure complete sulphide oxidation and prevent localised corrosion.

ACKNOWLEDGEMENT

The authors wish to thank the National University of Malaysia and the Centre for Research and Instrumentation (CRIM) for financial support through the Research University Grant (GUP-2024-104).

REFERENCES

Abd Rahman, H., Sedaralit, M. F., Zainal, S., & de Rezende, J. R. (2022). Modelling reservoir souring mitigation strategy based on dynamic microorganisms interactions. In *Abu Dhabi International Petroleum Exhibition and Conference* (Paper SPE-211359-MS). Society of Petroleum Engineers. https://doi.

org/10.2118/211359-ms

AlAbbas, F. M., Kakpovbia, A., Olson, D. L., Mishra, B., & Spear, J. R. (2012). The role of bacterial attachment to metal substrate and its effects on microbiologically influenced corrosion (MIC) in transporting hydrocarbon pipelines. In R. Narayan, S. Bose, & A. Bandyopadhyay (Eds.), *Biomaterials science: Processing, properties and applications II* (pp. 131-144). Wiley. https://doi.org/10.1002/9781118511466.ch14

Alamri, A. H. (2020). Localized corrosion and mitigation approach of steel materials used in oil and gas pipelines – an overview. *Engineering Failure Analysis, 116*, Article 104735. https://doi.org/10.1016/j.engfailanal.2020.104735

Al-Humam, A., Khanfar, H., Al-Ahmari, M., & Al-Sudairi, N. (2022). Nitrate application in mitigation of SRB activity and biogenic sulfide production. *Journal of Materials Science and Engineering B, 12*(4-6), 58-70. https://doi.org/10.17265/2161-6221/2022.4-6.003

Alkan, H., Kögler, F., Namazova, G., Hatscher, S., Jelinek, W., & Amro, M. (2024). Assessment of the biogenic souring in oil reservoirs under secondary and tertiary oil recovery. *Energies, 17*(11), Article 2681. https://doi.org/10.3390/en17112681

Al-Tamimi, W. H., & Mehdi, K. H. (2017). Inhibition of biogenic hydrogen sulfide produce by sulfate reducing bacteria isolated from oil fields in Basra by nitrate based treatment. *Journal of Petroleum Research and Studies, 7*(3), 88-106. https://doi.org/10.52716/jprs.v7i3.163

Ambepitiya, H., Rathnayaka, S., Perera, Y., Jayathilake, C., Ferdinandez, H., Herath, A., Sanjula, U., Rathnayake, A., Basnayaka, C., & Fernando, E. (2025). Mitigating microbiologically influenced corrosion of iron caused by sulphate-reducing bacteria using ZnO nanoparticles. *Processes, 13*(10), Article 3239. https://doi.org/10.3390/pr13103239

Ayodele, A., Olusesan, O. O., Ojo, T. J., Yoonus, A. O., & Olarewaju, V. O. (2025). *Microbiologically influenced corrosion of galvanised steel by sulphate-reducing bacteria in marine environments* [Preprint]. SSRN. https://doi.org/10.2139/ssrn.5547798

Azhar, L. Y., Jarni, H. H., Alauddin, S. M., Othman, N. K., & Yaakob, N. (2025). Hydrogen sulfide concentration influence toward uniform and localized corrosion rate of API 5L X65 steel at elevated temperature. *Corrosion, 81*(1), 96-106. https://doi.org/10.5006/4645

Azhar, L. Y., Ramlan, D. G., Shahrizal, N. A. A., & Yaakob, N. (2026). Transitional corrosion behavior of API 5L X65 steel under marginally sour environment. *Journal of Materials Engineering and Performance, 35*(21), 21551-21567. https://doi.org/10.1007/s11665-026-13151-3

Beeder, J., Andersen, T. R., Liengen, T., Drønen, K., & Torsvik, T. (2007). Corrosion as a side effect during nitrate treatment of produced water and aquifer water injection. In *CORROSION 2007* (Paper NACE-07512). NACE International. https://doi.org/10.5006/c2007-07512

Bhagwat, G., O'Connor, W., Grainge, I., & Palanisami, T. (2021). Understanding the fundamental basis for biofilm formation on plastic surfaces: Role of conditioning films. *Frontiers in Microbiology, 12*, Article 687118. https://doi.org/10.3389/fmicb.2021.687118

Bødtker, G., Thorstenson, T., Lillebø, B.-L. P., Thorbjørnsen, B. E., Ulvøen, R. H., Sunde, E., & Torsvik, T. (2008). The effect of long-term nitrate treatment on SRB activity, corrosion rate and bacterial community

composition in offshore water injection systems. *Journal of Industrial Microbiology & Biotechnology, 35*(12), 1625-1636. https://doi.org/10.1007/s10295-008-0406-x

Bolaji, T. A., Oti, M. N., Abu, G. O., & Onyekonwu, M. O. (2019). Biogenic sulfide production in offshore petroleum reservoirs undergoing waterflooding. *Petroleum and Coal, 61*(5), 1144-1157. https://www.vurup.sk/wp-content/uploads/2019/09/PC-X-2019_Bolaji_103_rev1.pdf

Bourceau, O. M., Ferdelman, T., Lavik, G., Mussmann, M., Kuypers, M. M. M., & Marchant, H. K. (2023). Simultaneous sulfate and nitrate reduction in coastal sediments. *ISME Communications, 3*(1), Article 17. https://doi.org/10.1038/s43705-023-00222-y

Buckingham, K. M., & Eckert, R. B. (2021). Elemental sulfur corrosion: A case study. In R. B. Eckert & T. L. Skovhus (Eds.), *Failure analysis of microbiologically influenced corrosion* (pp. 211-223). CRC Press. https://doi.org/10.1201/9780429355479-12

Callbeck, C. M., Dong, X., Chatterjee, I., Agrawal, A., Caffrey, S. M., Sensen, C. W., & Voordouw, G. (2011). Microbial community succession in a bioreactor modeling a souring low-temperature oil reservoir subjected to nitrate injection. *Applied Microbiology and Biotechnology, 91*(3), 799-810. https://doi.org/10.1007/s00253-011-3287-2

Cui, P., Wan, N., Li, C., Zou, L., Ma, M., Du, J., & Jiang, Y. (2024). Comparative analysis of sulfur-driven autotrophic denitrification for pilot-scale application: Pollutant removal performance and metagenomic function. *Bioresource Technology, 413*, Article 131433. https://doi.org/10.1016/j.biortech.2024.131433

Dall'Agnol, L. T., & Moura, J. J. G. (2014). Sulphate-reducing bacteria (SRB) and biocorrosion. In T. Liengen, D. Féron, R. Basséguy, & I. B. Beech (Eds.), *Understanding biocorrosion* (pp. 77-106). Woodhead Publishing. https://doi.org/10.1533/9781782421252.1.77

Davidova, I., Hicks, M. S., Fedorak, P. M., & Suflita, J. M. (2001). The influence of nitrate on microbial processes in oil industry production waters. *Journal of Industrial Microbiology and Biotechnology, 27*(2), 80-86. https://doi.org/10.1038/sj.jim.7000166

Deng, S., Wang, B., Su, S., Sun, S., She, Y., & Zhang, F. (2022). Dynamics of microbial community and removal of hydrogen sulfide (H2S) using a bio-inhibitor and its application under the oil reservoir condition. *Energy & Fuels, 36*(23), 14128-14135. https://doi.org/10.1021/acs.energyfuels.2c02912

Deng, S., Wang, B., Zhang, C., Wang, C., Sun, S., You, Q., Huo, Q., She, Y., & Zhang, F. (2025). Microbiologically influenced corrosion and community dynamics of oil-well produced water at different temperatures. *Physics of Fluids, 37*(1), Article 011904. https://doi.org/10.1063/5.0245135

Dolfing, J., & Hubert, C. R. J. (2017). Using thermodynamics to predict the outcomes of nitrate-based oil reservoir souring control interventions. *Frontiers in Microbiology, 8*, Article 2575. https://doi.org/10.3389/fmicb.2017.02575

Drønen, K., Roalkvam, I., Beeder, J., Torsvik, T., Steen, I. H., Skauge, A., & Liengen, T. (2014). Modeling of heavy nitrate corrosion in anaerobe aquifer injection water biofilm: A case study in a flow rig. *Environmental Science & Technology, 48*(15), 8627-8635. https://doi.org/10.1021/es500839u

Eckford, R. E., & Fedorak, P. M. (2004). Using nitrate to control microbially-produced hydrogen sulfide in oil field waters. In R. Vazquez-Duhalt & R. Quintero-Ramirez (Eds.), *Petroleum biotechnology: Developments and perspectives* (pp. 307-340). Elsevier. https://doi.org/10.1016/s0167-2991(04)80152-6

Fida, T. T., Voordouw, J., Ataeian, M., Kleiner, M., Okpala, G., Mand, J., & Voordouw, G. (2018). Synergy of sodium nitroprusside and nitrate in inhibiting the activity of sulfate reducing bacteria in oil-containing bioreactors. *Frontiers in Microbiology, 9*, Article 981. https://doi.org/10.3389/fmicb.2018.00981

Fu, Q., Xu, J., Wei, B., Qin, Q., Bai, Y., Yu, C., & Sun, C. (2022). Mechanistic diversity between nitrate and nitrite on biocorrosion of X80 pipeline steel caused by *Desulfovibrio desulfuricans* and *Pseudomonas stutzeri*. *Corrosion Science, 207*, Article 110573. https://doi.org/10.1016/j.corsci.2022.110573

Gao, P., & Fan, K. (2023). Sulfur-oxidizing bacteria (SOB) and sulfate-reducing bacteria (SRB) in oil reservoir and biological control of SRB: A review. *Archives of Microbiology, 205*(5), Article 162. https://doi.org/10.1007/s00203-023-03520-0

Gieg, L. M., Jack, T. R., & Foght, J. M. (2011). Biological souring and mitigation in oil reservoirs. *Applied Microbiology and Biotechnology, 92*(2), 263-282. https://doi.org/10.1007/s00253-011-3542-6

Gittel, A. (2010). Problems caused by microbes and treatment strategies monitoring and preventing reservoir souring using molecular microbiological methods (MMM). In C. Whitby & T. L. Skovhus (Eds.), *Applied microbiology and molecular biology in oilfield systems: Proceedings from the International Symposium on Applied Microbiology and Molecular Biology in Oil Systems (ISMOS-2), 2009* (pp. 103-107). Springer. https://doi.org/10.1007/978-90-481-9252-6_12

Gittel, A., Sørensen, K. B., Skovhus, T. L., Ingvorsen, K., & Schramm, A. (2009). Prokaryotic community structure and sulfate reducer activity in water from high-temperature oil reservoirs with and without nitrate treatment. *Applied and Environmental Microbiology, 75*(22), 7086-7096. https://doi.org/10.1128/aem.01123-09

Greene, E. A., Hubert, C., Nemati, M., Jenneman, G. E., & Voordouw, G. (2003). Nitrite reductase activity of sulphate-reducing bacteria prevents their inhibition by nitrate-reducing, sulphide-oxidizing bacteria. *Environmental Microbiology, 5*(7), 607-617. https://doi.org/10.1046/j.1462-2920.2003.00446.x

Grigoryan, A. A., Cornish, S. L., Buziak, B., Lin, S., Cavallaro, A., Arensdorf, J. J., & Voordouw, G. (2008). Competitive oxidation of volatile fatty acids by sulfate- and nitrate-reducing bacteria from an oil field in Argentina. *Applied and Environmental Microbiology, 74*(14), 4324-4335. https://doi.org/10.1128/aem.00419-08

Grigoryan, A., & Voordouw, G. (2008). Microbiology to help solve our energy needs: Methanogenesis from oil and the impact of nitrate on the oil-field sulfur cycle. *Annals of the New York Academy of Sciences, 1125*(1), 345-352. https://doi.org/10.1196/annals.1419.004

Guo, C., Shao, L., Zhao, W., Han, Y., & Zhang, X. (2024). Removal capacities and varying characteristics of substrates and microbial community structure in simultaneous sulfide and nitrate biological removal process. *Chemosphere, 352*, Article 141518. https://doi.org/10.1016/j.chemosphere.2024.141518

Guo, Y., Wen, C., Duan, M., & Lan, G. (2026). A combined glutaraldehyde and denitrifying bacteria strategy for enhanced control of SRB-induced corrosion in shale gas infrastructure. *Processes, 14*(2), Article 334. https://doi.org/10.3390/pr14020334

Hagar, H. S., Foroozesh, J., Kumar, S., Zivar, D., Banan, N., & Dzulkarnain, I. (2022). Microbial H2S generation in hydrocarbon reservoirs: Analysis of mechanisms and recent remediation technologies. *Journal of Natural Gas Science and Engineering, 106*, Article 104729. https://doi.org/10.1016/j.jngse.2022.104729

Halim, A., Gubner, R., & Watkin, E. (2011). Preliminary study on nitrate injection to control souring problem in oil reservoir: Benefits and side effects on steel material (UNS S31603). In *CORROSION 2011* (Paper NACE-11229). NACE International. https://doi.org/10.5006/c2011-11229

Herath, A. (2019). *The influence of nitrate on the short-term corrosion of steel catalyzed by marine microorganisms* [Doctoral dissertation, University of Oklahoma].

Hubert, C., Nemati, M., Jenneman, G., & Voordouw, G. (2005). Corrosion risk associated with microbial souring control using nitrate or nitrite. *Applied Microbiology and Biotechnology, 68*(2), 272-282. https://doi.org/10.1007/s00253-005-1897-2

Hubert, C., & Voordouw, G. (2007). Oil field souring control by nitrate-reducing *Sulfurospirillum* spp. that outcompete sulfate-reducing bacteria for organic electron donors. *Applied and Environmental Microbiology, 73*(8), 2644-2652. https://doi.org/10.1128/aem.02332-06

Ihsan, Y. N., Fellatami, K., Permana, R., Mingguo, J., & Pribadi, T. D. K. (2022). The optimization of sulphide oxidizing bacteria (SOB) for oil corrosivity reduction at Indramayu coast, the northern coastal area of West Java. *Jurnal Ilmiah Perikanan dan Kelautan, 14*(2), 360-368. https://doi.org/10.20473/jipk.v14i2.33462

Iino, T., Shono, N., Ito, K., Nakamura, R., Sueoka, K., Harayama, S., & Ohkuma, M. (2021). Nitrite as a causal factor for nitrate-dependent anaerobic corrosion of metallic iron induced by *Prolixibacter* strains. *MicrobiologyOpen, 10*(4), Article e1225. https://doi.org/10.1002/mbo3.1225

Jahanbani Veshareh, M., & Nick, H. M. (2019). A sulfur and nitrogen cycle informed model to simulate nitrate treatment of reservoir souring. *Scientific Reports, 9*(1), Article 7546. https://doi.org/10.1038/s41598-019-44033-5

Janssen, A. J. H., Meijer, S., Bontsema, J., & Lettinga, G. (1998). Application of the redox potential for controling a sulfide oxidizing bioreactor. *Biotechnology and Bioengineering, 60*(2), 147-155. https://doi.org/10.1002/(sici)1097-0290(19981020)60:2<147::aid-bit2>3.0.co;2-n

Javaherdashti, R. (2008). *Microbiologically influenced corrosion: An engineering insight*. Springer. https://doi.org/10.1007/978-1-84800-074-2

Javaherdashti, R., & Singh Raman, R. K. (2001). Microbiologically-influenced corrosion of stainless steels in marine environments: A materials engineering approach. In E. V. Pereloma & K. Raviprasad (Eds.), *Engineering materials 2001* (pp. 299-304). Institute of Materials Engineering Australasia.

Jenneman, G. E., & Gevertz, D. (1999). Identification, characterization and application of sulfide-oxidizing bacteria in oil fields. In C. R. Bell, M. Brylinsky, & P. Johnson-Green (Eds.), *Microbial biosystems: New frontiers*. Atlantic Canada Society for Microbial Ecology. https://plato.acadiau.ca/isme/Symposium12/jenneman.PDF

Jia, R., Yang, D., Xu, D., & Gu, T. (2017). Electron transfer mediators accelerated the microbiologically influence corrosion against carbon steel by nitrate reducing *Pseudomonas aeruginosa* biofilm. *Bioelectrochemistry, 118*, 38-46. https://doi.org/10.1016/j.bioelechem.2017.06.013

Jian, C., Hao, Y., Liu, R., Qi, X., Chen, M., & Liu, N. (2023). Mixotrophic denitrification process driven by lime sulfur and butanediol: Denitrification performance and metagenomic analysis. *Science of The Total Environment, 903*, Article 166654. https://doi.org/10.1016/j.scitotenv.2023.166654

Jones, L., Salta, M., Skovhus, T. L., Thomas, K., Illson, T., Wharton, J., & Webb, J. (2024). Dual anaerobic reactor model to study biofilm and microbiologically influenced corrosion interactions on carbon steel. *npj Materials Degradation, 8*(1), Article 125. https://doi.org/10.1038/s41529-024-00542-x

Kebbouche-Gana, S., & Gana, M. L. (2012). Biocorrosion of carbon steel by a nitrate-utilizing consortium of sulfate-reducing bacteria obtained from an Algerian oil field. *Annals of Microbiology, 62*(1), 203-210. https://doi.org/10.1007/s13213-011-0247-0

Khan, M. A. A., Hussain, M., & Djavanroodi, F. (2021). Microbiologically influenced corrosion in oil and gas industries: A review. *International Journal of Corrosion and Scale Inhibition, 10*(1), 80-106. https://doi.org/10.17675/2305-6894-2021-10-1-5

Kiss, H., Lang, E., Lapidus, A., Copeland, A., Nolan, M., Glavina Del Rio, T., Chen, F., Lucas, S., Tice, H., Cheng, J.-F., Han, C., Goodwin, L., Pitluck, S., Liolios, K., Pati, A., Ivanova, N., Mavromatis, K., Chen, A., Palaniappan, K., … Klenk, H.-P. (2010). Complete genome sequence of *Denitrovibrio acetiphilus* type strain (N2460T). *Standards in Genomic Sciences, 2*(3), 270-279. https://doi.org/10.4056/sigs.892105

Korte, H. L., Saini, A., Trotter, V. V., Butland, G. P., Arkin, A. P., & Wall, J. D. (2015). Independence of nitrate and nitrite inhibition of *Desulfovibrio vulgaris* Hildenborough and use of nitrite as a substrate for growth. *Environmental Science & Technology, 49*(2), 924-931. https://doi.org/10.1021/es504484m

Kumaraswamy, R., Ebert, S., Gray, M. R., Fedorak, P. M., & Foght, J. M. (2011). Molecular- and cultivation-based analyses of microbial communities in oil field water and in microcosms amended with nitrate to control H2S production. *Applied Microbiology and Biotechnology, 89*(6), 2027-2038. https://doi.org/10.1007/s00253-010-2974-8

Lahme, S., Enning, D., Callbeck, C. M., Menendez Vega, D., Curtis, T. P., Head, I. M., & Hubert, C. R. J. (2019). Metabolites of an oil field sulfide-oxidizing, nitrate-reducing *Sulfurimonas* sp. cause severe corrosion. *Applied and Environmental Microbiology, 85*(3), Article e01891-18. https://doi.org/10.1128/aem.01891-18

Lee, W., Lewandowski, Z., Okabe, S., Characklis, W. G., & Avci, R. (1993). Corrosion of mild steel underneath aerobic biofilms containing sulfate-reducing bacteria part I: At low dissolved oxygen concentration. *Biofouling, 7*(3), 197-216. https://doi.org/10.1080/08927019309386254

Lei, Y. H., Yin, Y. S., Liu, C. H., Chang, X. T., Chen, Y., & Cheng, S. (2009). Electrochemical corrosion behavior of Cu-15Ni-8Sn alloy in marine microbial medium. *Advanced Materials Research, 79-82*, 1099-1102. https://doi.org/10.4028/www.scientific.net/amr.79-82.1099

Liu, B., Fan, E., Jia, J., Du, C., Liu, Z., & Li, X. (2021). Corrosion mechanism of nitrate reducing bacteria on X80 steel correlated to its intermediate metabolite nitrite. *Construction and Building Materials, 303*, Article 124454. https://doi.org/10.1016/j.conbuildmat.2021.124454

Liu, F., Wang, S., Zhang, X., Qian, F., Wang, Y., & Yin, Y. (2021). Nitrate removal from actual wastewater by coupling sulfur-based autotrophic and heterotrophic denitrification under different influent concentrations. *Water, 13*(20), Article 2913. https://doi.org/10.3390/w13202913

Liu, P., He, J., Lian, H., Fan, Y., Wang, F., & Xu, D. (2026). Microbiologically influenced corrosion of 316L stainless steel by marine *Sunxiuqinia elliptica* via extracellular electron transfer. *Corrosion Communications, 24*, 117-128. https://doi.org/10.1016/j.corcom.2025.11.002

Liu, S., Chen, Y., & Xiao, L. (2021). Metagenomic insights into mixotrophic denitrification facilitated nitrogen removal in a full-scale A2/O wastewater treatment plant. *PLOS ONE, 16*(4), Article e0250283. https://doi.org/10.1371/journal.pone.0250283

Liu, W., Zhang, Y., Dong, Y., & Min, J. (2025). Revisiting corrosion models: The influence of nitrate and nitrate-reducing bacteria on the corrosion of steel in marine and transitional environments [Retracted article]. *Regional Studies in Marine Science, 92*, Article 104617. https://doi.org/10.1016/j.rsma.2025.104617

Londry, K. L., & Suflita, J. M. (1999). Use of nitrate to control sulfide generation by sulfate-reducing bacteria associated with oily waste. *Journal of Industrial Microbiology and Biotechnology, 22*(6), 582-589. https://doi.org/10.1038/sj.jim.2900668

MacDonald, D. D., Roberts, B., & Hyne, J. B. (1978). The corrosion of carbon steel by wet elemental sulphur. *Corrosion Science, 18*(5), 411-425. https://doi.org/10.1016/s0010-938x(78)80037-7

Machuca, L., Bailey, S., Gubner, R., Watkin, E., Ginige, M., & Kaksonen, A. (2011). *Bacterial community structure in natural marine biofilms and the corrosion of carbon steel in seawater* [Conference paper]. 18th International Corrosion Congress, Perth, Australia. https://publications.csiro.au/publications/publication/PIcsiro:EP112931

Magot, M. (2005). Indigenous microbial communities in oil fields. In B. Ollivier & M. Magot (Eds.), *Petroleum microbiology* (pp. 21-33). ASM Press. https://doi.org/10.1128/9781555817589.ch2

Makama, Z., Celikkol-Aydin, S., Aydin, E., Bonifay, V., Sunner, J., & Beech, I. (2016). The effect of a complex planktonic and biofilm bacterial consortia on marine corrosion of 1020 carbon steel [Conference abstract]. *ECS Meeting Abstracts, MA2016-01*(15), 934. https://doi.org/10.1149/ma2016-01/15/934

Mand, J., Shen, Y., Hoffmann, H., & Voordouw, G. (2017). MIC under conditions of oxygen or nitrate ingress. In T. L. Skovhus, D. Enning, & J. S. Lee (Eds.), *Microbiologically influenced corrosion in the upstream oil and gas industry* (pp. 309-323). CRC Press. https://doi.org/10.1201/9781315157818-15

Miller, R. B., Lawson, K., Sadek, A., Monty, C. N., & Senko, J. M. (2018). Uniform and pitting corrosion of carbon steel by *Shewanella oneidensis* MR-1 under nitrate-reducing conditions. *Applied and Environmental Microbiology, 84*(12), Article e00790-18. https://doi.org/10.1128/aem.00790-18

Mohammadi, M. H., Kulakhmetovna, Y. A., & Joia, R. (2024). An overview of oil recovery techniques: From primary to enhanced oil recovery methods. *Journal for Research in Applied Sciences and Biotechnology, 3*(1), 291-301. https://doi.org/10.55544/jrasb.3.1.48

Mohanakrishnan, J., Kofoed, M. V. W., Barr, J., Yuan, Z., Schramm, A., & Meyer, R. L. (2011). Dynamic microbial response of sulfidogenic wastewater biofilm to nitrate. *Applied Microbiology and Biotechnology, 91*(6), 1647-1657. https://doi.org/10.1007/s00253-011-3330-3

Nemati, M., Jenneman, G. E., & Voordouw, G. (2001). Impact of nitrate-mediated microbial control of souring in oil reservoirs on the extent of corrosion. *Biotechnology Progress, 17*(5), 852-859. https://doi.org/10.1021/bp010084v

Nemati, M., & Voordouw, G. (2000). Identification and characterization of sulfate-reducing bacteria involved in microbially influenced corrosion in oil fields. In *CORROSION 2000* (Paper NACE-00126). NACE International. https://doi.org/10.5006/c2000-00126

Nicoletti, D., Sharma, M., & Gieg, L. M. (2022). Assessing microbial corrosion risk on offshore crude oil production topsides under conditions of nitrate and nitrite treatment for souring. *Microorganisms, 10*(5), Article 932. https://doi.org/10.3390/microorganisms10050932

Okpala, G. N., Chen, C., Fida, T., & Voordouw, G. (2017). Effect of thermophilic nitrate reduction on sulfide production in high temperature oil reservoir samples. *Frontiers in Microbiology, 8*, Article 1573. https://doi.org/10.3389/fmicb.2017.01573

Procópio, L. (2022). Microbially induced corrosion impacts on the oil industry. *Archives of Microbiology, 204*(2), Article 138. https://doi.org/10.1007/s00203-022-02755-7

Reinsel, M. A., Sears, J. T., Stewart, P. S., & McInerney, M. J. (1996). Control of microbial souring by nitrate, nitrite or glutaraldehyde injection in a sandstone column. *Journal of Industrial Microbiology, 17*(2), 128-136. https://doi.org/10.1007/bf01570056

Rempel, C. L., Evitts, R. W., & Nemati, M. (2006). Dynamics of corrosion rates associated with nitrite or nitrate mediated control of souring under biological conditions simulating an oil reservoir. *Journal of Industrial Microbiology & Biotechnology, 33*(10), 878-886. https://doi.org/10.1007/s10295-006-0142-z

Rizk, T. Y. (2016). Effect of nitrate treatment on corrosion, and guidelines for its implementation. *Materials Performance, 55*(2), 36-40. https://doi.org/10.5006/mp2016_55_2-36

Salgar-Chaparro, S. J., Tarazona, J., & Machuca, L. L. (2022). Corrosion of carbon steel by *Shewanella chilikensis* DC57 under thiosulphate and nitrate reducing conditions. *Frontiers in Bioengineering and Biotechnology, 10*, Article 825776. https://doi.org/10.3389/fbioe.2022.825776

Sampara, P., Luo, Y., Lin, X., & Ziels, R. M. (2022). Integrating genome-resolved metagenomics with trait-based process modeling to determine biokinetics of distinct nitrifying communities within activated sludge. *Environmental Science & Technology, 56*(16), 11670-11682. https://doi.org/10.1021/acs.est.2c02081

Schmitt, G. (1991). Effect of elemental sulfur on corrosion in sour gas systems. *Corrosion, 47*(4), 285-308. https://doi.org/10.5006/1.3585257

Street, S. R., Xu, W., Amri, M., Guo, L., Glanvill, S. J. M., Quinn, P. D., Mosselmans, J. F. W., Vila-Comamala, J., Rau, C., Rayment, T., & Davenport, A. J. (2015). The effect of nitrate on salt layers in pitting corrosion of 304L stainless steel. *Journal of The Electrochemical Society, 162*(9), C457-C464. https://doi.org/10.1149/2.0421509jes

Streets, M. D. T., & Eden, R. D. (2024). A strategic approach to significant cost reduction through advanced souring forecasting modelling and high-pressure bioreactors in H2S production management. In *SPE Symposium and Exhibition—Production Enhancement and Cost Optimisation* (Paper SPE-220660-MS). Society of Petroleum Engineers. https://doi.org/10.2118/220660-ms

Sun, M.-Q., Yang, J., Wang, Z.-B., & Zheng, Y.-G. (2024). Effect of coexistence of sulfate reducing bacteria and nitrate reducing bacteria on the under-deposit corrosion of carbon steel. *Corrosion Science, 231*, Article 111958. https://doi.org/10.1016/j.corsci.2024.111958

Sun, Z., Wu, J., Zhang, D., Li, C., Zhu, L., & Li, E. (2022). Influence of nitrate concentrations on EH40 steel corrosion affected by coexistence of *Desulfovibrio desulfuricans* and *Pseudomonas aeruginosa* bacteria. *Journal of Oceanology and Limnology, 40*(4), 1448–1461. https://doi.org/10.1007/s00343-021-1247-y

Sunde, E., Lillebø, B.-L. P., Bødtker, G., Torsvik, T., & Thorstenson, T. (2004). H2S inhibition by nitrate injection on the Gullfaks field. In *CORROSION 2004* (Paper NACE-04760). NACE International. https://doi.org/10.5006/c2004-04760

Suri, N., Voordouw, J., & Voordouw, G. (2017). The effectiveness of nitrate-mediated control of the oil field sulfur cycle depends on the toluene content of the oil. *Frontiers in Microbiology, 8*, Article 956. https://doi.org/10.3389/fmicb.2017.00956

Tao, J., Wang, D., Sun, Y., & Xie, F. (2025). The interaction mechanisms of sulfate-reducing bacteria with other corrosive microorganisms and synergistic corrosion prevention strategies. *World Journal of Microbiology and Biotechnology, 41*(12), Article 501. https://doi.org/10.1007/s11274-025-04718-8

ter Heijne, A., de Rink, R., Liu, D., Klok, J. B. M., & Buisman, C. J. N. (2018). Bacteria as an electron shuttle for sulfide oxidation. *Environmental Science & Technology Letters, 5*(8), 495-499. https://doi.org/10.1021/acs.estlett.8b00319

Tian, H., Gao, P., Qi, C., Li, G., & Ma, T. (2024). Nitrate and oxygen significantly changed the abundance rather than structure of sulphate-reducing and sulphur-oxidising bacteria in water retrieved from petroleum reservoirs. *Environmental Microbiology Reports, 16*(2), Article e13248. https://doi.org/10.1111/1758-2229.13248

Uhlig, H. H. (Ed.). (1948). *The corrosion handbook*. Wiley.

