University of Hertfordshire UH

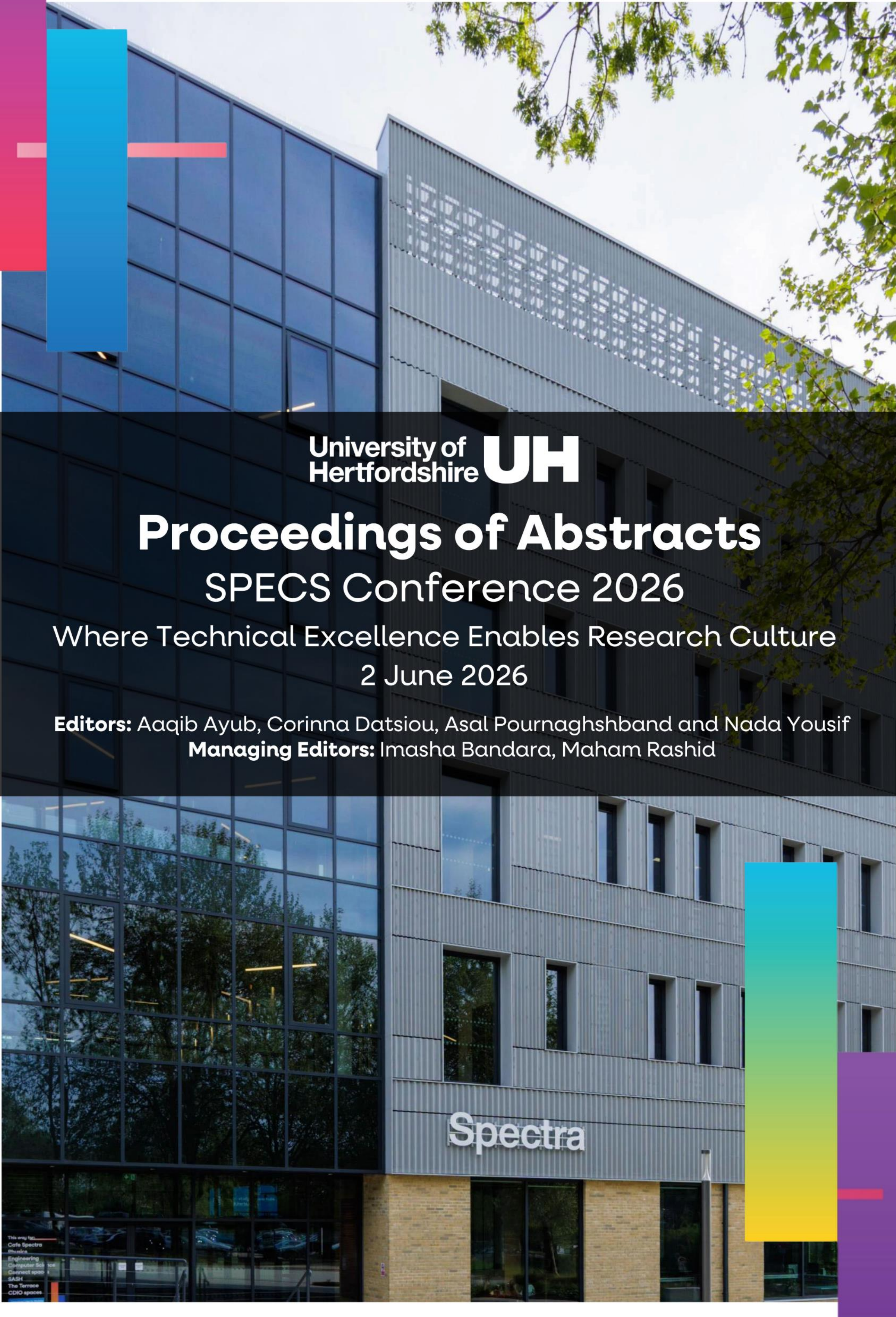
# Proceedings of Abstracts

## SPECS Conference 2026

### Where Technical Excellence Enables Research Culture

2 June 2026

**Editors:** Aaqib Ayub, Corinna Datsiou, Asal Pournaghshband and Nada Yousif
**Managing Editors:** Imasha Bandara, Maham Rashid



https://doi.org/10.18745/pb.27307

***Proceedings of Abstracts: School of Physics, Engineering and Computer Science Conference, 2nd June 2026, University of Hertfordshire***

**Editors:** Aaqib Ayub, Corinna Datsiou, Asal Pournaghshband, Nada Yousif

**Managing Editors:** Imasha Bandara, Maham Rashid

**Published online:** UH Research Archive, 2nd June 2026

**Published by**

School of Physics, Engineering and Computer Science

University of Hertfordshire

College Lane, Hatfield, AL10 9AB, United Kingdom

**RIS ID:** 81533947

**Title:** Proceedings of Abstracts: School of Physics, Engineering and Computer Science Conference 2026

**DOI:** https://doi.org/10.18745/pb.27307

**UHRA handle:** https://uhra.herts.ac.uk/id/eprint/27307/

**Citation:** Ayub, A., Datsiou, K. C., Pournaghshband, A., & Yousif, N. (2026). Proceedings of Abstracts: School of Physics, Engineering and Computer Science Conference 2026. *University of Hertfordshire*. https://doi.org/10.18745/PB.27307

**Contact**

Aaqib Ayub

Technical Team Leader (Research)

School of Physics, Engineering and Computer Science,

University of Hertfordshire, College Lane, Hatfield, Herts, AL10 9AB, UK

Email: a.ayub5@herts.ac.uk

https://doi.org/10.18745/pb.27307

# Preface

Dear Colleagues,

Welcome to the SPECS Conference 2026: Where Technical Excellence Enables Research Culture at the University of Hertfordshire.

This one-day event brings together researchers, technical professionals, students, industry partners, and research-enabling staff from across Physics, Astronomy, Mathematics, Engineering, and Computer Science. It reflects the strength of interdisciplinary research and the increasing recognition that technical expertise, specialist facilities and research infrastructure are fundamental for enabling high-quality research and innovation.

SPECS 2026 has attracted exceptional engagement, with more than 100 abstract submissions and over 300 attendees expected. This year's theme, Building Sustainable Research Environments Through Technical Expertise, highlights the vital role of technical professionals and research-enabling communities in shaping inclusive, collaborative, and sustainable research cultures.

The programme features keynote lectures, plenary sessions with oral presentations, flash talks, posters, and the Inside Spectra Labs showcase. Colleagues from across the School of Physics, Engineering and Computer Science will present our research, laboratories, advanced facilities, and technical capabilities that support discovery, innovation, and teaching. These are complemented by contributions from national technical initiatives, learned societies, and our industry partners and sponsors. We are particularly honoured to welcome Professor Dame Ottoline Leyser and Dr Neil Robinson as our keynote speakers, together with invited speakers and panellists from across academia, industry, and the wider research sector.

We would like to thank all speakers, presenters, exhibitors, sponsors, organisers, volunteers, and attendees for contributing to SPECS Conference 2026. We hope that you enjoy the day, find our research stimulating, and discussions engaging and leave with new ideas and collaborations.

With best wishes,

**Aaqib Ayub**

Conference Chair

**Corinna Datsiou**

Conference Chair

https://doi.org/10.18745/pb.27307


## Organising Committee



## Scientific Committee



## Reviewers Panel

https://doi.org/10.18745/pb.27307


## Sponsors

**University of Hertfordshire**

*Grateful acknowledgement is made to the University of Hertfordshire for its support through the Enhancing Research Culture Grant.*

**UK RAS STEPS**

*Sincere appreciation is extended to UK RAS STEPS for supporting research technical professionals through travel grants.*

**RSC-PGIC**

*Grateful acknowledgement is made to the RSC-PGIC for its sponsorship of prizes awarded at the event.*

**NTPMI**

Sincere thanks are extended to NTPMI for its support through travel and accommodation grants for technicians.

## Exhibitors

SHIMADZU Excellence in Science · EVIDENT · JEOL · RS

KEYENCE · Anton Paar · INSTRON

UniGreenScheme

https://doi.org/10.18745/pb.27307

# Awards

Many congratulations to all our delegates for their excellent contributions.

**Best Paper Award**

- Rob Yates - From stardust to us: Predicting planet habitability in the Universe from computational simulations of galaxy chemical evolution.

**Best Oral Presentation Award**

- Alessandro Angrilli Muglia - Modelling Extreme Nitrogen Emitting Galaxies from Cosmic Dawn to the Present.
- Davide Bernasconi - Reactivity of Calcined CDW Fines and Mine Tailings in Blended Cements through Integrated Experimental and Modelling Approaches.
- Milad Heidari-Koochi - Characterisation of the aerosol transport evolution in a bioaerosol chamber using low-cost sensors.

**Best Poster Award**

- Preethy Sethuraman - Effects of Slots on Frame on the Performance of Millimeter-Wave Dual-Polarized Antenna.
- Tanojit Paul - Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy.
- Quang Nam Phi - Robotis OP3 Joystick Locomotion: Mujoco Playground Towards Sim-To-Real.

**Best Flash Talk Award**

- Mengqing Huang - AI-Powered Learning Transformation for Inclusive and Secure Education.

**Early Career Excellence Award**

- Nibaldo Navarro-Castro - Post-fire analysis of aluminium alloys 6082-T6, 6005A-T6 and 7020-T6: Experimental tests and material modelling.

**Impact Award**

- Maria Quiroz – Machine-Learning-Based Damage Detection in a Small-Scale Wind Turbine Blade Using Data-Driven Features.
- Otis Lau - Experiencing University While Still at School.

**Research Technical Professional Excellence Award**

- Beatriz Ribeiro - Effects of Airflow Incidence on Propeller Thrust Across Varying Freestream Velocities.
- Eugen Guritanu - Data Driven Virtual Reality Tensile Testing Laboratory Development.

**Best UG/PG Student Contribution Award**

- Muhammad Irfan - TempFAS: Video-Level Face Anti-Spoofing Using CNN-LSTM Temporal Modelling

https://doi.org/10.18745/pb.27307

# Table of Contents

https://doi.org/10.18745/pb.27307

# Programme

### Opening Remarks (9:00 – 9:10)

Prof. Anthony Woodman, Vice Chancellor - Welcome

Prof. Pandelis Kourtessis – Introduction to SPECS Research

### Keynote 1 (9.10 – 9.55)

Dr Neil Robinson - The critical role of technical expertise in enabling research, innovation and impact.

### Plenary Sessions (10.00 – 11.20)

#### 1. Advanced Functional Materials and Computational Mechanics

Chair: Dr Pieter Desnerck

Judges: Dr Nusrat Rashid, Prof. Andreas Chrysanthou

Invited Speaker

- Dr Anand Sreeram - Developments in Biobased Road Construction Materials at the Nottingham Transportation Engineering Centre (NTEC).

Oral Talks

- Nibaldo Navarro-Castro - Post-fire analysis of aluminium alloys 6082-T6, 6005A-T6 and 7020-T6: Experimental tests and material modelling.
- Austin Migua Njagi - Experimental Investigation of Pcm-Enhanced Heat Sinks for Passive Electronics Cooling.
- Beatriz Ribeiro - Effects of Airflow Incidence on Propeller Thrust Across Varying Freestream Velocities.

#### 2. Intelligent Systems for Vision, Learning, and Data-Driven Decision Making

Chair: Dr Nada Yousif

Judges: Dr Mahmoud Eltaweel, Dr Mehroze Iqbal

Invited Speaker

- Prof. Nikhil Vasdev - How AI-led Robotic Surgical Research at UH is Transforming Urological Care

Oral Talks

- María E. Quiroz - Machine-Learning-Based Damage Detection in a Small-Scale Wind Turbine Blade Using Data-Driven Features
- Ojikutu Damilola - Multimodal Biometric Authentication
- Muhammad Irfan - TempFAS: Video-Level Face Anti-Spoofing Using CNN-LSTM Temporal Modelling
- Govardhan Reddy Bommana - Class Imbalance Correction in Deep Learning for Telecom Customer Churn Prediction: A Comparative Study of ANN and TabNet

https://doi.org/10.18745/pb.27307

### 3. Sensing the Invisible: Airborne Particles and Bio-Detection Technologies

Chair: Prof. Loic Couldron

Judges: Dr Etelka Chung, Dr Lanka Weerasiri

Invited Speaker

- Dr Nikolay Dimov - Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device

Oral Talks

- Muhammad Sohaib Ikram - Low-Flowrate Collection and Quantification of Airborne Human DNA in an Occupied Meeting Room
- Rob Lewis - Respirable Fibre Detection from Light Scattering Patterns
- Milad Heidari-Koochi - Characterisation of the aerosol transport evolution in a bioaerosol chamber using low-cost sensors
- Shamil Rafeeq - Performance evaluation of bioaerosol samplers on capturing fungal spores in a controlled environment

### 4. Modelling the Universe: Astrophysics, Simulations, and Observational Systems

Chair: Dr Sophie Koudmani

Judges: Dr Krupa Kuruvila, Dr Ragesh Thelakkadan Puthiyaveettil

Invited Speaker

- Dr Lily Whitler - New Eyes on the Sky: Insights into the Early Universe from the James Webb Space Telescope.
- Dr Will Coulton - Glimpses into the Universe’s Youth: A Search for New Information on the Primordial Universe.

Oral Talks

- Rob Yates - From stardust to us: Predicting planet habitability in the Universe from computational simulations of galaxy chemical evolution.
- Supriyo Ghosh - Enhancing Dynamic Range of an sCMOS Astronomical Camera through Row-Readout Sequencing.
- Alessandro Angrilli Muglia - Modelling Extreme Nitrogen Emitting Galaxies from Cosmic Dawn to the Present.

### 5. Interactive Intelligent Systems: Robotics, XR, and Adaptive Control

Chair: Prof. Salvatore Livatino

Judges: Dr Javed Khan, Anna Ojinnaka

Invited Speaker

- Dr Caroline Whyatt - Human-Robot Interaction and Adaptive Intelligent Systems.

https://doi.org/10.18745/pb.27307

Oral Talks

- Klinsmann Agyei - Attention-Based Soft Supervisory for Controlling Multiple Systems.
- Ali Fallahi - Manipulating Perceived Agency in Human-Robot Interaction: A Dialogue and Behaviour Design Framework for Pepper and Fetch Robots.
- Eugen Guritanu - Data Driven Virtual Reality Tensile Testing Laboratory Development.
- Bente Riegler - Bold Hearts: Developing new generations of humanoid robots.

### 6. Resilient Infrastructure: Materials, Modelling, and Decision Systems

Chair: Dr Asal Pournaghshband

Judges: Dr Antonis Kanellopoulos, Dr Mike Otieno

Invited Speaker

- Dr Antonis Kanellopoulos - Global Mineral Waste Streams and Their Upcycling Potential.

Oral Talks

- Bernasconi D - Reactivity of Calcined CDW Fines and Mine Tailings in Blended Cements through Integrated Experimental and Modelling Approaches.
- I. Onwochei - Non-Destructive Evaluation of Bond Integrity in Adhesively Bonded Joints in Composite Structures Using Acousto-Ultrasonics.
- Sajjad Karimi - Debonding-on-demand for structural silicone joints in façade applications.
- Farzad Piadeh - Optimising Green Building Design Using BREEAM and Artificial Intelligence Modelling (BREE-AIM).

### 7. Applied Science and Industrial Innovation

Chair: Denis Smith

Judges: Ricky Linforth, Anju Viswanath

Invited Speaker

- Mel Disher - The Less-Travelled Paths of Particle Characterisation.

Oral Talks

- Andy Bonney (RS Pro) - ESG & Sustainable Product Framework.
- Clement Eliott (Shimadzu) - Combining Two Analytical Techniques to Provide Complementary Molecular Information.
- Jane Bratherton (Evident Scientific) - Evident Industrial (Materials) Microscopy.

## KTP (Knowledge Transfer Partnership) Flash talks (11.40 – 12.00)

- Akhila Anil - Virtual Assessor for Occupational Health.
- Hafiz Hassan - AI-Powered Home Care Management System.
- Justin Morton - Advancing Fire Safety Through Riser Shaft Design.
- Wasim Beig - Advanced Asset Tracking for EMS.
- Pan Cao - Advanced mmWave 4D Radar Design for Complex Transport Monitoring.
- Mengqing Huang - AI-Powered Learning Transformation for Inclusive and Secure Education.
- Otis Lau - Experiencing University While Still at School. (Non KTP Project)

https://doi.org/10.18745/pb.27307

## Poster session (12.00 – 13.30)

Approximately 70 research posters were displayed across the Spectra floors, highlighting diverse research innovations and interdisciplinary contributions. The exhibition fostered meaningful academic discussions, networking opportunities, and knowledge exchange among participants.

## Keynote 2 (13.30 – 14.45)

Prof. Ottoline Leyser - Science: Changing the story

## Inside Spectra Labs (14.15 – 15.45)

Step inside Spectra Labs and explore a world of innovation featuring over 40 advanced laboratory facilities and research platforms. Spanning materials characterisation, advanced manufacturing, robotics, intelligent systems, simulation, aerospace, automotive, and immersive digital technologies, Spectra Labs showcases the breadth of cutting-edge research and engineering taking place across the University. Throughout the labs, application scientists, technical specialists, researchers, and industry experts will be on hand to demonstrate equipment, discuss research capabilities, and share insights into how these facilities are being used to address real-world engineering and scientific challenges.

## Panel Discussion (16.00 – 17.00)

### Building Sustainable Research Environment through Technical Excellence

Panel Members

- Prof. Daniel McCluskey (chair)
  Dean, School of Physics, Engineering and Computer Science, University of Hertfordshire
- Mrs Karen Withers
  Assistant Director HR, University of Hertfordshire
- Dr Jeremy Cohen
  Advanced Research Fellow, Imperial College London
- Prof. Tim Newton
  Dean of Research Culture, King's College London
- Dr Neil Robinson
  Head of Research Infrastructure, Engineering and Physical Sciences Research Council (EPSRC)
- Dr Han Wu
  Research Laboratories Manager, University College London
- Mr Chris Williamson
  Chair, Royal Society of Chemistry

## Closing Remarks and Prizes (17.00 – 17.30)

## Plenary Sessions

https://doi.org/10.18745/pb.27307

## 1. Advanced Functional Materials and Computational Mechanics

Chair: Dr Pieter Desnerck

Judges: Dr Nusrat Rashid, Prof. Andreas Chrysanthou

https://doi.org/10.18745/pb.27307

## Invited Speaker – Dr Anand Sreeram

Bio: Dr Anand Sreeram is an Assistant Professor in Transportation Engineering at the University of Nottingham and the Nottingham Transportation Engineering Centre (NTEC). Previously, he was a Research Fellow at the University of Cambridge (UK) and completed postdoctoral training at the Centre for Transportation Research, The University of Texas at Austin (USA). He received his PhD in Civil and Environmental Engineering from The Hong Kong Polytechnic University (Hong Kong) and holds a BEng in Chemical Engineering and an MSc in Environmental Engineering from The Hong Kong University of Science and Technology (HKUST). He has an internationally recognised research track record in pavement engineering has published over 70 peer-reviewed journal articles. In 2026, he was appointed an Associate Editor of the International Journal of Pavement Engineering, a leading journal in the field.

**Title: Developments in Biobased Road Construction Materials at the Nottingham Transportation Engineering Centre (NTEC).**

This talk presents recent advancements at the Nottingham Transportation Engineering Centre (NTEC) in developing a comprehensive analytical framework to assess biobased materials as sustainable alternatives in asphalt pavements. Asphalt, a composite of mineral aggregates and bitumen, remains the dominant material in road construction. However, its dependence on non-renewable resources raises concerns regarding cost, environmental impact, and long-term sustainability. To address these challenges, NTEC researchers are exploring renewable, bio-derived materials such as lignin and biochar that can partially or fully substitute bitumen in pavement applications. A key barrier to the widespread adoption of these biobased alternatives is the variability in their chemical composition, which depends heavily on feedstock source and processing methods. These variations can significantly influence the compatibility, mechanical performance, and durability of asphalt mixtures. To support the selection and optimisation of these materials, our work integrates techniques such as solubility science and other advanced material characterisation methods to evaluate the physicochemical compatibility between bitumen and biobased modifiers. In particular, the talk will highlight work on investigating chemical interactions between various additives and bitumen (and its subfractions), using Hansen Solubility Parameters (HSP). The insights gained from this approach will ultimately contribute to the design of more sustainable, high-performance asphalt and provide a scientific basis for the effective integration of biobased materials into road infrastructure.

https://doi.org/10.18745/pb.27307

# Post-fire analysis of aluminium alloys 6082-T6, 6005A-T6 and 7020-T6: Experimental tests and material modelling

**Nibaldo Navarro-Castro[1*], Gabriel Barros dos Santos[2], and Andreas Chrysanthou[3]**
*[1]PhD candidate, University of Hertfordshire, nibaldo.navarro-castro@herts.ac.uk*
*[2]Senior Lecturer, University of Hertfordshire, g.santos@herts.ac.uk*
*[3]Professor, University of Hertfordshire, a.chrysanthou@herts.ac.uk*
*Corresponding author: Nibaldo Navarro-Castro

Structural behaviour at the post-fire stage is crucial for making informed decisions, from evacuation strategies to designing for retrofitting or reuse. This paper describes the changes in material behaviour after exposure to simulated fire scenarios in aluminium alloys 7020-T6, 6005A-T6 and 6082-T6. The experimental programme considered different heating rates, exposure durations and cooling methods. Finally, materials model recommendations are provided based on post-fire tests.

Keywords: Aluminium alloys; Post-fire behaviour; Reduction factors; Microstructural analysis; Material modelling.

## Introduction and background

In recent years, structural behaviour at post-fire stages has been investigated to improve the assessment of structures after fire events and to support decisions on retrofitting or reuse. However, the behaviour of some aluminium alloys used in civil engineering construction remains unexplored. The aim of this study is to examine the post-fire mechanical properties of three aluminium alloys: 6082-T6, 6005A-T6 and 7020-T6.

Post-fire tests consist of heating a sample at a specific heating rate, keeping the temperature constant during the exposure time, and cooling using a process that usually involves either water quenching or natural cooling inside or outside the furnace. Mechanical properties are then assessed [1-2]. The most commonly used heating rate and exposure time for post-fire tests are 15 °C/min and 30 min [3-4], however, a review of full-scale structural fire tests showed a wider range of heating rates, and cooling rates [5-7]. Specifically, fire tests on aluminium structures exhibit heating rates ranging from 5-155 °C/min and cooling rates ranging from 20-31°C/min.

## Methodology

The experimental campaign of post-fire tests was divided into two phases: phase 1 focused on the effect of temperature exposure from 20 °C to 500 °C and two cooling methods (water quenching and natural cooling inside the furnace), and phase 2 examined the effects of heating rates (15, 80, 135 °C/min), exposure times (20, 60, 120 min), and cooling methods (water quenching, air inside the furnace, and air outside the furnace) at an exposure temperature of 400 °C. The specimen geometry complies with the ASTM E8/E8M standard [2], with a total length of 200mm and a gauge length of 50mm. Finally, using stress-strain behaviour data, two materials models i.e., the one and two-stage Ramberg-Osgood models [8-9] were assessed, along with predictive equations for the strain hardening exponents “n” and “m”.

## Results and discussion

The reduction factors for proof strength for AA6082-T6, AA6005A-T6 and AA7020-T6 cooled in water and in air inside the furnace are shown in Figure 1. AA6082-T6 and AA6005-T6 exhibit nearly constant proof strength between 20–200 °C, followed by a significant drop between 300–400 °C, reaching a minimum reduction factor at 400 °C. Strength recovery is then observed at 500 °C for the samples cooled in water, while samples cooled inside the furnace continue to show a decrease in the proof strength. For AA7020-T6, the proof strength remains nearly unchanged between 20–200 °C followed by a significant drop at 300 °C. Strength recovery is observed at 400 °C for samples cooled in water, followed by a further decrease in strength. Post-fire test data from other studies for similar aluminium alloys are also given in Figure 1. The heating rate and exposure time do not show a significant effect on the alloys studied in terms of strength.

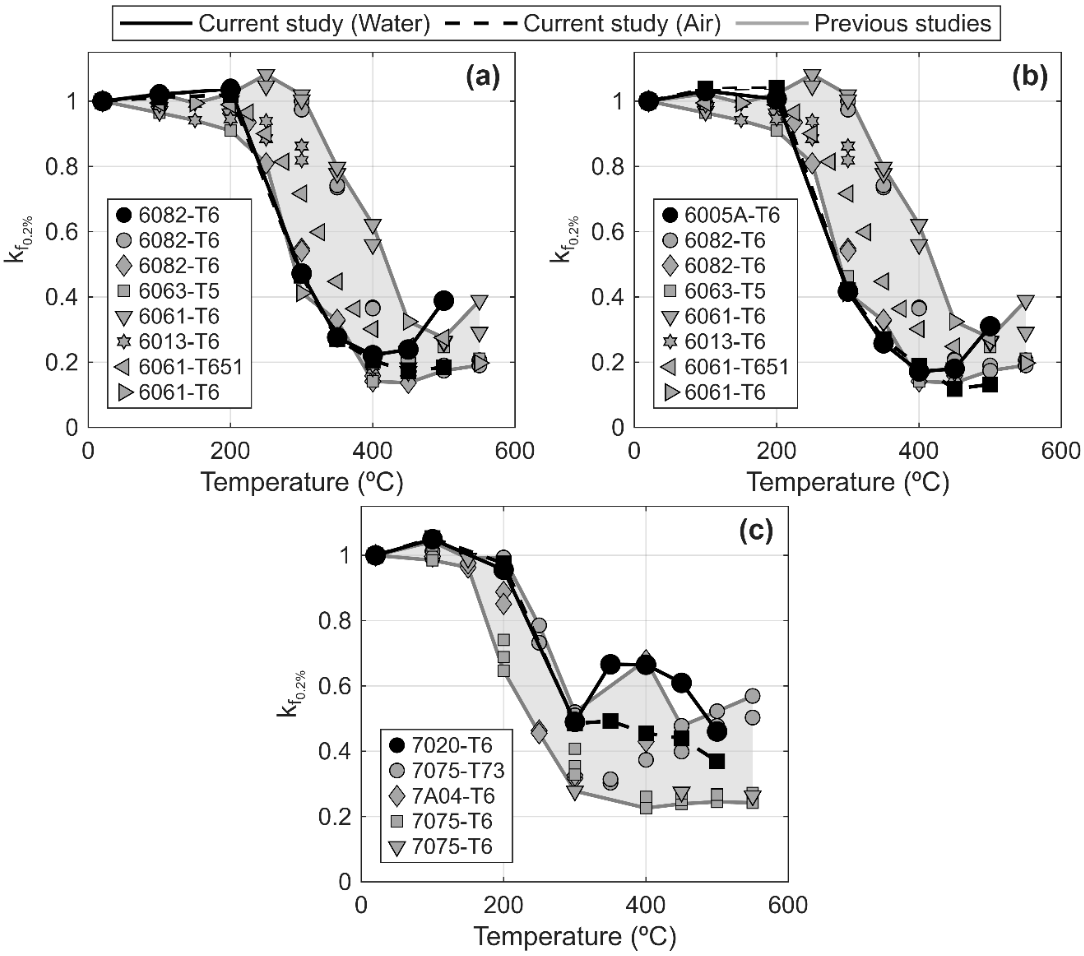



*Figure 2. Post-fire reduction factor of proof strength ($f_{0.2\%}$) considering two cooling methods: Water and air inside the furnace. (a) AA6082-T6. (b) AA6005A-T6. (c) AA7020-T6*

Using the stress–strain response of the three aluminium alloys under post-fire conditions, the one and two-stage Ramberg–Osgood models were assessed by fitting them to the data and minimising the Mean Square Error (MSE). Figure 2 shows two stress–strain responses along with the performance of each model. It can be observed that the two-stage Ramberg–Osgood model better represents the stress–strain behaviour of AA6005A-T6 after 400 °C exposure, which is also reflected in the MSE values. This improved performance is also observed for the rest of the cases studied. This improvement can be attributed to the formulation of the two-stage Ramberg-Osgood model which uses two hardening exponents: one for the first stage ($f<f_{0.2\%}$) and another for the second stage ($f>f_{0.2\%}$), enabling the model to better capture both the initial nonlinear response and the subsequent strain-hardening behaviour.

## Conclusions

Initial results show distinct post-fire behaviour for aluminium alloys, with the reduction factors for 6xxx (AA6082-T6 and AA6005A-T6) and 7xxx alloys (A7020-T6) differing due to their chemical composition and precipitate evolution. The effect of the cooling method is more pronounced for temperatures higher than 400°C for AA6082-T6 and AA6005A-T6, whereas for AA7020-T6 a more significant effect is observed above 300°C based on experimental results. From the model assessment it was concluded that the two-stage Ramberg-Osgood model better predicts the post-fire behaviour of aluminium alloys.

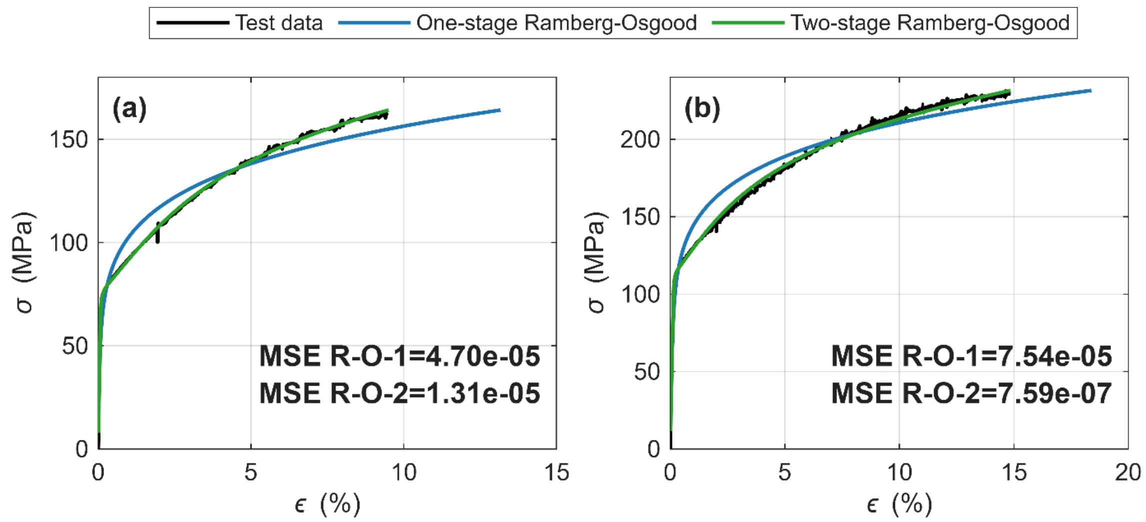



*Figure 3. Material model assessment. (a) AA6005A-T6, T=500°C, cooling method=water. (b) AA7020-T6, T=500°C, cooling method = water*

## References

[1] British Standards Institution, "EN ISO 6892-1:2019 Metallic materials - Tensile testing Part 1: Method of test at room temperature," *British Standard*, 2019.

[2] ASTM, "ASTM E8/E8M standard test methods for tension testing of metallic materials," *Annual Book of ASTM Standards*, no. C, 2021.

[3] Z. Chen, J. Lu, H. Liu, and X. Liao, "Experimental investigation on the post-fire mechanical properties of structural aluminum alloys 6061-T6 and 7075-T73," *Thin-Walled Structures*, vol. 106, 2016, doi: 10.1016/j.tws.2016.05.005.

[4] Y. Sun, K. Zhang, and G. Gong, "Material properties of structural aluminium alloys after exposure to fire," *Structures*, vol. 55, 2023, doi: 10.1016/j.istruc.2023.07.027.

[5] T. Lennon and D. Moore, "The natural fire safety concept—full-scale tests at Cardington," *Fire Saf. J.*, vol. 38, no. 7, pp. 623–643, Nov. 2003, doi: 10.1016/S0379-7112(03)00028-6.

[6] L. Yin, Z. Ni, F. Fan, P. Qiu, Q. Kan, and Y. Ou-yang, "Temperature field characteristics of cylindrical aluminum alloy reticulated roof system under localized fire," *Fire Saf. J.*, vol. 121, p. 103267, May 2021, doi: 10.1016/j.firesaf.2020.103267.

[7] S. Zhu, X. Guo, S. Jiang, S. Zong, and C. Chen, "Experimental Study on the Fire-Induced Collapse of Single-Layer Aluminum Alloy Reticulated Shells with Gusset Joints," *Journal of Structural Engineering*, vol. 146, no. 12, Dec. 2020, doi: 10.1061/(ASCE)ST.1943-541X.0002819.

[8] W. Ramberg and W. R. Osgood, "Description of stress-strain curves by three parameters," *National Advisory Committee For Aeronautics*, 1943.

[9] E. Mirambell and E. Real, "On the calculation of deflections in structural stainless steel beams: An experimental and numerical investigation," *J. Constr. Steel Res.*, vol. 54, no. 1, 2000, doi: 10.1016/S0143-974X(99)00051-6.

# Experimental Investigation of Pcm-Enhanced Heat Sinks for Passive Electronics Cooling

**Austin Migua Njagi**[1*]

[1] *BEng in Mechanical Engineering Student of University of Hertfordshire*

* Corresponding author: austo342@gmail.com

**This study investigates a Phase Change Material (PCM)-integrated heat sink for passive electronic cooling using experimental testing and CFD simulation. Three PCMs, paraffin wax, sodium sulphate decahydrate, and sodium thiosulphate pentahydrate, were evaluated, with sodium sulphate decahydrate achieving the best temperature control below 50 °C. A validated ANSYS Fluent model using the enthalpy–porosity method was employed to analyse four fin geometries. Circular pin fins showed the highest PCM utilisation, and further optimisation using hollow fins improved thermal performance. Results confirm that PCM selection and fin design significantly enhance passive cooling efficiency in electronic systems.**

Keywords: Phase Change Material (PCM); passive cooling; heat sink; enthalpy–porosity method; CFD; ANSYS Fluent; thermal management; electronic cooling; fin optimisation; liquid fraction.

## Introduction

The increasing power density of modern electronic systems, combined with reduced component size, has led to significant heat generation in compact devices. This is particularly evident in high-performance computing, artificial intelligence (AI) systems, and data centres, where effective thermal management is essential to ensure reliability and prevent performance degradation [1].

Conventional cooling methods primarily rely on air-cooled heat sinks due to their simplicity and cost-effectiveness. These systems enhance heat dissipation through extended surfaces; however, their performance is limited by the low thermal conductivity and heat capacity of air, making them less effective under transient and high heat flux conditions [2].

In applications such as solid-state drives (SSDs), thermal management remains a critical challenge. For instance, testing of the Samsung 990 EVO SSD has shown temperatures approaching 50 °C under load, which may lead to thermal throttling if not properly controlled [3]. This highlights the need for improved passive cooling strategies.

Phase change materials (PCMs) offer a promising solution by absorbing heat through latent heat during phase transition, enabling temperature regulation without significant temperature rise. This makes them particularly effective for transient thermal management applications [4]. In addition, heat sink geometry plays a key role in thermal performance, as fin design influences heat distribution and PCM melting behaviour.

This study investigates a PCM-integrated heat sink using a combined experimental and numerical approach. The work focuses on evaluating different PCM materials, validating a CFD model, analysing fin geometries, and optimising heat sink design for improved thermal performance.

## Experimental and Simulation Setup

The experimental investigation was conducted using a custom-fabricated aluminium heat sink integrated with a Phase Change Material (PCM) chamber. A cartridge heater was embedded at the base of the heat sink to simulate a constant heat input representative of electronic components. The overall experimental setup and thermocouple arrangement (Tb1–Tb4) within the PCM domain are shown in Figure 1. Three PCMs—paraffin wax, sodium sulphate decahydrate, and sodium thiosulphate pentahydrate—were experimentally evaluated under identical heating and cooling conditions. Temperature–time data were recorded to analyse phase change behaviour, including melting duration, peak temperature, and cooling characteristics.

https://doi.org/10.18745/pb.27307

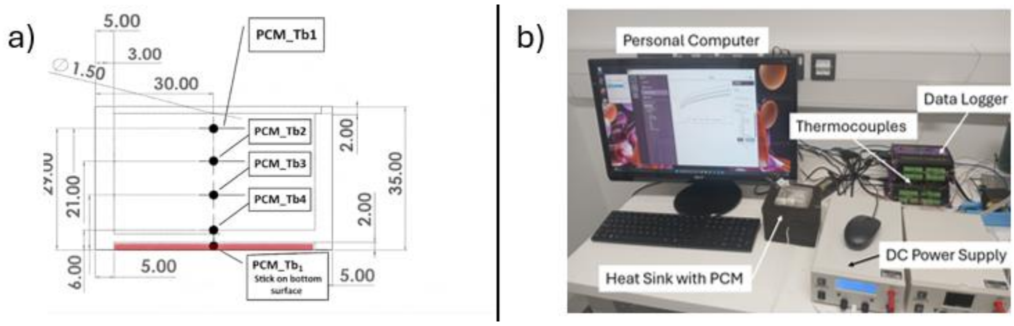



*Figure 4: This figure has subfigures of a) Thermocouple positions and b) The Experimental Rig*

A transient CFD model was developed in ANSYS Fluent to simulate heat transfer and PCM melting behaviour within the heat sink. The enthalpy–porosity method was employed to model solid–liquid phase change, incorporating latent heat effects and treating the mushy region as a porous medium [5], [6], [7]. Temperature-dependent PCM properties were implemented, and the governing conservation equations for mass, momentum, and energy were solved. A mesh convergence study was performed to ensure mesh-independent results, with an optimal element size of 0.0007m selected for the simulations. The numerical model was validated against experimental temperature data before being used to investigate circular, rectangular, V-shaped, and tree-shaped fin geometries, followed by optimisation studies using modified circular fin configurations.

## Results and discussion

The experimental temperature–time responses for the three PCM materials are shown in Figure 2. The empty heat sink exhibited a rapid temperature rise and exceeded the required operating limit, confirming the limitations of air cooling. In contrast, all PCM-integrated cases demonstrated improved performance through reduced peak temperature and delayed temperature rise due to latent heat absorption.