Valdés, E., Gabriel, D., González, D., Munz, G., & Polizzi, C. (2023). Integrating thermodynamics and mathematical modelling to investigate the stoichiometry and kinetics of sulphide oxidation-nitrate reduction with a special focus on partial autotrophic denitrification. *Chemosphere, 339*, Article 139605. https://doi.org/10.1016/j.chemosphere.2023.139605

Veshareh, M. J., Kjeldsen, K. U., Findlay, A. J., Nick, H. M., Røy, H., & Marietou, A. (2021). Nitrite is a more efficient inhibitor of microbial sulfate reduction in oil reservoirs compared to nitrate and perchlorate: A laboratory and field-scale simulation study. *International Biodeterioration & Biodegradation, 157*, Article 105154. https://doi.org/10.1016/j.ibiod.2020.105154

Vik, E. A., Janbu, A. O., Garshol, F., Henninge, L. B., Engebretsen, S., Kuijvenhoven, C., Oilphant, D., & Hendriks, W. P. (2007). Nitrate-based souring mitigation of produced water—side effects and challenges from the Draugen produced-water reinjection pilot. In *SPE International Symposium on Oilfield Chemistry* (Paper SPE-106178-MS). Society of Petroleum Engineers. https://doi.org/10.2118/106178-ms

Voordouw, G. (2008). Emerging oil field biotechnologies: Prevention of oil field souring by nitrate injection. In J. D. Wall, C. S. Harwood, & A. L. Demain (Eds.), *Bioenergy* (pp. 377-388). ASM Press. https://doi.org/10.1128/9781555815547.ch29

Wang, B., Deng, S., Su, S., Sun, S., Chen, C., Xu, H., Ma, H., Banat, I. M., She, Y., & Zhang, F. (2024). Microbial characteristics and potential mechanisms of souring control for a hypersaline oil reservoir. *Petroleum Science and Technology, 42*(25), 4644-4659. https://doi.org/10.1080/10916466.2023.2241508

Wang, T., Wang, H., Ran, X., & Wang, Y. (2024). Salt stimulates sulfide−driven autotrophic denitrification: Microbial network and metagenomics analyses. *Water Research, 257*, Article 121742. https://doi.org/10.1016/j.watres.2024.121742

Wang, X., Zhang, Y., Zhang, T., & Zhou, J. (2016). Effect of dissolved oxygen on elemental sulfur generation in sulfide and nitrate removal process: Characterization, pathway, and microbial community analysis. *Applied Microbiology and Biotechnology, 100*(6), 2895-2905. https://doi.org/10.1007/s00253-015-7146-4

Xie, P., Xu, X.-J., Zhang, Q., Hou, Y.-Y., Fan, K.-L., Zhang, R.-C., Chen, C., Wang, A.-J., Lee, D.-J., & Ren, N.-Q. (2025). Potent and selective inhibition of sulfate-reducing bacteria by neutral red. *Environmental Science & Technology, 59*(12), 6115-6125. https://doi.org/10.1021/acs.est.4c09915

Xu, J., Chen, T., Dai, J., Liu, C., Zhou, R., Wang, J., Zhai, X., & Guan, H. (2026). Stabilization of biogenic elemental sulfur in simultaneous nitrogen and sulfur removal via early-stage inhibition. *Environmental Technology, 47*(4), 612-625. https://doi.org/10.1080/09593330.2025.2592740

Xue, Y., & Voordouw, G. (2015). Control of microbial sulfide production with biocides and nitrate in oil reservoir simulating bioreactors. *Frontiers in Microbiology, 6*, Article 1387. https://doi.org/10.3389/fmicb.2015.01387

Yang, Z., Zhu, D. Z., Yu, T., Shypanski, A., Zhang, G., & Zhou, Y. (2021). Effect of ferric iron and nitrate on hydrogen sulfide control in lab-scale reactors. *Environmental Science: Water Research & Technology, 7*(10), 1806-1818. https://doi.org/10.1039/d1ew00460c

Youssef, N., Elshahed, M. S., & McInerney, M. J. (2009). Microbial processes in oil fields: Culprits, problems, and opportunities. *Advances in Applied Microbiology, 66*, 141-251. https://doi.org/10.1016/s0065-2164(08)00806-x

Yuk, S., Kamarisima, Azam, A. H., Miyanaga, K., & Tanji, Y. (2020). The contribution of nitrate-reducing bacterium *Marinobacter* YB03 to biological souring and microbiologically influenced corrosion of carbon steel. *Biochemical Engineering Journal, 156*, Article 107520. https://doi.org/10.1016/j.bej.2020.107520

Zhou, L., Gao, J., Li, Y.-X., Wu, J., Wang, B., Liu, Y.-F., Yang, S.-Z., Gu, J.-D., & Mu, B.-Z. (2025). A new control strategy of sulfate-reducing microorganisms and corrosion rate from a high-temperature oil reservoir by nitrate/nitrite coupling with nitrate-reducing bacteria. *International Biodeterioration & Biodegradation, 201*, Article 106065. https://doi.org/10.1016/j.ibiod.2025.106065

Zulkafli, R., Othman, N. K., & Yaakob, N. (2023). Localised corrosion of API 5L X65 carbon steel in marine environments: The role of sulfate-reducing bacteria (SRB). *Journal of Bio- and Tribo-Corrosion, 9*(1), Article 12. https://doi.org/10.1007/s40735-022-00730-9

Zuo, Z., Xing, Y., Lu, X., Liu, T., Zheng, M., Guo, M., Liu, Y., & Huang, X. (2024). Nitrite-dependent microbial utilization for simultaneous removal of sulfide and methane in sewers. *Water Research X, 24*, Article 100231. https://doi.org/10.1016/j.wroa.2024.100231

Pertanika J. Sci. & Technol. 34 (5): 3109 - 3130 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Mobile-Based Inventory Counting Using 2D Object Detection and Layer-Based Inference

Gan Heng Lai[1], Goh Eg Su[1,2]*, and Johanna Ahmad[1]

[1]Faculty of Computing, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia
[2]Media and Game Innovation Centre of Excellence (MaGICX), Institute of Human Centred Engineering (iHumEn), Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia

ABSTRACT

The term "2D" refers to layer-based quantity inference from a spatially stacked arrangement and does not indicate direct 2D bounding-box regression or depth reconstruction. YOLOv8 is a deep learning framework that enables the development of exceedingly fast computer vision applications. It is adept at object detection and classification, including multi-object detection. Computer Vision explores methods for reconstructing, interpreting, and understanding a scene's three-dimensional structure from its two-dimensional images, leveraging the scene's inherent structure. It involves modelling and simulating human vision using both hardware and software. In this project, YOLOv8 was used as the processing algorithm, integrated with other tools and libraries to build a mobile-based inventory-counting application using 2D object detection and layer-based inference. It is a design that will be used for stock checking or dispatching goods from stock. The input will be the mobile phone's front camera for capturing the boxes. The scene will be captured from the surrounding area of the stock area, where the rack is located. To train the model, a custom dataset of 1,500 images containing three cardboard box sizes was utilised, annotated via the Roboflow platform. Separate YOLOv8 models were trained for the Small, Middle, and Big box datasets. Bounding-box detection was evaluated on labelled validation images using mAP@0.5 and mAP@0.5:0.95, while the complete mobile inventory-counting application was evaluated separately through 250 controlled trials for each box-size model. Finally, the system demonstrated its effectiveness in application. It accurately recognised and measured various box types, assuming there were no vacant spaces within a 3 × 3 × 3 stacking configuration. The variation in box size enables an in-depth evaluation of the device's effectiveness across a range of in-stock items.

Keywords: Artificial intelligence, computer vision, inventory system, object detection, YOLOv8

ARTICLE INFO

Article history:
Received: 30 July 2026
Accepted: 05 October 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.23

E-mail addresses:
ganhenglai1030@gmail.com (Gan Heng Lai)
eg.su@utm.my (Goh Eg Su)
johanna@utm.my (Johanna Ahmad)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

The application of vision technology for computers across many businesses has brought a revolutionary change to traditional methods. It has provided novel solutions for processes such as inventory management, object detection, and visual data analysis over the past few years (Javaid et al., 2022). Using advanced algorithms to process camera-captured visual data, computer vision systems can automate and improve processes that were previously reliant on human intervention and error-prone (Radanliev et al., 2021). This technology was widely used across industries such as retail, logistics, manufacturing, and healthcare. Its integration with automated stock management systems has brought about significant changes and is having a major impact on how businesses manage, monitor, and control their inventory levels (Verhoef et al., 2021; Koh & Yuen, 2022).

There are real-world applications demonstrating the impact of applying computer vision technology in managing stock and inventory for large retail and logistics companies. They could be demonstrated by the example of one of the world's retail giants, Walmart, which is famous for its technological advancement and passion for innovation. Automated systems have been implemented at Walmart using computer vision technology in-store to scan inventory levels, monitor shelf elements, and automate restocking (Collins, 2026). It could improve operating efficiency by detecting stock shortages, identifying shelf gaps, and optimising product placement through the deployment of cameras with machine learning algorithms (Toneguzzi, 2022; Gama et al., 2025).

Dalsey, Hillblom and Lynn (DHL), like many other businesses, has suffered from poor stock management because they are holding onto a lot of inventory (DHL Supply Chain, n.d.). This had been especially challenging for the organisation due to low visibility and low stock accuracy. This led to the duplication of valuable commodities across multiple sites (DHL Supply Chain, n.d.). It will lead to excess costs in terms of finance and consumer satisfaction. Thus, DHL chose to apply computer vision technology to address these challenges, recognising the need for a more precise and effective approach to managing their stock (DHL Supply Chain, n.d.). DHL improved the accuracy of its stock counting method, reduced product duplication, and gained better visibility into its stock levels by using high-speed cameras and enhanced image recognition technology (Villegas-Ch et al., 2024).

LITERATURE REVIEW

The integration of digital transformation technologies into inventory management represents a major shift toward automated, error-free operational environments (Sengupta & Dreyer, 2023). This literature review highlights and discusses technologies and studies relevant to the current inventory system's operation and to the limitations of human operation. Central to this transformation is Computer Vision, a field of artificial intelligence

that enables computers to derive meaningful information from digital images and videos. Specifically, the You Only Look Once (YOLO) architecture, and its recent YOLOv8 iteration, have revolutionised real-time object detection by predicting bounding boxes and class probabilities in a single neural network evaluation (Li et al., 2022; Jocher et al., 2023). Exploring these technologies provides the necessary context for identifying the existing gaps in manual inventory systems and evaluating the efficacy of proposed automated alternatives (Gao et al., 2020; Terven & Cordova-Esparza, 2023).

The Process of the Inventory System

The inventory system process begins with purchasing goods and delivering those goods to the company's storage facility. This will start the inventory flow by addressing physical goods, which involve acquiring products needed for operations (Büyüközkan & Göçer, 2018). Determining the appropriate order quantity can be tricky, as the quantity recorded in one's own system may lead to excess stock due to overordering (Fildes et al., 2009). Once purchased, goods are sent directly to the company's warehouse area.

The next step in the inventory system involves evaluating, sorting and storing goods in the assigned stock area. Carefully measuring the number of products purchased was essential at this stage to maintain the accuracy of inventory management (Chen & Patel, 2024; Villegas-Ch et al., 2024). Sorting is grouping products by size, type, and demand frequency, among other criteria. To maintain optimal stock levels and avoid operational disruptions, inventory must be closely monitored. Through periodic physical inventory counts, businesses can closely track stock quantities in sequence, identify stock discrepancies in a timely manner, such as deadstock or stockouts, and then take swift corrective actions.

When it comes to order placement, businesses rely on a mix of internal processes and external suppliers. Once orders are placed, businesses use their tracking systems to monitor shipment status. The key to this phase was attentively checking the number of orders made to ensure the order was completed accurately. When goods were delivered from the warehouse to the customer, inventory records were promptly updated to reflect the latest stock levels. Its purpose was to identify potential stock shortages and adjust their replenishment strategies accordingly (Belyh, 2022).

Limitation of Current System

One major drawback of the current inventory system is the additional complexity it imposes on the employees responsible for managing it (Liboni et al., 2019). Many businesses underestimate how long it takes for new inventory management procedures to be completely understood and accepted by their employees. Inventory management becomes difficult when process leaders are unfamiliar with the new rules, instruments, and frameworks.

Businesses will be required to invest much time and resources in training employees on the specifics of order processing and inventory updating (Nguyen & Tran, 2024).

Another significant limitation is the lack of efficient order and stock management practices. The responsibilities of order management include receiving, monitoring, and updating stock levels in response to customer orders. Businesses risk overselling products if they don't establish an effective order management process. In the absence of appropriate supervision and management protocols, companies would find it difficult to sustain sufficient stock levels. It could result in client discontent and decreased earnings (Belyh, 2022).

Comparison Between Existing Systems and Proposed System

Inventory systems in use today predominantly rely on stationary arrangements, such as fixed barcode scanners or permanently affixed factory cameras along conveyor belts. A primary limitation of these current approaches is that they capture stock only from a single, direct line of sight and usually only the front face of a pallet. Consequently, users are compelled to manually estimate the remaining hidden stock behind the initial item, a process that is highly error-prone and prone to record inconsistencies. However, this proposed system solution will be utilised with small components that can be integrated into any inventory system, offering flexibility and compatibility across various operational setups.

Table 1 presents a comparison between the proposed system with existing systems regarding mobility and cost. It is crucial to study existing systems like Walmart Canada's Out-of-Stock Detection and CountThings to establish a comparative baseline for performance, cost, and mobility in modern inventory operations. Walmart Canada represents a large-scale, enterprise-grade deployment utilising fixed infrastructural cameras, highlighting the high costs and lack of mobility inherent in permanent setups. Conversely, CountThings exemplifies a mobile, template-based software approach that introduces subscription costs and dimensional limitations. Analysing these opposing ends of the technological spectrum justifies the need for the proposed mobile-based inventory-counting

Table 1.
Comparison of The Proposed System with Existing Systems.

System	Mobility	Cost
Walmart Canada's Out-of-Stock Detection (Wassel, 2022)	An operation that operates in a fixed location cannot be migrated.	Depending on the size of the retail operation
CountThings from Photos (CountThings, n.d.)	Count on one side only; it cannot compute quantities in three dimensions	$120 per month or RM1200 per year
2D Object Detection Used for an Inventory System	High flexibility; can be installed on most Android mobile phones.	Low-cost device.

application, which uses 2D object detection and layer-based inference to merge the mobility of smartphone applications with the robust, two-dimensional detection capabilities typically reserved for enterprise AI, while remaining cost-effective for smaller businesses (Redmon et al., 2016; Wang et al., 2023).

Walmart Canada implemented an automated artificial intelligence system in all its locations to provide consumers and associates with a better shopping experience. As they recognise that product unavailability can be frustrating for customers, they utilise Focal Systems to solve the problem by automating out-of-stock detection with cameras. Focal Systems is a technique that utilises the most advanced and proven AI technology in retail, known as FocalOS. It automates and optimises all current human-made decisions. Focal Systems has developed a durable and cost-effective Wi-Fi camera that scans shelves hourly to provide real-time counts of shelf contents throughout the day. Focal can boast accurate computer vision technology because it is trained on over a billion labelled images from global businesses. (*Focal Systems*, n.d.). This system was installed in high-traffic areas throughout Walmart stores. The camera integrated Walmart's existing inventory system to automatically detect stock issues and then provide alerts for out-of-stock items (Toneguzzi, 2022).

CountThings is an automatic counting application for industrial and digital transformation. It consists of techniques for counting objects from photos and videos. This will automatically count the items in the photo on mobile devices and recognise, categorise, and initiate actions using computer vision. This development aims to reduce the error-prone task of manual counting in industrial operations and make it an efficient, effective process. This application requires a model template to allow the app to count the items. It does not work without a template for the things we want to detect. The procedure for this application starts by inserting the counting template before taking a new picture, and then tapping the "count" button summarises the detection result for the user. There is a constraint in this application: the camera's position. The user must make sure the objects are directly across from the camera, not angled toward it. This application only allows free access for the top seven days of a new account. The user must pay in USD for the license on a daily, monthly, or annual basis (CountThings, n.d.).

This system aims to address and complement the weaknesses of existing solutions, such as Walmart's fixed-location camera system and the CountThings application. This proposed solution offers greater versatility than Walmart's system, as it can be deployed anywhere in the inventory area, whereas Walmart's system is limited to fixed locations. Even though CountThings was good at identifying and counting objects when viewed from a single angle, it was much more expensive when used in isolation from inventory management systems in small businesses or companies and was unable to calculate numbers when viewed from three dimensions. This proposed solution can reduce costs and integrate easily into the current inventory system.

The "You Only Look Once" (YOLO) framework is a state-of-the-art family of object detection models in computer vision. Unlike region-based convolutional neural networks (R-CNNs), which require thousands of network evaluations per image, YOLO is a single-shot algorithm that directly classifies objects and predicts bounding boxes in a single forward pass. This project utilises YOLOv8, which introduces an anchor-free detection mechanism. By eliminating predefined anchor boxes, YOLOv8 improves generalisation and learning speed for custom datasets. Furthermore, YOLOv8 incorporates a C2f module and a decoupled head, which significantly enhances the model's ability to distinguish densely packed objects.

METHODOLOGY

The extended waterfall model was chosen in this case because it requires all features or functions to be completed before testing. It eliminates the need for user interaction or communication throughout the project's development stage. Additionally, clarification will be provided on the current state of development for each phase of the extended waterfall methodology. There are five stages in the waterfall methodology: requirements, design, implementation, testing, and maintenance. However, the process in this project was covered only up to the testing phase, as shown in Figure 1. The difference between an extended waterfall and a standard waterfall methodology is that the system is adjusted if accuracy is not met at the evaluation stage. It was evaluated again after the adjustment was completed.

Analysis Phase

The project starts with the Analysis and Requirements Gathering Phase, which discusses case studies related to the project's inventory system. Companies face significant challenges due to inventory discrepancies, including inaccurate item counts and operational inefficiencies. The application was designed for inventory managers and warehouse workers, with a focus on capturing photos from multiple angles to ensure precise counts. The scope was restricted to a single user using the device's front camera to detect three distinct box sizes and shapes within the marked space on a single rack level. There was a user requirement collection process in

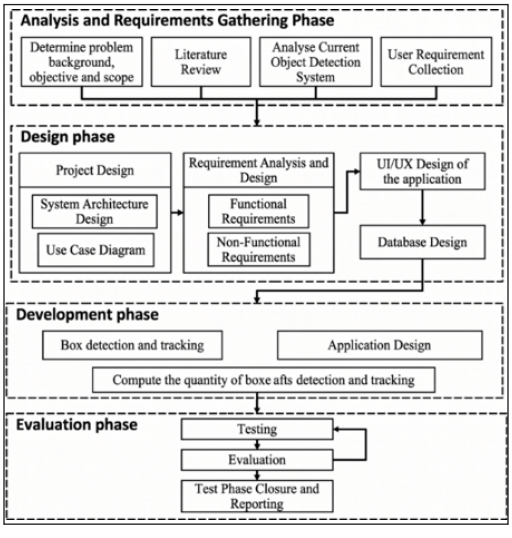


Figure 1. Project methodology phases overview

this analysis stage. A structured approach using Google Forms was used to collect user requirements from the staff members of Purple Fashion Sdn. Bhd. to guide the inventory management application's development.

Design Phase

The interaction between the mobile application and the backend program is the main emphasis of the user-device architecture. The application will be developed using the Flutter framework to capture images of cardboard boxes via the device's camera. The images will be stored and used for object detection processing. The detection result will be stored again in a cloud-based database, Firebase, to store the inventory data. Each inventory record will contain details such as the item ID, the quantity detected, and timestamps. The primary use case is user authentication, allowing users to log in via the mobile app. Image capture and upload is another important use case that allows customers to take pictures of box stores and process them for object detection. The program can identify and count boxes in real-time, showing the user the findings.

Development Phase

The first step in the development stage will be detecting cardboard boxes using YOLOv8, implemented in Python within a Jupyter Notebook environment. The detection process begins with the YOLOv8 backbone extracting low-, medium-, and high-level feature maps from the captured warehouse images. The model then utilises anchor-free detection to predict the midpoint of each cardboard box, drawing precise bounding boxes around them. The C2f module integrates contextual information with high-level features to improve detection accuracy and model performance. By processing the spatial coordinates of these bounding boxes, the system categorises the detected objects into varying hierarchical layers. It was decided to use only the essential object detection and tracking components of YOLOv8 (Zhao et al., 2019; Vijayakumar & Vairavasundaram, 2024). It achieves greater accuracy with deep learning techniques than traditional computer vision algorithms. This improved method dispenses with the need for supplementary algorithms that were to be developed in C++.

Testing Phase

This stage included evaluating the system's functionality and assessing the detector's accuracy across a range of samples. The project primarily focused on accuracy percentage testing, as system accuracy is highly dependent on the arrangement of boxes and the angles at which images are captured. A specialised user testing session will be conducted in collaboration with Purple Fashion Sdn. Bhd. to evaluate the system in an active retail environment. A total of 5 participants from the inventory staff will be invited to test the

application. The specific use case focuses on autonomously managing the remaining stock on the racks, eliminating the need for manual counting. The number of participants was limited to one individual to maintain strict control over the restricted stock arrangement order, ensuring the box patterns were placed sequentially without unwanted repetition that could skew the algorithm's performance data. These participants will be given specific use cases and guidance to arrange the boxes and capture images to test the system's detection capabilities.

DESIGN

The various modules in the use case model of this research-prototype are designed to handle specific functionalities in Figure 2. The main modules were the User Authentication module, Image Processing module and Object Detection module, which consist of several use case functions that establish the interactions between the user and the system.

The primary use case within the User Authentication module is the Login Account function. This scenario enables users to verify their identity by inputting their username and password. The Image Processing module manages all aspects associated with capturing and processing box images. It involves various essential use cases such as Select Template, Take Photo, and Check Result. The Select Template use case lets you choose from a range of box detection templates offered by the system. The system provided multiple detection algorithms designed to accommodate different box types, ensuring flexibility in box detection.

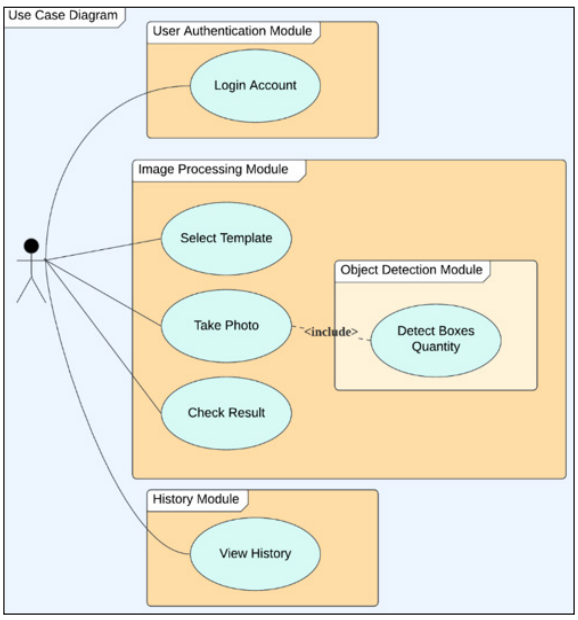


Figure 2. Use case diagram of the proposed research prototype

The third module was the Object Detection module, which is responsible for detecting and computing cardboard boxes in images. It consists of a use case, Detect Boxes Quantity, to perform actual detection and counting of boxes in the captured image. By leveraging the YOLOv8 algorithm's capabilities, the system was able to analyse the image and accurately identify and count the boxes. The system processes the image with the chosen detection template and calculates the number of boxes based on the detected features. The last module was the History module, used for the history review display to the users. The primary use case for this module is the view history function. This function was integrated to allow the user to review the box detection result saved from the previous operation. This allowed multiple users to check the latest updated stock quantity status.

IMPLEMENTATION

This section describes the justification for Open-Source Computer Vision Library (OpenCV) techniques and the implementation of a new box detection algorithm, model training, and application development.

Justification For Deprecating OpenCV Techniques

Two standard OpenCV methods were tested for the box computation algorithm: Harris Corner Detection and FLANN-based feature matching (Bradski & Kaehler, 2008). The Harris Corner Detection algorithm was first created to detect the corners of cardboard boxes (OpenCV, n.d.-b). It served as a clear marker for counting boxes and marked the corners with blue-coloured points. However, its use within real warehouse environments showed a few major shortcomings. The algorithm showed trouble distinguishing true corners on the box from bordering background features in Figure 3. While human viewers can easily notice the high visual contrast between the container edges and backgrounds, the algorithm frequently struggled to distinguish corners from neighbouring features. Real cardboard containers in warehouse environments led to high false-detection rates for corners. This Harris Corner Detection algorithm will also mistakenly detect scores of areas that were not corners, especially within backgrounds containing complex textures.

The FLANN was tested as one possible technique for detecting identical boxes imaged from different camera angles, as shown in Figure 4 (*OpenCV*, n.d.-a). However, the experiment's outcome showed significant limitations to its practical application. While aligning features from



Figure 3. Harris corner detection result

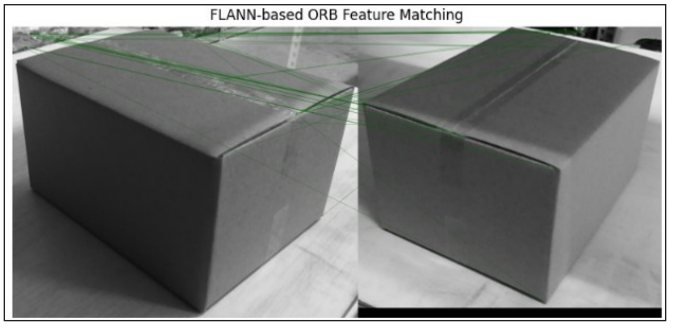


Figure 4. FLANN-based matching result

two images of the same box taken from different angles, the algorithm often misidentified corners as features that were anything but corners, generating false correspondences. Features from one box often ended up being matched to features from different positions on the box, or even from the background, creating erroneous box identifiers.

The comparison between Harris Corner Detection and FLANN-Based Matching showed that, though both are theoretically grounded, they did not meet the practical requirements for accurate identification and counting of boxes in warehouse environments. Their high rate of false positives and variability in feature correspondences would have led to significant errors in the procedure for counting the boxes. These preliminary tests were used for feasibility assessment rather than as formal performance baselines because the OpenCV methods do not directly generate equivalent box-level and layer-level detection outputs.

Implementation of Enhanced Box Detection Algorithm

Based on previous studies on the existing algorithms for both Harris Corner Detection and FLANN-based algorithms, an enhanced algorithm is implemented by eliminating the need for corner detection and feature matching by focusing on a crucial insight into the holistic layout of boxes in warehouse environments. The basic premise of the suggested algorithm is to detect and count boxes based on their spatial position within a hierarchical system. The proposed algorithm classifies the detected boxes into three different layers: Layer 1, Layer 2, and Layer 3, rather than attempting to find every individual box. These layers correspond to ground-level, middle-level, and top-level boxes, respectively, within the hierarchical system. The merit of this proposed algorithm lies in its ability to accurately detect the topmost layer of boxes, which follow stacking principles, thereby enabling the overall estimation of the total number present within the entire stack.

The algorithm implements a structured decision-making process based on predefined layer classes. When the arrangement has 9 or fewer boxes, they are classed as Layer 1 boxes, and the system then calculates the total number of objects detected directly. When

the arrangement involves about 10 to 18 boxes, they were classed as Layer 2 boxes, which then indicates the presence of an overlying layer. In this model, the total count of boxes was calculated by adding recognised Layer 2 boxes to a base of 9, the minimum count for the stacked first-layer configuration. For example, it calculates a total count of 11 boxes if the system recognises two Layer 2 boxes (9 base boxes and 2 Layer 2 boxes). If the arrangement ranges from about 19 to 27 boxes, the system detects Layer 3 boxes, which indicate the presence of a third tier of stacking. When the system detects Layer 3 boxes, it first assumes an overall total of 18, then adds the number of detected Layer 3 boxes to this total to determine the final count. For instance, the system calculates an overall total of 22 when detecting four Layer 3 boxes (18 base boxes plus 4 detected Layer 3 boxes). The minimal pseudocode is: if Layer 3 is detected, set total to 18 + Layer 3 count. Else if Layer 2 is detected, total = 9 + Layer 2 count. Else, total = Layer 1 count.

However, several assumptions are considered to guarantee the precision of this calculation, as shown in Figure 5. Firstly, it was necessary to arrange the boxes so that the bottom layer was fully occupied before placing any boxes on the upper layers. Secondly, there should be no gaps between the boxes, as they may lead to inaccurate volume calculations. Lastly, the boxes should be arranged in rows to avoid isolated boxes in a corner. The boxes should be arranged to connect to each other and cannot be a box standing alone in a corner. The 3 × 3 × 3 configuration represents the supported operating domain of the current prototype and should not be interpreted as a general assumption about all warehouse arrangements. Irregular stacks, vacant positions, isolated boxes, and partially filled lower layers are outside the scope of the present layer-based calculation.

Implementation of Model Training

During the implementation phase, the YOLOv8 model was trained specifically for box detection and the corresponding box labels, but not for detecting arbitrary features or corners. This was done to enhance the model by distinguishing the boxes placed at varying levels, even when they had the same appearance and shape. During training, the model was

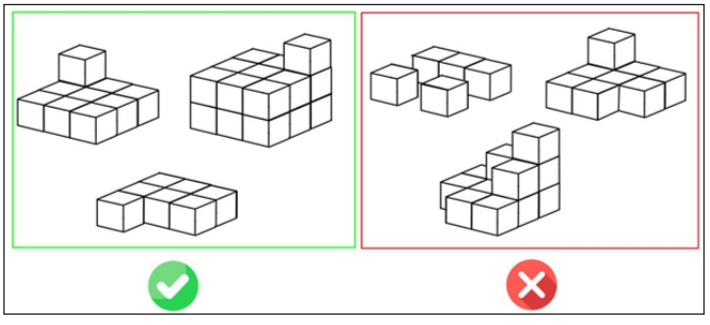


Figure 5. Box arrangement assumptions example

trained on images containing different box configurations, with each box labelled based on its location across multiple layers. It helped the neural network detect ground-level and higher-level placed boxes.

The model was trained on three different datasets using the same methodology. The same code structure and training pipeline were applied to the three models, with the only difference being the dataset download step. There are 500 pictures for each different type of cardboard box model training. Those dataset images were recognised and labelled using the Roboflow platform (*Roboflow*, n.d.). Each box-sized dataset was divided into 80% training images and 20% validation images. The validation split was used to select the best.pt checkpoint and to assess bounding-box detection; it was not an independent test set. The small variant of YOLOv8 was initialised using the pretrained yolov8s.pt checkpoint and trained separately on each of the three box-size datasets. The dataset paths and Layer 1–Layer 3 class definitions were specified in the corresponding data.YAML files. Each model was trained for a maximum of 100 epochs using an input resolution of 640 pixels, a batch size of 16, automatic optimiser selection (optimizer=auto), an initial learning-rate parameter (lr0) of 0.01, a weight decay of 0.0005, and a random seed of 0. Training was capped at 100 epochs. Although the recorded training configuration set patience=100, this setting did not provide an effective early stopping condition within the 100-epoch schedule. The checkpoint with the best validation performance, best.pt, was retained for application testing.

During application testing, the best. The PT checkpoint was used to process input images with a nominal image size of 640 pixels. No custom confidence threshold was specified in the prediction configuration; therefore, the default Ultralytics version 8.3.39 confidence threshold of 0.25 was applied. The non-maximum-suppression IoU threshold was 0.70, and the maximum number of detections per image was 300. Predictions were generated for all three-layer classes without class filtering.

Roboflow was used to annotate and export the three datasets in YOLOv8 format. No additional offline augmentation was applied during the Roboflow export process. Because no custom augmentation parameters were specified in the model-training command, the default online augmentation pipeline of the installed Ultralytics YOLOv8 version was used. This pipeline applied transformations such as HSV colour variation, translation, scaling, horizontal flipping, and mosaic augmentation to the training images.

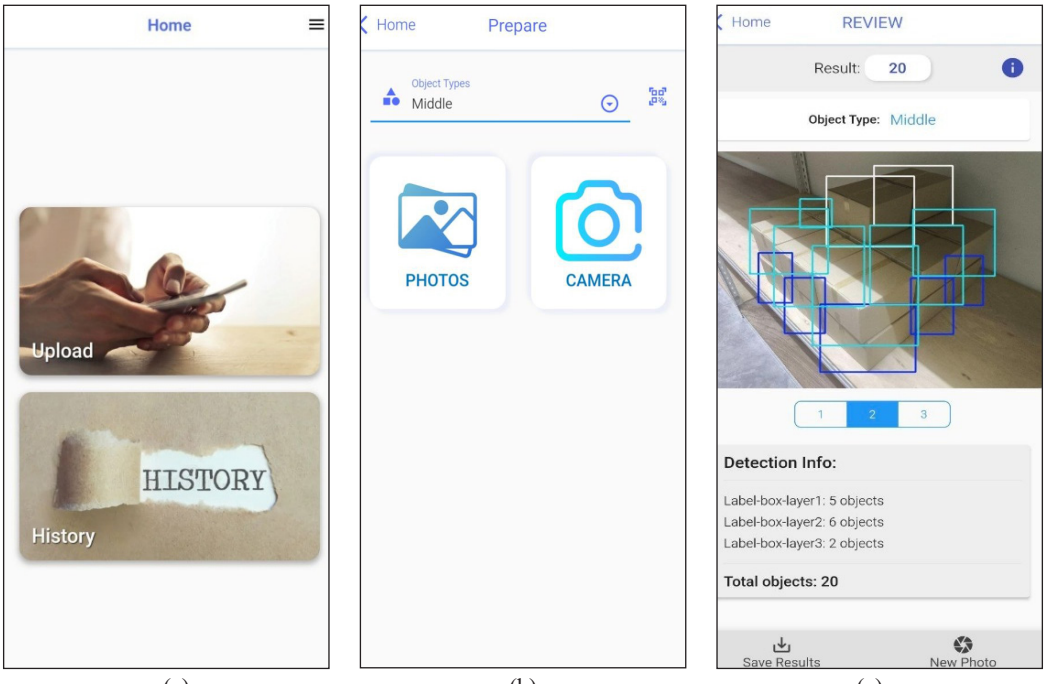
The use of three separate models was an experimental design choice intended to isolate performance across the predefined box dimensions. It is not proposed as a scalable production strategy for arbitrary SKUs. A future implementation could train a unified multi-class model on a substantially larger dataset that includes multiple packaging types, dimensions, and warehouse environments.

Implementation of Application Interface

The home page provided two operational options: Take Photo and View History. The photo option allows the user to compute the number of boxes by capturing images, while the view history option provides a history of the detection process the user has made. The template selection interface provides three options corresponding to the three different box sizes to be detected. After selecting either template, the system proceeded to the image capture page. After the user captures images of the stock area, the user proceeds to the next interface, the result page, by clicking the proceed button. The result interface displayed the detection outcome and the computed box count. The number of boxes detected is displayed at the top of the screen, followed by the captured image. The user can click the save button to store the result in the database. The system interfaces are shown in Figure 6.

Implementation of Database

The box detection and counting process involves creating a database specifically to store important information about users, images, and detection outcomes, as shown in Figure 7. Below is a breakdown of the information provided, including the primary key and the foreign key that will be stored in the database.



(a) (b) (c)

Figure 6. Application interface: (a) Home page, (b) Template selection page, (c) Result page

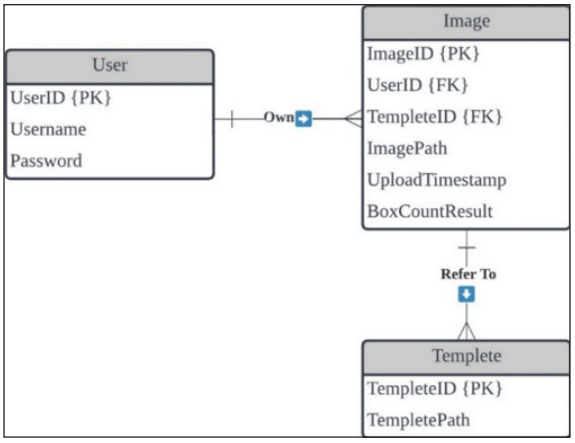


Figure 7. Database diagram

The User table was created to store crucial user details, such as UserID, Username, and Password. UserID serves as a unique identifier for each user within the system and database. The Image table stores information about the images uploaded by users and the results processed by the system. The primary key for this table was ImageID as a unique identifier for each detection. There were two foreign keys in this table: UserID and TemplateID, which are references from the User table and the Template table, respectively. These were the linking of images to their uploaders and the indication of the template used to process the images. The Template table was used to store details about the various templates accessible for box detection. The primary key of this table was TemplateID, which serves as a unique identifier assigned to each template. TemplatePath was the location of the file where the image processing template was stored.

RESULT AND DISCUSSIONS

The experiment was conducted in a warehouse stock room with goods on the racks and on the floor, creating a visually complex background for object detection. An empty shelf was used to arrange the test boxes. During testing, the operator stood approximately 0.5–1.0 m from the shelf and captured images at an angle of approximately 45°, as shown in Figure 8. Maintaining the same position throughout the experiment ensured consistent testing conditions and reflected the typical posture warehouse workers use during stock checking. All available room lights were switched on to provide sufficient illumination for the smartphone camera.

This experiment had only a single tester perform 250 trials for each model trained, rather than distributing testing across multiple individuals. This decision was rooted in the unique nature of this research. This procedure aligns with the nature of the research, in which performance measures are influenced by patterns and the sequential placement of

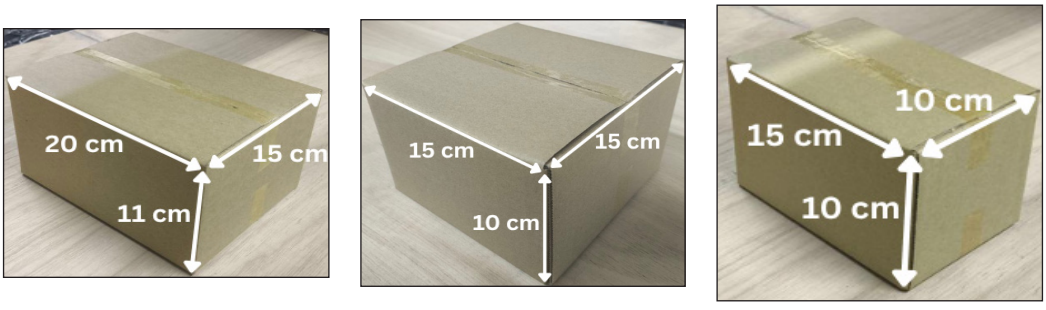


Figure 8. Experimental environment setup

boxes on the rack. The use of a single tester not only provided a deeper understanding of the earlier boxes' arrangements but also enabled more purposeful adjustments to prevent unwanted pattern repetition and further reduce the risk of compromising the authenticity of the experimental data. The resulting images were grouped into three categories: true positives (cases in which boxes were identified with accurate layer detection), false positives (cases in which boxes were found but with incorrect layer categorisations), and false negatives (cases in which boxes were missed by the model). True negative cases are considered irrelevant in this experimental setup due to the nature of the detection problem.

Experiment Record

The experiment required an Android phone with the proposed research prototype installed and 27 cardboard boxes in three sizes: 20 × 15 × 11 cm, 15 × 15 × 10 cm, and 15 × 10 × 10 cm, as shown in Figure 9. Testing was conducted in an operational warehouse stockroom to evaluate the proposed system's performance in a realistic environment.



(a) (b) (c)

Figure 9. Cardboard box type: (a) Big size, (b) Middle size, (c) Small size

Bounding-Box Detection Validation

The selected checkpoint for each box-size model was assessed on its labelled validation split. Table 2 reports the mean average precision at an intersection-over-union threshold of 0.50 (mAP@0.5) and averaged across thresholds from 0.50 to 0.95 (mAP@0.5:0.95). These metrics evaluate predicted bounding boxes against annotated boxes. They are reported separately from the subsequent 250 application-level counting trials for each model.

Two evaluations were conducted. Bounding-box detection was assessed on labelled validation images using mAP@0.5 and mAP@0.5:0.95, as reported in Table 2. The integrated mobile application was assessed separately through 250 controlled counting trials for each box-size model. Figures 10 to 14 report the recorded application-trial outcomes and their derived measures. These trial results illustrate how the complete counting workflow is carried out. They do not measure bounding-box localisation and are not a substitute for detector-level mAP results.

	A	B	C	D	E
1	Box Types	Iteration	Actual number	Number of boxes detected	Accuracy Status (✓/✗)
2	Middle	1	13	14	×
3		2	13	14	×
4		3	15	15	×
5		4	22	21	×
6		5	23	21	×
7		6	26	27 (Layer Error)	×
8		7	26	27 (Layer Error)	×
9		8			
10		9			
11		10			
12	Big	1	4	19	×
13		2	4	11	×
14		3	5	11	×
15		4	5	11	×
16		5	15	14	×
17		6	17	16	×
18		7	22	22 (Layer Error)	×
19		8	22	21	×
20		9	22	22 (Layer Error)	×
21		10	23	24	×
22		11	23	21	×
23		12	23	23 (Layer Error)	×
24		13	24	24 (Layer Error)	×
25		14	25	25 (Layer Error)	×
26		15	25	25 (Layer Error)	×
27		16	25	25 (Layer Error)	×
28		17	26	26 (Layer Error)	×
29	Small	1	11	12	×
30		2	14	15	×
31		3	15	16	×
32		4	21	20	×
33		5	21	20	×
34		6	25	25 (Layer Error)	×
35		7	22	22 (Layer Error)	×
36		8	22	22 (Layer Error)	×
37		9	22	22 (Layer Error)	×
38		10	23	24	×
39		11	23	23 (Layer Error)	×
40		12	25	25 (Layer Error)	×
41		13	24	24 (Layer Error)	×
42		14	22	21	×

Figure 10. Fail record

Figure 10 presents the table used to record unsuccessful trials during the experiment. Because 250 trials were conducted for each model, this table facilitated the identification and subsequent analysis of counting and layer-classification errors.

The recorded failures were primarily associated with incorrect layer classification, over- or undercounting due to viewing angle, and violations of the assumed gap-free stacking arrangement. Low illumination may further reduce the visibility of box boundaries.

Table 2
Bounding Box Detection Results on Validation Splits

Box-size model	Variant	mAP@0.5	mAP@0.5:0.95
Small	YOLOv8s	0.99500	0.99273
Middle	YOLOv8s	0.99500	0.99159
Big	YOLOv8s	0.99500	0.99345

Result of Middle Box Type Model Experiment

The Middle box type model showed significant performance across the 250 test iterations. As shown in Figure 11, it achieved 243 true positives, 7 false positives, and no false negatives. This distribution indicated a strong detection capacity, but with the presence of occasional errors in its classification.

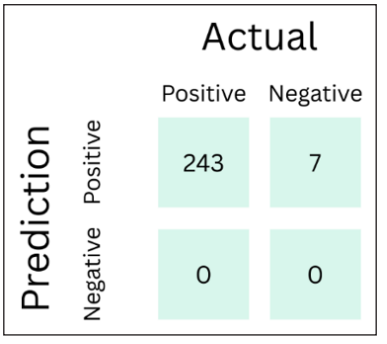


Figure 11. Middle type results classification

The performance metrics for the middle box model are calculated as follows:

$$\text{Accuracy} = \frac{243}{250} \times 100\% = 97.2\% \quad [1]$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% = \frac{243}{250} \times 100\% = 97.2\% \quad [2]$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% = \frac{243}{243} \times 100\% = 100\% \quad [3]$$

$$\text{F1 Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% = 2 \times \frac{0.972}{1.972} \times 100\% = 98.6\% \quad [4]$$

The Middle box model exhibits excellent overall performance with errorless recall. These performance measurement parameters were near to 100%. This proved that the middle-box-type model could label all boxes in the test cases.

Result of Big Box Type Model Experiment

The Big box model showed significant performance across the 250 test iterations. It achieved 233 true positives, 14 false positives, and 3 false negatives, as shown in Figure 12. The distribution reflects good overall performance with periods of both misclassification and missed instances.

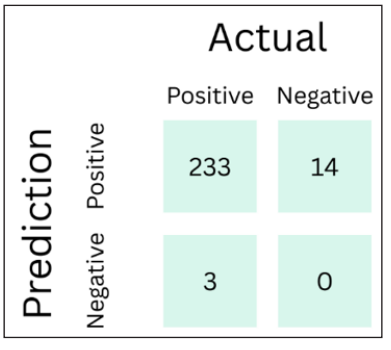


Figure 12. Big type results classification

The performance metrics for the big box model are calculated as follows:

$$\text{Accuracy} = \frac{233}{250} \times 100\% = 93.2\% \quad [5]$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% = \frac{233}{247} \times 100\% = 94.3\% \quad [6]$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% = \frac{233}{236} \times 100\% = 98.7\% \quad [7]$$

$$\text{F1 Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% = 2 \times \frac{0.931}{1.930} \times 100\% = 96.5\% \quad [8]$$

The big box model achieves optimal performance, but it performs marginally worse than the box box model. The performance measurement parameters near 100% proved that this big box model could label all boxes in the test cases.

Result of Small Box Type Model Experiment

The Small box model showed high performance across the 250 test iterations. As shown in Figure 13, it achieved 236 true positives, 14 false positives, and no false negatives. This trend indicates high detection quality, even with some false positives.

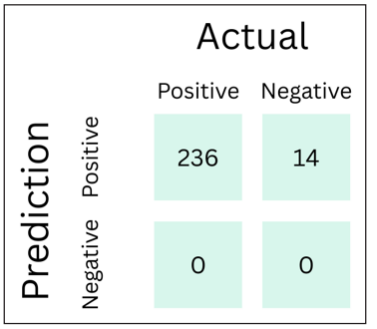


Figure 13. Small type results classification

The performance metrics for the small box model are calculated as follows:

$$\text{Accuracy} = \frac{236}{250} \times 100\% = 94.4\% \quad [9]$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% = \frac{236}{250} \times 100\% = 94.4\% \quad [10]$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% = \frac{236}{236} \times 100\% = 100\% \quad [11]$$

$$\text{F1 Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% = 2 \times \frac{0.944}{1.944} \times 100\% = 97.1\% \quad [12]$$

The small box-type model exhibits excellent overall performance, with optimal recall. These performance measurement parameters were near to 100%. This proved that the small box-type model can label all boxes in the test cases, and that there are opportunities to improve classification accuracy.

DISCUSSION

The analysis of the experimental results reveals significant performance trends across the three model types, as shown in Figure 14. Table 2 presents detector-level bounding-

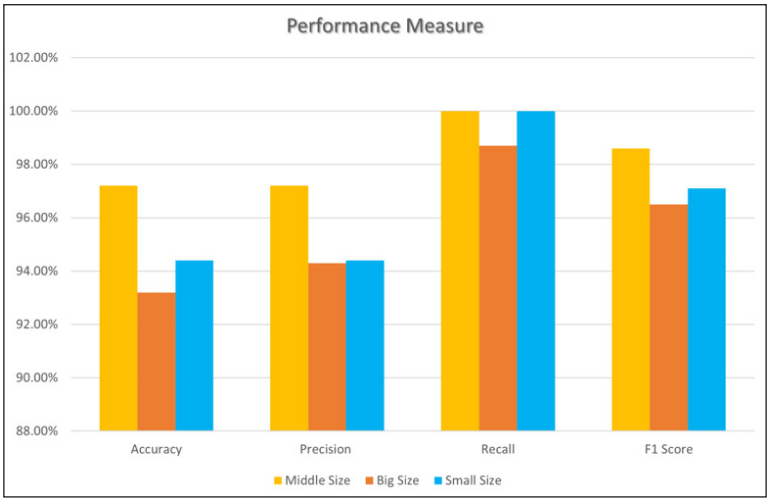


Figure 14. Model performance measure

box performance, whereas Figure 14 presents application-level trial outcomes; the two evaluations measure different aspects of the system. The highest-performing model was the Middle type, with 97.2% accuracy and an F1 score of 98.6%. It maintained an excellent balance between precision and recall. The Small-type model followed closely, with an accuracy of 94.4% and an F1 score of 97.1%. But it faces slightly greater difficulty with classification than the Middle type. This model correctly picked up most of the positive instances but missed a few, as evidenced by the recall being less than 100%. The Big type of model, with its relatively satisfactory result of 93.2% in terms of accuracy, along with an F1 of 96.3%. That is, the model accurately identifies all actual positive cases without omitting any and is highly accurate in its positive predictions.

Study Limitations and Future Work

This system is a proof of concept evaluated under a constrained 3 × 3 × 3 stacking arrangement. Its counting rule assumes that lower layers are fully occupied and boxes are placed in contiguous rows. Consequently, the present method cannot reliably count stacks with vacant positions, partially filled lower layers or irregular arrangements. This fixed-layout assumption is a key barrier to deployment in general warehouse settings. Future work must revise the counting logic to accommodate such arrangements and validate it using more diverse warehouse scenes.

Lack of Multi-Operator Usability Testing

One operator performed 250 trials for each of the three box-size models. Using one operator reduced variation in image-capture conditions during this initial system test, but the results

do not establish whether other warehouse workers would achieve comparable performance. Future work should assess multiple operators with different experience levels, capture styles, and report variation between operators.

CONCLUSION

The proposed solution successfully identified and enumerated the number of boxes. It has high potential to solve the major challenges faced by warehouse and stock management systems. The achievements indicated that by employing advanced computer vision methods and machine learning algorithms, the process of warehouse stock enumeration can be automated and simplified. There are some limitations that have been recognised as potentially influencing the system's performance and usability in certain settings. The system is currently confined to recognising three-dimensional square cardboard boxes. The application's detection accuracy is heavily dependent on environmental conditions, specifically lighting and image-taking angles. Extremely low-light conditions might immediately reduce detection performance. There must also be certain prerequisites to get the best results. The layer-based sorting algorithm also takes several important assumptions that must be maintained for accurate counting. The algorithm also requires the boxes to be positioned without gaps and arranged in touching rows rather than discrete locations. There are multiple areas where the system's effectiveness can be improved and its limitations overcome. The system's current tri-layer detection system can be expanded to encompass five or six layers. This aims to support a wider range of stacking scenarios commonly found in today's warehouses. The accuracy of detection results can be significantly improved by collecting and integrating large training sets across varied lighting conditions, including dim lighting, artificial light common in warehouses, natural light, and combinations thereof. The additional dataset should also include images that represent complex backgrounds with varied warehouse equipment, shelving units, and other inventory items. Future improvements should focus on models skilled at detecting various types of stock items. For example, cylindrical containers, plastic containers, metal objects, fabric containers, and unusual items are commonly found in warehouses. The development of these multi-format detection capabilities would transform the system from a box-counting device to a comprehensive inventory management system. Future work should remove the fixed stacking constraint by incorporating training examples with gaps, partially occupied layers, irregular arrangements, and larger stacking configurations.

ACKNOWLEDGMENT

This research is supported by Media and Game Innovation Centre of Excellence (MaGICX), Institute of Human Centred Engineering (iHumEn), Universiti Teknologi Malaysia (UTM) and Faculty of Computing, UTM.

REFERENCES

Belyh, A. (2022, December 6). *What is inventory management process? Flow, challenges, implementation*. FounderJar. https://www.founderjar.com/inventory-management-process-flow/

Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media.

Büyüközkan, G., & Göçer, F. (2018). Digital supply chain: Literature review and a proposed framework for future research. *Computers in Industry, 97*, 157-177. https://doi.org/10.1016/j.compind.2018.02.010

Chen, Y., & Patel, K. (2024). The impact of automated sorting on warehouse logistics and inventory accuracy. *International Journal of Supply Chain Automation, 15*(2), 200-215.

Collins, G. (2026, March 1). *Walmart's AI automation push sets new retail benchmark*. AI Magazine. https://aimagazine.com/news/walmarts-ai-automation-push-sets-new-retail-benchmark

CountThings. (n.d.). *CountThings from photos*. https://countthings.com/

DHL Supply Chain. (n.d.). *Improving field inventory management for a medical devices provider*. https://www.dhl.com/content/dam/dhl/global/csi/documents/pdf/glo-csi-inventory-optimization-for-global-medical-devices-customer-2023.pdf

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: An empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting, 25*(1), 3-23. https://doi.org/10.1016/j.ijforecast.2008.11.010

Focal Systems. (n.d.). *Focal Systems*. https://focal.systems/

Gama, L., et al. (2025). Transforming enterprise structures through digitalization and advanced machine learning. *Technology in Society, 72*, 101-115.