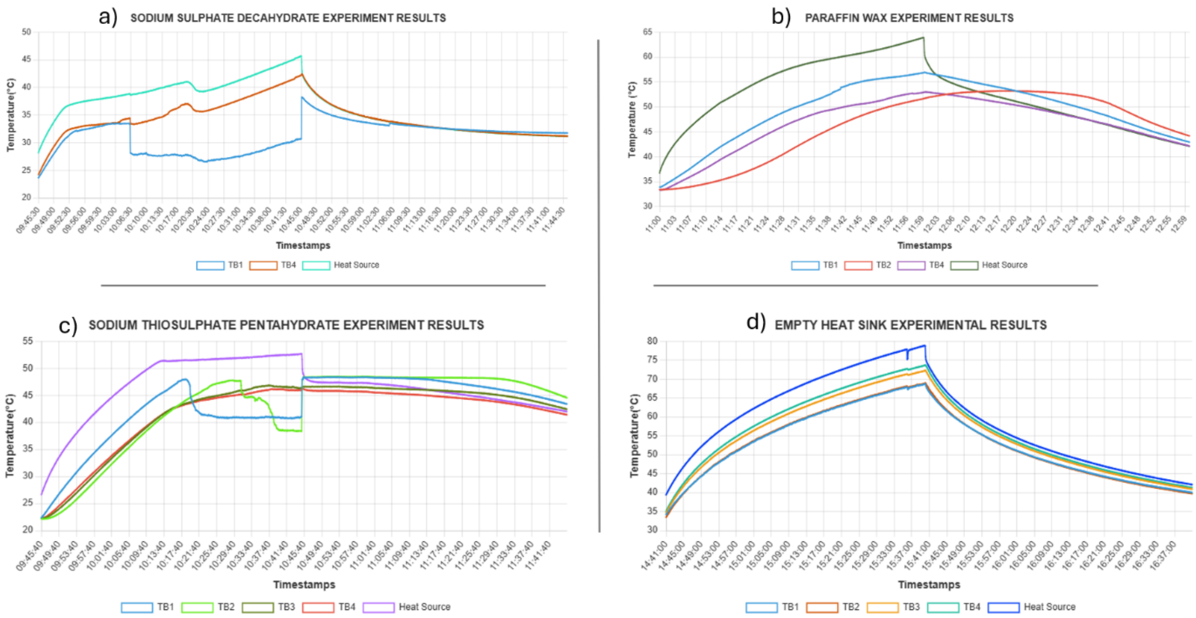



*Figure 5: Temperature-curve graph of a) Sodium Sulphate Decahydrate, b) Paraffin wax, c) Sodium Thiosulphate Pentahydrate and d) Empty Heat sink*

Among the materials tested, sodium sulphate decahydrate achieved the lowest peak temperature and maintained operation below 50 °C, making it the most effective for temperature control. Paraffin wax showed stable and predictable behaviour but resulted in higher peak temperatures. Sodium thiosulphate pentahydrate exhibited the longest phase change duration, indicating superior heat storage capability, although its peak temperature slightly exceeded the required limit. These results highlight the trade-off between temperature regulation and heat storage capacity.

The numerical results for different fin geometries are presented in Figure 3. The circular pin-fin geometry achieved the highest liquid fraction, indicating the most effective PCM utilisation. The rectangular and V-shaped fins showed similar performance, while the tree-shaped geometry exhibited less uniform melting and lower liquid fraction despite providing extended heat conduction paths.

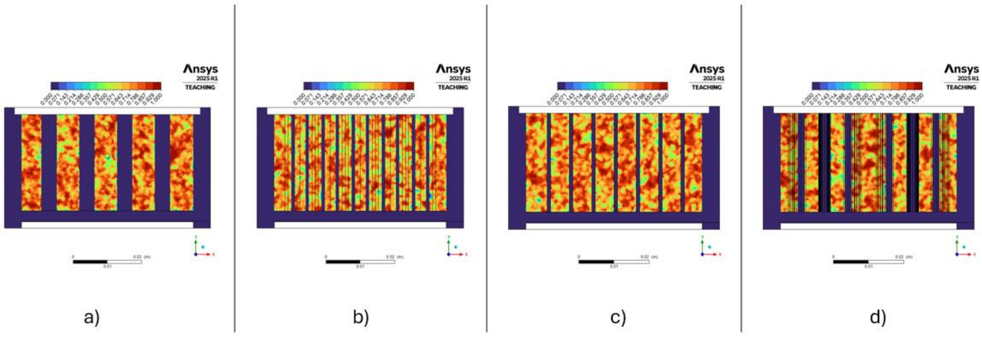



*Figure 6: Figure of the Liquid fraction contour namely, a) Circle fins, b) V-shape fins, c) Rectangular fins, d) Tree shape fins*

Further optimisation of the circular pin-fin design showed that the hollow fin configuration provided the best overall performance, achieving both higher liquid fraction and lower PCM temperature compared to the original and staggered designs. This demonstrates that both fin geometry and PCM distribution significantly influence thermal performance.

## Conclusion

This study demonstrated that integrating Phase Change Materials (PCMs) into heat sinks significantly improves passive cooling performance by reducing peak temperature and delaying temperature rise. Among the materials evaluated, sodium sulphate decahydrate provided the best temperature control, while sodium thiosulphate pentahydrate showed superior heat storage capability. Numerical analysis identified circular pin fins as the most effective geometry, with further optimization using hollow fins enhancing PCM utilization and overall thermal performance. The results confirm that both PCM selection and heat sink design are critical for improving thermal management in electronic applications.

## References

[1] F. P. Incropera, D. P. DeWitt, T. L. Bergman, and A. S. Lavine, *Fundamentals of Heat and Mass Transfer*, 7th ed. Hoboken, NJ, USA: John Wiley & Sons, 2017.
[2] V. Bianco, M. De Rosa, and K. Vafai, “Phase-change materials for thermal management of electronic devices,” *Applied Thermal Engineering*, vol. 214, p. 118839, 2022.
[3] J. Harris, “Samsung 990 EVO Review & Benchmarks,” GeekaWhat, 2025. [Online]. Available: GeekaWhat Samsung 990 EVO Review
[4] A. Sharma, V. V. Tyagi, C. R. Chen, and D. Buddhi, “Review on thermal energy storage with phase change materials and applications,” *Renewable and Sustainable Energy Reviews*, vol. 13, no. 2, pp. 318–345, 2009.
[5] A. D. Brent, V. R. Voller, and K. J. Reid, “Enthalpy-porosity technique for modeling convection-diffusion phase change,” *Numerical Heat Transfer*, vol. 13, no. 3, pp. 297–318, 1988.
[6] V. R. Voller and C. Prakash, “A fixed grid numerical modelling methodology for convection–diffusion mushy region phase-change problems,” *International Journal of Heat and Mass Transfer*, vol. 30, no. 8, pp. 1709–1719, 1987.
[7] ANSYS Inc., *ANSYS Fluent Theory Guide, Release 2023 R1*. Canonsburg, PA, USA: ANSYS Inc., 2023.

https://doi.org/10.18745/pb.27307

# Effects of Airflow Incidence on Propeller Thrust Across Varying Freestream Velocities

**Abhinav Balaji[1], Beatriz Ribeiro[2*]**
[1]*Independent Researcher*
[2]*Technical Services Group, SPECS, University of Hertfordshire, Hatfield, United Kingdom*
* Corresponding author: b.da-mata-ribeiro@herts.ac.uk

**This study experimentally investigates how airflow incidence affects propeller thrust across various freestream velocities. Using a closed-section wind tunnel, a centrally mounted propeller–motor system measured thrust via a sting load cell. Data spanned yaw angles from 0° to 45° at velocities of 0, 5, 10, 15, and 20 m/s, with results normalised against static baselines for comparison.**

**The results show a reduction in thrust as yaw angle increases, with a clearer trend observed at 10 and 15 m/s. At 20m/s, deviations suggest additional nonlinear aerodynamic effects. The findings highlight the sensitivity of propeller performance to flow incidence, providing a basis for further investigation.**

Keywords: Aerodynamics; Asymmetric Blade Loading; Propeller Thrust; Flow Incidence; Wind Tunnel Testing;

## Introduction

Propellers are widely used across a range of engineering applications, from small UAVs to larger propulsion systems. Propeller behaviour and thrust generation is typically characterised assuming alignment between the incoming air and the propeller axis. Within these conditions, their performance is generally well established within aerodynamic theory [1][2]. However, in practical scenarios, variations in flight conditions, crosswinds, and manoeuvres introduce significant misalignment between the incoming flow and the propeller axis. This results in an off-axis inflow condition, where the velocity distribution across the propeller disk becomes non-uniform, potentially affecting thrust generation and efficiency [3].

Previous studies have shown that off-axis inflow can lead to asymmetric blade loading and variations in induced velocity, which may result in a reduction in thrust and overall performance [4][5]. In addition, similar behaviour has been observed in related rotating systems such as wind turbines operating under yawed conditions, where performance degradation and flow asymmetry are present [6]. However, the extent of these effects depends on several factors, including freestream velocity, propeller geometry, and operating conditions, making it difficult to generalise behaviour across different cases [7].

The aim of this study is to experimentally investigate the effect of airflow incidence on propeller thrust across a range of freestream velocities. A controlled wind tunnel setup is used to systematically vary the yaw angle and measure the corresponding thrust. By normalising the results against a static baseline, comparisons can be made across different conditions to identify underlying trends. This work focuses on establishing these trends through experimental observation and providing a basis for further investigation into more complex flow conditions, such as multi-axes incidence (pitch and yaw).

## Experimental Methodology

The experimental work was carried out in an AEROLAB closed return wind tunnel. The tunnel is capable of freestream velocities up to 62 m/s and provides measurements of airspeed, temperature, humidity and air density. A small-scale propeller–motor assembly was mounted centrally to the test section on a sting balance. In this study, the thrust is measured from the axial force generated by the propeller on this load cell balance.

https://doi.org/10.18745/pb.27307

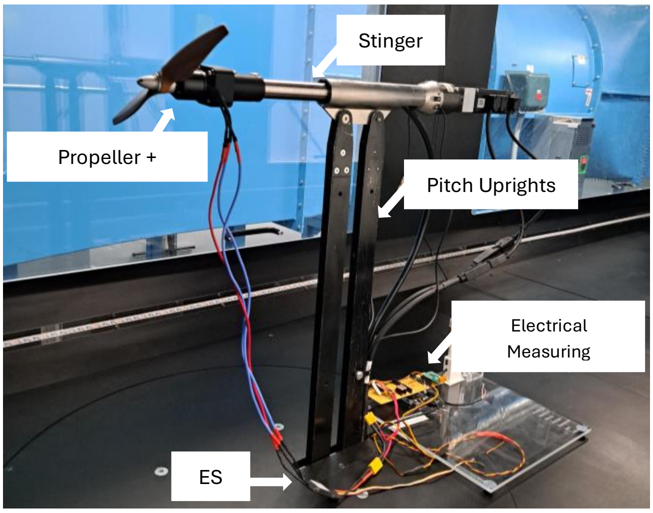



*Fig. 1. Experimental Setup.*

As seen in Figure 1, the electrical measurement system was assembled independently from the wind tunnel instrumentation. Motor current was measured using a TLI4970 current sensor, while voltage was measured using a custom voltage divider circuit. The motor control signal was also recorded and converted to a normalised 0–100% PWM scale based on the experimentally determined minimum and maximum pulse-width limits used during testing. Thrust data (tunnel DAQ) and electrical parameters (measured using an Arduino-based system) were synchronised via timestamp-based alignment.

Tests were carried out at freestream velocities of 5, 10, 15 and 20 m/s. The propeller was tested over a yaw range from 0° to 45° in 1° increments, with 0 m/s included as a static baseline case. The propeller used in this study was a 12 × 6 configuration driven by an electric motor.

## Results and Discussion

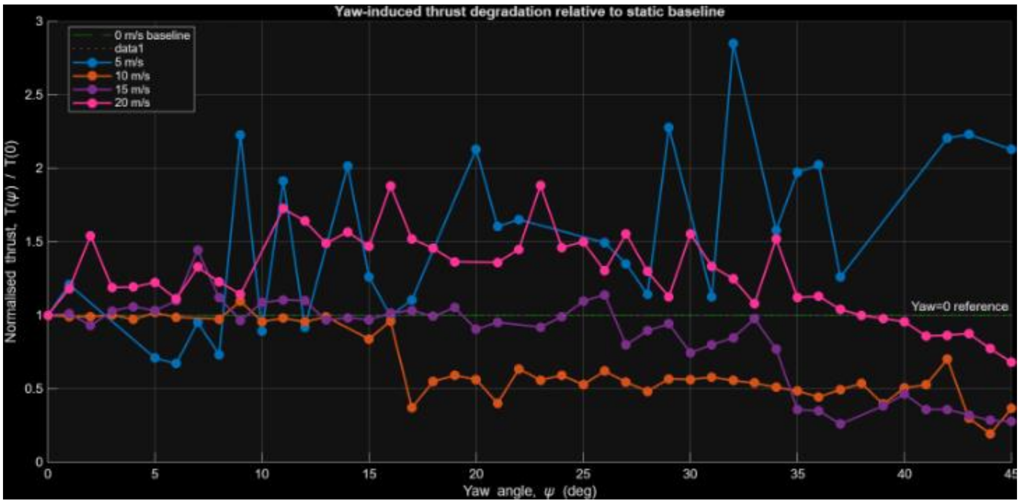



*Fig. 2. Normalised Thrust $T(\psi)/T(0)$ plotted over Yaw Angle (deg) for 5 (blue), 10 (orange), 15 (purple) and 20 (pink) m/s.*

The graphs obtained for yaw-induced thrust degradation show a clear trend of reduction in thrust as the yaw angle increases. However, this phenomenon is velocity dependent. In Figure 2, a value of 1 represents the results in terms of normalised thrust relative to the zero-yaw condition.

As seen in Figure 2, at 10 m/s, a relatively smooth decreasing trend can be observed. The thrust remains close to the baseline at low yaw angles, before gradually decreasing beyond approximately 15 degrees. The linear trendline fits the data reasonably well ($R^2 \approx 0.73$), indicating that the relationship between yaw angle and thrust degradation can be approximated as linear in this case. This reduction in thrust can be attributed to the misalignment between the propeller axis and the incoming flow, where the effective axial component of velocity decreases as yaw increases.

A similar trend is observed at 15 m/s, the reduction in thrust is slightly more pronounced, suggesting a stronger sensitivity to yaw. The overall behaviour follows a decreasing trend, with a moderate linear fit ($R^2 \approx 0.66$). This further supports the idea that yaw effects become more significant with increasing freestream velocity.

For 20 m/s, the behaviour becomes less consistent. The data is close to the baseline only at low yaw angles followed by a high deviation. While there is still an overall reduction in thrust at higher yaw angles, the low $R^2$ value ($\approx 0.21$), indicates that a simple linear model is not sufficient to describe the behaviour at this speed. This suggests that additional aerodynamic effects may be present at higher velocities, such as non-uniform inflow across the propeller disk or asymmetric blade loading, although further testing would be required to confirm this.

The 5 m/s case shows significant variability, with large fluctuations in the measured thrust values. The trendline in this case shows a slight increase in thrust with yaw angle. At low-thrust producing speeds, any noise in the measurement system becomes more significant and is heavily influenced by setup limitations, making it difficult to obtain reliable data.

Notably, the experimental setup used in this study, while sufficient for initial data collection, has several limitations. Furthermore, multiple measurements can be taken for each test condition and averaged to reduce the effect of noise in the data. Despite this, a general trend of yaw-induced thrust degradation was captured, particularly at moderate freestream velocities, which provides a strong basis for further investigation.

## Conclusion

This study investigated the effect of airflow incidence on propeller thrust across a range of freestream velocities using a controlled wind tunnel setup. The results show a general reduction in thrust with increasing yaw angle at moderate velocities (10–15 m/s). At lower and higher velocities, the results were varied suggesting more complex aerodynamic phenomena. The findings indicate that propeller performance is sensitive to off-axis flow conditions, and that this sensitivity varies with operating velocity. Future work will involve refining the experimental setup to reduce noise and improve measurement accuracy, as well as expanding the test matrix to include multi-axes variations and diverse propeller geometries for a more comprehensive aerodynamic characterisation.

## References

[1] J. D. Anderson Jr., *Fundamentals of Aerodynamics*, 5th ed., New York: McGraw-Hill, 2011.
[2] J. G. Leishman, *Principles of Helicopter Aerodynamics*, 2nd ed., Cambridge: Cambridge University Press, 2006.
[3] V. B. Nguyen, D. Rozehnal, J. Hnidka, and V. U. Pham, "Influence of yaw on propeller aerodynamic characteristics," *MATEC Web of Conferences*, 2018.
[4] Y. Yang, Y. Liu, Y. Li et al.,"Aerodynamic and aeroacoustic characteristics of a multicopter propeller during forward flight," *2018 Joint Propulsion Conference*, Cincinnati, Ohio, USA, 9-11 July, 2018.
[5] Q. R. Wald, "A Summary of Propeller Design Procedures and Data," NASA, Washington, D.C., NASA-RP-1309, May 1993.
[6] Y. Tu, K. Zhang, Z. Han, D. Zhou, and O. Bilgen, "Aerodynamic characterization of two tandem wind turbines under yaw misalignment control using actuator line model," *Ocean Engineering*, Vo. 281, 114992, 2023.
[7] S. D'Angelo, F. Berardi, and E. Minisci, "Aerodynamic performance of propellers with parametric considerations on optimal design," *Politecnico di Torino*, 2014.

https://doi.org/10.18745/pb.27307

## 2. Intelligent Systems for Vision, Learning, and Data-Driven Decision Making

Chair: Dr Nada Yousif

Judges: Dr. Mahmoud Eltaweel, Dr Mehroze Iqbal

https://doi.org/10.18745/pb.27307

## Invited Speaker – Prof. Nikhil Vasdev

Bio: Professor Nikhil Vasdev is a leading figure in Robotic Urological Surgery and Oncology, blending clinical excellence, innovation and academic leadership. His career spans advanced surgical techniques, global mentorship, research leadership and influential academic roles. He is a Full Professor and Chair of Robotic Surgery at the Hertfordshire Medical School, University of Hertfordshire. In May 2020, Professor Vasdev was awarded an "Doctor of Science (DSc)" by the awards committee at the University of Hertfordshire in recognition for Professor Vasdev's contribution in Robotics in Surgery. He is a member of the medical advisory committee for the soon to open Hertfordshire Medical School. Prof Vasdev is a Consultant Urological and Robotic Surgeon and Associate Medical Director for Cancer (Cancer Lead) at the Lister Hospital, East and North Herts Teaching NHS Trust, UK. Professor Vasdev also holds an international professorship at the Apollo Hospitals Education and Research Foundation, India. He holds certification in AI from Harvard Medical School and has completed a National Clinical Leadership programme at Prostate Cancer UK.