Gao, F., He, Y., Wang, J., Hussain, A., & Zhou, H. (2020). Anchor-free convolutional network with dense attention feature aggregation for ship detection in SAR images. *Remote Sensing, 12*(16), Article 2619. https://doi.org/10.3390/rs12162619

Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2022). Artificial intelligence applications for industry 4.0: A literature-based study. *Journal of Industrial Integration and Management, 7*(1), 83-111. https://doi.org/10.1142/s2424862221300040

Jocher, G., Chaurasia, A., & Qiu, J. (2023). *Ultralytics YOLOv8* (Version 8.0.0) [Computer software]. Ultralytics. https://github.com/ultralytics/ultralytics/tree/v8.0.0

Koh, L. Y., & Yuen, K. F. (2022). Emerging competencies for logistics professionals in the digital era: A literature review. *Frontiers in Psychology, 13*, Article 965748. https://doi.org/10.3389/fpsyg.2022.965748

Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Ke, Z., Li, Q., Cheng, M., Nie, W., Li, Y., ong, B., Liang, Y., Zhou, L., Xu, X., Chu, X., Wei, X., & Wei, X. (2022). *YOLOv6: A single-stage object detection framework for industrial applications* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2209.02976

Liboni, L. B., Cezarino, L. O., Jabbour, C. J. C., Oliveira, B. G., & Stefanelli, N. O. (2019). Smart industry and the pathways to HRM 4.0: Implications for SCM. *Supply Chain Management: An International Journal, 24*(1), 124-146. https://doi.org/10.1108/scm-03-2018-0150

OpenCV. (n.d.-a). *Feature matching with FLANN*. https://docs.opencv.org/4.10.0/d5/d6f/tutorial_feature_flann_matcher.html

OpenCV. (n.d.-b). *Harris corner detection*. https://docs.opencv.org/4.10.0/dc/d0d/tutorial_py_features_harris.html

Radanliev, P., De Roure, D., Nicolescu, R., Huth, M., & Santos, O. (2021). Artificial intelligence and the Internet of Things in industry 4.0. *CCF Transactions on Pervasive Computing and Interaction, 3*(3), 329-338. https://doi.org/10.1007/s42486-021-00057-3

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779-788). IEEE. https://doi.org/10.1109/CVPR.2016.91

Roboflow. (n.d.). *Roboflow*. https://roboflow.com/

Sengupta, S., & Dreyer, H. (2023). Realizing zero-waste value chains through digital twin-driven S&OP: A case of grocery retail. *Computers in Industry, 148*, Article 103890. https://doi.org/10.1016/j.compind.2023.103890

Terven, J., & Cordova-Esparza, D. (2023). *A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond* [Preprint]. arXiv. https://arxiv.org/abs/2304.00501v1

Toneguzzi, M. (2022, September 21). *Walmart Canada introduces AI tech to avoid out-of-stock in stores [interview]*. Retail Insider. https://retail-insider.com/retail-insider/2022/09/walmart-canada-introduces-ai-tech-to-avoid-out-of-stock-in-stores-interview/

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research, 122*, 889-901. https://doi.org/10.1016/j.jbusres.2019.09.022

Vijayakumar, A., & Vairavasundaram, S. (2024). YOLO-based object detection models: A review and its applications. *Multimedia Tools and Applications, 83*(35), 83535-83574. https://doi.org/10.1007/s11042-024-18872-y

Villegas-Ch, W., Maldonado Navarro, A., & Sanchez-Viteri, S. (2024). Optimization of inventory management through computer vision and machine learning technologies. *Intelligent Systems with Applications, 24*, Article 200438. https://doi.org/10.1016/j.iswa.2024.200438

Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 7464-7475). IEEE. https://doi.org/10.1109/CVPR52729.2023.00721

Wassel, B. (2022, September 20). *Walmart Canada rolls out AI-powered out-of-stock detection to all locations*. Retail TouchPoints. https://www.retailtouchpoints.com/news/walmart-canada-rolls-out-ai-powered-out-of-stock-detection-to-all-locations/117730/

Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems, 30*(11), 3212-3232. https://doi.org/10.1109/tnnls.2018.2876865

Pertanika J. Sci. & Technol. 34 (5): 3131 - 3162 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Bridging the Ocean Literacy Gap: Teacher Needs and Design Requirements for Virtual Reality-Based Sustainable Maritime Education for Coastal High Schools in Tanjungpinang, Indonesia

Nurul Asikin[1]*, Arinda Eka Lidiastuti[2], Adam Fernando[1], and Rahmania Pamungkas[3]

[1]Department of Biology Education, Faculty of Teacher Training and Education, Universitas Maritim Raja Ali Haji, Tanjungpinang 29124, Indonesia
[2]Department of Biology Education, Faculty of Economics, Social Sciences, Humanities and Science Technology, Universitas Kristen Cipta Wacana, Malang 65149, Indonesia
[3]Department of Biology Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri, Yogyakarta 55281, Indonesia

ABSTRACT

Ocean literacy is an essential competence for understanding the reciprocal relationship between humans and the ocean and supporting sustainable development, particularly in coastal communities. However, including marine-related content in Biology education does not necessarily ensure the systematic integration of maritime contexts, authentic learning experiences, or multidimensional ocean literacy assessment. This study aimed to identify Biology teachers' instructional, contextual, technological, and assessment needs and to formulate evidence-based design requirements for developing Virtual Reality (VR) supported Sustainable Maritime Education (SME) in coastal high schools. We employed a convergent mixed-methods approach involving 26 Biology teachers from nine secondary schools in Tanjungpinang City, Riau Islands Province, Indonesia. We analysed quantitative data using descriptive statistics and examined open-ended questionnaire responses using a hybrid inductive-deductive thematic analysis. We integrated the two strands through a joint display to identify convergence, complementarity, and potential divergence, and to formulate empirically informed design requirements. Most teachers (84.6%) reported teaching marine and coastal ecosystem content and using educational technology, but only 15.4% frequently integrated maritime contexts into Biology learning. VR equipment was unavailable in 69.2% of the participating schools, and 84.6% of teachers indicated that a specific ocean literacy assessment instrument was needed or highly

ARTICLE INFO

Article history:
Received: 24 August 2026
Accepted: 07 October 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.24

E-mail addresses:
nurul_asikin@umrah.ac.id (Nurul Asikin)
arindaekal1@gmail.com (Arinda Eka Lidiastuti)
adamfernando@umrah.ac.id (Adam Fernando)
rahmaniapamungkas@uny.ac.id (Rahmania Pamungkas)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

needed. The integrated findings yielded seven context-specific design propositions: contextual authenticity, immersive experiential learning, pedagogical integration, ocean literacy alignment, accessibility and feasibility, realistic curriculum-aligned content, and action-oriented sustainability. These findings provide preliminary design propositions for future development, validation, and testing of SME–VR's effectiveness in strengthening ocean literacy among secondary school students.

Keywords: Biology education, coastal education, ocean literacy, sustainable maritime education, virtual reality

INTRODUCTION

The ocean is a fundamental component of the Earth system, covering approximately 70% of the planet's surface and playing critical roles in climate regulation, carbon and water cycles, biodiversity conservation, food provision, livelihoods, and the ecosystem services that underpin human wellbeing (Asikin et al., 2026; Kelly et al., 2022; Watson et al., 2020). This relationship is reciprocal, as human activities also exert substantial pressure on the ocean through resource exploitation, pollution, habitat alteration, and climate change (Fielding et al., 2019; Kelly et al., 2022). Ocean sustainability is therefore both an ecological and a social issue that requires people to understand the reciprocal relationship between humans and the ocean and make responsible decisions (Fauville et al., 2021; Mendler de Suarez et al., 2014). Within the framework of sustainable development, education has a strategic role in building such capacity through Sustainable Development Goal (SDG) 4 on quality education and SDG 14 on the conservation and sustainable use of the oceans (Asikin et al., 2023; Diz et al., 2018). UNESCO-IOC also positions ocean literacy as an important means of developing a society that understands how the ocean influences people and how human activities influence the ocean, and that can translate such understanding into responsible action (Francesca et al., 2017; Pinardi, 2019). This relevance is particularly pronounced in Indonesia as an archipelagic country whose ecological, social, and economic systems are closely connected to coastal and marine environments. This condition positions education as a strategic instrument for developing ocean literacy at the secondary education level.

Ocean literacy (OL) refers to the capacity to understand the reciprocal relationship between the ocean and humans and to apply that understanding in decision-making and responsible action toward the marine environment (Asikin et al., 2023; Francesca et al., 2017; Kelly et al., 2022). The concept of OL has evolved from an initial emphasis on marine knowledge toward a more multidimensional construct encompassing knowledge, awareness, attitudes, behaviours, experience, emotional connectedness, communication, and the capacity to act (Costa et al., 2021; Koutsopoulos & Stel, 2021; Mokos et al., 2020; Zielinski et al., 2022). Thus, knowing about marine ecosystems does not necessarily mean an individual has attitudes and behaviours that support ocean sustainability (Asikin et al., 2026). In education, OL therefore needs to be developed through learning experiences

that connect scientific understanding with environmental contexts, values, attitudes, and actions. A systematic review of OL teaching in science education indicates that curriculum integration, project-, problem-, and inquiry-based learning, the use of technology, and the development of assessment instruments that address knowledge, attitudes, and behaviours are important requirements for marine education (Asikin et al., 2026). This issue matters in coastal communities because proximity to the ocean does not automatically lead to greater understanding of, or responsible behaviour toward, the marine environment (Heck et al., 2018; Kasten et al., 2021; Pazoto et al., 2022). Low levels of ocean literacy indicate that students need learning experiences that go beyond transmitting knowledge about the ocean and connect that knowledge with attitudes and behaviours toward the marine environment (Asikin et al., 2023).

At the secondary education level, Biology provides relevant opportunities to develop OL through topics such as ecosystems, biodiversity, organism–environment interactions, environmental change, and conservation (Lucrezi et al., 2018; Ward et al., 2022). However, the presence of marine-related content in the curriculum does not guarantee that maritime contexts are systematically integrated into students' learning experiences (Asikin et al., 2025; Shi & Fan, 2021). Learning may remain focused on conceptual transmission and rely on representations such as text, images, or videos without providing sufficient opportunities for students to explore marine environments contextually (Sahriana et al., 2020; Shi & Fan, 2021; Shih, 2024). This challenge is further complicated by the fact that marine ecosystems are relatively difficult to access directly. Phenomena such as coral reef structures, seagrass communities, organism interactions, habitat changes, and the impacts of human activities on coastal ecosystems cannot always be routinely observed by students (Asikin et al., 2025; Shi & Fan, 2021). Field activities can provide authentic experiences, but site accessibility, cost, time, weather conditions, safety, and school facilities may constrain their implementation. Furthermore, assessments used in learning do not always measure OL multidimensionally (Lin et al., 2020; O'Brien et al., 2023). Systematic evidence indicates that OL assessment instruments still need to be developed to measure not only knowledge but also attitudes and behaviours toward the ocean (Asikin et al., 2026). These conditions indicate a gap between the availability of marine content in Biology education and the availability of learning experiences that can develop ocean literacy in contextual and multidimensional ways.

Virtual Reality (VR) technology offers an opportunity to address some of these limitations in learning experiences (Asikin et al., 2025; Liu et al., 2020). VR enables learners to interact with three-dimensional digital environments and experience immersive representations of objects or environments that are difficult to access directly (Hite, 2022; Huang et al., 2019). In marine education, these characteristics matter because virtual environments can visualise and support exploration of underwater environments, organisms,

habitat structures, and ecological processes (Fauville et al., 2021, 2026; Frank et al., 2026). Research on underwater virtual reality shows that immersive experiences can generate high levels of presence and evoke affective responses such as awe, empathy, and flow, potentially reducing students' psychological distance from marine environments (Fauville et al., 2026). Research in science education has also shown that VR can provide interactive learning experiences, and systematic evidence identifies VR as a technology with the potential to support OL learning (Asikin et al., 2026). However, this potential does not mean that using VR automatically produces better learning outcomes. VR's pedagogical value depends on alignment among technological characteristics, learning objectives, student activities, environmental context, and assessment. Therefore, VR should be positioned as part of a learning-experience design rather than as a substitute for pedagogy or merely as a medium for displaying three-dimensional objects (Frank et al., 2026; Gomez, 2020; Hite, 2022). VR can make meaningful pedagogical contributions only when designed as part of a structured, contextualised, reflective learning experience aligned with ocean literacy objectives.

Although research on OL and VR in education continues to expand, an important gap remains in learning design. Previous studies have examined OL development in science education and explored VR's potential to provide immersive marine experiences. However, few studies have systematically integrated teachers' needs, local maritime contexts, pedagogical constraints, technological readiness, and OL assessment practices into VR-based learning design (Asikin et al., 2025). Recent evidence indicates that technologies such as VR have been used to create interactive experiences in OL education, while also emphasising the need to integrate curriculum, innovative pedagogy, technology, and assessment appropriately (Asikin et al., 2026). Even underwater VR studies that explicitly connect VR with OL have largely focused on technological potential and users' psychological experiences. Further research is therefore needed to translate these potentials into learning-design requirements that are appropriate for school contexts and teachers' needs (Fauville et al., 2026; Frank et al., 2026). This gap matters for secondary schools in Indonesia's coastal regions because effective learning design must account for local environmental characteristics, Biology content, students' abilities and needs, teacher readiness, technological infrastructure, and assessment practices. Therefore, before developing VR and testing its effectiveness, an evidence-based needs analysis is required to identify the pedagogical, contextual, technological, and assessment requirements that can serve as the foundation for Sustainable Maritime Education (SME).

Although research on ocean literacy (OL) and VR in education continues to expand, a key gap remains in understanding how teachers' locally situated needs, instructional practices, technological constraints, and assessment priorities can inform the design of maritime learning. VR can represent underwater environments and processes that may be difficult to access directly, but it is not the only approach. Resources such as 360-degree

media, simulations, multimedia, virtual field trips, and direct field experiences may also address some learning needs, separately or in combination. The present study did not compare these alternatives. Instead, it examined teacher-reported needs and constraints within a research programme that proposes VR-supported SME as the development direction. Therefore, the study does not seek to establish VR as the optimal or superior solution; it seeks to identify context-specific conditions to consider when developing VR for this setting.

Based on this gap, this study aimed to identify Biology teachers' instructional, contextual, technological, and assessment needs and to integrate quantitative and qualitative evidence to inform provisional design propositions for VR-supported learning within Sustainable Maritime Education in coastal secondary schools. The analysis focused on current Biology teaching practices related to marine content, maritime contextualisation, perceived learning difficulties and experiential needs, technology and VR readiness, priority marine content, pedagogical expectations, and ocean literacy assessment practices. The study contributes context-specific evidence from 26 teachers across nine secondary schools in Tanjungpinang City. Because students did not participate in this phase, references to student learning needs represent teachers' perceptions rather than direct evidence from students. The propositions are intended to guide subsequent development and validation, not to constitute a validated design framework.

THEORETICAL BACKGROUND

Ocean Literacy as a Multidimensional Educational Construct

Ocean literacy (OL) was initially developed to enhance public understanding of the ocean and the reciprocal relationship between humans and the marine environment. Conceptually, OL refers to understanding how the ocean influences humans and how human activities influence the ocean, including the ability to understand marine-related information and use such knowledge to make responsible decisions (Cava, 2005; Francesca et al., 2017; NOAA, 2020). Over time, OL has evolved beyond the acquisition of factual knowledge about marine ecosystems and is now conceptualised as a multidimensional construct encompassing knowledge, awareness, attitudes, behaviours, experiences, emotional connectedness, communication, and the capacity to act (Asikin et al., 2026; Bastos et al., 2026; McKinley et al., 2023). This conceptual expansion suggests that individuals may know marine concepts without necessarily developing concern for, or taking actions that support, ocean sustainability. Therefore, OL education needs to give students opportunities to connect scientific understanding, environmental experiences, values, attitudes, and action (Kelly et al., 2022; McKinley et al., 2022). In the present study, knowledge, attitude, and behaviour are the three primary dimensions because they represent the relationships among cognitive understanding, affective orientation, and behavioural tendencies toward

the marine environment. Selecting these three dimensions also provides an operational basis for developing and evaluating secondary-school learning.

Place-Based and Contextual Environmental Education

Place-based education provides a theoretical foundation for explaining why local environmental contexts are important in environmental education (Almeida & Cutter-Mackenzie, 2011; Asikin et al., 2026; Otto & Pensini, 2017). This approach positions place, environment, community, and learners' lived experiences as sources of learning, ensuring that acquired knowledge remains connected to the environmental conditions students know (Fauville et al., 2021; Kao et al., 2026; McPherson et al., 2018). In marine education, a place-based approach is particularly relevant because students' relationships with the ocean are shaped not only by scientific knowledge but also by their experiences, geographical proximity, local identity, and understanding of environmental issues in their surrounding areas (Bastos et al., 2026; Freitas et al., 2026; Mothershed & Stilwell, 2025). Research on place-based marine education indicates that connections with local environments can strengthen students' understanding, perceptions, and relationships with marine environments and support the development of concern for the ocean (Bastos et al., 2026; Freitas et al., 2026; Mothershed & Stilwell, 2025). This framework has direct implications for SME development because maritime learning is insufficient when it relies solely on generic ocean representations disconnected from students' lives. Content such as mangroves, seagrass, coral reefs, marine biodiversity, and coastal pollution has greater pedagogical value when situated within environmental contexts that are relevant to students' lived experiences (Asikin et al., 2026). Therefore, this study positions place-based environmental education as a principle for ensuring that SME–VR achieves contextual authenticity, namely the capacity to represent marine ecosystems and maritime issues relevant to the coastal contexts in which students live and learn.

Experiential Learning for Marine Environmental Education

Experiential learning provides a theoretical basis for understanding how experience can develop environmental knowledge and competencies. From the perspective of Kolb & Kolb (2017), learning occurs through a cycle involving concrete experience, reflective observation, abstract conceptualisation, and active experimentation. This framework is particularly relevant to marine education because ocean-related objects and phenomena are often difficult to understand solely through verbal descriptions or two-dimensional visual representations (Asikin et al., 2023; Williams, 2024). Students need opportunities to observe phenomena, interpret information, connect observations with scientific concepts, and apply their understanding to analyse environmental problems. In marine education, direct experiences through field observations, coastal exploration, organism observation,

or conservation activities can provide concrete experiences that strengthen connections between knowledge and the environment (Ditrich & Lachmair, 2026; Frank et al., 2026; Zhang et al., 2026). However, direct experiences with marine ecosystems cannot always be conducted routinely because of geographical limitations, safety considerations, weather conditions, costs, time constraints, site accessibility, and available facilities. Technology can therefore complement direct experiences, particularly when learning objects or environments are difficult to access (Ditrich & Lachmair, 2026; Fauville et al., 2026; Hetherington et al., 2026). In this context, VR can provide technology-mediated experiences that allow students to explore environments that are difficult to observe directly (Fauville et al., 2026; Frank et al., 2026). Thus, within SME–VR, virtual experiences are not intended to replace real environmental experiences but rather to extend access to marine experiences when direct experiences are difficult to provide.

Immersive Learning and the Pedagogical Affordances of VR

Virtual Reality has characteristics that distinguish it from conventional learning media, as it allows users to interact with three-dimensional digital environments and experience a sense of presence within them (Elme et al., 2022; Makransky et al., 2020; Petersen et al., 2020). In education, these characteristics offer several pedagogical affordances, including representing environments that are difficult to access, facilitating exploration from multiple perspectives, and providing more interactive learning experiences (Di Natale et al., 2020; Han, 2020; Jensen & Konradsen, 2018). Bower & Jong (2020) showed that immersive VR can support various learning activities, particularly when virtual environments are designed around learning objectives and appropriate activities. Hamilton et al. (2021) Likewise, reported generally positive evidence regarding the effects of immersive VR on several learning outcomes. However, its effectiveness is not automatic and depends on the characteristics of the intervention and the research design. More specifically, the Cognitive Affective Model of Immersive Learning (CAMIL) provides a framework for explaining the mechanisms underlying learning in immersive VR (Makransky & Petersen, 2021). Makransky & Petersen (2021) proposed that immersive VR characteristics can influence presence and agency, which in turn relate to cognitive and affective processes during learning. Thus, immersion should not be regarded as a learning outcome, but as a condition that may facilitate cognitive and affective experiences (Cheng & Tsai, 2020; Gomez, 2020; Makransky & Petersen, 2021). Marine contexts provide a particularly relevant setting for these affordances because underwater environments, habitat structures, organism interactions, and ecosystem changes are not always directly observable by students (Kelly et al., 2022; Markowitz et al., 2018). Therefore, VR has potential as a marine learning environment; however, educators must translate that potential into practice through effective pedagogical design.

Pedagogical Integration: From VR Experience to Meaningful Learning

VR in education cannot be separated from instructional design, which guides students in interpreting and reflecting on their virtual experiences (Asikin et al., 2025; Han, 2020; Makransky & Petersen, 2021). Immersive experiences without clear learning objectives, investigative questions, analytical activities, or reflection may remain primarily visual experiences and fail to produce deep conceptual understanding (Liu et al., 2020; Moro et al., 2017; Yahaya, 2007). Therefore, VR should be positioned as part of a learning sequence that connects experience with cognitive processes, social interaction, and action (Adams et al., 2021; Armougum et al., 2019; Paulsen et al., 2024). In the SME–VR context, pedagogical integration can follow a sequence of problem orientation, VR exploration, investigation, analysis, discussion and reflection, decision-making, and sustainability-oriented action. During the exploration stage, students experience marine environments through VR; during the investigation stage, they identify relevant phenomena or problems; during the analysis stage, they connect their observations with Biology concepts; and during reflection and decision-making, they consider the ecological and social consequences of alternative actions. The final stage can connect virtual experiences with real-world activities such as environmental campaigns, waste management, mangrove conservation, or coastal ecosystem protection (Asikin et al., 2025; McKinley et al., 2023). Thus, the central principle of SME–VR is not simply to use VR to teach about the ocean, but to use VR as part of a learning experience designed to understand, reflect on, and respond to marine issues. This positioning frames VR as a pedagogical affordance that must be aligned with learning objectives. It explains why teachers' needs extend beyond a VR application itself to include instructional modules, investigative activities, discussion, reflection, and ocean-literacy assessment.

Sustainable Maritime Education as an Integrative Framework

Building on these theoretical foundations, this study positions Sustainable Maritime Education (SME) as an integrative educational framework that connects maritime contexts, learning experiences, immersive technology, ocean literacy, and sustainability-oriented learning. SME is not treated as a standalone theory but as a conceptual framework that integrates several complementary educational perspectives. Ocean literacy defines the competencies to be developed (Asikin et al., 2025); place-based environmental education determines the learning context (Almeida & Cutter-Mackenzie, 2011; Otto & Pensini, 2017); experiential learning explains the importance of experience as a source of knowledge construction (Kolb & Kolb, 2017; Roland, 2017); while immersive learning and CAMIL explain how VR can provide experiences that are difficult to obtain directly (Makransky & Petersen, 2021). These components then support the goals of environmental education:

developing learners' capacity to understand environmental issues, develop responsible attitudes, and take actions that support sustainability. This integrative framework matters because sustainability-oriented education requires more than transmitting environmental knowledge. It requires learning experiences that connect knowledge with context, experience, reflection, values, decision-making, and action. Within this framework, SME–VR integrates what students need to learn through ocean literacy, where learning takes place in local maritime contexts, learning experiences are constructed through experiential and immersive learning, and learning is directed toward sustainability-oriented action. Accordingly, SME–VR provides the conceptual basis for translating the needs-analysis findings into concrete learning-design requirements, including contextual authenticity, immersive experiential learning, pedagogical integration, ocean-literacy alignment, accessibility and feasibility, realistic curriculum-aligned content, and action-oriented sustainability.

METHODOLOGY

Research Design

This study used a convergent mixed methods design to investigate Biology teachers' needs, perceptions, current instructional practices, and contextual constraints related to integrating Virtual Reality (VR) into Sustainable Maritime Education (SME) to develop ocean literacy in coastal senior high schools. Quantitative and qualitative data were collected during the same needs-analysis phase from the same 26 participants using a single questionnaire. Closed-ended items provided quantitative data, while open-ended responses provided qualitative material. The strands were analysed separately, then integrated through a joint display. This design allowed the open-ended responses to contextualise and expand patterns in the closed-ended data; however, because both strands shared the same instrument and participants, they do not constitute independent sources of corroboration. The resulting integrated interpretations are therefore treated as context-specific meta-inferences rather than as independent validation of a design framework. This approach aligns with established mixed-methods principles, which call for explicit integration rather than presenting quantitative and qualitative findings as two unrelated strands (Fetters & Tajima, 2022; Han, 2020).

Research Context

The study was conducted in Tanjungpinang City, Riau Islands Province, Indonesia, a coastal and maritime region where marine and coastal environments are an important part of the local geographical and socio-environmental context. The study focused on secondary school Biology education because Biology curricula offer opportunities to address marine

ecosystems, biodiversity, environmental change, and human–environment interactions. The study treated the coastal setting as an educational context rather than only as a sampling location. The study examined how maritime contexts were incorporated into Biology learning and which technology-supported learning approaches could provide meaningful access to marine environments. The findings are therefore interpreted primarily within the participating Tanjungpinang context, while the resulting design requirements may provide propositions for examination in other coastal settings.

Participants and Sampling

The participants were 26 Biology teachers from nine secondary schools in Tanjungpinang City. The researchers purposively recruited participants because they had direct experience teaching Biology at the senior high school level and could provide informed perspectives on marine-related Biology content, instructional practices, technology use, and perceived student learning needs. The participant group included 25 female teachers (96.2%) and one male teacher (3.8%). Eighteen teachers (69.2%) had more than 10 years of teaching experience, six (23.1%) had 5–10 years, and two (7.7%) had less than five years. Participation was voluntary. Before completing the questionnaire, participants were informed about the study purpose, voluntary participation, intended data use, and confidentiality. We anonymised schools and teachers in reporting and identified qualitative quotations using participant codes R1–R26.

Data Sources and Instruments Development

We collected data using a structured needs-analysis questionnaire developed specifically for this study, which gathered both quantitative and qualitative evidence to inform the development of SME–VR. The questionnaire was developed from the research objectives, relevant literature on ocean literacy, contextual and experiential learning, immersive learning, and VR-supported science education, as well as the biology curriculum and the coastal educational context of Tanjungpinang. The instrument was organised into six domains: (1) existing Biology learning practices; (2) maritime contextualization; (3) marine learning needs and difficult-to-teach content; (4) technology and VR readiness; (5) priority content and learning requirements for VR; and (6) ocean literacy assessment practices. The questionnaire included categorical and multiple-response items, rating-scale items, and open-ended questions. Closed-ended items identified the distribution and priority of teachers' responses, while open-ended items sought contextual explanations of perceived learning difficulties, implementation constraints, expectations for VR-supported learning, and ocean literacy assessment needs.

Before administration, the research team discussed the questionnaire, and three experts with relevant expertise in Biology education, educational technology/VR-

supported learning, and marine or environmental education reviewed it. The experts reviewed the items' relevance and clarity, their alignment with the study objectives and intended domains, and their suitability for Biology teachers in the coastal educational context. We revised the instrument based on the expert review before administration. We did not conduct a separate pilot test with an independent sample. This limitation is acknowledged explicitly. We designed the questionnaire as a multidomain needs-analysis instrument rather than a unidimensional psychometric scale. Because its items represented heterogeneous categorical, multiple-response, rating-scale, and open-ended formats and were not intended to measure a single latent construct, the questionnaire did not have a single internal consistency coefficient, such as Cronbach's alpha. The three-expert review, therefore, served as evidence of content relevance and clarity rather than a substitute for psychometric reliability.

Data Collection Procedure

Data collection was conducted during the initial needs-analysis phase of the SME–VR research program. Participants completed the questionnaire individually and were encouraged to provide detailed responses to the open-ended questions. The same participant group contributed both quantitative and qualitative data, allowing direct integration between the two strands. Quantitative items identified the distribution and relative priority of teachers' responses, while open-ended items explained why they reported needs, challenges, or preferences. For example, the quantitative pattern of difficulty in teaching marine ecosystems was interpreted alongside qualitative explanations of limited access to underwater environments, constraints on field observation, and difficulty visualising complex ecological interactions. This structure enabled the qualitative data to provide contextual depth rather than merely supplementary quotations. This approach matters in mixed-methods research because integration should generate understanding beyond what either quantitative or qualitative analysis can provide independently (Boda & Brown, 2020; Fetters & Tajima, 2022; Han, 2020).

Quantitative Data Analysis

We analysed quantitative data using descriptive statistics. Frequencies (n) and percentages (%) were calculated for categorical and multiple-response items. The analysis focused on current instructional practices, maritime contextualisation, technology and VR readiness, priority marine content, and ocean literacy assessment practices and needs.

Because the sample was purposive, geographically restricted, and relatively small, the study did not use inferential tests to make population-level generalisations. We interpreted the quantitative results as context-specific evidence from the participating coastal schools.

Qualitative Data Analysis

We analysed the open-ended questionnaire responses using hybrid inductive-deductive thematic analysis. The deductive component drew on sensitising concepts from the theoretical framework, particularly contextual learning, experiential learning, pedagogical integration, technology readiness, ocean literacy, and sustainability-oriented learning. These concepts guided attention to theoretically relevant issues but were not treated as predetermined final themes. This hybrid approach is consistent with methodological work showing that theory-informed and data-driven coding can be combined within mixed-methods research (Fereday & Muir-Cochrane, 2006; Proudfoot, 2023). The inductive component allowed additional codes to emerge directly from teachers' responses. The analysis involved repeated familiarisation with the responses, identifying meaningful segments, generating initial codes, comparing and refining codes, grouping related codes into broader categories, developing and reviewing themes, and defining final themes. The initial coding framework included 16 codes covering issues such as limited authentic maritime experience, the need for realistic and immersive visualisation, contextual and active pedagogy, technology and infrastructure constraints, gaps in ocean literacy assessment, and sustainability-oriented action. The team then organised these codes into six broader themes.

To strengthen the research team's analytic credibility, the team reviewed the initial coding framework and thematic structure. The team discussed and resolved coding differences through consensus, and refined code definitions as needed. We maintained an audit trail through the codebook, respondent-level coding records, thematic structure, and joint display. We report no inter-coder reliability coefficient because the study did not use independent parallel coding to calculate one. Because the qualitative data came from open-ended questionnaire responses rather than individual interviews or focus groups, the qualitative strand was intended to provide contextual explanations of quantitative patterns rather than claim the depth associated with extended qualitative interviewing. Representative quotations were translated into English while preserving their intended meaning and were identified using anonymous participant codes R1–R26.

Mixed-Methods Integration & Development of Design Requirements

The quantitative and qualitative findings were integrated through a joint display. Joint displays can function as analytic tools for examining relationships between quantitative and qualitative evidence and for developing integrated interpretations (Haynes-Brown & Fetters, 2021; Reeping & Edwards, 2022). The joint display in this study contained four components: quantitative findings, qualitative explanations, integrated inferences, and design implications. Four criteria determined whether an integrated finding could be translated into a design requirement. First, the finding had to demonstrate clear empirical

relevance to teachers' reported needs, challenges, priorities, or instructional practices. Second, it had to be supported or meaningfully expanded by the qualitative evidence rather than relying on an isolated response. Third, it had to have a direct and operational implication for SME–VR design. Fourth, the implication had to be consistent with the educational purposes of Biology learning, ocean literacy, and sustainability-oriented maritime education. We treated convergence as confirmation when both strands supported the same interpretation. When qualitative evidence added context to a quantitative pattern, we treated the relationship as complementary or expansive. We examined potentially divergent evidence before formulating an integrated inference. This procedure moved beyond parallel reporting toward meta-inferences grounded in both strands. The final seven SME–VR design requirements were therefore not treated as predetermined outcomes. The team developed them iteratively by linking theoretical sensitising concepts, empirical findings, thematic interpretations, and integrated meta-inferences. The requirements are consequently interpreted as empirically informed design requirements for subsequent development and validation, not as validated intervention principles.