**Title: How AI Led Robotic Surgical Research at UH is transforming Urological Cancer Care**

Artificial intelligence (AI) is increasingly shaping the future of robotic surgery and urological care by enabling more precise, data-driven, and personalized approaches to diagnosis, treatment planning, and surgical intervention. Advances in machine learning, medical imaging, biomarker analysis, and predictive modeling are improving the ability to identify disease patterns, assess patient risk, guide biopsies, and optimize surgical strategies. By integrating large volumes of clinical, imaging, and biological data, AI can support more accurate decision-making, enhance surgical precision, and contribute to better patient outcomes across a range of urological conditions, including prostate and bladder cancer.

The growing use of AI also extends beyond the operating room, with applications in postoperative monitoring, outcome prediction, and remote patient care through wearable devices and patient-generated health data. These technologies offer opportunities to detect complications earlier, monitor recovery more effectively, and tailor follow-up care to individual patient needs. Realizing these benefits, however, depends on the availability of high-quality data, robust model development, rigorous clinical validation, and appropriate regulatory oversight. Continued collaboration between clinicians, researchers, data scientists, and industry partners will be essential to ensure that AI systems are safe, reliable, and capable of delivering meaningful improvements in healthcare while complementing, rather than replacing, clinical expertise.

https://doi.org/10.18745/pb.27307

# Machine-Learning-Based Damage Detection in a Small-Scale Wind Turbine Blade Using Data-Driven Features

**María E. Quiroz*, Miguel A. Bravo-Haro**
*City St George's, University of London, London, United Kingdom.*
*Corresponding author: maria.quiroz-roa@city.ac.uk

**Data-driven structural health monitoring (SHM) enables damage detection without requiring a pristine baseline, making it particularly suitable for ageing structures where labelled data are scarce. This study evaluates the robustness of several feature-extraction strategies under both supervised and unsupervised learning, using experimental vibration data from a small-scale wind-turbine blade. By analysing responses across multiple damage severities and temperature conditions, we assess the generality of the extracted features and examine how different learning paradigms influence detectability and localisation. The results highlight key trade-offs between robustness and sensitivity, offering practical guidance for feature selection in data-driven SHM.**

Keywords: Structural health monitoring; vibration-based analysis; data-driven methods; machine learning; feature extraction

## Introduction

Vibration-based structural health monitoring (SHM) faces a long-standing challenge: the discriminative power of commonly used features varies widely with structural type, damage mechanism and severity, and environmental or operational conditions [1]. Although numerous feature-extraction methods have been proposed—from simple statistical descriptors to autoregressive models [2-3], their relative performance under consistent evaluation settings remains poorly understood. Existing comparative studies offer valuable insights but typically examine individual features, focus on a single structure or learning paradigm, or neglect realistic environmental variability [4]. As a result, the current understanding of feature-set performance across different systems remains limited, and it is still unclear which descriptors are consistently reliable for detecting and localising damage.

This study addresses that gap by benchmarking a broad suite of time-, frequency-, and time–frequency-domain features in a laboratory-tested composite wind-turbine blade subjected to temperature variations. Using both supervised classifiers and unsupervised novelty-detection algorithms, applied to accelerometer and strain-gauge data, the analysis evaluates feature separability, robustness, and sensitivity to structural transitions. By focusing on output-only methods that require minimal prior knowledge, the study aims to provide practical guidance for feature selection, highlighting trade-offs between detectability and localisation and supporting the design of resilient SHM strategies.

## Methodology

This study evaluates several output-only feature-extraction approaches for data-driven SHM under both supervised and unsupervised learning. Seven feature families are considered guided by their widespread use in the SHM literature, their computational efficiency, and their demonstrated relevance in previous studies: power spectral density (PSD), short-time Fourier transform (STFT), discrete wavelet transform (DWT), continuous wavelet transform (CWT), temporal statistics, and univariate and multivariate autoregressive models (AR and MAR). A feature-selection step is applied to remove redundant or highly correlated descriptors, ensuring a compact and informative representation.

The unsupervised framework applies the methodology from [5], based on Principal Component Analysis-based normalisation to suppress components dominated by environmental and operational variability, followed by Mahalanobis-distance fusion and automatic K-means clustering. Performance is quantified using false-positive and false-negative ratios under a consistent 75/25 train–test split.

For supervised analysis, support vector machines (SVM), K-nearest neighbours (KNN), and Random Forest classifiers are trained using min–max normalisation, with hyperparameters selected via grid search and fivefold

https://doi.org/10.18745/pb.27307

cross-validation. Accuracy, precision, recall, and F1-score are used to assess performance, with balanced class weights applied where appropriate.

The experimental case study uses a 1.75 m composite wind-turbine blade tested in a climate-controlled chamber across −15 °C to +40 °C [6]. Eight accelerometers and 21 strain gauges (see Fig. 1) recorded the response under white-noise excitation as controlled cracks of increasing length were introduced at three locations. Data were collected for one healthy and nine damaged states (see Table 1), providing a rich dataset for evaluating feature-extraction strategies and damage-detection performance under varying environmental and structural conditions.

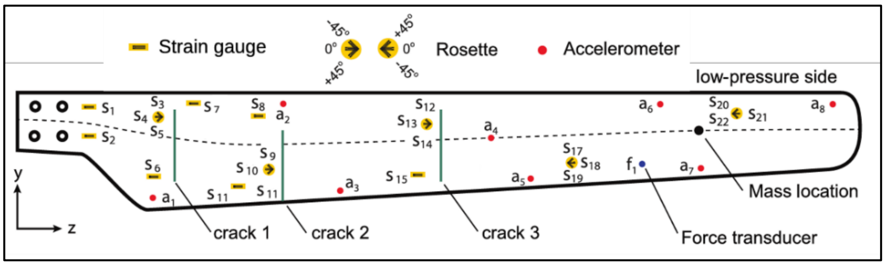



*Fig. 1. Spatial distribution of the sensors installed along the blade structure (modified from [6]).*

*Table 1. Damage configurations considered in the blade structure.*

| Group | Description | States |
|---|---|---|
| Reference | Undamaged blade | R |
| Low damage | 5 cm cracks in 1–3 locations | D–F |
| Moderate damage | 10 cm cracks in 1-3 locations | G–I |
| High damage | 15 cm cracks in 1-3 locations | J–L |

## Results and discussion

Supervised learning achieved consistently high accuracy across all feature families, confirming that labelled data allow most descriptors to separate healthy and damaged states effectively. In contrast, the unsupervised online framework successfully detected structural changes without prior labels, but revealed clearer differences in robustness across feature sets as the baseline updated with each new state. Figure 2 presents the heatmaps of the dissimilarity metrics for the various feature-extraction methods using strain data. Dissimilarity is consistently higher in sensors located near the damage site for all transitions, where a “damage transition” refers to the comparison between two consecutive structural conditions.
Strain-gauge features exhibited strong local sensitivity, reliably highlighting the sensors closest to the emerging crack through pronounced Mahalanobis-distance shifts. Autoregressive models, by contrast, provided more stable global detection but occasionally misidentified the damage location due to their broader spatial footprint. Overall, the blade results demonstrate that global features such as AR/MAR are highly effective for detecting the presence of damage, while locally sensitive features—particularly those derived from strain gauges—offer superior spatial resolution. This underscores the complementary value of combining global and local descriptors in practical SHM deployments, where both robustness and localisation accuracy are essential.

## Conclusion

This study benchmarked several vibration-based feature-extraction methods for data-driven damage identification in a temperature-varying small-scale wind-turbine blade. Features derived from accelerometers and strain gauges were evaluated under both supervised and unsupervised learning. All feature families performed strongly in supervised classification—exceeding 94% accuracy—with autoregressive and multivariate autoregressive models showing the most consistent robustness across learning paradigms. Temporal features produced the largest errors when using accelerometer data, while Mahalanobis-distance analysis demonstrated that most feature sets correctly localised damage using strain gauges, except for two transitions where AR-based features misidentified the location. Overall, the findings highlight the strengths and limitations of different feature families and point toward future opportunities in hybrid feature fusion to support efficient and reliable damage-detection methodologies.

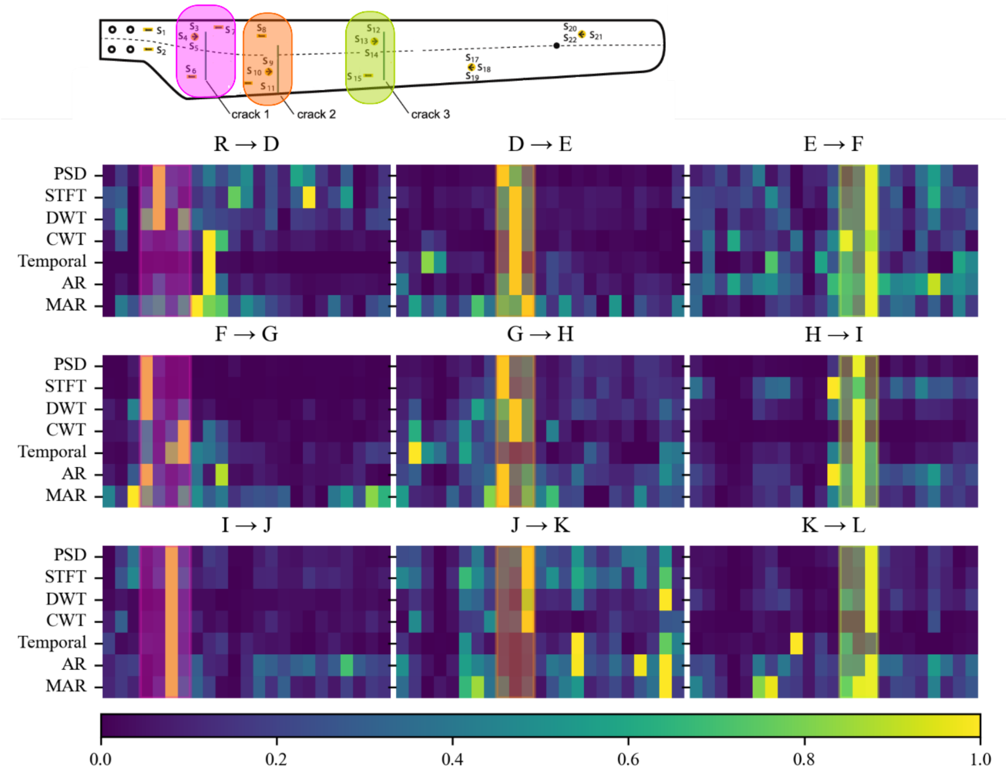



*Fig. 2. Normalised IMI heatmaps for the unsupervised analysis of the blade dataset. The sensor regions closest to the emerging crack in each damage transition are highlighted.*

## References

[1] Farrar, C. R., & Worden, K. (2012). Structural health monitoring: a machine learning perspective. John Wiley & Sons.

[2] Ahmadi, H. R., Mahdavi, N., & Bayat, M. (2021). A novel damage identification method based on short time Fourier transform and a new efficient index. Structures, 33, 3605–3614. https://doi.org/10.1016/j.istruc.2021.06.081

[3] Huang, Z., & Liu, J. (2025). Fast adaptive time-varying window length STFT for compound track short-wave defects identification. Journal of Vibration and Control, 31(5–6), 651–663. https://doi.org/10.1177/10775463241237088

[4] Buckley, T., Ghosh, B., & Pakrashi, V. (2023). A feature extraction & selection benchmark for structural health monitoring. Structural Health Monitoring, 22(3), 2082–2127. https://doi.org/10.1177/14759217221111141

[5] Quiroz, M. E., Kuncham, E., Ponnusami, S. A., & Bravo-Haro, M. A. (2026). Data-Driven Damage Detection on a Temperature-Varying Turbine Blade Using Multiple Sensor Modalities. Structural Control and Health Monitoring. 2026(1), 7943746. https://doi.org/10.1155/stc/7943746

[6] Ou, Y., Tatsis, K. E., Dertimanis, V. K., Spiridonakos, M. D., & Chatzi, E. N. (2021). Vibration-based monitoring of a small-scale wind turbine blade under varying climate conditions. Part I: An experimental benchmark. Structural Control and Health Monitoring, 28(6). https://doi.org/10.1002/stc.2660

https://doi.org/10.18745/pb.27307

# Multimodal Biometric Authentication

**Ojikutu Damilola, Li Meng*, Na Helian**
*Corresponding author: do25aan@herts.ac.uk

**This work presents a lightweight multimodal biometric system combining face and voice recognition for edge deployment. MobileFaceNet and a distilled Emphasized Channel Attention, Propagation and Aggregation (ECAPA) model are used for efficient feature extraction, while the lightweight BlazeFace enables real-time face detection. Three fusion strategies are evaluated, including attention-based approaches with and without temporal modeling. Our experimental results show that multimodal fusion significantly improves performance over unimodal systems when lightweight models are utilized, with attention-based fusion without Long Short-Term Memory (LSTM) achieving the best accuracy. The study demonstrates that lightweight architectures, combined with attention mechanisms, can maintain strong performance under computational constraints, though robustness to degraded inputs remains a challenge.**

Keywords: Multimodal biometrics; lightweight models; attention-based fusion; knowledge distillation; edge computing

## Introduction

Biometric recognition systems traditionally rely on a single modality such as face or voice, making them vulnerable to environmental challenges including noise, occlusion, and poor lighting conditions. These limitations reduce reliability in real-world deployments, especially in uncontrolled environments.

Multimodal biometric systems address this issue by combining complementary modalities. Face and voice are particularly suitable due to their distinct physical origins and ease of simultaneous acquisition. While facial features capture geometric and textural identity, voice encodes vocal tract characteristics, allowing one modality to compensate for the failure of the other.

Despite their advantages, multimodal systems are often computationally expensive and unsuitable for deployment on edge devices. Recent advances in lightweight deep learning architectures [1, 2] and knowledge distillation techniques [4] have enabled the development of efficient models without significant drop in performance. This project focuses on developing a lightweight multimodal framework that maintains competitive performance while operating under strict resource constraints, leveraging attention-based fusion and temporal modeling. Recent advances in attention-based and transformer-driven multimodal fusion have further improved the effectiveness and adaptability of biometric systems [6– 10].

## Experiments

The proposed biometric authentication framework integrates lightweight models for both modalities. Facial embeddings are extracted using MobileFaceNet [1], while face detection is performed using BlazeFace [2]. BlazeFace has been selected for its real-time inference capability and efficiency compared to multi-stage detectors such as Multi-task Cascaded Convolutional Neural Network (MTCNN).

For voice recognition, a compressed ECAPA model is trained through knowledge distillation, where a pretrained large scale Emphasized Channel Attention, Propagation and Aggregation - Time Delay Neural Network (ECAPA-TDNN) model [3, 4] acts as a teacher to transfer its knowledge to the compressed ECAPA student model. The training objective function combines the student's cross-entropy loss against the ground truth with cosine similarity loss plus mean squared errors between the embeddings produced by the teacher and the student. Such a combined loss function aligns the embeddings from the student with those from the teacher while preserving the overall discriminative power of the student.

The experimental dataset is derived from a subset of VoxCeleb2 [5], consisting of over 67,000 processed video samples with synchronized audio-visual data. Audio is converted into log-Mel spectrograms, while facial frames are processed into feature embeddings through the face pipeline.

Three fusion strategies are evaluated: Feature concatenation followed by LSTM, Attention-based fusion with LSTM, and Attention-based fusion with masked mean pooling (no LSTM). These system configurations facilitate the analysis of multimodal interactions and the necessity of additional temporal modeling, providing a natural extension to recent advances in attention-based fusion for biometric frameworks [6=8].

https://doi.org/10.18745/pb.27307

## Results and discussion

The performance of the proposed system is evaluated using verification accuracy, which measures its ability to correctly determine whether two samples originate from the same individual. The results for different system configurations are summarized in **Table 1**, with a visual comparison provided in **Fig. 1**.

*Table 1. Quantitative evaluation of verification accuracy across unimodal and multimodal frameworks. The face modality consistently uses the lightweight MobileFaceNet model as the feature extractor, while the voice modality always employs the compressed ECAPA model.*

| Model Configuration | Accuracy |
|---|---|
| **Face Modality Only** | 0.8830 |
| **Voice Modality Only** | 0.5827 |
| **Multimodal (Concat + LSTM)** | 0.9222 |
| **Multimodal (Attention + LSTM)** | 0.9304 |
| **Multimodal (Attention, No LSTM)** | 0.9364 |

The results clearly show that **multimodal fusion outperforms unimodal approaches**. The unimodal face recognition model achieves strong performance with an accuracy of 0.8830, demonstrating that MobileFaceNet retains high discriminative capability despite being lightweight. In contrast, the voice model achieves a significantly lower accuracy of 0.5827, which can be attributed to the limited training data and the reduced model capacity following distillation.

Despite the relatively weak performance of the voice model, combining both modalities leads to substantial performance improvements. The concatenation-based fusion with LSTM achieves an accuracy of 0.9222, confirming that even simple fusion strategies benefit from complementary information.

More advanced fusion strategies further improve performance. The attention-based fusion with or without LSTM both achieve an accuracy above 0.93, indicating that attention mechanisms enhance the interaction between modalities by selectively focusing on informative features, consistent with findings in recent multimodal fusion studies [6, 9].

The best performance is achieved by the **attention-based fusion without LSTM**, which records an accuracy of 0.9364. As illustrated in **Fig. 1**, this model outperforms all other configurations. This result suggests that explicit temporal modeling through LSTM is not necessary for short sequences, aligning with recent transformer-based approaches that rely on attention mechanisms without sequential modeling [7], [8]. Instead, attention mechanisms combined with masked mean pooling are sufficient to capture cross-modal relationships while reducing computational complexity.

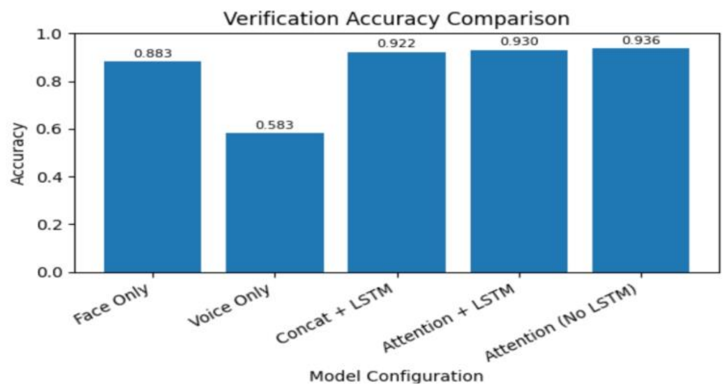



*Fig. 1. Comparative visualization highlighting performance differences between unimodal and multimodal approaches*

*Table 2. Quantitative evaluation of face recognition accuracy under varying image degradation conditions using the lightweight MobileFaceNet model as the facial feature extractor. The results demonstrate a substantial decline in recognition performance when moderate degradation is introduced, with accuracy decreasing from 0.8830 under normal conditions to 0.0958 under moderate degradation, highlighting the sensitivity of unimodal facial recognition systems to image quality deterioration.*

| Experiment | Accuracy |
|---|---|
| **Face Recognition (FaceOnlyModel)** | 0.8830 |
| **Face Recognition — Mild degradation** | 0.1008 |
| **Face Recognition — Moderate Degradation** | 0.0958 |
| **Face Recognition — Severe Degradation** | 0.0866 |

In addition to overall performance, additional robustness analysis using low-quality image inputs reveals a key limitation of the face recognition model. Under degraded conditions, accuracy drops drastically from 0.8830 down to 0.1008, 0.0958, and 0.0866 for mild, moderate, and severe degradation, indicating that the lightweight face model is highly sensitive to noise, occlusion, and illumination changes.

This finding reinforces the advantages of multimodal systems. While the face modality performs well under ideal conditions, its reliability deteriorates significantly under degraded conditions. Incorporating voice introduces a complementary source of identity information, thereby improving system resilience in real-world scenarios.

Overall, the results demonstrate that **lightweight multimodal architectures can achieve competitive performance**, particularly when enhanced with attention-based fusion. However, improvements in robustness and more effective compression of the voice model remain essential for practical deployment.

## Conclusion

This study demonstrates the feasibility oflightweight multimodal biometric systems for edge deployment. By combining efficient architectures such as MobileFaceNet [1], BlazeFace [2], and distilled ECAPA models [3-4] with attention-based fusion, the proposed framework achieves strong performance under computational constraints, in line with recent advances in lightweight multimodal and attention-based architectures [6-10].

Multimodal fusion consistently improves recognition accuracy, and attention-based fusion without LSTM provides the best trade-off between efficiency and performance. Nevertheless, challenges remain in improving robustness to degraded inputs and enhancing the performance of lightweight voice models.

https://doi.org/10.18745/pb.27307

Future work could focus on improving the training strategies of knowledge distillation and enhancing robustness under degraded conditions to support reliable deployment in real-world environments.

## References

[1] S. Chen, Y. Liu, X. Gao, and Z. Han, "MobileFaceNets: Efficient CNNs for accurate real-time face verification on mobile devices," in *Proc. Chinese Conf. Biometric Recognition (CCBR)*, 2018, pp. 428–438.
[2] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, "BlazeFace: Sub-millisecond neural face detection on mobile GPUs," *arXiv preprint arXiv:1907.05047*, 2019.
[3] B. Desplanques, J. Thienpondt, and K. Demuynck, "ECAPA-TDNN: Emphasized channel attention, propagation and aggregation in TDNN-based speaker verification," in *Proc. Interspeech*, 2020, pp. 3830–3834.
[4] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, 2015.
[5] A. Nagrani, J. S. Chung, and A. Zisserman, "VoxCeleb2: Deep speaker recognition," in *Proc. Interspeech*, 2018, pp. 1086–1090.
[6] A. Prakash, A. Nambiar, and A. Bernardino, "Multimodal adaptive fusion of face and gait features using attention-based deep neural networks for human identification," *arXiv preprint arXiv:2303.13814*, 2023.
[7] R. Yang, Q. Zhang, and L. Meng, "AuthFormer: Adaptive multimodal biometric authentication transformer for middle-aged and elderly people," *arXiv preprint*, 2024.
[8] K. O. Yang, J. Koh, and J. W. Choi, "Unified contrastive fusion transformer for multimodal human action recognition," *arXiv preprint arXiv:2309.05032*, 2023.
[9] S. Dong, W. Xie, D. Yang, Y. Li, J. Zhang, J. Tian, and J. Lei, "SeaDATE: Remedy dual-attention transformer with semantic alignment via contrast learning for multimodal object detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 35, no. 5, pp. 4713–4726, 2024.
[10] L. Lin, Y. Zhao, J. Meng, and Q. Zhao, "A federated attention-based multimodal biometric recognition approach in IoT," *Sensors*, vol. 23, no. 13, p. 6006, 2023.

https://doi.org/10.18745/pb.27307

# TempFAS: Video-Level Face Anti-Spoofing Using CNN-LSTM Temporal Modelling

**Muhammad Irfan, Dr. Lily Meng*, Prof. Iosif Mporas**

SPECS, University of Hertfordshire, Hatfield, UK

*Corresponding author: l.1.meng@herts.ac.uk

**Face anti-spoofing (FAS) systems must reliably distinguish genuine facial presentations from spoofing attacks across diverse attack modalities. This paper presents TempFAS, a video-level FAS framework combining a ResNet18 convolutional backbone with a 3-layer Long Short-Term Memory (LSTM) temporal aggregator, trained and evaluated under SiW-Mv2 Protocol I encompassing 14 distinct attack types spanning 2D presentation attacks, 3D masks, cosmetic makeup, and partial coverings. TempFAS achieves 1.69% Average Classification Error Rate (ACER), outperforming the dataset-provided Protocol I baseline of 2.6% ACER reported by SRENet [8], demonstrating the effectiveness of deep temporal sequence modelling for multi-attack face anti-spoofing.**

**Keywords:** face anti-spoofing; temporal modelling; LSTM; CNN; SiW-Mv2

## Introduction

Face recognition has become a widely adopted biometric technology in applications such as mobile authentication, access control, and digital identity verification [1], [2]. Despite its convenience, face recognition systems are vulnerable to presentation attacks, where an attacker attempts to impersonate a genuine user using printed photographs, replayed videos, or display-based media. Face anti-spoofing (FAS) aims to determine whether a presented facial sample originates from a live subject or an artificial source. Early approaches relied on handcrafted texture descriptors and traditional classifiers, while recent methods predominantly employ deep learning frameworks to learn discriminative features directly from data [3].

A fundamental limitation of frame-level FAS methods is their reliance on spatial texture cues visible in individual frames. Modern FAS benchmarks such as SiW-Mv2 [8] encompass 14 distinct attack types across four categories: 2D presentation attacks (Replay, Paper), 3D masks (Mannequin, Silicone, Mask_TransparentMask, Mask_PaperMask, Mask_HalfMask), cosmetic makeup (Makeup_Cosmetic, Makeup_Impersonation, Makeup_Obfuscation), and partial coverings (Partial_Eye, Partial_FunnyeyeGlasses, Partial_Mouth, Partial_PaperGlasses). Many of these attacks preserve natural skin texture and facial geometry at the frame level, making spatial appearance an unreliable discriminator. Existing frameworks such as SRENet [8] address this through pixel-wise spoof region estimation. However, the distinguishing characteristics of many attacks are temporal: 3D masks lack the micro expression variability of real faces, transparent masks produce temporally invariant specular highlights, and obfuscation makeup creates abnormally low temporal variance relative to surrounding skin [4]. Long Short-Term Memory (LSTM) networks [5] capture these sequence-level cues by processing ordered frame features from a convolutional backbone [6].

This paper presents TempFAS, evaluated under SiW-Mv2 Protocol I [8], a Known Spoof Attack Detection protocol with a fixed train/test partition enabling direct comparison against published baselines. A depth ablation over LSTM layers demonstrates that a 3-layer LSTM achieves 1.69% ACER, outperforming the SRENet baseline of 2.6% ACER and improving substantially on the transparent mask attack category compared to shallower variants.

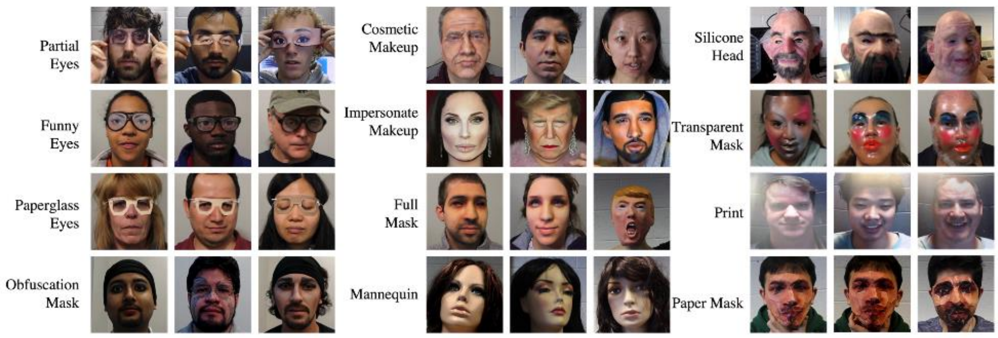



*Fig 1. illustrates representative examples from each attack category in SiW-Mv2. The visual similarity between genuine faces and several attack types at the frame level underscores the need for temporal sequence modelling.*

https://doi.org/10.18745/pb.27307

## Experiment

The SiW-Mv2 dataset [8] contains high-resolution videos of live subjects and 14 distinct attack types as described above. After deduplication of repeated entries in the protocol file, the training set comprised 537 unique attack videos and 524 live videos; the test set comprised 362 attack videos and 261 live videos. Videos are recorded at 35fps with an average duration of 6 seconds (approximately 210 frames). Frames were extracted at every 10th frame (3.5fps effective), producing approximately 21 frames per video, resized to 224x224 pixels. Each training sample is a sequence of 10 consecutive frames. Videos with fewer than 10 extracted frames were excluded, yielding 971 training clips and 572 test clips.

Fig 2 illustrates the TempFAS architecture. ResNet18 [7] was selected as the spatial encoder for its 512-dimensional output well suited as LSTM input, its efficiency of 11.2M parameters and 1.82~GFLOPs per frame, and its wide adoption in FAS research [2],[8]. The final FC layer is removed and each frame is processed independently producing 512-d feature vectors. A 3-layer LSTM with hidden size 256 and inter-layer dropout processes the ordered sequence, producing a 256-d temporal summary at the final time step. A classification head comprising Dropout (0.3) followed by Linear (256, 2) produces binary logits. Total parameters are approximately 13.02M. Training uses cross-entropy loss, Adam optimiser ($lr = 1e-4$, weight $decay = 1e-4$), cosine annealing for 20 epochs, and gradient clipping at $max\_norm = 5.0$.

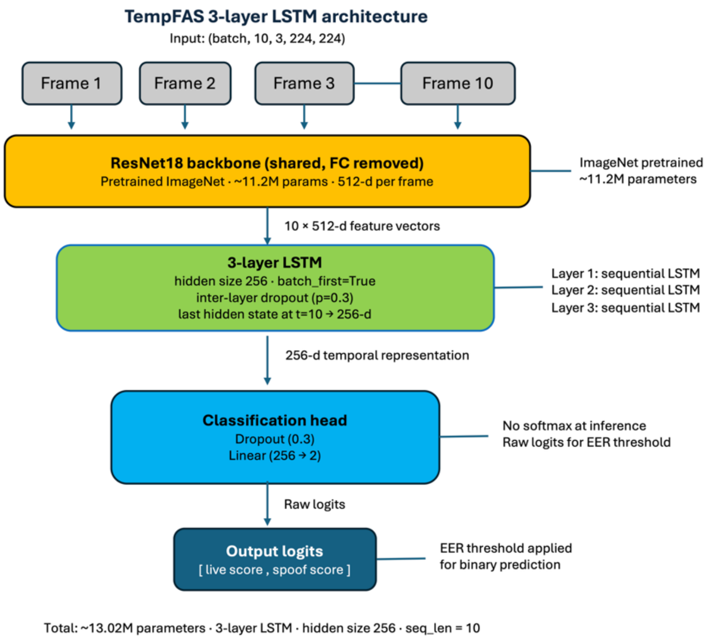



*Fig 2. TempFAS framework with 3-layer LSTM architecture. ResNet18 extracts 512-d features per frame. The 3-layer LSTM aggregates the sequence producing a 256-d temporal representation passed to the classification head.*

*Evaluation follows Protocol I benchmark standards. Raw logit scores are used without softmax at inference. Video-level scores are produced by averaging the top 30% of per-frame scores. The Equal Error Rate threshold is computed from the test-set ROC curve. The primary metric is ACER, the mean of APCER and BPCER.*

## Results and Discussion

TempFAS with a 3-layer LSTM achieves 1.69% overall ACER on SiW-Mv2 Protocol I, outperforming the SRENet Protocol I baseline of 2.6% ACER [8]. A depth ablation confirms the 3-layer model is optimal: the 1-layer baseline achieves 2.46% ACER and the 2-layer variant 2.89% ACER, with the 2-layer regression attributable to training instability at the available dataset size. The 3-layer model benefits from inter-layer dropout regularisation that stabilises convergence.

Fig 3 presents per-attack results for TempFAS against the SRENet baseline. TempFAS achieves lower overall ACER (1.69% vs 2.6%), driven by strong performance on 2D attacks, partial coverings, and impersonation makeup. SRENet retains lower per-attack ACER on several 3D mask and makeup categories where it reports near-zero error, highlighting the complementary nature of spatial and temporal detection approaches. The 3-

layer LSTM reduces the transparent mask error substantially compared to the 1-layer baseline, confirming that deeper temporal abstraction captures higher-order specular highlight dynamics across the sequence.

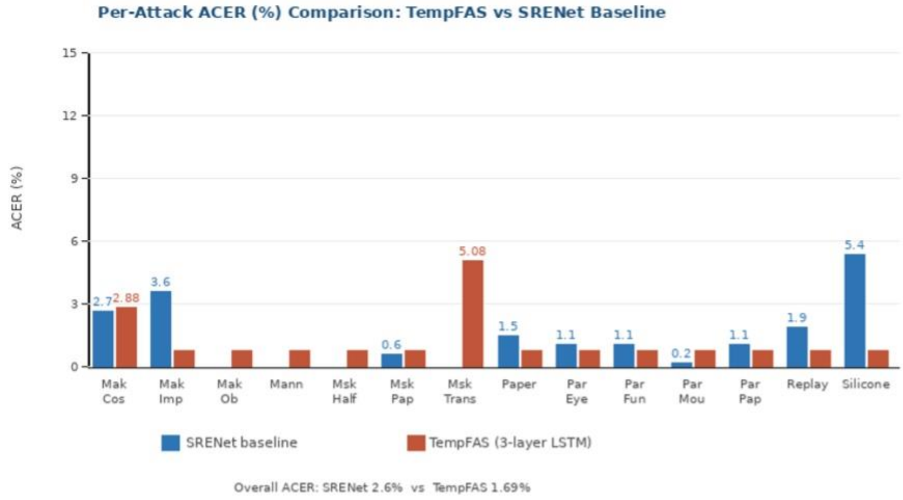



*Fig 3. Per-attack ACER (%) comparison: TempFAS 3-layer LSTM vs SRENet baseline. TempFAS achieves lower overall ACER (1.69% vs 2.6%) through strong performance across 2D attacks and partial coverings. SRENet retains advantage on several 3D mask categories*

The LSTM captures the abnormal rigidity of mask-covered regions and the temporally uniform appearance of obfuscation makeup across the 10-frame sequence, temporal cues absent in any single frame. The complementary strengths of spatial localisation (SRENet) and temporal sequence modelling (TempFAS) suggest that combining both approaches may yield further improvements, a direction explored in future work alongside Bidirectional LSTM and attention-based frame weighting.

## Conclusion

This paper presented TempFAS, a video-level FAS framework combining a ResNet18 backbone with a 3-layer LSTM temporal aggregator, achieving 1.69% ACER on SiW-Mv2 Protocol I across 14 attack types and outperforming the SRENet Protocol I baseline of 2.6% ACER [8]. A depth ablation demonstrates that 3 LSTM layers are necessary to capture higher-order temporal patterns, particularly for the transparent mask attack category. SRENet retains a per-attack advantage on several 3D mask categories through spatial spoof region localisation, highlighting the complementary nature of the two approaches. Future work will investigate Bidirectional LSTM to exploit both past and future temporal context, spatial attention on per-frame CNN features, and temporal attention across LSTM hidden states to replace fixed score aggregation with a learned frame-weighting scheme.

## References

[1] A. Howard et al., "Searching for MobileNetV3," in Proc. IEEE/CVF ICCV, 2019, pp. 1314-1324.
[2] Z. Yu et al., "Searching Central Difference Convolutional Networks for Face Anti-Spoofing," in Proc. IEEE/CVF CVPR, 2020, pp. 5294-5304.
[3] Z. Wang et al., "Deep Spatial Gradient and Temporal Depth Learning for Face Anti-Spoofing," in Proc. IEEE/CVF CVPR, 2020, pp. 5041-5050.
[4] A. Jourabloo, Y. Liu, and X. Liu, "Face De-Spoofing: Anti-Spoofing via Noise Modelling," in Proc. ECCV, 2018, pp. 290-306.
[5] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
[6] X. Li et al., "Generalized Face Anti-Spoofing by Learning to Fuse Features," in Proc. IEEE ICPR, 2021.
[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE CVPR, 2016, pp. 770-778.
[8] X. Guo, Y. Liu, A. Jain, and X. Liu, "Multi-domain Learning for Updating Face Anti-Spoofing Models," in Proc. ECCV, 2022.
[9] D. Wen, H. Han, and A. K. Jain, "Face Spoof Detection with Image Distortion Analysis," IEEE Trans. IFS, vol. 10, no. 4, pp. 746-761, 2015.
[10] J.-D. Lin et al., "Lightweight Face Anti-Spoofing Network for Telehealth Applications," IEEE J. BHI, vol. 26, no. 5, pp. 1987-1996, 2022.

https://doi.org/10.18745/pb.27307

# Class Imbalance Correction in Deep Learning for Telecom Customer Churn Prediction: A Comparative Study of ANN and TabNet

**Govardhan Reddy Bommana[1], and Okechi Onuoha[1*]**
[1] *School of Physics, Engineering and Computer Science, University of Hertfordshire, UK*
* Corresponding author: o.onuoha@herts.ac.uk

**Abstract**

**This study investigates the effectiveness of three synthetic class imbalance correction techniques, SMOTE, ADASYN, and SMOTE-Tomek applied to two deep learning architectures, Artificial Neural Network (ANN) and TabNet, for telecom customer churn prediction. Using the IBM Telco Customer Churn dataset (7,043 records, 26.54% churn rate), eight experimental configurations are evaluated under identical conditions with recall as the primary performance evaluation metric. The results show that all oversampling techniques improve minority-class detection over uncorrected baselines. TabNet combined with ADASYN achieves the best performance with recall of 0.7941, F1-score of 0.6312, and ROC-AUC of 0.8340, outperforming all other configurations.**

Keywords: Telecom Customer Churn Prediction; Class Imbalance; SMOTE; ADASYN; TabNet

## Introduction

Customer churn prediction is one of the most financially critical challenges in the telecommunications industry. Retaining existing customers is significantly more cost-effective than acquiring new ones, making early identification of "at-risk" customers is essential for proactive retention strategies [1]. With the growing availability of behavioural and transactional customer data, machine learning and deep learning approaches have become central to automated churn management systems [2].

A fundamental obstacle in churn prediction is class imbalance, where churned customers represent a minority of the overall dataset. Models trained on raw skewed distributions tend to favour the majority non-churn class, producing deceptively high overall accuracy while failing to correctly identify customers who matter most. Without any correction, an AdaBoost classifier achieved 84.9% accuracy but identified only 13 out of 980 actual churn cases a recall of just 0.013 [3], illustrating why accuracy alone is an inadequate evaluation metric in imbalanced settings.

This study makes two key contributions. First, it provides a controlled comparison of SMOTE, ADASYN, and SMOTE-Tomek within a unified deep learning experimental framework. Second, it introduces TabNet, a deep learning architecture designed for structured tabular data with built-in sequential attention for feature selection and interpretability into the telecom churn domain where it has not previously been evaluated alongside systematic imbalance correction.