Trustworthiness and Quality Assurance

For the quantitative strand, the team checked data for completeness, response consistency, and coding accuracy before analysis. For the qualitative strand, the team reviewed themes against the original responses to ensure traceability between interpretations and participants' accounts. The research team reviewed the coding framework and thematic definitions and retained an audit trail through respondent-level coding, the codebook, thematic structure, and joint display. For mixed-methods integration, the quantitative and qualitative strands were first analysed separately, then integrated through a joint display. The integration explicitly considered convergence, complementarity, and potential divergence before formulating design implications. These procedures made the analytical pathway from raw responses to design requirements transparent.

Ethical Considerations

Participation in the study was voluntary. Before completing the questionnaire, participants were informed about the study's purpose and procedures, the voluntary nature of participation, the confidentiality of their responses, and the use of their data for research purposes. Formal ethics approval was not required for this study under the applicable institutional procedures because it involved a non-interventional educational needs analysis based on voluntary teacher participation and did not involve clinical procedures, interventions, or the collection of personally identifying or sensitive information. Informed consent was obtained from all participants before their participation. Participants could decline participation without consequence. No personally identifying information was

reported, and anonymous participant codes (R1–R26) were used for quotations. The study data were handled confidentially and used solely for research purposes.

RESULT AND DISCUSSIONS

Characteristics of the Participants

A total of 26 Biology teachers participated in the needs analysis. The participants represented nine secondary schools in Tanjungpinang City, Riau Islands Province, Indonesia. As presented in Table 1, Most participants were female (25, 96.2%), while one participant was male (3.8%). In terms of teaching experience, 18 teachers (69.2%) had more than 10 years, 6 (23.1%) had 5 to10 years, and 2 (7.7%) had less than 5 years. The sample was predominantly composed of experienced Biology teachers. Given the purposive sample and local setting, these characteristics should be considered when interpreting the findings.

Existing Biology Learning Practices and Maritime Contextualisation

The quantitative findings indicate that teachers had already adopted student-centred learning approaches. As shown in Table 2, Problem-Based Learning (PBL) was the most frequently reported model (17, 65.4%), followed by Project-Based Learning (PBL) and Discovery Learning (each 3, 11.5%). However, teachers had not yet consistently integrated maritime contexts into Biology instruction. Although 22 teachers (84.6%) reported teaching marine and coastal ecosystem content, only four teachers (15.4%) reported frequently integrating maritime contexts into their instruction. Most teachers (18, 69.2%) reported

Table 1
Participant Characteristics (N = 26)

Characteristic	Category	n	%
Gender	Female	25	96,2
	Male	1	3,8
Teaching experience	<5 years	2	7,7
	5–10 years	6	23,1
	>10 years	18	69,2
School	School 1	8	30,8
	School 2	5	19,2
	School 3	4	15,4
	School 4	3	11,5
	School 5	2	7,7
	School 6	1	3,8
	School 7	1	3,8
	School 8	1	3,8
	School 9	1	3,8

Table 2
Existing Biology Learning Practices and Maritime Contextualisation

Indicator	Category	n	%
Instructional model	Problem-Based Learning (PBL)	17	65.4
	Project-Based Learning (PBL)	3	11.5
	Discovery Learning	3	11.5
	Other approaches	3	11.5
Teaching marine/coastal ecosystem content	Yes	22	84.6
	No	4	15.4
Frequent integration of maritime context	Yes	4	15.4
	No/limited	22	84.6

integrating maritime contexts only occasionally, while four (15.4%) reported doing so rarely. The qualitative findings explain this pattern: teachers described difficulty connecting textbook-based materials with local marine environments and bringing authentic maritime experiences into the classroom. One participant stated, "It is difficult to bring authentic experiences into the classroom." (R2)

These findings distinguish between teaching marine content and systematically contextualising Biology learning within maritime environments. The quantitative pattern shows the extent of this gap, while the qualitative responses indicate that limited authentic access is one reason for it.

Priority Learning Content and Experiential Learning Needs

The findings indicate that marine ecosystem topics were among the biology topics considered relatively difficult to teach. As shown in Table 3, Teachers identified marine ecosystems as the most difficult Biology content to teach, reported by 18 teachers (69.2%), followed by marine biodiversity (15, 57.7%), marine food chains (7, 26.9%), and marine pollution (5, 19.2%). These responses were multiple-response items, so percentages do not sum to 100%. The open-ended responses provided contextual explanations for these difficulties. Teachers repeatedly referred to limited opportunities for direct observation, restricted access to marine environments, and difficulty visualising complex ecological processes. One teacher stated, "The main limitation is the difficulty of seeing the marine ecosystem directly" (R13). Participants repeatedly identified underwater exploration as difficult to conduct in real-world settings. Other challenging activities included observing interactions among marine organisms, identifying marine biodiversity, examining food webs, and monitoring environmental changes. These qualitative findings complement the quantitative pattern by indicating that difficulty is associated not only with conceptual complexity but also with limited opportunities for authentic observation and experiential engagement.

Table 3
Priority Learning Content Considered Difficult to Teach

Priority Learning Content Considered Difficult to Teach	n	%
Marine ecosystems	18	69.2
Marine biodiversity	15	57.7
Marine food chains	7	26.9
Marine pollution	5	19.2

Priority Content for VR Development

Teachers identified several marine topics as priorities for development into VR-based learning content. As presented in Table 4, seagrass was the most frequently selected topic, identified by 17 teachers (65.4%). Mangroves and marine biota were each selected by 16 teachers (61.5%), followed by coral reefs, selected by 15 teachers (57.7%). The findings indicate that teachers prioritised marine topics that are both ecologically important and relevant to the coastal environment surrounding students. Teachers also emphasised that VR content should be realistic, natural, locally relevant, and aligned with the biology curriculum.

Teachers specifically requested realistic visualisations rather than predominantly animated representations. One participant stated:

> The visual presented through the VR device should be real video footage of the actual condition being studied.
>
> (R16)

Another participant emphasised:

> The development should use real videos rather than animation, so students can actually see and experience the system being studied.
>
> (R19)

Teachers also indicated that VR should not be developed as a stand-alone virtual tour. Instead, the VR experience should be integrated into student worksheets, teaching modules, investigations, discussions, reflections, and assessments.

Table 4
Priority Marine Content for VR Development

Rank	Marine content	n	%
1	Seagrass	17	65,4
2	Mangrove	16	61,5
2	Marine biota	16	61,5
4	Coral reef	15	57,7
5	Marine pollution	13	50,0

Technology Readiness and Implementation Constraints

The quantitative findings revealed a discrepancy between teachers' general experience with educational technology and the availability of VR infrastructure in schools. As shown in Table 5, 22 teachers (84.6%) reported having previously used technology in their teaching. However, 18 teachers (69.2%) reported that VR devices were not available at their schools. This finding indicates that general digital experience does not automatically translate into institutional readiness for VR-based learning.

Teachers identified several constraints that could affect implementation, including device availability, cost, infrastructure, instructional time, and teacher readiness. The qualitative findings reinforced these concerns. Teachers emphasised that VR implementation should consider device availability and supporting infrastructure, affordability, teacher competence, and classroom time. One teacher identified "permission, time, cost, and safety" as constraints on marine field activities (R5). These findings suggest that SME–VR should be designed as an accessible, feasible learning solution rather than assuming schools have dedicated VR infrastructure.

Ocean Literacy Assessment Practices

Teachers' practices for assessing students' ocean literacy showed a substantial gap. Of the 26 participating teachers, 17 (65.4%) reported previously assessing ocean literacy without a structured instrument. Eight teachers (30.8%) had never assessed students' ocean literacy, whereas only one teacher (3.8%) reported using a structured instrument specifically designed for ocean literacy. Regarding integration into Biology assessment, 13 teachers (50.0%) reported fairly integrating ocean literacy, 10 teachers (38.5%) reported limited integration, and 3 teachers (11.5%) reported no integration. Table 6 presents the teachers' practices and needs related to ocean literacy assessment.

The need for a specific ocean literacy assessment instrument was particularly evident. Eight teachers (30.8%) considered such an instrument highly needed, while 14 (53.8%) considered it needed. Thus, 22 teachers (84.6%) expressed a need for a specific instrument

Table 5
Technology Readiness and VR Implementation

Indicator	Category	n	%
Previous use of educational technology	Yes	22	84.6
	No	4	15.4
Availability of VR devices	Not available	18	69.2
	Limited availability	5	19.2
	Available	2	7.7
	Highly available	1	3.8

Table 6
Ocean Literacy Assessment Practices and Needs

Indicator	Category	n	%
Previous ocean literacy assessment	Structured assessment	1	3.8
	Previously assessed, but unstructured	17	65.4
	Never assessed	8	30.8
Integration into Biology assessment	Fairly integrated	13	50.0
	Less integrated	10	38.5
	Not integrated	3	11.5
Need for a specific instrument	Highly needed	8	30.8
	Needed	14	53.8
	Less needed	3	11.5
	Not needed	1	3.8

to assess students' ocean literacy. The teachers' responses also indicate that ocean literacy should be assessed beyond factual knowledge. Knowledge of marine and coastal ecosystems was identified by 21 teachers (80.8%) as an important assessment component, followed by marine environmental behaviour (19 teachers, 73.1%), attitudes toward the marine environment (16 teachers, 61.5%), critical thinking about marine issues (16 teachers, 61.5%), and marine-related problem-solving ability (16 teachers, 61.5%). These findings indicate that teachers perceive ocean literacy as a multidimensional competence that extends beyond students' cognitive understanding of marine ecosystems. The qualitative responses further explain this assessment gap. One teacher stated, "Ocean literacy is not only about knowledge, but also attitudes, concern, and behaviour toward marine and coastal ecosystems. This makes it difficult to measure using written tests" (R2). Another teacher identified the absence of "standardised and structured instruments" as a major constraint (R7). These responses reinforce the quantitative findings that existing assessment practices remain predominantly unstructured and that teachers require an assessment instrument capable of capturing multiple dimensions of ocean literacy.

Qualitative Themes and Subthemes

The thematic analysis generated six major themes describing teachers' experiences, perceived challenges, and expectations regarding the development of SME–VR. The themes emerged from recurring patterns in the open-ended responses and were then linked to the quantitative findings. Table 7 presents the qualitative themes, subthemes, and representative evidence derived from the teachers' responses.

The qualitative themes demonstrate that teachers' needs were interconnected. Limited access to authentic marine environments created a need for realistic, immersive visualisation. However, teachers expected these experiences to be embedded in active,

Table 7
Qualitative Themes, Subthemes, and Representative Evidence

Theme	Key subthemes	Representative evidence
Limited authentic maritime experience	Difficulty bringing authentic marine experiences into classrooms; limited access to real marine environments	"It is difficult to bring authentic experiences into the classroom." (R2)
Need for realistic and immersive visualisation	Underwater exploration; realistic marine environments; visualisation of complex ecological processes	"The visual presented through the VR device should be real video footage of the actual condition being studied." (R16)
Need for contextual and active pedagogy	PBL/PjBL; inquiry; investigation; student worksheets; discussion and reflection	Teachers requested integrating VR with worksheets, modules, assessments, and investigative activities.
Technology and implementation constraints	Device availability; cost; infrastructure; time; teacher readiness	Teachers emphasised the availability of devices, cost, infrastructure, and teacher readiness as implementation considerations.
Ocean literacy assessment gap	Lack of specific instruments; difficulty assessing attitudes and behaviour; cognitive-dominant assessment	"It is difficult to measure [ocean literacy] using written tests." (R2)
Sustainability-oriented action	Beach clean-ups; mangrove planting; waste management; conservation campaigns	Teachers proposed beach clean-ups, mangrove planting, waste management, and environmental campaigns.

contextual learning activities. At the same time, technological constraints influenced implementation feasibility, while the assessment gap indicated the need for explicit alignment between SME–VR and multidimensional ocean literacy.

Integration of Quantitative and Qualitative Findings

The study integrated the quantitative and qualitative findings through a joint display. Quantitative results established the magnitude and priority of needs, whereas qualitative findings explained or expanded the contextual reasons underlying those patterns. The joint display served as an analytic step rather than only a presentation device. The integrated evidence showed convergence across the two strands for the main design issues. Qualitative evidence generally elaborated on the quantitative patterns by explaining why teachers experienced difficulties or prioritised certain areas. No major contradiction requires discarding a quantitative or qualitative interpretation. Table 8 presents a joint display that integrates the quantitative and qualitative findings and their implications for SME–VR design.

Table 8
Joint Display of Integrated Quantitative and Qualitative Findings

Integrated issue	Quantitative finding	Qualitative explanation	Integrated inference	Design implication
Maritime contextualization gap	84.6% teach marine/coastal ecosystem content, but only 15.4% frequently integrate maritime contexts	Teachers reported limited authentic examples and difficulty connecting textbook content with local marine environments	Marine content is present, but systematic maritime contextualization remains limited	Contextual maritime learning
Experiential learning constraint	69.2% identify marine ecosystems as difficult content	Underwater exploration, direct observation, complex interactions, and long-term changes are difficult to conduct	Learning difficulty is partly related to limited experiential access	Immersive experiential learning
Priority content	Seagrass 65.4%; mangrove/marine biota 61.5%; coral reef 57.7%	Teachers requested realistic, natural, local, and curriculum-aligned content	VR content should be selected according to contextual relevance and learning difficulty	Localised realistic VR content
Technology readiness gap	84.6% have used technology, but 69.2% report no VR availability	Cost, devices, infrastructure, time, and teacher readiness constrain implementation	Digital familiarity does not automatically imply VR readiness	Accessible and feasible VR design
Pedagogical integration	PBL is used by 65.4%	Teachers requested inquiry, projects, student worksheets, discussion, and reflection around VR	VR should function within a pedagogical sequence rather than as a stand-alone medium	Pedagogically integrated VR
Assessment gap	84.6% need a specific ocean literacy instrument	Teachers reported difficulty assessing attitudes, behaviour, and environmental concern	Ocean literacy assessment needs to become multidimensional and structured	Knowledge–attitude–behaviour assessment

SME–VR Design Requirements

The integrated findings were subsequently translated into seven design requirements for VR-supported Sustainable Maritime Education (SME-VR), as presented in Table 9.

Table 9
Design Requirements for VR-Based Sustainable Maritime Education

Requirement	Evidence from the findings	Proposed design specification
DR1. Contextual authenticity	Maritime content is taught, but has not been consistently contextualised	Use local coastal and marine environments and authentic maritime issues
DR2. Immersive experiential learning	Direct observation and underwater exploration are difficult	Provide immersive exploration and simulation of difficult-to-access environments
DR3. Pedagogical integration	Teachers support PBL, PjBL, inquiry, student worksheets, discussion, and reflection	Integrate VR into a structured learning sequence
DR4. Ocean literacy alignment	Knowledge, behaviour, attitudes, critical thinking, and problem-solving are prioritised	Explicitly align learning activities and assessment with ocean literacy dimensions
DR5. Accessibility and feasibility	VR availability, cost, infrastructure, time, and readiness are constraints	Develop a practical, accessible, and teacher-friendly implementation model
DR6. Realistic and curriculum-aligned content	Teachers request realistic, natural, and curriculum-relevant visualisations	Use authentic visual content aligned with Biology learning objectives
DR7. Action-oriented sustainability	Teachers emphasise conservation and environmental action	Connect VR experiences with reflection, problem-solving, and real-world conservation actions

The seven propositions translate the reported needs and constraints into possible specifications for subsequent development. DR1 addresses the gap in systematic maritime contextualization; DR2 responds to limitations in students' access to direct marine experiences; DR3 addresses teachers' expectations for pedagogical integration; DR4 responds to the need for multidimensional ocean literacy assessment; DR5 addresses technological and institutional constraints; DR6 addresses teachers' demand for realistic and curriculum-aligned content; and DR7 connects learning experiences with sustainability-oriented action. Overall, the findings indicate that teachers' needs extend beyond providing VR technology itself. These links are design interpretations informed by both theory and local evidence, not proof that the proposed specifications are effective or feasible. The study did not test an SME–VR intervention or compare VR with alternative approaches.

DISCUSSION

From Teaching Marine Content to Contextualised Maritime Learning

The findings distinguish between teaching marine-related content and implementing contextualised maritime learning. Although 84.6% of participating teachers reported teaching marine and coastal ecosystems, only 15.4% reported frequently integrating

maritime contexts. Because the participants were 26 teachers from nine secondary schools in Tanjungpinang, this result should be interpreted as a local instructional pattern rather than a representation of all coastal schools in Indonesia. The open-ended responses provide contextual explanations for this pattern, but their depth is limited compared with interviews or focus groups. The finding supports the need to examine contextualised maritime learning in this setting; it does not, on its own, establish which technology would best address that need. Nevertheless, this finding aligns with research indicating that ocean literacy education requires stronger alignment among curriculum, pedagogy, technology, and assessment (Asikin et al., 2026; Bastos et al., 2026). The qualitative evidence helps explain this contextualization gap. Teachers described difficulty bringing authentic marine experiences into classrooms and connecting textbook materials with local marine environments. In practical design terms, teachers can address contextual authenticity by using locally relevant marine settings and issues, such as seagrass beds, mangrove areas, coral reefs, marine biota, and coastal pollution, and by connecting these settings to Biology learning objectives. This interpretation aligns with place-based approaches to marine education, which emphasise relationships among learners, local environments, and environmental issues (Asikin et al., 2025).

VR Addresses the Experiential Access Gap, but Its Educational Value Depends on Pedagogical Design

Teachers' reports that marine ecosystems and biodiversity are difficult to teach align with the need for better access to observation and visualisation of underwater environments, organism interactions, and ecological processes. VR is one possible way to provide mediated access to environments that may be difficult to observe directly. However, this needs analysis did not compare VR with 360-degree media, simulations, multimedia, virtual field trips, or direct field experiences. It therefore cannot establish VR as the most appropriate or effective response. VR is discussed here as the proposed technological direction of the broader SME research programme; its feasibility, usability, and educational value in participating school contexts remain to be tested. Any future implementation should complement, rather than automatically replace, direct experiences where these are practical and safe. Immersive VR is relevant to this challenge because it can provide access to environments that are difficult to experience directly and allow students to explore complex phenomena from perspectives unavailable through conventional media (Hamilton et al., 2021; Petersen et al., 2020; Radianti et al., 2020). Systematic evidence indicates that immersive VR can positively affect learning outcomes, although these effects are not uniform and depend substantially on intervention characteristics and assessment design (Adams et al., 2021; Hamilton et al., 2021; Paulsen et al., 2024). Recent marine education research has also demonstrated the potential of underwater VR to reduce psychological distance from marine environments and evoke affective responses such as awe and empathy

(Fauville et al., 2021, 2026). Therefore, VR's contribution to SME should not be understood merely as providing attractive visualisation. VR needs to function as an experiential learning environment that enables observation, exploration, inquiry, and reflection on marine phenomena that are difficult for students to access within their physical environments.

SME–VR Requires Pedagogical Integration Rather Than Technology Substitution

The findings indicate that teachers did not view VR as an adequate technology when used independently; instead, they saw it as part of an overall learning design. Teachers expected VR experiences to be followed and reinforced through student worksheets, instructional modules, investigation activities, discussion, reflection, and assessment. This indicates that VR's pedagogical value depends on how immersive experiences connect to learning objectives and students' cognitive activities (Han, 2020; Makransky & Petersen, 2021; Petersen et al., 2020). This finding aligns with research showing that the effectiveness of VR-based learning depends not only on immersion but also on learning structure, task characteristics, learner interaction, and assessment strategies (Hamilton et al., 2021; Makransky & Petersen, 2021; Radianti et al., 2020). Therefore, immersive experiences need to be positioned within learning stages that allow students to observe phenomena, identify problems, interpret information, and construct understanding based on their experiences (Fauville et al., 2026; Lamb et al., 2019; Sunardi et al., 2022). In the context of SMEs, this pedagogical integration becomes increasingly important because learning objectives extend beyond understanding marine concepts to developing ocean literacy, encompassing knowledge, attitudes, and behaviour. VR experiences involving marine ecosystems should therefore be followed by activities that encourage students to connect virtual phenomena with the conditions of their surrounding coastal environments, analyse maritime problems, discuss alternative solutions, and reflect on their implications for marine environmental sustainability (Asikin et al., 2025; Warahmah et al., 2022). Research in marine education has also shown that VR visualisation and perspective-taking can help students understand complex marine phenomena that are difficult to observe directly (Fauville et al., 2021, 2026; Hetherington et al., 2026; Petersen et al., 2020).

Local Authenticity and Accessibility as Key Design Conditions

Teachers' prioritisation of seagrass, mangroves, marine biota, coral reefs, and marine pollution indicates that VR content needs to be locally meaningful while remaining relevant to the curriculum. This requirement aligns with developments in ocean-XR research showing that immersive technologies can support marine education through presence, interactivity, embodiment, and the representation of ocean literacy principles (Fauville et al., 2026; Hetherington et al., 2026; Mothershed & Stilwell, 2025; Petersen et al., 2020). However, contextual authenticity must align with implementation feasibility. The findings reveal a substantial gap between general digital experience and institutional

readiness for VR. Most teachers had experience using educational technologies, whereas most schools lacked dedicated VR devices. This suggests that general digital readiness does not automatically translate into readiness to implement VR. Systematic reviews likewise emphasise that immersive VR research needs to consider intervention characteristics, technological conditions, usability, and appropriate assessment, rather than if immersion alone guarantees improved learning outcomes (Hamilton et al., 2021; Luo, 2024; Matovu et al., 2022; Suh & Prophet, 2018). Therefore, SME–VR design needs to consider local authenticity alongside accessibility, device availability, cost, infrastructure, instructional time, and teacher readiness.

Ocean Literacy Assessment Needs to Move from Knowledge toward Multidimensional Competence

One of the most important findings concerns the limited use of structured ocean literacy assessment. Most teachers who reported assessing ocean literacy did so without a specific instrument, while only a small proportion used a structured assessment. This finding is consistent with the literature, indicating that ocean literacy education requires stronger alignment among learning objectives, instructional strategies, and assessment approaches (Asikin et al., 2026; McPherson et al., 2018). Teachers in this study prioritised dimensions beyond knowledge of marine ecosystems, including attitudes, environmental behaviour, critical thinking, and problem-solving. This multidimensional understanding aligns with contemporary conceptualisations of ocean literacy, which encompass knowledge, attitudes, awareness, behaviour, activism, emotional connection, and experience (Fielding et al., 2019; Hetherington et al., 2026; Kelly et al., 2022; Lin et al., 2020). Furthermore, knowledge acquisition should not automatically lead to sustainable behaviour. Stoll-Kleemann (2019) demonstrated that the relationship between environmental knowledge and behavioural change is relatively weak, indicating that ocean literacy education requires approaches that support behavioural transformation. Therefore, SME–VR needs to integrate assessment from the early stages of the design process, particularly by measuring knowledge, attitudes, and behaviour, rather than treating assessment as an additional activity implemented after the VR experience.

Toward an Empirically Grounded SME–VR Design Framework

Taken together, the findings indicate that the central design challenge is not simply the absence of marine content or educational technology. It aligns marine learning objectives, authentic learning experiences, technological feasibility, pedagogical structure, and ocean literacy assessment. The seven design requirements provide an empirical design framework for the next stage of SME–VR development: contextual authenticity; immersive experiential learning; pedagogical integration; ocean literacy alignment; accessibility and feasibility;

realistic curriculum-aligned content; and action-oriented sustainability. This interpretation extends previous research on ocean literacy education, emphasising the need to integrate curriculum, innovative pedagogy, technology, and assessment (Asikin et al., 2026) while responding to the VR literature demonstrating that immersion alone is insufficient without appropriate instructional design (Hamilton et al., 2021; Radianti et al., 2020; Stoll-Kleemann, 2019). Importantly, this study does not claim that VR has already improved students' ocean literacy because the research remains at the needs-analysis stage. Instead, the study identifies empirically derived design requirements that provide a foundation for subsequent development and effectiveness testing. The seven resulting requirements, contextual authenticity, immersive experiential learning, pedagogical integration, ocean literacy alignment, accessibility and feasibility, realistic curriculum-aligned content, and action-oriented sustainability, constitute this needs-analysis study's principal design contribution. The framework positions VR not as the goal of innovation but as a supporting technology that enables students to progress from immersive experience to understanding, reflection, decision-making, and sustainable action.

LIMITATIONS

Several limitations should be considered when interpreting the findings. First, the sample comprised 26 purposively selected Biology teachers from nine secondary schools in Tanjungpinang City, so the findings are context-specific and should not be generalised to all coastal schools. The participant group was predominantly female and experienced, which may shape reported needs. Second, the study relied on teacher self-report and did not collect data directly from students. Third, both quantitative and qualitative strands came from the same questionnaire and participants; open-ended responses offered contextual explanations but lacked the depth of interviews or focus groups. Fourth, three experts reviewed the questionnaire, but the authors did not pilot it with an independent sample before administration. Because it was designed as a heterogeneous needs-analysis instrument rather than a unidimensional scale, the authors did not estimate overall internal-consistency reliability. Fifth, the study did not compare VR with alternative technologies or direct field experiences, nor did it test implementation feasibility. Finally, the authors have not yet validated the seven design requirements through implementation. Subsequent development, expert validation, pilot implementation, and effectiveness research are needed to test their practical feasibility, usability, learning effects, and potential contribution to students' ocean literacy.

CONCLUSION

This study identified a gap between teaching marine-related Biology content and the systematic use of contextualised maritime learning in the participating coastal schools

in Tanjungpinang. Most teachers reported teaching marine and coastal ecosystems and using educational technology, but frequent maritime contextualization, institutional VR readiness, and structured ocean literacy assessment remained limited. Teachers' responses also indicated that teaching marine ecosystems and biodiversity was difficult, partly because students had limited access to direct observation of underwater environments and complex ecological processes. Integrating closed-ended and open-ended responses informed seven provisional design propositions for SME–VR: (1) contextual authenticity, (2) immersive experiential learning, (3) pedagogical integration, (4) ocean literacy alignment, (5) accessibility and feasibility, (6) realistic and curriculum-aligned content, and (7) action-oriented sustainability. These requirements were derived from the participating teachers' reported needs and explanations and should be interpreted as design propositions for subsequent development and validation. The study does not provide evidence that VR has improved students' ocean literacy because students did not participate, and no intervention was implemented. Its contribution is to provide a context-specific empirical foundation for the development and testing of SME–VR in coastal secondary education. Future research should validate the design with experts and teachers, pilot its usability and feasibility, and then examine its effects on students' ocean literacy using multidimensional measures of knowledge, attitudes, and behaviour.

ACKNOWLEDGEMENT

The authors sincerely thank the Directorate of Research, Technology, and Community Service (DPPM), Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kemendiktisaintek), for providing research funding and support through the 2026 Research Grant Program. The authors also gratefully acknowledge Universitas Maritim Raja Ali Haji (UMRAH) for institutional support throughout the research process. We sincerely appreciate the biology teachers who participated in the needs analysis and provided valuable information, perspectives, and feedback that helped develop the Sustainable Maritime Education (SME) framework. Their participation and contributions were essential to completing this study.

REFERENCES

Adams, A., Feng, Y., Liu, J. C., & Stauffer, E. (2021). Potentials of Teaching, Learning, and Design with Virtual Reality: An Interdisciplinary Thematic Analysis. In B. Hokanson, M. Exter, A. Grincewicz, M. Schmidt, & A. A. Tawfik (Eds.), *Intersections Across Disciplines: Interdisciplinarity and learning* (pp. 173-186). Springer International Publishing. https://doi.org/10.1007/978-3-030-53875-0_14

Almeida, S., & Cutter-Mackenzie, A. (2011). The Historical, Present and Future ness of Environmental Education in India. *Australian Journal of Environmental Education*, *27*(1), 122-133.