## Model Development and Implementation

The IBM Telco Customer Churn dataset [4] was used, containing 7,043 customer records. After removing customerID, 20 features were retained covering demographic, service, and account attributes, with Churn as the binary target. The dataset contained no missing values and no duplicates. The class distribution confirmed 5,174 non-churn (73.46%) and 1,869 churn (26.54%) customers.

Preprocessing involved converting TotalCharges to numeric, imputing missing values with the column median, label-encoding all categorical variables, and standardizing numerical features. A stratified 80:20 train-test split produced a training set of 5,634 and a test set of 1,409 samples, both preserving the 26.5% churn ratio. All oversampling was applied exclusively to training data to prevent data leakage.

https://doi.org/10.18745/pb.27307


*Fig. 1. Training set class distribution before and after resampling techniques*

SMOTE [5] generated synthetic minority samples through interpolation, producing 4,139 samples per class. ADASYN [6] adaptively concentrated synthesis in harder-to-learn boundary regions, yielding 4,164 churn and 4,139 non-churn samples. SMOTE-Tomek combined oversampling with Tomek link removal to clean borderline instances, resulting in 3,935 samples per class, as illustrated in Fig. 1.

The artificial neural network (ANN) comprised three hidden layers (128, 64, 32 neurons) with ReLU activation, Batch Normalisation, and Dropout regularisation (0.3, 0.3, 0.2), totaling 13,313 trainable parameters. TabNet [7] was configured with n_d=32, n_a=32, n_steps=5, and gamma=1.5, using sequential attention to dynamically select the most relevant features at each decision step. Early stopping (patience=10) based on validation AUC was applied to both models. All experiments were conducted in Google Colab using Python, TensorFlow/Keras, PyTorch-TabNet, Scikit-learn, and Imbalanced-learn.

## Results and Discussion

The full experimental results across all eight configurations are presented in below Table1.

*Table. 1. Performance comparison of ANN and TabNet across imbalance correction techniques*

| *Model* | *Technique* | *Recall* | *Precision* | *F1-Score* | *ROC-AUC* |
|---|---|---|---|---|---|
| *ANN* | *Baseline* | *0.5241* | *0.5886* | *0.5545* | *0.8211* |
| *ANN* | *SMOTE* | *0.6551* | *0.542* | *0.5932* | *0.811* |
| *ANN* | *ADASYN* | *0.7112* | *0.4935* | *0.5827* | *0.804* |
| *ANN* | *SMOTE-Tomek* | *0.6658* | *0.5366* | *0.5943* | *0.8067* |
| *TabNet* | *Baseline* | *0.4492* | *0.6316* | *0.525* | *0.8257* |
| *TabNet* | *SMOTE* | *0.7059* | *0.5156* | *0.5959* | *0.8202* |
| *TabNet* | *ADASYN* | *0.7941* | *0.5238* | *0.6312* | *0.834* |
| *TabNet* | *SMOTE-Tomek* | *0.7193* | *0.5019* | *0.5912* | *0.8039* |

Without imbalance correction, both models demonstrated the characteristic limitation of skewed training data. The ANN baseline identified only 52.41% of actual churners. TabNet baseline was even more conservative with recall of 0.4492, though it achieved the highest precision of 0.6316 reflecting its attention mechanism defaulting to a selective, high-precision strategy without balanced data.

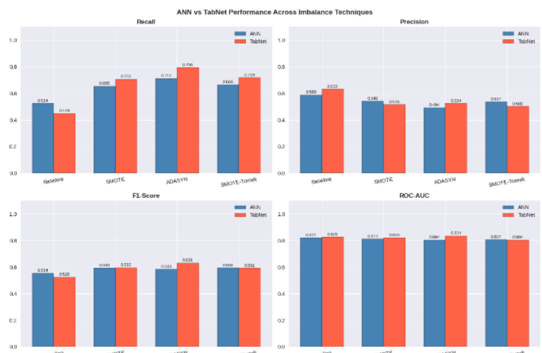



*Fig. 2. ANN vs TabNet performance comparison across imbalance correction techniques*

All oversampling techniques improved recall substantially. For the ANN, recall rose from 0.5241 to 0.6551 with SMOTE, 0.6658 with SMOTE-Tomek, and 0.7112 with ADASYN. TabNet responded even more strongly: SMOTE improved recall to 0.7059, SMOTE-Tomek to 0.7193, and ADASYN to 0.7941, the highest across the entire

https://doi.org/10.18745/pb.27307

experiment. TabNet + ADASYN also achieved the best F1-score (0.6312) and ROC-AUC (0.8340), making it the best-performing configuration overall, as shown in Fig 2. This suggests TabNet's sequential attention mechanism benefits particularly from ADASYN's boundary-focused synthetic samples, enabling more precise feature selection on discriminative minority-class regions.

SMOTE-Tomek consistently produced better precision than ADASYN for both models, confirming that Tomek link cleaning reduces false positives, though it does not maximise recall. The precision trade-off under ADASYN most notable for ANN (0.4935) reflects the adaptive technique's tendency to generate samples in ambiguous regions, introducing some noise into model learning.

A robustness checks across two additional seeds confirmed stability, with mean recall of 0.7179 (std = 0.0095) and mean ROC-AUC of 0.8260 (std = 0.0050). Contract type, tenure, and monthly charges emerged as top predictors via TabNet's attention weights, consistent with [8] and [9], and can directly inform retention interventions such as contract upgrade offers for high-risk customers.

## Conclusion

This study conducted a controlled experimental comparison of three synthetic class imbalance correction techniques SMOTE, ADASYN, and SMOTE-Tomek applied to ANN and TabNet deep learning architectures for telecom customer churn prediction using the IBM Telco dataset. All oversampling techniques improved minority-class recall over uncorrected baselines for both models, confirming the central hypothesis that synthetic correction meaningfully enhances churn detection in deep learning systems. TabNet combined with ADASYN emerged as the best overall configuration, achieving recall of 0.7941, F1-score of 0.6312, and ROC-AUC of 0.8340, outperforming all other combinations across the primary evaluation metrics. Robustness checks confirmed that these results are stable across different data splits. The findings provide evidence-based guidance for practitioners selecting imbalance correction strategies in telecom churn prediction, and establish TabNet as a strong candidate architecture for structured tabular churn data. Future work should explore cost-sensitive learning, ensemble combinations of ANN and TabNet, and evaluation on additional telecom benchmark datasets to further generalise these findings.

## References

[1] M. N. H. Mamun, "Advancements in machine learning for customer retention: a systematic literature review of predictive models and churn analysis," Journal of Sustainable Development and Policy, vol. 1, no. 01, pp. 250–284, 2025.
[2] M. Shahabikargar, A. Beheshti, X. Zhang, J. Foo, and A. Jolfaei, "A comprehensive survey on customer churn analysis studies," Journal of Information and Telecommunication, pp. 1–47, 2025.
[3] A. Chowdhury, S. Kaisar, M. M. Rashid, S. S. Shafin, and J. Kamruzzaman, "Churn prediction in telecom industry using machine learning ensembles with class balancing," IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), 2021. DOI: 10.1109/CSDE53843.2021.9718498
[4] IBM, "Telco-customer-churn-on-icp4d," GitHub repository, 2015. Available: https://github.com/IBM/telco-customer-churn-on-icp4d
[5] S. Yadav, "SMOTE in predictive modeling: a comprehensive evaluation of synthetic oversampling for class imbalance," International Journal of Innovative Research in Multidisciplinary Field, vol. 8, no. 4, pp. 1–9, 2020.
[6] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," IEEE International Joint Conference on Neural Networks, pp. 1322–1328, 2008.
[7] S. Ö. Arik and T. Pfister, "TabNet: Attentive interpretable tabular learning," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 8, pp. 6679–6687, 2021.
[8] S. K. Wagh et al., "Customer churn prediction in telecom sector using machine learning techniques," Results in Control and Optimization, vol. 14, 100342, 2024.
[9] A. Y. W. Chong, K. W. Khaw, W. C. Yeong, and W. X. Chuah, "Customer churn prediction of telecom company using machine learning algorithms," Journal of Soft Computing and Data Mining, vol. 4, no. 2, pp. 1–22, 2023.

https://doi.org/10.18745/pb.27307

## 3. Sensing the Invisible: Airborne Particles and Bio-Detection Technologies

Chair: Prof. Loic Couldron

Judges: Dr Etelka Chung, Dr Lanka Weerasiri

https://doi.org/10.18745/pb.27307

## Invited Speaker – Dr Nikoly Dimov

# Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device

**Nikolay Dimov[1*], Darryl Ethan Bernstein[2], Srinivasu Puttaswamy[1], Gloria Sikapite[2], Martin Ebon[2], Jameel M Inal[3,4], Ian David Johnston[1], and Nikhil Vasdev[2]**

[1]*School of Physics, Engineering & Computer Science, University of Hertfordshire, UK*
[2]*Hertfordshire and Bedfordshire Urological Cancer Centre, Department of Urology, Lister Hospital, East and North Hertfordshire Teaching NHS Trust, Stevenage, UK*
[3]*School of Life & Medical Sciences, University of Hertfordshire, UK*
[4]*School of Human Sciences, London Metropolitan University, UK*
*Corresponding author: n.dimov@herts.ac.uk

**Using Principal Component Analysis (PCA), this contribution investigates possible links between Prostate Cancer (PCa) patients' Gleason Scores and the flow behaviours of urine samples as they pass through an extracellular vesicle (EV) isolation stage. A bespoke mesofluidic tangential flow filtration system logged pressure cycles for each sample; the data were then cleaned, normalised, and reduced to three parameters capturing the dynamics of purification cycles as they progressed. Variance in the data is analysed, and two principal components are used to group the patients (N=17) in four clusters. Despite explaining much of the variance, no correlation was found between the Gleason Scores and the clustering.**

Keywords: Diagnostics; Prostate Cancer (PCa); Liquid Biopsy; Point-of-Care; Principal Component Analysis.

## Introduction

Liquid biopsy for early detection of prostate cancer (PCa) from urine extracellular vesicles (EVs) can reduce the time from test to diagnosis by side-stepping invasive, cumbersome and costly laboratory practices, reducing the socioeconomic impact by timely treatment. Developing such technology requires purifying low-concentration species from a relatively large volume of a challenging sample. A mesofluidic Tangential Flow Filtration system has successfully isolated EVs from urine samples from 31 patients, followed by shotgun proteomics to identify biomarkers [1]. However, machine learning (ML) and embedded data analysis will become part of PCa diagnosis. Here, we present a PCA analysis of the purification metadata, linking Gleason scores to the rheological behaviour of the samples.

## Experimental

To separate EVs based on size, the system (Figure 1a) operates at feed flow rates ranging from 0.56 to 25 mL/s, and patient sample volumes of 65.3±18.4 mL are used. Nanometre-sized EVs remain in the retentate, whilst larger, heavier proteins are deposited on the membrane; the urine collects in the permeate. The retentate is recycled through the system, and pressure is continuously logged for each sample.

Using data from 17 samples and Jupyter (Python 3), pressure data (Figure 1b) were scaled, normalised, and reduced to three parameters: average time to reach peak pressure, average drop time and the average amplitude. Two principal components were used to explain the variance in those three parameters. Finally, the samples were clustered into four groups and mapped to patients' Gleason scores.

https://doi.org/10.18745/pb.27307

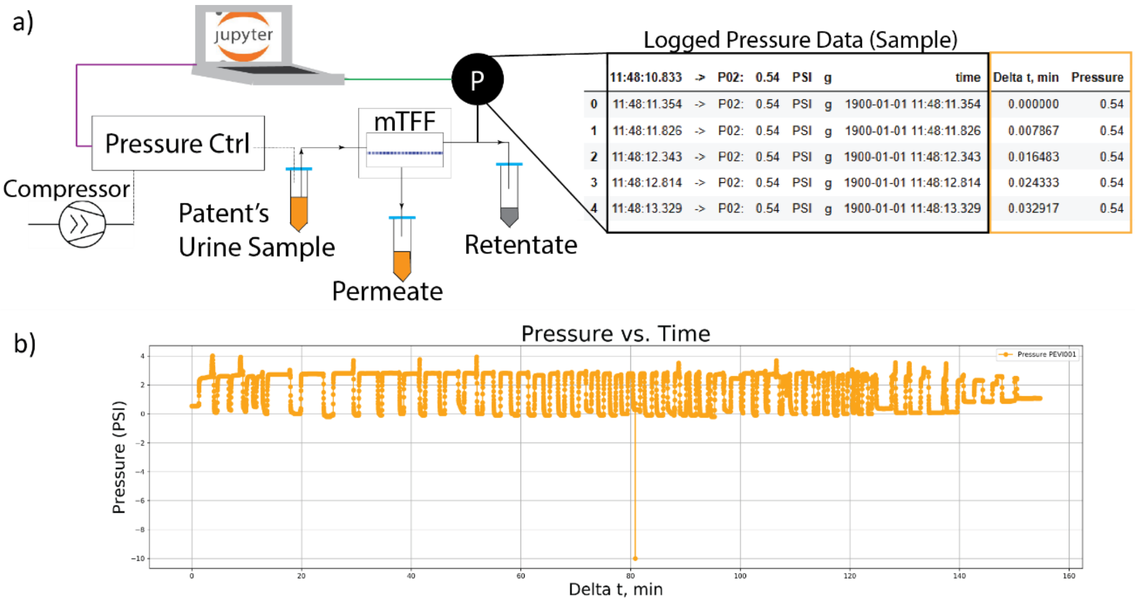



*Fig. 1. Schematic of the EV purification system consisting of a mesofluidic Tangential Flow Filtration (mTFF) device, splitting the urine samples in EV-containing retentate and waste permeate, whilst logging the pressure data continuously with a timestamp (A); and a plot of the recorded pressure versus the elapsed time (B).*

## Results and discussion

As the samples passed through the system, the membrane in the mTFF would begin to foul, gradually reducing permeate flux. This also affected the split ratio between retentate, and the pressure on the retentate side fluctuated in this process. Effectively, each consecutive cycle drove more fluid past the membrane than through it. Exploring the raw data of the process dynamics, pressure versus time, and comparing the pressure profiles did not indicate any significant differences. However, the normalised and truncated data (Fig. 2a) were useful for comparing like with like and sidestepping variations in the samples, e.g., initial volume. Timing the pressure rises and falls, and the amplitude of the signal were the features that led to two principal components explaining more than 95% of the variance, with PC1 and PC2 (Fig. 2b), an indication of PCA fidelity.

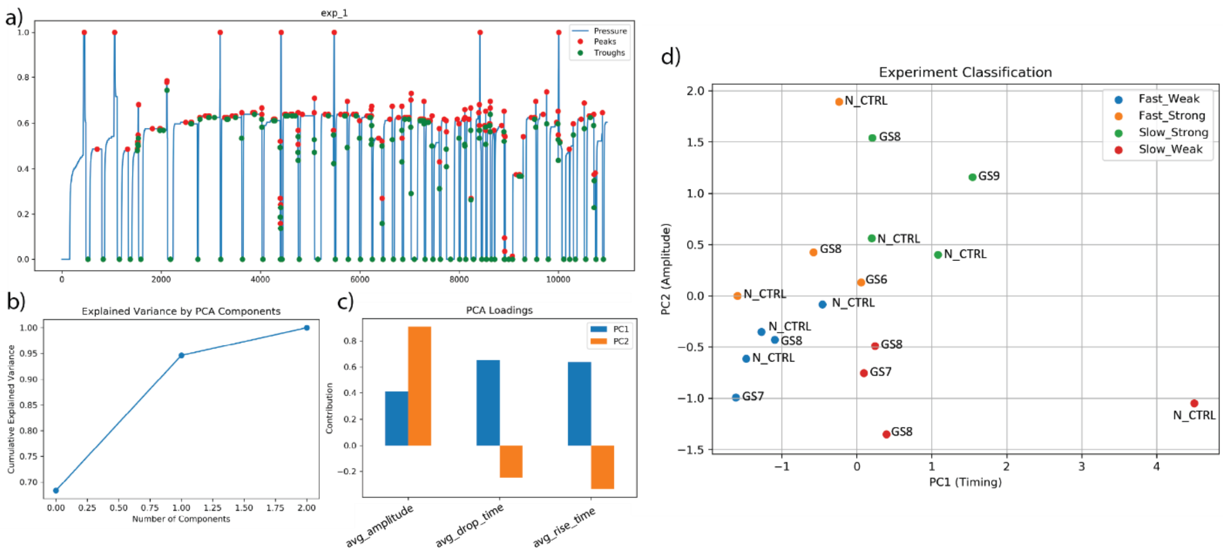



*Fig. 2. Normalised amplitudes for the pressure with identified peaks and through regions (a) are used in this analysis, leading to two principle components (PC) that explain the variance in the data (b); amplitude, time to rise and time to*

*drop were derived to reduce the data and observe their loadings (c); the samples are classified in four groups (d) and the Gleason Scores (GS 6 to 9) are mapped alongside negative controls.*

Overall loadings indicate that the principal components show a positive correlation with the average amplitude, with PC1 showing the highest loading of nearly 91%. These two components were used to identify two classes: speed and strength, and the combined classes were colour-coded in Figure 2d. However, these four classes did not match the Gleason scores. The scatter of data from healthy volunteers indicates that the model captures the sample properties in terms of purification dynamics, but these properties are not directly correlated to the disease. In this author's opinion, an automated system with identical sampling conditions and feed rates might close the gap between the two by reducing any initial variations and minimising the truncation of the raw data.

## Conclusion

In summary, successful modelling of metadata from pressure readings was performed using PCA. Classifying the samples according to that model showed a weak correlation between the flow dynamics through the system and the diagnosis of PCa. Partially, this could be due to initial volume differences, or equally valid, from the low number of samples considered in this paper. Further investigations and technical advances are required to improve our understanding of those variations; through modelling, combining and analysing metadata may improve noninvasive PCa diagnostics.

## References

[1] S. Lange, D.E. Bernstein, N. Dimov *et al*. "Urinary extracellular vesicle signatures as biomarkers in prostate cancer patients" *Int. J. Mol. Sci.*, vol. 26, no. 6895, 2025.

https://doi.org/10.18745/pb.27307

# Low-Flowrate Collection and Quantification of Airborne Human DNA in an Occupied Meeting Room

**Muhammad Sohaib Ikram[1]*, Richard Kaye[2], Ian Munro[2], Lanka Weerasiri[2], Shan Goh[2], Loic Coudron[2], Ian D Johnston[2], Nikolay Dimov[2]**

*School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK*

*Corresponding author: mi22aay@herts.ac.uk

**Airborne human DNA (hDNA) has strong forensic potential, but most reported air-collection studies rely on high-flow, noisy samplers. This work evaluated whether lower-flow systems can recover hDNA in a meeting room under controlled, stagnant-air conditions. A pre-sampling stage, one-hour occupied sampling stage and 30-minute post-sampling stage were monitored using OPC/APS aerosol instruments and multiple hDNA samplers. qPCR showed that C-PAS yielded the highest peak recovery, ESP provided broader spatial detection, Sterivex showed the most consistent temporal recovery, and SKC samples were undetermined. The study demonstrates that low-flow airborne hDNA recovery is feasible, but precision remains limited by low-template quantification and extraction efficiency.**

**Keywords:** airborne human DNA; forensic genetics; low-flow air sampling; electrostatic precipitation; Quantifiler Trio kit

## Introduction

Airborne human DNA (hDNA) is an emerging evidence type in forensic genetics because people continuously shed biological material into the surrounding environment through skin cells, breath and other bodily emissions. Earlier work on epidermal desquamation showed that humans shed large numbers of skin cells daily [1], and more recent forensic studies have shown that hDNA can be recovered from indoor air and dust and may be informative for reconstructing human occupancy [2]–[4]. However, most currently reported airborne hDNA studies use high-flow commercial samplers, which can be noisy, intrusive and less practical for real-world deployment in occupied rooms [2]–[4]. That leaves an important methodological gap: whether lower-flow and quieter devices can still recover amplifiable hDNA under indoor conditions.

The present study addressed this gap in a standard university meeting room under stagnant-air conditions. The objective was to evaluate the collection and quantification of hDNA using a bespoke electrostatic precipitator (ESP), a compact particle air sampler (C-PAS), and Sterivex filtration, while linking DNA recovery to aerosol behaviour measured by optical particle counters (OPCs) and an aerodynamic particle sizer (APS). In parallel, the work aimed to test a contamination-controlled extraction and Quantifiler™ Trio qPCR workflow suitable for low-template airborne samples, and to compare pre-occupancy, active occupancy and post-occupancy recovery.

## Experimental/Simulation

Air sampling was performed in a standard meeting room (approximately 6 m × 4 m × 2.5 m) in the Spectra Building, Room SP2024A, University of Hertfordshire, under UH Ethics Protocol Number: 1423 ST HSET 2025. Before each experiment, the room surfaces, cables and equipment were thoroughly cleaned with DNA Away© using sterilised Kim wipes. The study comprised three sequential stages: a 40 min pre-sampling phase during which only the researcher and principal investigator entered to set up the instruments; a 60 min occupied sampling phase with 19 individuals in total, including the researcher; and a 30 min post-sampling phase after the room was vacated. The participants represented mixed ages and body sizes and undertook normal meeting-room activity such as sitting, talking and working. Ventilation remained off throughout, so the room air was intentionally stagnant.

Air sampling employed six ESP devices operating at approximately 10–15 L/min with non-coated steel slides as the collection substrate, four C-PAS units operating at 2 L/min with mixed cellulose ester filters, one 0.45 µm Sterivex filter at 10 L/min, and one SKC liquid-based sampler at 10 L/min using 20 mL of 1% PBS. Six OPCs and one APS were arranged around the table to monitor spatial and temporal aerosol behaviour. The table was divided into left and right sides relative to the seated researcher, and the ESP and OPC devices were distributed along researcher-side, middle and screen-side positions, with the C-PAS units near the centre and the

https://doi.org/10.18745/pb.27307

Sterivex/SKC units placed close to the APS. Passive controls were also included at each stage using exposed filters, slides and nuclease-free water.


*Fig. 1. Meeting-room layout showing participant seating and the approximate placement of ESP, C-PAS, Sterivex/SKC, OPC and APS instruments around the table.*

All collected samples were sealed in sterile zip-lock bags, transferred to the laboratory and stored at −20°C. DNA extraction was carried out in a biosafety cabinet decontaminated with DNA Away©. Extraction followed the use of QIAamp© DNA Investigator kit [5], using the blood/saliva protocol for ESP slide washings and the paper/similar-material protocol for filters, with an overnight incubation at 56°C and 300 rpm used as a kit protocol modification. DNA was eluted in 50 µL of Elution Buffer. Quantification was then performed on a QuantStudio® 7 Flex instrument using the Quantifiler™ Trio kit, which multiplexes a small autosomal (SA), large autosomal (LA), Y-chromosome and internal PCR control (IPC) assay in a single reaction [6].

## Results and discussion

OPC monitoring showed a clear occupancy effect. Mean counts increased from about 5.2–7.4 counts/s during pre-sampling to 12.8–16.7 counts/s during the occupied stage and remained elevated at 15.0–19.3 counts/s during post-sampling, with the left side of the room consistently higher than the right. APS measurements supported these observations and indicated a stable fine-aerosol population during and after occupancy, with total concentrations of 18.46 and 18.06 particles/cm³ for the occupied and post-sampling stages, respectively, and mean aerodynamic diameters of approximately 0.90 µm in both stages.

The Quantifiler™ Trio qPCR standard curves were highly linear, with efficiencies of 92.3% for the LA target, 95.7% for the SA target and 100.6% for the Y target. IPC Ct values were tightly grouped at approximately 27.5–28.2 cycles, indicating the absence of substantial inhibition. Most negative controls were undetermined for SA, LA and Y, confirming effective contamination control. However, the lowest standard, corresponding to 0.5 pg/ul, showed the highest replicate standard deviation across the targets and therefore limited precision at the lower end of the assay.

The three active collection platforms showed distinct performance characteristics. C-PAS produced the highest peak hDNA recovery of the study, but this occurred predominantly at one occupied-stage location on the left side of the room, consistent with the higher aerosol counts recorded there. In contrast, other C-PAS positions yielded much lower quantities or long-target dropout, showing that recovery was strongly position dependent. The ESP array produced broader spatial detection across pre-, occupied- and post-sampling stages, generally in the $10^{-4}$ to low $10^{-3}$ ng/µL range, but with greater variability in degradation index because several low-template samples lost the LA target while retaining SA. Sterivex showed the most coherent temporal pattern, increasing

https://doi.org/10.18745/pb.27307

from low pre-sampling values to clear occupied-stage and post-sampling detections in the $10^{-3}$ ng/µL range, with comparatively stable degradation indices close to unity.

Overall, these results indicate that sampler type and spatial position both controlled apparent performance. C-PAS demonstrated the strongest peak recovery potential, ESP provided the widest spatial sensitivity, and Sterivex yielded the most stable pre-to-sampling-to-post response. SKC samples were not taken forward for qPCR interpretation because all results were undetermined. Across the positive samplers, degradation indices were often modest, but values above 10 were mainly associated with near-threshold samples in which the LA target failed rather than with uniformly severe biological degradation. Importantly, positive ESP and Sterivex detections in the post-sampling stage agreed with the aerosol data and support the conclusion that hDNA-bearing particles persisted in the stagnant room after the meeting ended.

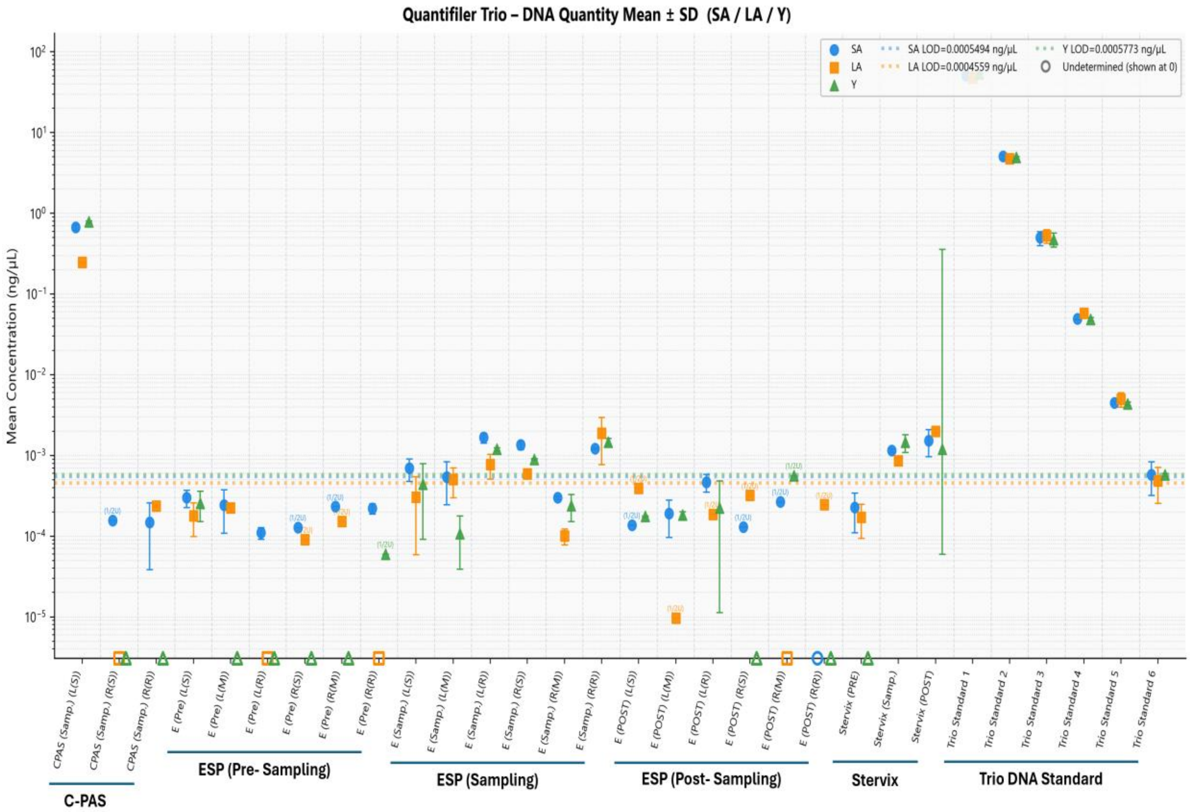



*Fig. 2. Combined Quantifiler™ Trio mean DNA quantities (SA, LA and Y) recovered from C-PAS, ESP and Sterivex samples relative to the assay limits of detection and the standard series.*

## Conclusion

This study demonstrates that airborne human DNA can be recovered from a meeting-room using lower-flow and comparatively non-intrusive air samplers. The aerosol measurements showed that occupancy generated a fine-particle cloud that persisted into the post-sampling stage, and the qPCR data confirmed that this persistence translated into recoverable hDNA. C-PAS delivered the highest individual recovery, ESP detected DNA across the broadest spatial range, and Sterivex provided the most consistent temporal profile. At the same time, the work highlights the analytical challenge of quantifying very low-template airborne DNA close to the lower end of the standard curve. Improved replication at 0.5–5 pg, more accurate pipetting and better extraction from steel slides and filters will be essential for strengthening future forensic interpretation.

## References

[1] D. Roberts and R. Marks, "The determination of regional and age variations in the rate of desquamation: a comparison of four techniques," J. Invest. Dermatol., vol. 74, no. 1, pp. 13–16, 1980.
[2] C. Fantinato, P. Gill, and A. E. Fonneløp, "Detection of human DNA in the air," Forensic Sci. Int. Genet. Suppl. Ser., vol. 8, pp. 282–284, 2022.
[3] C. Fantinato, A. E. Fonneløp, Ø. Bleka, M. D. Vigeland, and P. Gill, "The invisible witness: Air and dust as DNA evidence of human occupancy in indoor premises," Sci. Rep., vol. 13, no. 1, Art. no. 19059, 2023.
[4] E. Bibbo, D. Taylor, R. A. H. van Oorschot, and M. Goray, "Air DNA forensics: Novel air collection method investigations for human DNA identification," J. Forensic Sci., 2024.
[5] QIAGEN, QIAamp DNA Investigator Handbook, Jan. 2020.
[6] Thermo Fisher Scientific, Quantifiler™ HP and Trio DNA Quantification Kits User Guide, Rev. H, 2018.
[7] T. G. Foat et al., "A prototype personal aerosol sampler based on electrostatic precipitation and electrowetting-on-dielectric actuation of droplets," J. Aerosol Sci., vol. 95, pp. 43–53, 2016.

https://doi.org/10.18745/pb.27307

# Respirable Fibre Detection from Light Scattering Patterns

**Rob Lewis[1*], Richard Greenaway[1], and Chris Stopford[1]**
*[1]University of Hertfordshire*
*Corresponding author: r.lewis22@herts.ac.uk

**The detection of respirable fibres is crucial in the reduction or prevention of exposure that may lead to the development of various diseases. Standard counting methods are laborious and slow, whilst hitherto instrumentation show poor results and are now discontinued. A method using forward scattered light from individual fibres has been developed using modelled data to infer length and diameter measurements and tested using both simulation-based Monte Carlo and experimental investigations. Results show over 70% of the inferred dimensions were within 10% of the true dimensions for the Monte Carlo test and promising size distribution comparisons for the experimental tests.**

Keywords: Aerosol; light scattering; fibres; single particle detection; particle sizing

## Introduction

Respirable fibres deposited within the lung have been shown to cause serious damage and disease [1]; these include materials such as asbestos, carbon fibre, refractory ceramic fibres, and other mineral and vitreous fibres. The current method of detection is to filter sample the ambient air in the vicinity of a suspected exposure risk and analyse the filters using phase contrast optical microscopy (PCOM) and scanning electron microscopy (SEM)[2]. PCOM is used to count the fibres within the respirable range: a diameter less than 3 µm, a length greater than 5 µm, and an aspect ratio greater than 3. SEM is used to determine the fibre type. This process can take around 24 hours. Instrumentation developed for real-time counting using various techniques have been made [3,4,5] but compare poorly against the standard method (e.g. [6]) and have now been discontinued. Typically, a combination of light scattering and either electric- or magnetic-induced alignment is used and most rely on an ensemble of particles. This investigation used a combination of modelled light scattering data, using the Amsterdam Discrete Dipole Approximation (ADDA, [7]), and experimental data to develop and test a measurement system for single respirable fibres using only forward scattered light. The intention of this investigation is to be eventually developed into a low-cost, real-time fibre measurement system with a high-resolution particle size distribution.

## Experimental/Simulation

Using ADDA-modelled scattering data, features of the scattering data were identified and related to changes in the dimensions of the input objects. Figure 1a (left) shows a typical example of the forward scattered light from a fibre and the green and magenta boxes represent linear arrays of photodetectors, typical of instruments such as ALERT [5] and the Multiparameter Bioaerosol Sensor [8]. Figure 1a (top right) is the mean value of the magenta area, the location of the first intensity minimum (indicated by a red arrow) and the rate of change of intensity between 7.75° and 9° (blue shaded area) change with varying diameter. Figure 1a (bottom right) is the mean value of the green area, the location of the first intensity minimum and the full-width, half maximum (both indicated by red arrows) changes with varying length. This was developed into a method of inferring length and diameter using the forward scattered light and tested using both a Monte Carlo-based modelled experiment and the comparison of size distributions of microscopy data and scattering data, captured by the Particle Phase Discriminator mk. 2 (PPD2, [9]) in a series of aerosolisation experiments.

## Results and discussion

Figure 1b shows the results of the Monte Carlo-based modelled experiment, where the inferred "output" dimensions are plotted against the "input" dimensions. The red dotted line is the line of equality, the shaded blue area is 10% of that line, and each data point is coloured according to the input imaginary part of the refractive index. A clear trend is observed for both the top and bottom plots of Fig. 1b for length and diameter respectively; ~70.8% of the length outputs and ~77.2% of the diameter outputs were within 10% of the input dimensions. The outliers for the diameter measurements can be explained by the low imaginary part of the refractive index, where the increased effects of refraction and reflection can alter the shape of the resulting interference pattern. Figure 1c shows the size distribution comparisons for the aerosolisation experiments, where the dimensions found using the described method from the light scattering patterns, capture using PPD2, were compared to microscopy measurements from the down stream filter collection. The discrepancy for the diameter at around 1.2 µm is due to the crossover between the two methods used to infer the measurement.

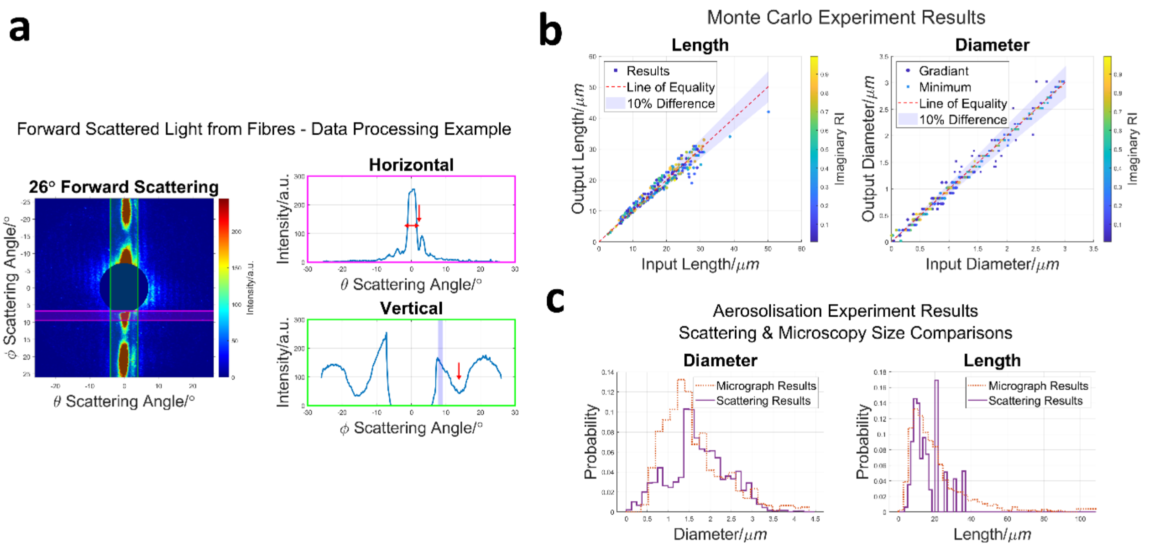



*Fig. 1. a) Left shows the forward scattering from a fibre and includes green and magenta areas, representing linear arrays of photodetectors. The right-hand side shows the mean values of the corresponding shaded areas and includes red arrows and a shaded blue area indicating measured features. b) Monte Carlo-based modelled experiment results for the length and diameter measurements (left and right respectively). The red dashed line is the line of equality, the blue shaded area is 10% of that value, and the data points are coloured depending on the imaginary part of their refractive index. c) Size distribution comparisons of the aerosolisation experiment for the diameter and length measurements (left and right respectively). The solid purple line shows the results using the scattering data and the dashed orange line shows the results from the microscopy data.*

## Conclusion

Both the modelled experiment and the aerosolisation experiment showed promisingly accurate results; most of the inferred measurements for the modelled data were correctly made within 10% and the size distributions of the aerosolisation experiment produced a reasonable match between measurement methods. However, further optimisation of the measurement algorithm is required. Additional testing using light scattering from fibres mounted on glass slides would provide robust verification into the validity of the algorithm. This is the first step in the development of a fibre measurement system for single fibrous aerosol that only relies on forward scattered light. By simplifying the instrument design, i.e., using only scattered light, the potential cost and size would be greatly reduced in comparison to previous or existing instrumentation. Furthermore, the measurement of single fibres will increase the particle size distribution resolution and accuracy in comparison with the previous instruments that relied on an ensemble of fibrous aerosol.