Armougum, A., Orriols, E., Gaston-Bellegarde, A., Marle, C. J. La, & Piolino, P. (2019). Virtual reality: A new method to investigate cognitive load during navigation. *Journal of Environmental Psychology*, *65*. https://doi.org/10.1016/j.jenvp.2019.101338

Asikin, N., Suwono, H., Bambang Sumitro, S., Dharmawan, A., & Qadri Tanjung, A. (2023). Trend ocean literacy research in Indonesia: A bibliometric analysis. *BIO Web of Conferences*, *79*. https://doi.org/10.1051/bioconf/20237913002

Asikin, N., Suwono, H., Dharmawan, A., & Qadri Tanjung, A. (2023). Trend oceanography research for enhancing ocean literacy to support sustainable development goals (SDGs): A systematic literature review. *BIO Web of Conferences*, *70*, 03013. https://doi.org/10.1051/bioconf/20237003013

Asikin, N., Suwono, H., Sumitro, S. B., & Dharmawan, A. (2026). Teaching ocean literacy in science education: a systematic review. *Environmental Education Research*, *32*(3), 812-837. https://doi.org/10.1080/13504622.2025.2490227

Asikin, N., Suwono, H., Sumitro, S. B., & Herak, R. (2025). VR Ocstal: Development and Validation of Virtual Reality Ocean-Coastal in Maritime Higher Education. *Salud, Ciencia y Tecnologia*, *5*. https://doi.org/10.56294/saludcyt20251660

Bastos, E., Clarke, G., Dodds, B., Higgs, S., Willcocks, K., & Bridge, N. (2026). Mapping ocean literacy educational journeys: a systematic scoping review. In *Environmental Education Research* (Vol. 32, Number 3, pp. 679-720). Routledge. https://doi.org/10.1080/13504622.2025.2512088

Boda, P. A., & Brown, B. (2020). Priming urban learners' attitudes toward the relevancy of science: A mixed-methods study testing the importance of context. *Journal of Research in Science Teaching*, *57*(4), 567-596. https://doi.org/10.1002/tea.21604

Bower, M., & Jong, M. S.-Y. (2020). Immersive virtual reality in education. *British Journal of Educational Technology*, *51*(6), 1981-1990. https://doi.org/https://doi.org/10.1111/bjet.13038

Cava, F. (2005). *Science Content and Standards for Ocean Literacy: A Report on Ocean Literacy*. http://www.cosee.net/fles/coseeca/OLit04-05FinalReport.pdf

Cheng, K.-H., & Tsai, C.-C. (2020). Students' motivational beliefs and strategies, perceived immersion and attitudes towards science learning with immersive virtual reality: A partial least squares analysis. *British Journal of Educational Technology*, *51*(6), 2139-2158. https://doi.org/10.1111/bjet.12956

Costa, R. L., Mata, B., Silva, F., Conceição, P., & Guimarães, L. (2021). Fostering Ocean-Literate Generations: The Portuguese Blue School. In K. C. Koutsopoulos & J. H. Stel (Eds.), *Ocean Literacy: Understanding the Ocean* (pp. 241-273). Springer International Publishing. https://doi.org/10.1007/978-3-030-70155-0_10

Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology*, *51*(6), 2006-2033. https://doi.org/https://doi.org/10.1111/bjet.13030

Ditrich, L., & Lachmair, M. (2026). A 'front-row seat' to catastrophe: testing the effect of immersive technologies on sympathy and pro-environmental behavior in the context of rising sea levels. *Environmental Education Research*, *32*(3), 569-587. https://doi.org/10.1080/13504622.2025.2459340

Diz, D., Johnson, D., Riddell, M., Rees, S., Battle, J., Gjerde, K., Hennige, S., & Roberts, J. M. (2018). Mainstreaming marine biodiversity into the SDGs: The role of other effective area-based conservation measures (SDG 14.5). *Marine Policy*, *93*, 251-261. https://doi.org/https://doi.org/10.1016/j.marpol.2017.08.019

Elme, L., Jørgensen, M. L. M., Dandanell, G., Mottelson, A., & Makransky, G. (2022). Immersive virtual reality in STEM: is IVR an effective learning medium and does adding self-explanation after a lesson improve learning outcomes? *Educational Technology Research and Development*, *70*(5), 1601-1626. https://doi.org/10.1007/s11423-022-10139-3

Fauville, G., Queiroz, A. C. M., Hambrick, L., Brown, B. A., & Bailenson, J. N. (2021). Participatory research on using virtual reality to teach ocean acidification: a study in the marine education community. *Environmental Education Research*, *27*(2), 254-278. https://doi.org/10.1080/13504622.2020.1803797

Fauville, G., Voşki, A., Mado, M., Bailenson, J. N., & Lantz-Andersson, A. (2026). Underwater virtual reality for marine education and ocean literacy: technological and psychological potentials. *Environmental Education Research*, *32*(3), 588-612. https://doi.org/10.1080/13504622.2024.2326446

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods*, *5*, 80-92. https://doi.org/https://doi.org/10.1177/160940690600500107

Fetters, M. D., & Tajima, C. (2022). Joint Displays of Integrated Data Collection in Mixed Methods Research. *International Journal of Qualitative Methods*, *21*. https://doi.org/10.1177/16094069221104564

Fielding, S., Copley, J. T., & Mills, R. A. (2019). Exploring our oceans: Using the global classroom to develop ocean literacy. *Frontiers in Marine Science*, *6*(JUN). https://doi.org/10.3389/fmars.2019.00340

Francesca, S., Selvaggia, S., Scowcroft, G., Fauville, G., & Tuddenham, P. (2017). *Ocean Literacy for All: a toolkit (IOC Manuals and Guides, 80)*. IOC/UNESCO & UNISCO Venice Office. ioc.unesco.org

Frank, A. L., Baldwin, J., Browning, M. H. E. M., Yuan, S., & Ahn, S. J. (Grace). (2026). Learning the dangers of storm surges in virtual reality: coastal leaders' attitudes toward VR as a risk communication tool. *Environmental Education Research*, *32*(3), 550-568. https://doi.org/10.1080/13504622.2025.2465713

Freitas, C., Hannigan, S., Bellgrove, A., Venzo, P., & Francis, P. (2026). Developing ocean literacy through a relational model: Exploring outcomes of a place-based ocean education programme in primary schools. *People and Nature*, *8*(4), 904-918. https://doi.org/10.1002/pan3.70264

Gomez, L. I. (2020). Immersive Virtual Reality for Learning Experiences. In D. Burgos (Ed.), *Radical Solutions and eLearning: Practical Innovations and Online Educational Technology* (pp. 183-198). Springer Singapore. https://doi.org/10.1007/978-981-15-4952-6_12

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, *8*(1), 1-32. https://doi.org/10.1007/s40692-020-00169-2

Han, I. (2020). Immersive virtual field trips in education: A mixed-methods study on elementary students' presence and perceived learning. *British Journal of Educational Technology*, *51*(2), 420-435. https://doi.org/https://doi.org/10.1111/bjet.12842

Haynes-Brown, T. K., & Fetters, M. D. (2021). Using Joint Display as an Analytic Process: An Illustration Using Bar Graphs Joint Displays From a Mixed Methods Study of How Beliefs Shape Secondary School Teachers' Use of Technology. *International Journal of Qualitative Methods*, *20*. https://doi.org/10.1177/16094069211993286

Heck, N., Lykkebo Petersen, K., Potts, D. C., Haddad, B., & Paytan, A. (2018). Predictors of coastal stakeholders' knowledge about seawater desalination impacts on marine ecosystems. *Science of the Total Environment*, *639*, 785-792. https://doi.org/10.1016/j.scitotenv.2018.05.163

Hetherington, L., Brandt, H., Chappell, K., Dillon, J., & Malmierca, M. R. (2026). Exploring the use of digital technologies for ocean literacy education using a material-dialogic theoretical framework. *Environmental Education Research*, *32*(3), 650–678. https://doi.org/10.1080/13504622.2025.2512974

Hite, R. (2022). Virtual Reality: Flight of Fancy or Feasible? Ways to Use Virtual Reality Technologies to Enhance Students' Science Learning. *American Biology Teacher*, *84*(2), 106-108. https://doi.org/10.1525/abt.2022.84.2.106

Huang, K.-T., Ball, C., Francis, J., Ratan, R., Boumis, J., & Fordham, J. (2019). Augmented versus virtual reality in education: An exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications. *Cyberpsychology, Behavior, and Social Networking*, *22*(2), 105-110. https://doi.org/10.1089/cyber.2018.0150

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515-1529. https://doi.org/10.1007/s10639-017-9676-0

Kao, R.-H., Chiang, N.-T., & Yang, Y.-L. (2026). Exploring the ocean literacy of Taiwanese college students: a hybrid qualitative and quantitative study. *Environmental Education Research*, *32*(3), 838-856. https://doi.org/10.1080/13504622.2025.2543004

Kasten, P., Jenkins, S. R., & Christofoletti, R. A. (2021). Participatory Monitoring—A Citizen Science Approach for Coastal Environments. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.681969

Kelly, R., Elsler, L. G., Polejack, A., van der Linden, S., Tönnesson, K., Schoedinger, S. E., Santoro, F., Pecl, G. T., Palmgren, M., Mariani, P., Glithero, D., Evans, K., Cvitanovic, C., Cook, J., Bartram, J., & Wisz, M. S. (2022). Empowering young people with climate and ocean science: Five strategies for adults to consider. In *One Earth* (Vol. 5, Number 8, pp. 861-874). Cell Press. https://doi.org/10.1016/j.oneear.2022.07.007

Kelly, R., Evans, K., Alexander, K., Bettiol, S., Corney, S., Cullen-Knox, C., Cvitanovic, C., de Salas, K., Emad, G. R., Fullbrook, L., Garcia, C., Ison, S., Ling, S., Macleod, C., Meyer, A., Murray, L., Murunga, M., Nash, K. L., Norris, K., … Pecl, G. T. (2022). Connecting to the oceans: supporting ocean literacy and public engagement. *Reviews in Fish Biology and Fisheries*, *32*(1), 123-143. https://doi.org/10.1007/s11160-020-09625-9

Kolb, A. Y., & Kolb, D. A. (2017). Experiential Learning Theory as a Guide for Experiential Educators in Higher Education. In *Experiential Learning & Teaching in Higher Education* (Vol. 7, Number 1). https://nsuworks.nova.edu/elthe/vol1/iss1/7

Koutsopoulos, K. C., & Stel, J. H. (2021). *Ocean Literacy: Understanding the Ocean*. https://doi.org/https://doi.org/10.1007/978-3-030-70155-0

Lamb, R. L., Etopio, E., Hand, B., & Yoon, S. Y. (2019). Virtual Reality Simulation: Effects on Academic Performance Within Two Domains of Writing in Science. *Journal of Science Education and Technology*, *28*(4), 371-381. https://doi.org/10.1007/s10956-019-09774-y

Lin, Y.-L., Wu, L.-Y., Tsai, L.-T., & Chang, C.-C. (2020). The beginning of marine sustainability: Preliminary results of measuring students' marine knowledge and ocean literacy. *Sustainability (Switzerland)*, *12*(17). https://doi.org/10.3390/su12177115

Liu, R., Wang, L., Lei, J., Wang, Q., & Ren, Y. (2020). Effects of an immersive virtual reality-based classroom on students' learning performance in science lessons. *British Journal of Educational Technology*, *51*(6), 2034-2049. https://doi.org/10.1111/BJET.13028

Lucrezi, S., Milanese, M., Danovaro, R., & Cerrano, C. (2018). 'Generation Nemo': motivations, satisfaction and career goals of marine biology students*. *Journal of Biological Education*, *52*(4), 391-405. https://doi.org/10.1080/00219266.2017.1385509

Luo, D. (2024). Research on practical teaching of Ideology and politics in higher education institutions based on VR technology in the information age. *Applied Mathematics and Nonlinear Sciences*, *9*(1). https://doi.org/10.2478/amns.2023.2.00303

Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. In *Educational Psychology Review* (Vol. 33, Number 3, pp. 937-958). Springer. https://doi.org/10.1007/s10648-020-09586-2

Makransky, G., Petersen, G. B., & Klingenberg, S. (2020). Can an immersive virtual reality simulation increase students' interest and career aspirations in science? *British Journal of Educational Technology*, *51*(6), 2079-2097. https://doi.org/10.1111/bjet.12954

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive Virtual Reality field trips facilitate learning about climate change. *Frontiers in Psychology*, *9*(NOV). https://doi.org/10.3389/fpsyg.2018.02364

Matovu, H., Ungu, D. A. K., Won, M., Tsai, C. C., Treagust, D. F., Mocerino, M., & Tasker, R. (2022). Immersive virtual reality for science learning: Design, implementation, and evaluation. *Studies in Science Education*. https://doi.org/10.1080/03057267.2022.2082680

McKinley, E., Burdon, D., & Shellock, R. J. (2023). The evolution of ocean literacy: A new framework for the United Nations Ocean Decade and beyond. *Marine Pollution Bulletin*, *186*. https://doi.org/10.1016/j.marpolbul.2022.114467

McKinley, E., Kelly, R., Mackay, M., Shellock, R., Cvitanovic, C., & van Putten, I. (2022). Development and expansion in the marine social sciences: Insights from the global community. *IScience*, *25*(8), 104735. https://doi.org/10.1016/j.isci.2022.104735

McPherson, K., Wright, T., & Tyedmers, P. (2018). Examining the Nova Scotia science curriculum for international ocean literacy principle inclusion. *International Journal of Learning, Teaching and Educational Research*, *17*(11), 1-16. https://doi.org/10.26803/ijlter.17.11.1

Mendler de Suarez, J., Cicin-Sain, B., Wowk, K., Payet, R., & Hoegh-Guldberg, O. (2014). Ensuring survival: Oceans, climate and security. *Ocean and Coastal Management*, *90*, 27-37. https://doi.org/10.1016/j.ocecoaman.2013.08.007

Mokos, M., cheimonopoulou, M. T., Koulouri, P., Previati, M., Realdon, G., Santoro, F., Mogias, A., Boubonari, T., Gazo, M., Satta, A., Ioakeimidis, C., Tojeiro, A., Chicote, C. A., Papathanassiou, M., & Kevrekidis, T. (2020). Mediterranean Sea Literacy: When Ocean Literacy becomes region-specific. *Mediterranean Marine Science*, *21*(3), 592-598. https://doi.org/10.12681/mms.23400

Moro, C., Stromberga, Z., & Stirling, A. (2017). Virtualisation devices for student learning: Comparison between desktop-based (Oculus Rift) and mobile-based (Gear VR) virtual reality in medical and health science education. *Australasian Journal of Educational Technology*, *33*(6), 1-10. https://doi.org/10.14742/ajet.3840

Mothershed, K., & Stilwell, M. D. (2025). Building Environmental Stewards: The Power of Place-Based Education in Gulf Coast Marine Literacy. *Current: The Journal of Marine Education*, *40*(1), 14-23. https://doi.org/10.5334/cjme.128

NOAA. (2020). *Ocean Literacy The Essential Principles and Fundamental Concepts of Ocean Sciences for Learners of All Ages Version 3: Februari 2020*. www.oceanliteracyNMEA.org

O'Brien, M., Freitas, C., Venzo, P., & Francis, P. (2023). Fostering ocean literacy through informal marine education programs. *Marine Pollution Bulletin*, *193*, 115208. https://doi.org/https://doi.org/10.1016/j.marpolbul.2023.115208

Otto, S., & Pensini, P. (2017). Nature-based environmental education of children: Environmental knowledge and connectedness to nature, together, are related to ecological behaviour. *Global Environmental Change*, *47*, 88-94. https://doi.org/https://doi.org/10.1016/j.gloenvcha.2017.09.009

Paulsen, L., Dau, S., & Davidsen, J. (2024). Designing for collaborative learning in immersive virtual reality: a systematic literature review. *Virtual Reality*, *28*(1). https://doi.org/10.1007/s10055-024-00975-4

Pazoto, C. E., Silva, E. P., & Duarte, M. R. (2022). Ocean literacy in Brazilian school curricula: An opportunity to improve coastal management and address coastal risks? *Ocean & Coastal Management*, *219*, 106047. https://doi.org/10.1016/J.OCECOAMAN.2022.106047

Petersen, G. B., Klingenberg, S., Mayer, R. E., & Makransky, G. (2020). The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology*, *51*(6), 2099-2115. https://doi.org/https://doi.org/10.1111/bjet.12991

Pinardi, N. (2019). The Joint IOC (of UNESCO) and WMO collaborative effort for met-ocean services. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00410

Proudfoot, K. (2023). Inductive/Deductive Hybrid Thematic Analysis in Mixed Methods Research. *Journal of Mixed Methods Research*, *17*(3), 308-326. https://doi.org/10.1177/15586898221126816

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers and Education*, *147*. https://doi.org/10.1016/j.compedu.2019.103778

Reeping, D., & Edwards, C. (2022). Integrating Using a Crossover Analysis With Formative Joint Displays in Mixed Methods Research. *Journal of Mixed Methods Research*, *16*(4), 458-477. https://doi.org/10.1177/15586898211047660

Roland, K. (2017). *Experiential Learning: Learning through Reflective Practice*.

Sahriana, N., Suminar, T., & Sugiyo Pranoto, Y. K. (2020). Development of Maritime Insight Learning Tools for Ocean Literacy in Children Aged 5-6 Years Old. *Journal of Primary Education*, *9*(5), 536-545. https://doi.org/10.15294/jpe.v9i5.43530

Shi, J., & Fan, L. (2021). Investigating Teachers' and Students' Perceptions of Online English Learning in a Maritime Context in China. *SAGE Open*, *11*(3). https://doi.org/10.1177/21582440211040800

Shih, Y. C. (2024). Ocean governance in practice: A study of the application of marine science and technology research techniques to maritime law enforcement in Taiwan. *Marine Policy*, *163*. https://doi.org/10.1016/j.marpol.2024.106081

Stoll-Kleemann, S. (2019). Feasible Options for Behavior Change Toward More Effective Ocean Literacy: A Systematic Review. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00273

Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, *86*, 77-90. https://doi.org/10.1016/j.chb.2018.04.019

Sunardi, Hidayanto, A. N., Meyliana, & Prabowo, H. (2022). Discipline, Impact, And Challenges Of Virtual Reality In Higher Education: A Systematic Literature Review. *2022 International Conference on Information Management and Technology (ICIMTech)*, 476-481. https://doi.org/10.1109/ICIMTech55957.2022.9915242

Warahmah, L. W., Putri, A. N., & Asikin, N. (2022). Interactive Multimedia on Basis Articulate Storyline and Maritime Oriented for Plantae Topic in Senior High School. *Atrium Pendidikan Biologi*, *7*(2), 175-183. https://doi.org/http://dx.doi.org/10.24036/apb.v7i2.12862

Ward, D., Melbourne-Thomas, J., Pecl, G. T., Evans, K., Green, M., McCormack, P. C., Novaglio, C., Trebilco, R., Bax, N., Brasier, M. J., Cavan, E. L., Edgar, G., Hunt, H. L., Jansen, J., Jones, R., Lea, M. A., Makomere, R., Mull, C., Semmens, J. M., … Layton, C. (2022). Safeguarding marine life: conservation of biodiversity and ecosystems. *Reviews in Fish Biology and Fisheries*, *32*(1), 65-100. https://doi.org/10.1007/s11160-022-09700-3

Watson, A. J., Schuster, U., Shutler, J. D., Holding, T., Ashton, I. G. C., Landschützer, P., Woolf, D. K., & Goddijn-Murphy, L. (2020). Revised estimates of ocean-atmosphere CO2 flux are consistent with ocean carbon inventory. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-18203-3

Williams, N. (2024). "Nature Environments Without Actually Being There": Virtual Recreation Experiences and Real-World Intentions. *Journal of Experiential Education*, *47*(2), 295-300. https://doi.org/10.1177/10538259231225621

Yahaya, R. A. (2007). *Immersive Virtual Reality Learning Environment: Learning Decision-Making skills in a Virtual Reality-enhanced Learning Environment*. Queensland University of Technology.

Zhang, Z., Jiang, W., Qiu, H., Xiao, Z., Ling, Z., & Song, J. (2026). Simulation of potential distribution of mangrove species in China and identification of restoration optimization sequences. *Ocean and Coastal Management*, *278*. https://doi.org/10.1016/j.ocecoaman.2026.108235

Zielinski, T., Kotynska-Zielinska, I., & Garcia-Soto, C. (2022). A Blueprint for Ocean Literacy: EU4Ocean. *Sustainability (Switzerland)*, *14*(2). https://doi.org/10.3390/su14020926

Pertanika J. Sci. & Technol. 34 (5): 3163 - 3181 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Antigenicity of Hydrolysed Casein in Goat's Milk Using Bromelain

Nurul Izza Nabilah Mohd Samsudin[1], Chai Jie Ying[1], Rashidah Sukor[1,2]*, Aliah Zannierah Mohsin[1], and Siva Raseetha[3]

[1]Department of Food Science, Faculty of Food Science and Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2]Laboratory of Food Safety and Food Integrity, Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[3]Department of Food Science and Technology, School of Industrial Technology, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

ABSTRACT

Milk is a major source of infant nutrition but is also a common allergen associated with cow's milk protein allergy (CMPA), which requires strict avoidance of milk proteins. Goat's milk has been proposed as an alternative due to its distinct protein composition; however, conserved casein epitopes may still trigger immune responses. While enzymatic hydrolysis is well established for reducing the antigenicity of cow's milk proteins, studies on goat's milk casein hydrolysis remain limited. This study evaluated the time-dependent hydrolysis of goat milk casein using bromelain (1:100, pH 5.5, 55 °C, 1-4 h) and its effects on protein degradation and antigenicity. The novelty of this study lies in the application of bromelain to hydrolyse goat milk casein, an area that remains less explored than the well-established hydrolysis of cow milk proteins. Degree of hydrolysis (DH), peptide content, SDS-PAGE, and ELISA were used for characterisation. Results showed increasing DH and peptide content with hydrolysis time ($p < 0.05$), with peptides mainly below 15 kDa. ELISA indicated that 1 h hydrolysis achieved the lowest antigenic reactivity. Overall, bromelain effectively reduced the antigenic reactivity of goat's milk casein under mild hydrolysis conditions, demonstrating its potential for modifying goat milk casein and supporting further investigation for specialised dairy and infant nutrition applications.

Keywords: antigenicity, bromelain, casein, ELISA, goat's milk, hydrolysis

ARTICLE INFO

Article history:
Received: 04 May 2026
Accepted: 10 August 2026
Published: 30 October 2026

DOI: https://doi.org/10.47836/pjst.34.5.25

izzanabilahsamsudin@gmail.com (Nurul Izza Nabilah Mohd Samsudin)
201903@student.upm.edu.my (Chai Jie Ying)
rashidah@upm.edu.my (Rashidah Sukor)
aliah_mohsin@upm.edu.my (Aliah Zannierah Mohsin)
raseetha@salam.uitm.edu.my (Siva Raseetha)
* Corresponding author

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

INTRODUCTION

Food allergies in both children and adults have increased significantly in recent years, with elimination diets remaining the primary management strategy. In Malaysia, approximately 5–15% of young children experience allergic symptoms in early life, with cow's milk protein allergy (CMPA) identified as one of the most common food allergies (Malaysian Paediatric Association & Malaysian Society of Allergy and Immunology, 2024). Major allergens such as milk, egg, fish, nuts, and cereals must be declared on food labels due to their potential to trigger hypersensitivity reactions (Food and Drug Administration, 2026). Cow's milk proteins, particularly casein and whey, are major allergens responsible for CMPA (Jensen et al., 2022). As cow's milk is often introduced early in infancy, its exclusion can lead to nutritional deficiencies, necessitating the use of alternative formulas such as extensively hydrolysed formulas (eHF), partially hydrolysed formulas (pHF), amino acid formulas (AAF), or soy-based formulas (SBF) (Ballabio et al., 2011). However, the suitability of these alternatives depends on clinical severity and patient tolerance.

Goat's milk has gained attention as a potential alternative due to its nutritional similarity to cow's milk and reportedly lower allergenicity. Some studies suggest that a proportion of CMPA patients may tolerate goat's milk, possibly due to its lower αs1-casein content (Benjamin-van Aalst et al., 2024). However, cross-reactivity between cow's and goat's milk proteins, particularly caseins, has been reported, indicating that goat's milk is not universally safe for CMPA patients (Mansor et al., 2023). Therefore, further modification may be required to reduce its antigenic reactivity.

Enzymatic hydrolysis is widely used to modify milk proteins and reduce their antigenic reactivity by breaking proteins into smaller peptides and altering antibody-recognised regions. The effectiveness of hydrolysis depends on factors such as enzyme type, pH, temperature, and reaction time, as different proteases can produce different cleavage patterns and peptide profiles (Medeiros et al., 2014). Proteases such as bromelain, papain, and alcalase have been investigated for milk protein hydrolysis, demonstrating the importance of enzyme selection in determining the extent of protein degradation and antigenic reactivity. Although enzymatic hydrolysis can reduce the antigenic reactivity of milk proteins, residual immunoreactive regions may remain (He et al., 2025).

Compared with the well-established research on cow's milk proteins, studies on the enzymatic hydrolysis of goat's milk proteins remain limited. In particular, the time-dependent hydrolysis of goat's milk casein using bromelain and its relationship with antigenic reactivity remain insufficiently characterised. Therefore, the present study advances current knowledge by evaluating the bromelain-mediated hydrolysis of goat's milk casein at different hydrolysis times and by relating the progression of hydrolysis to protein degradation and antigenic reactivity under mild hydrolysis conditions. This study aimed to evaluate the effect of bromelain hydrolysis time on the degree of hydrolysis and antigenicity of goat milk casein.

METHODOLOGY

All chemicals and reagents used in this study were of analytical grade. Sodium hydroxide (NaOH), potassium chloride (KCl), disodium hydrogen phosphate ($Na_2HPO_4 \cdot 2H_2O$), and potassium dihydrogen phosphate (KH_2PO_4) were purchased from R&M Chemicals (Essex, UK). Sulphuric acid and sodium tetraborate decahydrate ($Na_2B_4O_7 \cdot 10H_2O$) were obtained from Chemiz (Shah Alam, Selangor, Malaysia). Sodium sulphate anhydrous, methylene blue, glycine, and methanol were procured from Fisher Scientific (Leicestershire, UK), while ethanol was supplied by Systerm (Shah Alam, Selangor, Malaysia).

Phosphate-buffered saline (PBS) was prepared as a 10X stock solution containing 1.37 M NaCl, 27 mM KCl, 100 mM $Na_2HPO_4 \cdot 2H_2O$, and 18 mM KH_2PO_4, and adjusted to pH 7.4 using NaOH or HCl. Deionised water (18.2 MΩ·cm) was obtained from a Milli-Q Water Purification System (Merck KGaA, Burlington, MA, USA). Bovine serum albumin (BSA) was purchased from Amresco, LLC (Solon, OH, USA).

DL-dithiothreitol (DTT), β-mercaptoethanol, ammonium persulfate (APS), sodium dodecyl sulfate (SDS), Tris, Bradford reagent, 30% acrylamide/bis-acrylamide solution were obtained from Bio-Rad Laboratories Inc. (Hercules, CA, USA). Bromelain enzyme, o-phthaldialdehyde (OPA), and L-serine were purchased from Acros Organics (St. Louis, MO, USA). Tween 20, acetic acid, sulphuric acid, hydrochloric acid, Immobilon®-P PVDF membrane, N, N, N′, N′-tetramethylethylenediamine (TEMED), and NBT-BCIP were obtained from Merck (Kenilworth, NJ, USA). Bromophenol blue, boric acid, selenium dioxide, and Coomassie Brilliant Blue G-250 were purchased from Merck (Darmstadt, Germany). The 1-Step Ultra TMB-ELISA substrate was obtained from Thermo Fisher Scientific (Waltham, MA, USA).

The N-terminal (Nter) and C-terminal (Cter) immunoglobulin G (IgG) antibodies were generated against αs1-casein peptides derived from goat milk and produced in rabbits, as previously reported by Mohsin et al. (2020). Horseradish peroxidase (HRP)-conjugated AffiniPure goat anti-rabbit IgG (H + L) (streptavidin-HRP) was obtained from Jackson ImmunoResearch Laboratories (West Grove, PA, USA).

Isolation of Casein (CSN)

Goat milk was obtained from a single bulk batch collected from Saanen goats. Isolation of casein (CSN) from goat milk was performed as described by Hodgkinson et al. (2012), with minor modifications to the precipitate-washing step. Frozen goat milk samples were thawed and pooled at room temperature (24 °C), followed by heating to 45 °C. The milk was then centrifuged using a Sartorius 3-18K centrifuge (Göttingen, Germany) at 6000 rpm for 30 min at 4 °C.

After centrifugation, the fat layer was carefully removed, and the pH of the skimmed milk was adjusted to 4.2 using 1 M HCl to precipitate casein. The acidified skim milk was

stirred for 30 min and centrifuged again at 6000 rpm for 15 min at 20 °C to collect the casein precipitate.

The resulting precipitate was washed by resuspending it in purified water at ten times the initial milk volume and stirred for 1 h prior to re-centrifugation under the same conditions. The isolated casein was subsequently lyophilised and stored at −20 °C until further analysis.

Hydrolysis of Casein

Enzymatic hydrolysis of casein was performed as described by Zarei et al. (2015), with modifications to the time intervals and pH conditions. Lyophilised casein was hydrolysed using bromelain (EC 3.4.22.32; stem-derived; ≥2500 GDU/g; ≥5.0 FIP U/mg) at a substrate-to-enzyme ratio (S: E) of 100:1. The bromelain enzyme was used according to the activity specified by the manufacturer and stored under the recommended conditions prior to use. The hydrolysis conditions (pH 5.5 and 55°C) were chosen based on the optimal catalytic activity of bromelain reported in previous studies, enabling efficient hydrolysis while maintaining enzyme stability.

Briefly, 1.4 g (1 mM) of lyophilised casein was dissolved in 50 mM acetate buffer (45 mM sodium acetate and 5 mM acetic acid) at pH 5.5 and incubated at 55 °C. The casein solution was pre-heated in a water bath (Memmert GmbH + Co. KG, Buchenbach, Germany) at 80 °C for 20 min to denature the protein. The solution was subsequently cooled to 55 °C before adding bromelain (powdered form, 15 mg, 0.01 mM) to initiate hydrolysis.

Aliquots (2 mL) were withdrawn at 1-h intervals (1-4 h) into 2-mL microcentrifuge tubes to monitor the hydrolysis process. Enzyme activity was terminated by heating the samples in boiling water for 10 min. The samples were then centrifuged at 10,000 rpm for 10 min at 4 °C, and the resulting supernatant was collected and stored at −20 °C for further analysis.

Protein Quantification

Determination of Crude Protein

The crude protein content of each sample was determined using the Kjeldahl method in accordance with AOAC International Official Method 992.10 (2023). Milk samples (0.15 g) were weighed in triplicate (n = 3) and mixed with a catalyst mixture containing 96% sodium sulphate anhydrous, 3.5% copper sulphate, and 0.5% selenium dioxide.

Subsequently, 2.5 mL of concentrated sulphuric acid was added, and the mixture was digested using a protein digester (Turbotherm digestion unit, Gerhardt GmbH & Co. KG, Königswinter, Germany) for 30 min until the solution turned blue-green. After cooling for 1 h, 10 mL of distilled water and 10 mL of 45% NaOH were carefully added to the digested mixture.

Boric acid solution was prepared by mixing 10 mL of 2% boric acid with two drops of indicator solution (0.2% methyl red and 0.1% methylene blue in 96% ethanol). The digested sample was then steam-distilled into the boric acid solution using a distillation unit (Gerhardt GmbH & Co. KG, Königswinter, Germany) for 3 min.

The borate formed during distillation was titrated with 0.05 N sulphuric acid until the endpoint colour changed from green to purple. A blank sample was analysed using the same procedure. The percentage of nitrogen was calculated using the formula as shown in Equation 1:

$$\text{Nitrogen (\%)} = \frac{(Vs - Vb) \times N \times 1.4}{W} \quad [1]$$

Where:
Vs = Volume of H_2SO_4 used to titrate H_3BO_3 (sample)
Vb = Volume of H_2SO_4 used to titrate blank
N = Normality of H_2SO_4
W = Weight of sample (g)

The percentage of nitrogen obtained was multiplied by the dairy conversion factor (6.38) to determine the percentage of crude protein in the goat milk sample.

Bradford Assay

The Bradford protein assay (microplate method) was performed according to the protocol provided by Bio-Rad Laboratories to quantify total soluble protein in milk samples, lyophilised CSN, and hydrolysed CSN.

Serial dilutions of each sample (1:5, 1:10, 1:50, and 1:100) were prepared in phosphate-buffered saline (PBS). Aliquots (10 µL) of each diluted sample were pipetted into individual wells of a 96-well microtiter plate in triplicate (n = 3). Bradford reagent was diluted 1:4 (v/v) with deionised water, and 200 µL was added to each well. The contents were gently mixed to avoid bubble formation and allowed to react for 2 min at room temperature (24 °C).

Absorbance was measured at 595 nm (A595) using a microplate reader (Multiskan FC, Thermo Scientific, Waltham, MA, USA). Protein concentrations were calculated from a linear standard curve generated using bovine serum albumin (BSA) at concentrations of 0.125, 0.250, 0.375, 0.500, and 0.750 mg/mL.

Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Lyophilised CSN, purified αs1-CSN, and hydrolysed CSN samples were analysed by SDS-PAGE according to the method of Laemmli (1970).

A 12% separating gel was prepared by mixing 6.4 mL of 30% acrylamide/bis-acrylamide, 4 mL of 1.5 M Tris-HCl (pH 8.8), 160 µL of 10% SDS, 160 µL of 10% APS, 16 µL of TEMED, and 5.3 mL of deionised water.

Sample buffer consisted of 3.55 mL deionised water, 1.25 mL 0.5 M Tris-HCl (pH 6.8), 2.5 mL glycerol, 2.0 mL 10% SDS, and 0.2 mL 0.5% (w/v) bromophenol blue. Samples (10 µg/mL) were mixed with sample buffer at a 1:1 (v/v) ratio and heated at 95 °C for 4 min. A 10 µL volume per sample was loaded onto the gel and electrophoresed at 110 V for 80 min.

Gels were stained overnight using staining solution containing 0.05% (w/v) Coomassie Brilliant Blue G-250, 50% (v/v) methanol, and 10% (v/v) glacial acetic acid, followed by destaining for 2 h using 40% (v/v) methanol and 7% (v/v) acetic acid. Protein bands were identified using Precision Plus Protein Dual Colour Standards (Bio-Rad Laboratories Inc., Hercules, CA, USA).

Determination of Degree of Hydrolysis Using o-Phthaldialdehyde (OPA)

To determine complete hydrolysis, 0.5 g of casein was mixed with 15 mL of 6 N HCl and incubated at 110 °C for 24 h in an oven (Memmert UFE400 oven, Memmert GmbH + Co. KG, Germany). After cooling to room temperature (24 °C), samples were centrifuged at 5000 rpm for 10 min at 4 °C, and the supernatant was collected.

OPA reagent was freshly prepared from three solutions:

i. Solution A: 200 mg SDS and 7.62 g sodium tetraborate dissolved in 150 mL deionised water
ii. Solution B: 160 mg OPA dissolved in 4 mL of 96% ethanol
iii. Solution C: 400 µL β-mercaptoethanol diluted to 50 mL with deionised water

In a 96-well Maxisorp plate (Nunc, Thermo Scientific, Waltham, MA, USA), 36 µL of hydrolysed sample was mixed with 270 µL OPA solution. After 2 min incubation in the dark at room temperature, absorbance was measured at 340 nm using a microplate reader (Millipore, Darmstadt, Germany).

The degree of hydrolysis (DH, %) was calculated as in Equation 2:

$$\text{DH (\%)} = \frac{h_{hydrolysed} - h_{substrate}}{h_total} \times 100 \quad [2]$$

Where:

$h_h$ydrolysed = absorbance of hydrolysed casein
$h_s$ubstrate = absorbance of non-hydrolysed casein
$h_t$otal = absorbance after complete hydrolysis

An L-serine standard curve ($y = 3.9362x + 0.0009$; $R^2 = 0.9998$) was used to determine peptide concentration.

Enzyme-Linked Immunosorbent Assay (ELISA)

Antigenicity of hydrolysed casein at different hydrolysis times was evaluated using an in-house sandwich ELISA as described by Mohsin et al. (2023). Capture (Nter-IgG) and detection (Cter-IgG) antibodies were previously developed in rabbits (Mohsin et al., 2020).

Nter-IgG was diluted (1:1000 v/v) in carbonate buffer (pH 9.6) and coated onto a Nunc Maxisorp 96-well plate overnight (16 h) at 4 °C. Wells were blocked with 2.5% BSA in PBS for 2 h at room temperature using an orbital shaker at 150 rpm.

Samples (100 µL) were added and incubated for 1 h. Cter-IgG (1:2,500 v/v in PBS) was then added and incubated for 1 h. Streptavidin-horseradish peroxidase (Strep-HRP; 1:15,000 v/v in PBS) was added and incubated for 1 h under similar conditions. Between steps, wells were washed three times with 250 µL of 0.05% Tween 20 in 1X PBS for 1 min with agitation.

After adding 50 µL of 1-Step Ultra TMB ELISA substrate, colour development was monitored, and the reaction was stopped with 50 µL of 0.1 M H_2SO_4. Absorbance was measured at 450 nm using a Multiskan RC microplate reader (Thermo Fisher Scientific, Waltham, MA, USA)

Statistical Test

All absorbance values were measured in triplicate (n = 3). Absorbance data were analysed using mean and standard deviation. Minitab (Pty Ltd, Sydney, Australia) statistical software, Version 21, was used for the analysis of variance (ANOVA). A post hoc Tukey test was then conducted to compare pairs of means across treatment levels. Significant differences were calculated at $\alpha = 0.05$ (95% confidence interval).

RESULT AND DISCUSSION

Isolation of Casein (CSN)

Casein was isolated from goat milk by acid precipitation and subsequently lyophilised for further analysis. At acidic pH (approximately pH 4.6), the surface charge of κ-casein is neutralised, resulting in destabilisation of casein micelles. The disruption of intra- and inter-chain interactions prevents the proteins from remaining in the soluble phase. Consequently, micelles aggregate and precipitate, allowing separation from whey proteins (Corredig, 2019).

In this study, the total yield of lyophilised casein was 1.65 g per 100 mL of goat milk (1.65%). The crude protein content of goat milk was determined by the Kjeldahl method in accordance with AOAC International Official Method 992.10 (2023), yielding a total of 3.87%. A previous study by Mohsin et al. (2019) reported a protein content of 4.83% in Saanen goat milk. The lower protein content observed in the present study may be attributed

to methodological differences, as Mohsin et al. (2019) quantified amino acid composition, whereas the present study estimated total protein indirectly via nitrogen determination.

The concentration of total soluble protein in the isolated casein, determined using the Bradford assay, was 1.455 mg/mL. The assay is based on the binding of Coomassie Brilliant Blue G-250 dye to specific amino acid residues, with protein concentration determined from absorbance values. However, the Bradford assay may underestimate protein concentration if the protein contains a low proportion of aromatic residues, thereby limiting dye-binding efficiency (Martine & Vojtech, 2015).

The chemical composition of goat milk varies depending on breed, age, weight, diet, lactation stage, season, and environmental temperature (Mohsin et al., 2019). In this study, individual casein subunits were not quantified. According to Tziboula-Clarke (2002), total casein content in goat milk ranges from 2.33% to 4.63%, with β-casein representing the predominant fraction. The αs1-casein content in goat milk (0-28%) is significantly lower than that in cow milk (50-53.6%). Goat milk composition is highly influenced by lactation stage, with higher levels of primary components during early lactation, followed by a decline and a subsequent increase toward the end of lactation. Additionally, genetic polymorphism of casein within the same breed can contribute to compositional variability (Yadav et al., 2016). Therefore, the relatively low total casein yield observed in this study may be attributed to differences in lactation stage and to potential casein loss during whey protein removal.

Enzyme Hydrolysis

Degree of Hydrolysis (DH)

The degree of hydrolysis (DH) represents the percentage of peptide bonds cleaved by the enzyme. Casein was hydrolysed using bromelain at pH 5.5 and 55˚C for 1, 2, 3, and 4 h. The Michaelis-Menten constant (K_M) of casein for bromelain was reported at 0.138 mM, lower than that of haemoglobin (0.165 mM) (Corzo et al., 2012), indicating that casein is a suitable substrate for bromelain. Figure 1 shows the DH of casein hydrolysate over time, quantified using the o-phthaldialdehyde (OPA) method.

The DH of casein (CSN) hydrolysate ranged from 10.64% to 26.95%, with the highest DH observed after 4 h of hydrolysis (26.95%), indicating that prolonged hydrolysis time resulted in greater casein cleavage by bromelain ($p < 0.05$). The hydrolysis rate increased linearly up to 4 h and did not reach saturation. In comparison, hydrolysis of camel milk by alcalase, chymotrypsin, and papain plateaued after 2 h (Kumar et al., 2016). Although a similar enzyme-to-substrate (E:S) ratio of 1:100 was applied (Kumar et al., 2013), differences in enzyme specificity result in varying DH values.

Shu et al. (2018) reported a maximum DH of 20.4% for bromelain-treated goat milk, with the hydrolysis rate declining after 50 min due to limited availability of cleavable

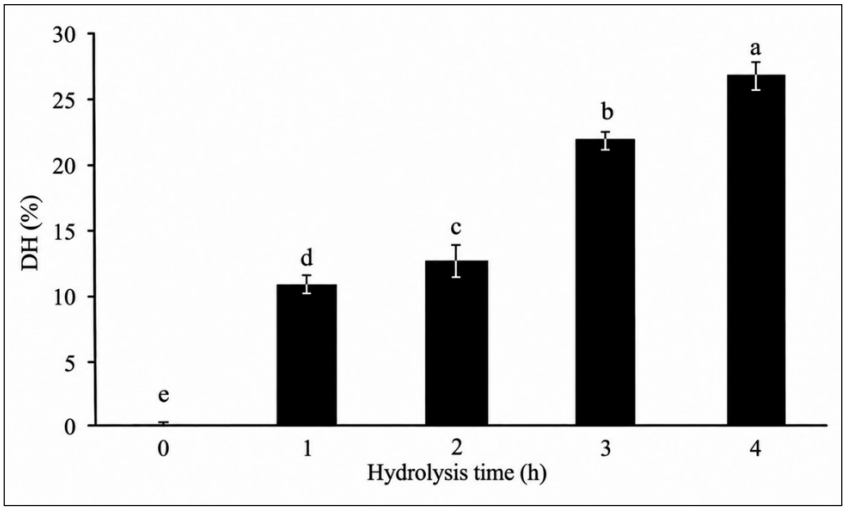


Figure 1. Degree of casein hydrolysis using bromelain. Different letters signify a significant difference ($p<0.05$) between the DH (280 nm) at different hydrolysis times. Data is presented as mean ± SD of triplicates (n=3)

peptide bonds, likely reaching saturation. They also suggested that performing hydrolysis at pH 6.0, close to the isoelectric point (pI) of caseins (4.6) and whey proteins (5.2), can promote protein aggregation, limiting peptide exposure for enzymatic cleavage.

The optimal hydrolysis conditions of the enzyme are crucial to achieving maximum activity. Bromelain exhibits optimal activity at 50°C and pH 6.5, remaining stable at pH 3-7 and 40-60°C (Mohapatra et al., 2013). Corzo et al. (2012) reported that over 90% of bromelain activity occurs at pH 6-7 with casein and sodium caseinate as substrates, and activity increases from 45°C to 59°C but declines significantly above 59°C. To enhance hydrolysis without denaturing the enzyme, this study was conducted at 55°C. For comparison, papain-treated camel casein achieved 16% DH after 5 h at 55°C and pH 6.5 (Kumar et al., 2016). This demonstrates that both optimal hydrolysis conditions and enzyme–substrate specificity influence enzymatic activity and DH.

Peptide Content in Casein Hydrolysate

Bromelain cleaves long polypeptide chains in casein into smaller peptides, releasing free amino acids that serve as an indicator of proteolytic activity. Figure 2 shows that peptide content increased with hydrolysis time, indicating that bromelain cleaves at multiple sites. Peptide concentration rose significantly ($p < 0.05$) from 0.14 to 3.90 mg/mL over the hydrolysis period, with the most substantial increase occurring during the first hour (0.14 to 1.96 mg/mL), consistent with the initial rise in DH (0.00 to 10.64%).

The results confirm that bromelain progressively hydrolyses casein, producing peptides in a time-dependent manner, with rapid cleavage occurring in the early stages of hydrolysis.

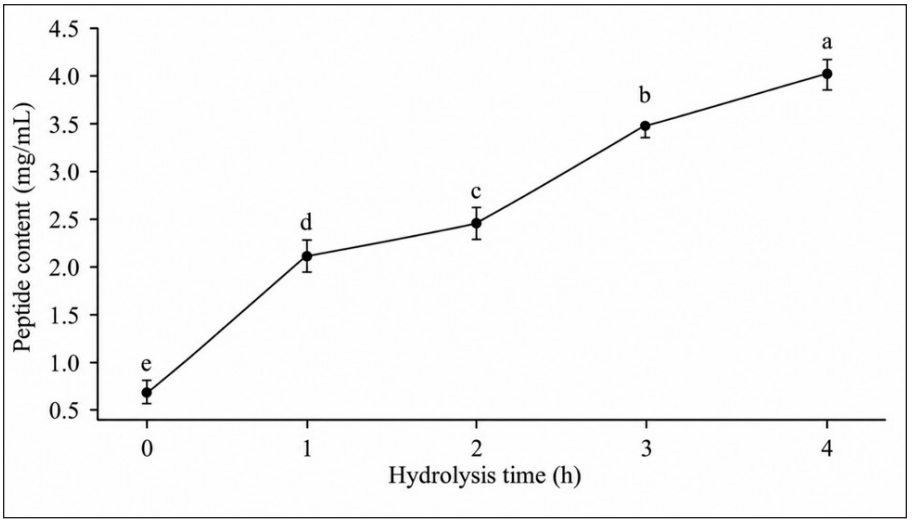


Figure 2. Peptide content of hydrolysed casein. Different letters indicate a significant difference ($p<0.05$) between the peptide content (280 nm) at different hydrolysis times

Proteolytic Activity and Molecular Weight Distribution of Hydrolysed Casein

The proteolytic activity of bromelain in this study ranged from 3 to 7 units/mg of protein. One unit of bromelain was defined as the amount of enzyme that releases 1.0 µmol of ρ-nitrophenol from N-α-Z-L-lysine ρ-nitrophenol per minute at pH 4.5 and 25°C (Medeiros et al., 2014). Bromelain cleaves peptide bonds adjacent to lysine, arginine, phenylalanine, and tyrosine residues (Mudgil et al., 2019). Medeiros et al. (2014) reported a preference for hydrolysing glutamic acid, aspartic acid, lysine, or arginine at the P1 site.

Substrates suggested for bromelain hydrolysis include azocasein, azoalbumin, casein, sodium caseinate, and haemoglobin. These substrates exhibit low K_M values (azoalbumin 0.026, azocasein 0.037, sodium caseinate 0.088, casein 0.138, haemoglobin 0.165 mM) and high V_{max}/K_M ratios (azoalbumin 0.096, azocasein 0.104, sodium caseinate 0.020, casein 0.022, haemoglobin 0.014), indicating efficient enzymatic activity (Corzo et al., 2012). Saptarini et al. (2019) also reported a V_{max} of 8.11 IU/mg·min and K_M of 1.00 mg/mL for bromelain degrading casein. A low K_M reflects high substrate affinity, confirming that casein is a suitable substrate for bromelain.

SDS-PAGE analysis of hydrolysed casein. Figure 3 showed a significant decrease in molecular weight compared to whole casein. After 1 hour of hydrolysis, casein was rapidly cleaved into peptides smaller than 15 kDa. With increasing hydrolysis time, the number of low-molecular-weight peptide bands increased, indicating progressive breakdown of larger peptides. These observations are consistent with Wang et al. (2013), who reported that enzymatic hydrolysis results in the gradual disintegration of high-molecular-weight

peptides into smaller fragments. This indicates that bromelain efficiently cleaves casein into low-molecular-weight peptides, which correlates with the observed increase in the degree of hydrolysis and peptide content over time.

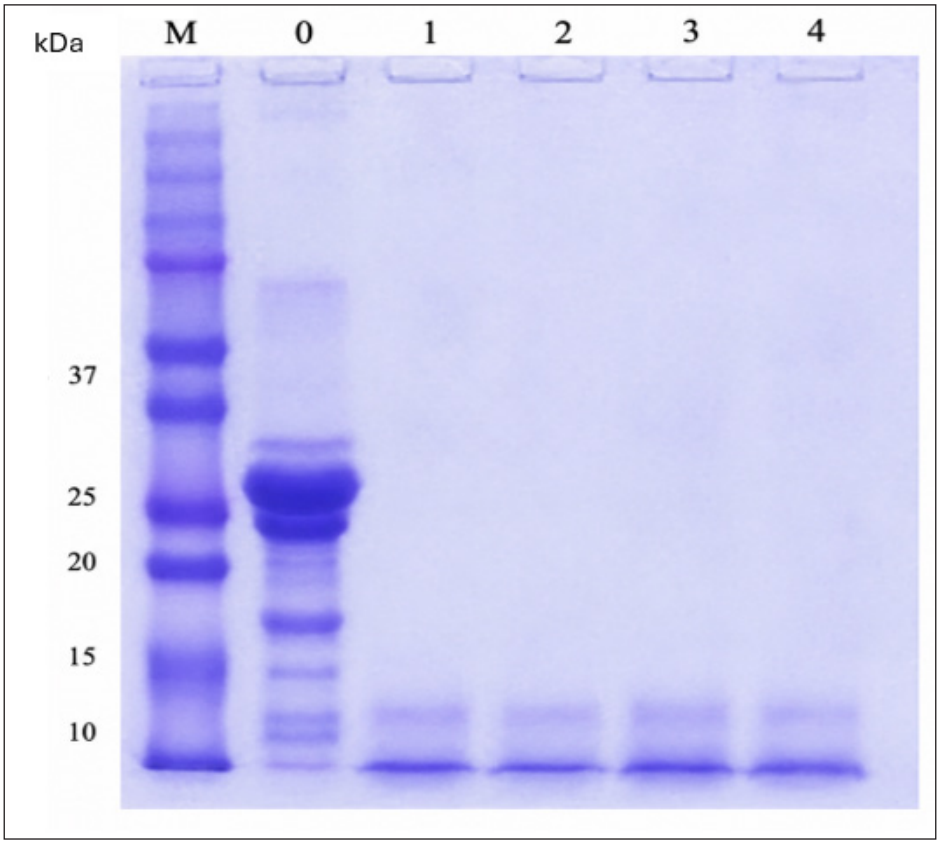


Figure 3. SDS-PAGE of casein hydrolysate. M: Marker; 0: unhydrolyzed casein; 1-4: Casein hydrolysed at 1, 2, 3, and 4 h. Marker was loaded in 10 µL. Each well was loaded with 10 ug/mL of protein. SDS-PAGE was stained using Coomasie Blue G-250

Implications of Molecular Weight on Antigenicity

Bromelain hydrolysis cleaved casein into peptides with molecular weights below 15 kDa after just 1 h. Enzymatic hydrolysis can yield either extensively or partially hydrolysed proteins, though these terms are not strictly defined. Extensively hydrolysed proteins often produce bitterness and astringency, which are undesirable in milk products. Partially hydrolysed peptides are preferred for sensory qualities; however, their antigenicity or allergenicity may persist (Lau et al., 2025).

The molecular weight of peptides is critical for immunogenicity. Peptides with lower molecular weights generally show reduced allergenic potential. For example, fractions

of whey protein concentrate (WPC) and β-lactoglobulin hydrolysates with molecular weights between 1,000 and 5,000 Da exhibited the lowest antigenicity (Svenning et al., 2000). Extensively hydrolysed casein-based formulas with molecular weights below 3,000 Da were administered to 30 CMPA patients, all of whom showed no adverse reactions, demonstrating safety for IgE-mediated cow's milk allergy (Terheggen-Largo et al., 2002).

Despite enzymatic hydrolysis, residual antigenicity may remain. Papain-treated bovine α- and β-casein exhibited reduced, but not fully eliminated, antigenicity (Liu et al., 2012). Similarly, buffalo milk αs1-casein hydrolysates treated with trypsin, α-chymotrypsin, or pepsin exhibited maximum reductions of 85%, 63%, and 60%, respectively, yet infants with CMPA still reacted to these hydrolysates (Ahmad et al., 2016). These studies underscore that enzymatic treatment can reduce, but not always completely abolish, the antigenicity of casein.

Sandwich Enzyme-Linked Immunosorbent Assay (ELISA)

To evaluate the antigenicity of the hydrolysed casein, the hydrolysates were tested for their binding to in-house αs1-casein polyclonal antibodies using a sandwich ELISA. In this assay, N-terminal (Nter) IgG antibodies were coated on high-binding 96-well plates to capture casein peptides. Detection was achieved using a secondary enzyme-conjugated antibody targeting a different epitope on the same antigen. This setup allows quantification of residual antigenicity in hydrolysed

Table 1 shows that ELISA coated with Cter-IgG exhibited a higher background signal compared to ELISA coated with Nter-IgG alone. This observation may be due to Cter-IgG binding directly onto the plate, resulting in non-specific binding (NSB) of antibodies. NSB can negatively affect both the accuracy and reproducibility of ELISA measurements. A sufficient blocking step is therefore essential to saturate unoccupied sites on the solid phase and minimise NSB (Jiang et al., 2021). Blocking has been identified as a major contributor to NSB; however, Güven et al. (2014) reported that extending blocking from 30 min to 1 h did not significantly reduce NSB. Common blocking solutions, such as BSA, skimmed milk, or casein, can inhibit hydrophobic and electrostatic interactions of proteins, thereby reducing background signal. These solutions may contain bovine IgG, which

Table 1
Negative controls (NC) were included in the ELISA test

Negative controls (NC)	Mean ± standard error of the mean (SEM)
NC 1[a]	0.411 ± 0.018
NC 2[b]	0.537 ± 0.001
NC 3[c]	0.569 ± 0.013

Note. Data are reported as the mean of triplicates ± standard error of the mean (SEM), calculated as SD/√n, where n = 3
[a]ELISA coated with Nter IgG, followed by PBS and Strep-HRP.[b] ELISA coated with Nter IgG, followed by PBS, Cter IgG, and Strep-HRP.[c] ELISA coated with Nter IgG, followed by bovine α-casein, Cter IgG, and Strep-HRP

can bind secondary antibodies and induce cross-reactivity. Insufficient washing with non-ionic detergents, such as Tween-20 or Triton X, can further contribute to NSB in ELISA assays (Buchwalow et al., 2011).

Table 2
Normalised absorbance values of untreated, treated, and hydrolysed caseins tested in the ELISA study

Caseins		Normalised absorbance values (mean ± SEM)
Untreated		0.464 ± 0.031^{a}
Heat treatment	60˚C	0.257 ± 0.037^{d}
	80˚C	0.315 ± 0.025^{d}
	90˚C	0.383 ± 0.022^{c}
Heat and chemical treatment	60˚C + 0.1% SDS	0.440 ± 0.016^{b}
	80˚C + 0.1% SDS	0.498 ± 0.022^{a}
	90˚C + 0.1% SDS	0.516 ± 0.019^{a}
Hydrolysed casein	1 h	0.141 ± 0.004^{g}
	2 h	0.178 ± 0.006^{f}
	3 h	0.208 ± 0.032^{f}
	4 h	0.268 ± 0.016^{e}

Note. Caseins were hydrolysed by bromelain at an S:E ratio of 100:1, pH 5.5, and 55˚C (Section 3.4). Different small letters indicate significant differences ($p<0.05$) between absorbance values at 450 nm

The antigenicity of whole casein and hydrolysed caseins was evaluated using indirect ELISA. Whole casein treated at 90˚C with 0.1% SDS produced the highest ELISA signal, indicating effective dissociation of casein micelles. Heat treatment is known to destabilise casein micelles by dissociating κ-casein (Dumpler et al., 2017), and at temperatures above 80˚C, β-lactoglobulin from whey binds to κ-casein via disulfide bonds, further destabilising micelles (Oka et al., 2022). SDS, an anionic surfactant, disrupts protein tertiary structure and cleaves disulfide bonds, thereby exposing epitopes for antibody binding (Anand et al., 2013). The combination of heat and SDS effectively dissociated casein more than heat alone.

Bromelain hydrolysis reduced the antigenic reactivity of goat milk casein, as indicated by the decreased ELISA response compared with untreated casein. ELISA results (Table 2) showed significant differences ($p < 0.05$) between the hydrolysed and untreated casein samples. Previous studies showed that the effectiveness of protein hydrolysis depends on the type of enzyme and the hydrolysis conditions. Liu et al. (2012) found that papain reduced the antigenicity of α-casein under optimised conditions of pH 5.6, 50°C, 2.5% E/S, and 90 min. Wang et al. (2013) reported that trypsin hydrolysis increased rapidly during the first 20 min and reached a DH of 45.2% after 45 min, showing active peptide formation during the early stage of hydrolysis. Medeiros et al. (2014) found that bromelain

could produce partially hydrolysed milk, while Ahmed et al. (2015) reported that different enzymes showed different levels of αs1-casein antigenicity reduction. Kumar et al. (2016) also showed that DH and hydrolysis time varied depending on the enzyme used. For goat's milk, Shu et al. (2016) reported that alcalase showed optimal hydrolysis at 150 min, 50°C, pH 8.0, and 2% E/S. Wang et al. (2023) further showed that alcalase–Neutrase hydrolysis achieved 26.45% DH and a 59.41% reduction in antigenicity. Overall, these studies show that enzyme type and hydrolysis time are important factors influencing peptide formation and the reduction of antigenicity.

Based on these findings, the present study further evaluated the effect of hydrolysis time on the antigenic reactivity of goat's milk casein using bromelain. ELISA results showed that 1 h hydrolysis produced the lowest antigenic reactivity (69.6%) among the hydrolysis times tested. The lower ELISA response suggests that bromelain hydrolysis reduced the ability of goat's milk casein to react with the antibodies. This may be due to the breakdown of protein regions recognised by the antibodies, resulting in the formation of smaller peptides. However, longer hydrolysis did not result in a further decrease in antigenic reactivity. This may be due to changes in the types of peptides formed during prolonged hydrolysis. Therefore, 1 h was identified as the most effective hydrolysis time for reducing antigenic reactivity under the conditions tested.

Affinity constants (Kd) indicated that Nter-IgG had stronger binding than Cter-IgG, while both antibodies had lower affinity for native αs1-casein compared to peptide-BSA conjugates (Mohsin et al., 2020). These observations highlight challenges in quantifying antigenicity by ELISA when both N- and C-terminal epitopes are present.

Previous studies align with these findings. Trypsin-treated buffalo αs1-casein exhibited reduced immunoreactivity, yet residual allergenicity persisted among CMPA patients (Ahmad et al., 2016). Papain-treated bovine α- and β-casein demonstrated that DH negatively correlated with binding inhibition, although high DH did not always eliminate antigenicity (Liu et al., 2012). These studies emphasise that enzyme specificity, peptide length, and epitope exposure significantly influence the observed antigenicity.

Industrial Feasibility and Practical Considerations

Although bromelain hydrolysis reduced the antigenic reactivity of goat's milk casein under the conditions investigated, several practical considerations need to be addressed before application at an industrial scale. Extensive proteolysis may generate small peptides that can contribute to bitterness or other undesirable sensory characteristics; therefore, sensory evaluation and bitterness assessment are necessary to determine consumer acceptability. In addition, the nutritional quality and digestibility of the resulting hydrolysates should be evaluated to ensure that hydrolysis does not adversely affect their nutritional value and to determine their suitability for use in specialised dairy products. From a hydrolysis

perspective, the scalability of bromelain hydrolysis should be investigated by considering enzyme dosage, reaction time, temperature control, mixing, enzyme inactivation, and batch-to-batch consistency. The economic feasibility of the process would also depend on bromelain cost, enzyme utilisation efficiency, process yield, and the costs associated with subsequent purification or drying steps. Furthermore, regulatory acceptance would require appropriate evidence of safety and efficacy for the intended application and in accordance with applicable food regulations. Therefore, while the present study provides a preliminary scientific basis for using bromelain to reduce the antigenic reactivity of goat milk casein, further sensory, nutritional, digestibility, techno-economic, scale-up, and regulatory studies are required before its suitability for commercial functional dairy products can be established.

CONCLUSION

In conclusion, bromelain hydrolysis at pH 5.5 and 55 °C effectively reduced the antigenicity of goat milk casein, as indicated by the highest degree of hydrolysis (DH) and peptide content, together with the lowest ELISA response. The antigenicity assessment in this study was performed using an in-house sandwich ELISA with rabbit IgG antibodies; therefore, the reduced ELISA response reflects reduced antigenic reactivity and does not establish clinical allergenicity or confirmed hypoallergenicity. Further investigation using clinically relevant immunological approaches is necessary to assess its relevance to specialised dairy and infant nutrition applications.

ACKNOWLEDGEMENT

This work was financially supported by Ministry of Higher Education Malaysia through Fundamental Research Grant Scheme [Grant number: FRGS/1/2025/STG02/UPM/02/1] and Universiti Putra Malaysia under Geran Putra [Grant number: GP/2020/9693800]. The authors would like to acknowledge University Driven Research Programs (UDRP) s: Functional Food Development and Future Food, Faculty of Food Science and Technology and Human Nutrition and Health: Food Safety (including Food Traceability and Authenticity), Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia for providing resources and facilities..

LIST OF ABBREVIATION

AAF	Amino Acid Formula
BSA	Bovine Serum Albumin
CMPA	Cow Milk Protein Allergy
CSN	Casein

Cter	C-terminal
DH	Degree of Hydrolysis
eHF	Extensively Hydrolysed Formula
ELISA	Enzyme-Linked Immunosorbent Assay
IgE	Immunoglobulin E
IgG	Immunoglobulin G
NC	Negative Control
NSB	Non-Specific Binding
Nter	N-terminal
OPA	o-Phthaldialdehyde
PBS	Phosphate-Buffered Saline
pHF	Partially Hydrolysed Formula
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
α-LA	Alpha-Lactalbumin
αs1-casein	Alpha-s1-Casein
β-LG	Beta-Lactoglobulin

REFERENCES

Ahmad, N., Imran, M., Khan, M. K., & un Nisa, M. (2016). Degree of hydrolysis and antigenicity of buffalo alpha S1 casein and its hydrolysates in children with cow milk allergy. *Food and Agricultural Immunology*, *27*(1), 87-98. https://doi.org/10.1080/09540105.2015.1079595

Ahmed, A. S., El-Bassiony, T., Elmalt, L. M., & Ibrahim, H. R. (2015). Identification of potent antioxidant bioactive peptides from goat milk proteins. *Food Research International*, *74*, 80-88. https://doi.org/10.1016/j.foodres.2015.04.032

Anand, U., & Mukherjee, S. (2013). Reversibility in protein folding: Effect of β-cyclodextrin on bovine serum albumin unfolded by sodium dodecyl sulphate. *Physical Chemistry Chemical Physics*, *15*(23), 9375-9383. https://doi.org/10.1039/C3CP50207D

AOAC International. (2023). *Official methods of analysis of AOAC International* (18th ed.).

Ballabio, C., Chessa, S., Rignanese, D., Gigliotti, C., Pagnacco, G., Terracciano, L., Fiocchi, A., Restani, P., & Caroli, A. M. (2011). Goat milk allergenicity as a function of αS1-casein genetic polymorphism. *Journal of Dairy Science*, *94*(2), 998-1004. https://doi.org/10.3168/jds.2010-3545

Benjamin-van Aalst, O., Dupont, C., van der Zee, L., Garssen, J., & Knipping, K. (2024). Goat milk allergy and a potential role for goat milk in cow's milk allergy. *Nutrients*, *16*(15), Article 2402. https://doi.org/10.3390/nu16152402

Bio-Rad. (n.d.). *Ion exchange chromatography*. https://www.bio-rad.com/en-my/applications-technologies/ion-exchange-chromatography?ID=MWHAY9ESH

Buchwalow, I., Samoilova, V., Boecker, W., & Tiemann, M. (2011). *Non-specific binding of antibodies in immunohistochemistry: Fakes and facts* [Preprint]. Nature Proceedings. https://doi.org/10.1038/npre.2011.5892.1

Corredig, M., Nair, P. K., Li, Y., Eshpari, H., & Zhao, Z. (2019). Invited review: Understanding the behavior of caseins in milk concentrates. *Journal of Dairy Science*, *102*(6), 4772-4782. https://doi.org/10.3168/jds.2018-15943

Corzo, C. A., Waliszewski, K. N., & Welti-Chanes, J. (2012). Pineapple fruit bromelain affinity to different protein substrates. *Food Chemistry*, *133*(3), 631-635. https://doi.org/10.1016/j.foodchem.2011.05.119

Dumpler, J., Wohlschläger, H., & Kulozik, U. (2017). Dissociation and coagulation of caseins and whey proteins in concentrated skim milk heated by direct steam injection. *Dairy Science and Technology*, *96*(6), 807-826. https://doi.org/10.1007/s13594-016-0304-3

Food and Drug Administration. (2026). *Food allergies*. https://www.fda.gov/food/nutrition-food-labeling-and-critical-foods/food-allergies

Güven, E., Duus, K., Lydolph, M. C., Jørgensen, C. S., Laursen, I., & Houen, G. (2014). Non-specific binding in solid phase immunoassays for autoantibodies correlates with inflammation markers. *Journal of Immunological Methods*, *403*(1-2), 26-36. https://doi.org/10.1016/j.jim.2013.11.014

Hodgkinson, A. J., McDonald, N. A., Kivits, L. J., Hurford, D. R., Fahey, S., & Prosser, C. (2012). Allergic responses induced by goat milk αS1-casein in a murine model of gastrointestinal atopy. *Journal of Dairy Science*, *95*(1), 83-90. https://doi.org/10.3168/jds.2011-4829

Jensen, S. A., Fiocchi, A., Baars, T., Jordakieva, G., Nowak-Wegrzyn, A., Pali-Schöll, I., Passanisi, S., Pranger, C. L., Roth-Walter, F., Takkinen, K., Assa'ad, A. H., Venter, C., & Jensen-Jarolim, E. (2022). World Allergy Organization (WAO) diagnosis and rationale for action against cow's milk allergy (dracma) guidelines update - iii - cow's milk allergens and mechanisms triggering immune activation. *World Allergy Organization Journal*, *15*(9), Article 100668. https://doi.org/10.1016/j.waojou.2022.100668

Jiang, X., Wu, M., Albo, J., & Rao, Q. (2021). Non-specific binding and cross-reaction of ELISA: A case study of porcine hemoglobin detection. *Foods*, *10*(8), Article 1708. https://doi.org/10.3390/foods10081708

Kumar, D., Chatli, M. K., Singh, R., Mehta, N., & Kumar, P. (2016). Enzymatic hydrolysis of camel milk casein and its antioxidant properties. *Dairy Science and Technology*, *96*(3), 391–404. https://doi.org/10.1007/s13594-015-0275-9

Kumar, S., Chouhan, V. S., Sanghi, A., & Teotia, U. V. S. (2013). Antioxidative effect of yak milk caseinates hydrolyzed with three different proteases. *Veterinary World*, *6*(10), 799-802. https://doi.org/10.14202/vetworld.2013.799-802

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680-685. https://doi.org/10.1038/227680a0

Lau, A. H. C., Tsui, V. T. K., Hung, K. I., Le, T. M. H., Liu, C. W. S., Tse, S. P. K., & Kong, H. K. (2025). Evaluating allergenicity of protein fragments and peptides in partially hydrolysed infant formulas available in Hong Kong. *Food Chemistry*, *487*, Article 144758. https://doi.org/10.1016/j.foodchem.2025.144758

Liu, X., Luo, Y., & Li, Z. (2012). Effects of pH, temperature, enzyme-to-substrate ratio and reaction time on the antigenicity of casein hydrolysates prepared by papain. *Food and Agricultural Immunology*, *23*(1), 69-82. https://doi.org/10.1080/09540105.2011.604770

Malaysian Paediatric Association & Malaysian Society of Allergy and Immunology. (2024). *Guidelines for the diagnosis and management of cow's milk protein allergy (CMPA) in children* (2nd ed.). https://mpaeds.my/wp-content/uploads/2024/07/240524_CMPA-Guidelines-2024_FA-Digital.pdf

Mansor, M., Al-Obaidi, J. R., Ismail, I. H., Abidin, M. A. Z., Zakaria, A. F., Lau, B. Y. C., Mohsin, A. Z., Sukor, R., Selamat, J., Mahmud, N. K., & Jambari, N. N. (2023). Cross-reactivity analysis of milk proteins from different goat breeds with cow's milk allergens using a proteomic approach. *Molecular Immunology*, *155*, 44-57. https://doi.org/10.1016/j.molimm.2022.12.016

Medeiros, V., Rainha, N., Paiva, L., Lima, E., & Baptista, J. (2014). Bovine milk formula based on partial hydrolysis of caseins by bromelain enzyme: Better digestibility and angiotensin-converting enzyme-inhibitory properties. *International Journal of Food Properties*, *17*(4), 806-817. https://doi.org/10.1080/10942912.2012.675607

Mohapatra, A., Rao, V. M., & Ranjan, M. (2013). Comparative study of the increased production & characterization of bromelain from the peel, pulp & stem pineapple (*Anannus commas*). *International Journal of Advancements in Research & Technology*, *2*(8), 249-279.

Mohsin, A. Z., Sukor, R., Selamat, J., Meor Hussin, A. S., & Ismail, I. H. (2019). Chemical and mineral composition of raw goat milk as affected by breed varieties available in malaysia. *International Journal of Food Properties*, *22*(1), 815-824. https://doi.org/10.1080/10942912.2019.1610431

Mohsin, A. Z., Sukor, R., Selamat, J., Meor Hussin, A. S., Ismail, I. H., Azri, F. A., Mustafa, R. R., & Mansor, M. (2023). Development of biotin-streptavidin amplified peptide antibody-based ELISA for quantification of αS1-casein in goat's milk. *Food Control*, *145*, Article 109263. https://doi.org/10.1016/j.foodcont.2022.109263

Mohsin, A. Z., Sukor, R., Selamat, J., Meor Hussin, A. S., Ismail, I. H., Jambari, N. N., & Jonet, A. (2020). A highly selective two-way purification method using liquid chromatography for isolating αS2-casein from goat milk of five different breeds. *Journal of Chromatography B*, *1160*, Article 122380. https://doi.org/10.1016/j.jchromb.2020.122380

Mohsin, A. Z., Sukor, R., Selamat, J., Meor Hussin, A. S., Ismail, I. H., Jambari, N. N., & Mustaffa-Kamal, F. (2020). Generation of high affinity anti-peptide polyclonal antibodies recognizing goat αs1-Casein. *Molecules*, *25*(11), Article 2622. https://doi.org/10.3390/molecules25112622

Mudgil, P., Baby, B., Ngoh, Y. Y., Vijayan, R., Gan, C. Y., & Maqsood, S. (2019). Identification and molecular docking study of novel cholesterol esterase inhibitory peptides from camel milk proteins. *Journal of Dairy Science*, *102*(12), 10748-10759. https://doi.org/10.3168/jds.2019-16520

Oka, D., Ono, W., Tamaki, S., Noguchi, T., & Takano, K. (2022). Effects of the thermal denaturation degree of a whey protein isolate on the strength of acid milk gels and the dissociation of κ-casein. *Journal of Dairy Research*, *89*(1), 104-108. https://doi.org/10.1017/s0022029922000103

Saptarini, N. M., Rahayu, D., & Kusuma, S. A. F. (2019). Protease activity and characterization of bromelain extract of pineapple (*Ananas comusus* (l.) merr) crown from Subang, Indonesia. *Rasayan Journal of Chemistry*, *12*(4), 2074-2081. https://doi.org/10.31788/rjc.2019.1245319

Shu, G., Huang, J., Bao, C., Meng, J., Chen, H., & Cao, J. (2018). Effect of different proteases on the degree of hydrolysis and angiotensin i-converting enzyme-inhibitory activity in goat and cow milk. *Biomolecules*, *8*(4), Article 101. https://doi.org/10.3390/biom8040101

Svenning, C., Brynhildsvold, J., Molland, T., Langsrud, T., & Elisabeth Vegarud, G. (2000). Antigenic response of whey proteins and genetic variants of β-lactoglobulin — the effect of proteolysis and processing. *International Dairy Journal*, *10*(10), 699-711. https://doi.org/10.1016/S0958-6946(00)00101-1

Tziboula-Clarke, A. (2002). Goat milk. In H. Roginski, J. W. Fuquay, & P. F. Fox (Eds.), *Encyclopedia of Dairy Sciences* (pp. 1270-1279). Academic Press. https://doi.org/10.1016/B0-12-227235-8/00193-0

Wang, J., Su, Y., Jia, F., & Jin, H. (2013). Characterization of casein hydrolysates derived from enzymatic hydrolysis. *Chemistry Central Journal*, *7*, Article 62. https://doi.org/10.1186/1752-153X-7-62

Wang, Y., Yang, F., Zheng, Y., Xue, Y., Kang, Z., Zulewska, J., Li, H., Li, H., & Yu, J. (2023). Relationship between antigenicity and allergenic epitopes of extensive hydrolyzed whey protein: LC-MS/MS-based analysis. *Food Bioscience*, *54*, Article 102870. https://doi.org/10.1016/j.fbio.2023.102870

Yadav, A., & Naidu, R. (2015). Clinical manifestation and sensitization of allergic children from malaysia. *Asia Pacific Allergy*, *5*(2), 78–83. https://doi.org/10.5415/apallergy.2015.5.2.78

Zarei, M., Forghani, B., Ebrahimpour, A., Abdul-Hamid, A., Anwar, F., & Saari, N. (2015). *In vitro* and *in vivo* antihypertensive activity of palm kernel cake protein hydrolysates: Sequencing and characterization of potent bioactive peptides. *Industrial Crops and Products*, *76*, 112-120. https://doi.org/10.1016/j.indcrop.2015.06.040

REFEREES FOR THE PERTANIKA

JOURNAL OF SCIENCE & TECHNOLOGY

Vol. 34 (5) Oct. 2026

The Editorial Board of the Pertanika Journal of Science and Technology wishes to thank the following:

Abdul Shibghatullah
(UCSI, Malaysia)

Ahmad Syakirin Ismail @ Rosdi
(UMPSA, Malaysia)

Alaa Jaber
(UOT, Iraq)

Alaa M. Rashad
(SU, Saudi Arabia)

Arfat Ahmad Khan
(KKU, Thailand)

Chi Wee Tan
(TAR UMT, Malaysia)

Christopher Heng Yii Sern
(USM, Malaysia)

David Perez-Jorge
(ULL, Spain)

Dayang Azra Awang Mat
(UNIMAS, Malaysia)

Eg Su Goh
(UTM, Malaysia)

Han Bing Chua
(CURTIN, Malaysia)

Hesam Kamyab
(KU, South Korea)

Husnul Azan Tajarudin
(USM, Malaysia)

Jafar A. Alzubi
(BAU, Jordan)

K. Kannan
(SASTRA, India)

K. Senthil Kumar
(REC, India)

Kazim Yildiz
(MU, Türkiye)

Keng-Hoong Ng
(UTAR, Malaysia)

Kian Fe Yee
(XMU, Malaysia)

Leonard Agamez Polo
(UNISINÚ, Colombia)

M. Muthu Krishnan
(KIT, India)

Mahmoud Ebrahimi
(UMAR, Iran)

Masnida Hussin
(UPM, Malaysia)

Mohamed Thariq Hameed Sultan
(UPM, Malaysia)

Mohammad Khoveyni
(IAU, Iran)

Mohd Fairouz Mohd Yousof
(UTHM, Malaysia)

Mohd Khair Hassan
(UPM, Malaysia)

Muhammad Ammirrul Atiqi Mohd Zainuri
(UKM, Malaysia)

Nor Hana Mamat
(TATIUC, Malaysia)

Norlaila Mohd Zanuri
(USM, Malaysia)

Pavol Bozek
(STUBA, Slovakia)

Pradeep Johnson
(HICET, India)

Purnawarman Musa
(UG, Indonesia)

Pushpa Rani K.
(MLRIT, India)

Saeid Amini
(UT, Iran)

Sales Aribe Jr.
(BukSU, Philippines)

Scheherazed Dakhmouche Djekrif
(UFMC1, Algeria)

Sheikh Mohd Ghazali Sheikh Ahmad Izaddin
(UiTM, Malaysia)

Siew Ling Lee
(UTM, Malaysia)

Siti Rabiatull Aisha Idris
(UMPSA, Malaysia)

Sivaram Rajeyyagari
(SU, Saudi Arabia)

Stephanie Chua
(UNIMAS, Malaysia)

Suleyman Adak
(MAU, Türkiye)

Suresh Kallam
(JU, India)

Syahrun Nizam Md Arshad
(UniMAP, Malaysia)

Thulasi Bikku
(AMRITA, India)

Wan Saidatul Syida Wan Kamarudin
(UiTM, Malaysia)

Wei Jer Lim
(SC, Malaysia)

Yasin Kaya
(ATU, Türkiye)

AMRITA - *Amrita Vishwa Vidyapeetham*
ATU - *Adana Alparslan Türkeş Science and Technology University*
BAU - *Al-Balqa Applied University*
BukSU - *Bukidnon State University*
CURTIN - *Curtin University Malaysia*
HICET - *Hindusthan College of Engineering and Technology*
IAU - *Islamic Azad University*
JU - *JAIN Deemed to be University*
KIT - *Kalaignar Karunanidhi Institute of Technology*
KKU - *Khon Kaen University*
KU - *Korea University*
MAU - *Mardin Artuklu University*
MLRIT - *MLR Institute of Technology*
MU - *Marmara University*
REC - *Rajalakshmi Engineering College*
SASTRA - *SASTRA Deemed to be a University*
SC - *Southern College*
STUBA - *Slovak University of Technology in Bratislava*
SU - *Shaqra University*
TAR UMT - *Tunku Abdul Rahman University of Management and Technology*
TATIUC - *TATI University College*
UCSI - *UCSI University*
UFMC1 - *Université Frères Mentouri Constantine 1*
UG - *Gunadarma University*
UiTM - *Universiti Teknologi MARA*
UKM - *Universiti Kebangsaan Malaysia*
ULL - *Universidad de La Laguna*
UMAR - *University of Maragheh*
UMPSA - *Universiti Malaysia Pahang Al-Sultan Abdullah*
UniMAP - *Universiti Malaysia Perlis*
UNIMAS - *Universiti Malaysia Sarawak*
UNISINÚ - *Universidad del Sinú*
UOT - *University of Technology, Iraq*
UPM - *Universiti Putra Malaysia*
USM - *Universiti Sains Malaysia*
UT - *University of Tehran*
UTAR - *Universiti Tunku Abdul Rahman*
UTHM - *Universiti Tun Hussein Onn Malaysia*
UTM - *Universiti Teknologi Malaysia*
XMU - *Xiamen University Malaysia*

While every effort has been made to include a complete list of referees for the period stated above, however if any name(s) have been omitted unintentionally or spelt incorrectly, please notify the Chief Executive Editor, *Pertanika* Journals at executive_editor.pertanika@upm.edu.my

Any inclusion or exclusion of name(s) on this page does not commit the *Pertanika* Editorial Office, nor the UPM Press or the university to provide any liability for whatsoever reason.

Diseases of Vegetable Plant Recognition Based on Enhanced SqueezeNet 2981
Kifaa Hadi Thanoon, Akram A. Abdulrahman, Ali Ibrahim Ahmed, Omar Ibrahim Alsaif, and Ibrahim Ahmed Saleh

A Novel Fusion of Computer Vision Capabilities and Deep Learning Techniques for OCT Retinal Disease Diagnosis 3003
Nagham Salim Mohammed, Omar S. Almolaa, and Omar Ibrahim Alsaif

Effects of Elevated CO_2-Grown Napier Grass and Cassava Starch Binder Ratios on Biomass Briquette Properties 3025
Muhammad Zulhilmi Mohd Nasirudin and Siti Zaharah Sakimin

Automated News Categorisation Using Transformer-Based Deep Learning Models: A Step toward Intelligent Media Classification 3047
B. Rupa Devi, Malyala Gayatri, Rajesh Lingam, Penubaka Balaji, S. Beatrice, and S. Athinarayanan

Review Article 3075
Nitrate-reducing Sulphide-Oxidising Bacteria (NR-SOB) in Corrosion Control: A Review
Aishah Zafre, Norinsan Kamil Othman, Najmiddin Yaakob, Azman Jalar, Maria Abu Bakar, and Rabiahtul Zulkafli

Mobile-Based Inventory Counting Using 2D Object Detection and Layer-Based Inference 3109
Gan Heng Lai, Goh Eg Su, and Johanna Ahmad

Bridging the Ocean Literacy Gap: Teacher Needs and Design Requirements for Virtual Reality-Based Sustainable Maritime Education for Coastal High Schools in Tanjungpinang, Indonesia 3131
Nurul Asikin, Arinda Eka Lidiastuti, Adam Fernando, and Rahmania Pamungkas

Antigenicity of Hydrolysed Casein in Goat's Milk Using Bromelain 3163
Nurul Izza Nabilah Mohd Samsudin, Chai Jie Ying, Rashidah Sukor, Aliah Zannierah Mohsin, and Siva Raseetha

Capacity-Limited Cooling Explains Simultaneous Energy Waste and Thermal Discomfort in Two University Buildings in the Southern Philippines 2715
Fejaycris O. Pillodar, Dioniela Mae C. Ellos, Jiza Gay P. Jarabe, John Russel A. Dalogdog, Monaliza Joy Z. Magsayo, Charlene Mae Z. Gaab, Manuel Anthony P. Jondonero, Melgie A. Alas, Rovick Pelos P. Tarife, Noel R. Estoperez, Mark June L. Aporador, Napoleon A. Enteria, Kim Emissary C. Magarin, Ryan Lloyd R. Carpentero, Hernando P. Bacosa, and Peter D. Suson

Regenerative Braking: A Range Extender for Electric Motorcycles in Hilly Terrain 2739
Mawada Ahmed, Siti Fauziah Toha, Khairul Affendy, Muhammad Abdullah, Salmiah Ahmad, Mohammad Osman Tokhi, and Ahmad Syahrin Idris

An Integrated Multi-Source Framework for Explainable Android Malware Detection with SOC Integration 2765
Pallavi M. Shimpi and Nitin Namdeo Pise

Time-dependent Polyethylene Terephthalate (PET) Biodegradation by *Aspergillus niger* and *Aspergillus oryzae* under Solid State Fermentation 2791
Latha Shankar, Norazlina Idris, Nurhafizah Ibrahim, Norakma, M. N, Rodiah, M. H, Wan Zarina, W. M, Noor Fazreen, D, and Nur Ain Izzati Mohd Zainudin

Fractional SIRM Epidemic Model with Microplastic-Mediated Environmental Transmission 2811
R. Sengothai, P. Balaganesan, D. Adhil Basha, and J. Renuka

Review Article 2835
Wireless Sensor Network and Leveraging IoT on Agro Farming: A Review
Muhammad Azwan Nasirudin, Mohammad Nishat Akhtar, and Farzad Ismail

Grey-Fuzzy Based Optimization of Machining Parameters in Turning of AA 7068/TiC MMCs 2887
Syed Altaf Hussain, Sudhamadhuri K, Upendra Rajak, and Abhishek Dasore

Review Article 2907
A Systematic Review of Hybrid Deep Learning and Optimization Techniques to Address Class Imbalance and Malware Detection in IoT Network
Ernest Ezema, and Julia Juremi

Context-Aware RFID-IoT Counterfeit Prevention: A Simulation-Based Proof of Concept Using Layered RBAC, MAC, ABAC, and ECA 2929
Ravichandran and Irene Anastasia Patric

Privacy-Preserving Federated Learning with Explainable AI for Multi-Site Autism Spectrum Disorder Biomarker Discovery 2947
Anusha Ravankol, and Nandula Anuradha

Pertanika Journal of Science & Technology
Vol. 34 (5) Oct. 2026

Content

Foreword xiii
Mohamed Thariq Hameed Sultan

Data-Driven System Identification Modelling for the Reactor TRIGA PUSPATI Using Real Operational Data *2545*
Nor Arymaswati Abdullah, Azura Che Soh, Siti Marhainis Othman, Ribhan Zafira Abdul Rahman, Samsul Bahari Mohd Noor, and Julia Abdul Karim

Pneumonia Screening and Prediction on Images using a Convolutional Neural Network Deep Learning Model 2569
Shuria Saaidin, Murizah Kassim, I Gede Pasek Suta Wijaya, and Nur Amirah Kamaluddin

Artificial Neural Network-Based UHF Partial Discharge Localisation in Mineral Oil with Datasets based on TDOA 2597
Zhi Yang Lim, Norhafiz Azis, Hafiz Rashidi Ramli, Mohd Amran Mohd Radzi, Nor Mohd Haziq Norsahperi, and Azrul Mohd Ariffin

Investigation of Assistive Grouser Traction Force in Toyoura Sand: Effects of Grouser Geometry and Contact Angle 2619
Mochamad Edoward Ramadhan, Ahmad Najmuddin Ibrahim, and Yasuhiro Fukuoka

Formulation of Regulatory Sandbox for Nanotechnology Innovation 2649
Nor Akhmal Hasmin, Juan Matmin, and Ainul Hafiza Zainudin

The Use of FSR-based Plantar Pressure Patterns for Deception Detection 2675
Gerelbat Batgerel, and Chunki Kwon

Urban Flooding as a Hydrological Connectivity Problem: An Entropy-Based GIS Assessment in a Low-Relief Tropical City 2691
Fachruddin Abdul Chaidir, Masha Nur Salsabiela Menhat, Mohd Azhafiz Abdullah, Muhammad Azraie Abdul Kadir, Effi Helmy Ariffin, Khairul Nizam Abdul Maulud, Ahmad Shakir Mohd Saudi, and Hasdinar Umar

Pertanika Editorial Office, Journal Division,
Putra Science Park,
1st Floor, IDEA Tower II,
UPM-MTDC Center,
Universiti Putra Malaysia,
43400 UPM Serdang,
Selangor Darul Ehsan
Malaysia

http://www.pertanika.upm.edu.my
Email: executive_editor.pertanika@upm.edu.my
Tel. No.: +603- 9769 1622

PENERBIT UPM
UPM PRESS

http://penerbit.upm.edu.my
Email: penerbit@upm.edu.my
Tel. No.: +603- 9769 3606

e-ISSN 2231-8526
9 772231 852000