## References

[1] Sen, D. (2015). Working with asbestos and the possible health risks. In *Occupational Medicine* (Vol. 65, Number 1, pp. 6–14). Oxford University Press. https://doi.org/10.1093/occmed/kqu175
[2] World Health Organisation. (1985). *REFERENCE METHODS FOR MEASURING AIRBORNE MAN-MADE MINERAL FIBRES (MMMF) WHO/EURO MMMF Reference Scheme Monitoring Concentration using a Phase Contrast Optical Microscope Determining Size using a Scanning Electron Microscope prepared by the*

*WHO/EURO Technical Committee for Monitoring and Evaluating Airborne MMMF WORLD HEALTH ORGANIZATION Regional Office for Europe COPENHAGEN 1985*.
[3] Baron, P. A. (1994). Direct-reading Instruments for Aerosols. *Analyst*, *119*.
[4] Kauffer, E., Martin, P., Grzebyk, M., Villa, M., & Vigneron, J. C. (2003). Comparison of two direct-reading instruments (FM-7400 and Fibrecheck FC-2) with phase contrast optical microscopy to measure the airborne fibre number concentration. *Annals of Occupational Hygiene*, *47*(5), 413–426. https://doi.org/10.1093/annhyg/meg055
[5] Stopford, C. (2010). *Ice Crystal Classification Using Two Dimensional Light Scattering Patterns*. http://uhra.herts.ac.uk/handle/2299/4797
[6] Iles, P. J., & Shenton-Taylor, T. (1986). COMPARISON OF A FIBROUS AEROSOL MONITOR (FAM) WITH THE MEMBRANE FILTER METHOD FOR MEASURING AIRBORNE ASBESTOS CONCENTRATIONS. In *Ann. occup. Hyg* (Vol. 30, Number 1). https://academic.oup.com/annweh/article/30/1/77/150676
[7] Yurkin, M. A., & Hoekstra, A. G. (2011). The discrete-dipole-approximation code ADDA: Capabilities and known limitations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *112*(13), 2234–2247. https://doi.org/10.1016/j.jqsrt.2011.01.031
[8] Ruske, S., Topping, D. O., Foot, V. E., Kaye, P. H., Stanley, W. R., Crawford, I., Morse, A. P., & Gallagher, M. W. (2017). Evaluation of machine learning algorithms for classification of primary biological aerosol using a new UV-LIF spectrometer. *Atmospheric Measurement Techniques*, *10*(2), 695–708. https://doi.org/10.5194/amt-10-695-2017

https://doi.org/10.18745/pb.27307

# Characterisation of the aerosol transport evolution in a bioaerosol chamber using low-cost sensors

**Milad Heidari-Koochi*, Richard Baxter, Richard Kaye, Loic Coudron, Ian Johnston**
*School of Physics, Engineering and Computer Science, University of Hertfordshire, College Lane, Hatfield AL10 9AB, United Kingdom*
*Corresponding author: m.heidari-koochi@herts.ac.uk

**Understanding aerosol dispersion in indoor environments remains a key challenge in environmental and exposure science. This study investigates the spatiotemporal dynamics of aerosol transport within a controlled 8 m³ bioaerosol chamber using a distributed network of low-cost sensors. Nebulised polystyrene microspheres (3 µm) are used as model aerosols and a three-dimensional array of optical particle counters enables reconstruction of time-resolved concentration fields following localised release of particles. The results provide new insight into aerosol transport and mixing and demonstrate the potential of low-cost sensor networks for resolving indoor aerosol behaviour.**

Keywords: aerosol transport; bioaerosol chamber; low-cost sensors; indoor air quality; sensor networks

## Introduction

Understanding the effects of indoor air quality on health is increasingly important [1], [2]. In developed countries people spend 80–90% of their time indoors, yet aerosol transport processes in enclosed environments remain insufficiently characterised. Indoor aerosol behaviour is governed by a combination of particle properties, source characteristics, airflow, and environmental conditions. Experimental studies typically rely on controlled aerosol chambers to provide reproducible environments for environmental, biological, and aerosol research, including instrument calibration and sampler evaluation. However, the spatiotemporal distribution of aerosols within such systems is often overlooked. Typically, single-point measurements are used, assuming entire domain representation, which introduces uncertainty into exposure estimates [3]. In reality, aerosol transport in confined environments is governed by a complex interplay of gravitational settling [4], turbulent and convective transport [5], electrostatic effects, and flow perturbations.

As a result, spatially resolved measurements are required to accurately capture aerosol dynamics. Recent advances in compact optical particle counters (OPC) [6] enable the deployment of distributed measurement networks that were previously impractical. This study uses such a network to quantify how aerosol spatial distributions evolve following nebulisation and how this evolution depends on internal flow conditions.

## Experimental setup

This study employs a large 8 m³ bioaerosol chamber (Fig. 1a) as a controlled and reproducible experimental environment representing a simplified indoor space, to investigate the spatiotemporal evolution of aerosols released from a localised source. The chamber features HEPA filtration, bio and non-bio aerosolisation capabilities, and real-time aerosol monitoring. Three-micron polystyrene microspheres are used as the test aerosol due to their well-defined physical properties and suitability for repeatable aerosol generation. Aerosols are introduced using a Hudson Updraft II nebuliser, and their subsequent dispersion is monitored using a distributed network of OPCs (Alphasense OPC-N3). A total of 9 OPCs are arranged in a body-centred cubic lattice with cube edge length of 1 m (Fig. 1b), enabling volumetric sampling of aerosol concentration. All sensors are time-synchronised to capture transient behaviour following aerosol release, allowing resolution of both spatial and temporal evolution of particle concentration. The chamber is equipped with extraction fans operating continuously at 10 % capacity to maintain a baseline airflow. In addition, an oscillating internal fan can be activated to enhance mixing, allowing investigation of the effects of forced convection on the aerosol transport regime.

Two main conditions were studied: (i) the internal fan switched off (fan-off condition), and (ii) the internal fan operating (fan-on condition). For each condition, aerosol generation was maintained for 30 minutes, followed by full extraction of the chamber.

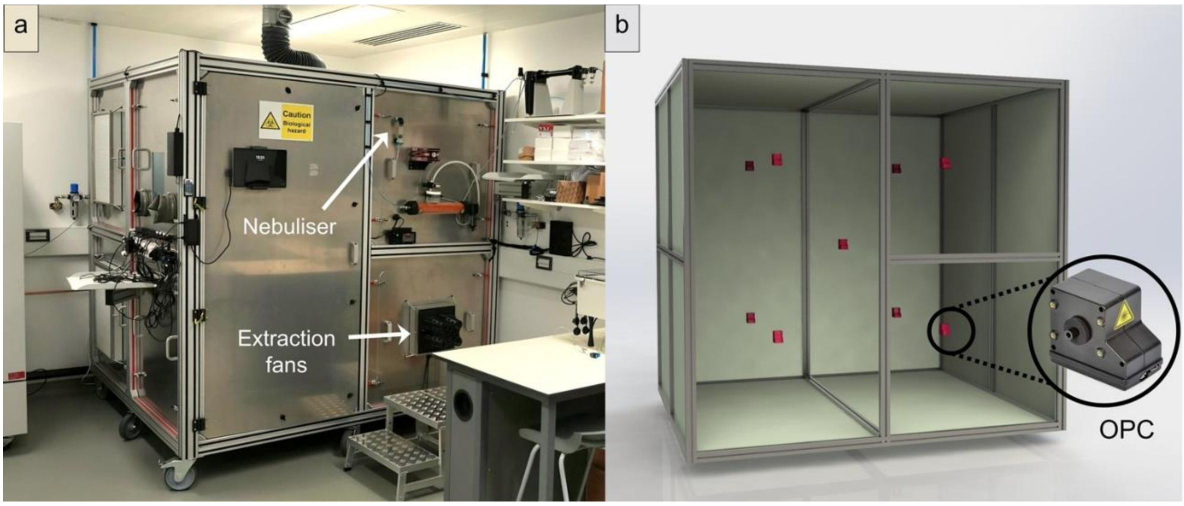



*Fig. 1. (a) Controlled 8m³ bioaerosol chamber. (b) Three-dimensional rendering of the sensor grid.*

## Results and discussion

The datasets acquired from the distributed OPC network were used to reconstruct three-dimensional, time-resolved concentration fields using natural neighbour interpolation and boundary extrapolation methods in MATLAB. The results reveal pronounced spatial and temporal variability in aerosol distribution. Fig. 2 presents slice views of a 10-minute mean particle number concentration for fan-off and fan-on conditions. Under the fan-off condition (Fig. 2a), aerosol transport is governed primarily by diffusion and gravitational settling. The reconstructed fields exhibit strong spatial gradients, with elevated concentrations persisting near the source region (y<100 cm). Stratification is evident along the Y-axis, with decreasing concentration as Y increases, indicating strong settling effects at low y values, i.e. close to the source. These observations highlight the persistence of non-uniform concentration fields over extended timescales, challenging the assumption of well-mixed conditions.

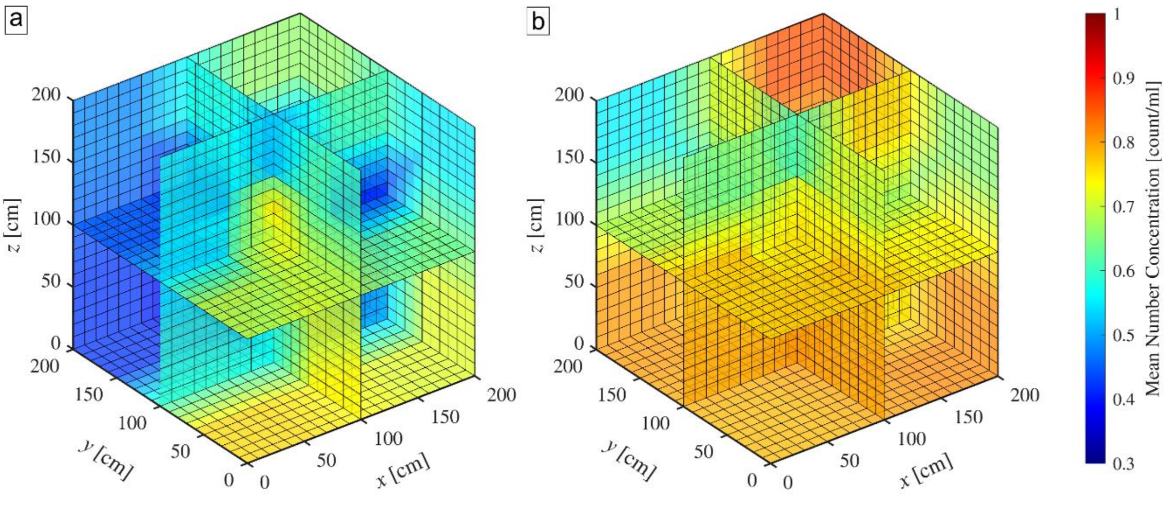



*Fig. 2. Slice views of the 10-minute mean particle number concentration field for (a) internal fan switched off and (b) internal fan switched on.*

In contrast, when internal mixing is introduced (fan-on condition), aerosol transport is significantly enhanced by forced convection (Fig. 2b). The concentration fields become more spatially uniform, with reduced gradients

https://doi.org/10.18745/pb.27307

indicating improved mixing throughout the chamber, especially for the z<100 cm region. This is corroborated by the coefficient of variation of the mean concentration across the chamber, which decreases from 0.21 (fan-off) to 0.09 (fan-on), confirming a noticeable reduction in spatial variability and enhanced homogeneity. Having said that, spatial asymmetry remains evident in regions where y>100 cm and z>100 cm, with higher concentrations observed for x>100 cm compared to x<100 cm. This behaviour is attributed to the off-centre location of the aerosol source (approximately at x=120 cm), imposed by chamber design constraints, in combination with the position of the oscillating fan (at x=100 cm). The interaction between source placement and induced airflow leads to preferential transport towards one side of the chamber, resulting in elevated concentrations in one region and corresponding depletion in the opposite region. It is also worth mentioning that the z<100 cm region reveals relatively good mixing, with more uniform concentration fields compared to z>100 cm region. This behaviour is linked to the combined effects of gravitational settling and the oscillatory motion of the internal fan, which promotes periodic sweeping and enhanced redistribution of particles in the lower region, reducing spatial variability over time.

These findings highlight the limitations of single-point measurements for representing chamber-wide conditions, particularly under weakly mixed regimes. They also underscore the importance of spatially resolved sensing approaches for accurately characterising aerosol transport and exposure, and for improving the interpretation of chamber-based experiments.

## Conclusion

This study demonstrates the capability of distributed low-cost sensor networks to resolve the spatiotemporal dynamics of aerosol transport within a controlled indoor environment. By reconstructing three-dimensional, time-resolved concentration fields, significant spatial heterogeneity was observed under fan-off conditions, where transport is governed by diffusion and gravitational settling. In contrast, the introduction of internal mixing enhanced transport through forced convection, leading to reduced spatial gradients and improved homogeneity, as quantified by a decrease in the coefficient of variation. Despite this, residual spatial asymmetry was identified, highlighting the influence of source placement and flow configuration on aerosol distribution. These findings emphasise that aerosol concentrations within enclosed environments cannot be assumed to be spatially uniform, particularly under weakly mixed conditions. Overall, the results demonstrate that spatially resolved measurements are essential for accurately characterising aerosol behaviour and can provide a robust experimental framework for improving the design, interpretation, and validation of indoor aerosol studies.

## Acknowledgements

This research was supported by the Biodetection Technologies Hub which is funded by Research England's Expanding Excellence in England (E3) fund.

## References

[1] A. Cincinelli and T. Martellini, "Indoor Air Quality and Health," *International Journal of Environmental Research and Public Health 2017, Vol. 14,* vol. 14, no. 11, Oct. 2017.
[2] A. C. Lewis *et al.*, "Indoor Air Quality," Nov. 2022.
[3] D. Licina, Y. Tian, and W. W. Nazaroff, "Inhalation intake fraction of particulate matter from localized indoor emissions," *Build. Environ.*, vol. 123, pp. 14–22, Oct. 2017.
[4] S. K. Friedlander, "Similarity considerations for the particle-size spectrum of a coagulating, sedimenting aerosol," *Journal of Atmospheric Sciences*, vol. 17, no. 5, pp. 479–483, 1960.
[5] J. Richmond-Bryant, A. D. Eisner, L. A. Brixey, and R. W. Wiener, "Transport of airborne particles within a room," *Indoor Air*, vol. 16, no. 1, pp. 48–55, 2006.
[6] L. R. Crilley *et al.*, "Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring," *Atmos. Meas. Tech.*, vol. 11, no. 2, pp. 709–720, Feb. 2018.

https://doi.org/10.18745/pb.27307

# Performance Evaluation of Bioaerosol Samplers on Capturing Fungal Spores in a Controlled Environment

**Shamil Rafeeq[1*], Loic Coudron[1], Jonathan S. West[2], Nikolay Dimov[1], Richard Kaye[1], Ian Munro[1], Ian Johnston[1] and Cristina Barrero Sicilia[3]**
[1]*Department of Engineering, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK.*
[2]*Rothamsted Research, Harpenden, Hertfordshire*
[3]*Department of Clinical, Pharmaceutical and Biological Sciences, School of Life and Medical Sciences, University of Hertfordshire, Hatfield, UK.*
*Corresponding author: s.rafeeq@herts.ac.uk

**The SKC BioSampler, electrostatic precipitator (ESP), Rotorod, and Burkard volumetric spore trap were evaluated in a controlled environment, representing devices developed for liquid-based bioaerosol collection, electrostatic precipitation-based capture, airborne pollen/fungal sampling respectively. The devices were compared for capturing efficiency of *Botrytis Cinerea* fungal spores. Burkard and Rotorod sampler achieved higher efficiency in fungal spore capture, while the electrostatic precipitator-electrowetting on dielectric (ESP-EWOD) system proved suitable for recovery of spores from the ESP collection slide to micro volumes of liquid.**

Keywords: Bioaerosol sampling; Chamber study; DNA recovery; Fungal Spores; Aerosolisation

## Introduction

Airborne fungal spores are major constituents of agricultural bioaerosol and are vital to spread diseases in crops [1]. Reliable samplers are required to capture these airborne fungal spores efficiently. The Rotorod and Burkard samplers are widely used in agritech and aerobiological monitoring. Whereas the electrostatic precipitator (ESP) sampler is suitable for automated sample recovery and the SKC Biosampler is liquid-based and often used in bioaerosol studies. A direct comparison therefore helps to assess their suitability for both conventional monitoring and integrated on chip electrowetting on dielectric (EWOD) detection workflows. The aerosolisation of larger particles such as fungal spores is harder to achieve due to size, hence a speaker and a rotating fan were utilised to aid in fungal spore aerosolisation [2]. This study compared the bioaerosol samplers for the collection performance of *Botrytis Cinerea* spores in a controlled environment.

## Experimental Setup

An 8 $m^3$ aerosol chamber (sealed, controlled test environment to generate and sample airborne particles under repeatable conditions) was used in the study on performance of bioaerosol samplers including SKC Biosampler® (SKC Inc., Eighty Four, PA, USA), ESP (UH developed) sampler, Burkard 7-day spore trap (Burkard Manufacturing Co. Ltd., Rickmansworth, UK) and Rotorod (Agri Samplers Ltd, UK) in collecting fungal spores. The spore releaser setup included a speaker (100 mm dia, 30W) mounted with a fully grown fungi agar plate without the lid. The speaker was connected to an amplifier and a function generator, and square waves with frequencies ranging from 20 and 80 Hz were used to trigger spore release. The aerosol chamber set up for fungal spores is illustrated in fig. 1.A rotating fan was used to aid in the release of fungal spores from the agar plate and to keep them airborne. The chamber is equipped with four mixing fans. Optical Particle Counters (OPCs) and an Aerodynamic Particle Sizer (APS) were used to record chamber particle concentration. The samplers, OPCs (AlphaSense N3), and APS (TSI 3320) were positioned at equal distances from the spore releaser and evenly spaced from one another. The sampler positions were rotated in trials to minimise the positional bias. Once the chamber achieved a suitable airborne concentration, the samplers were switched ON and run for 30 minutes. The chamber is equipped with an extractor fan and a HEPA filter, which was used to clean the air before and after the sampling. At least nine replicate trials were conducted for each sampler.

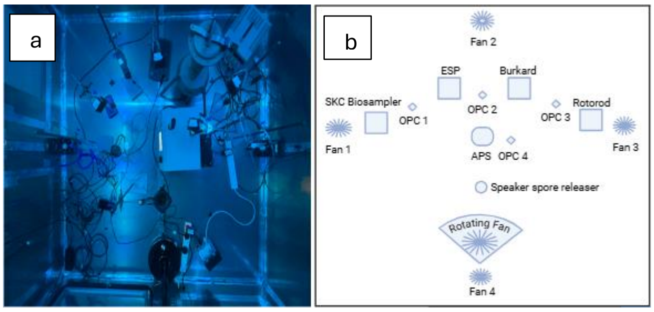



*Fig. 1. Experimental setup: (a) photograph of the aerosol chamber setup used for fungal spore aerosolization; (b) schematic representation of the setup.*

Collected spores were recovered from each sampler's collection medium or substrate and transferred to extraction tubes for DNA extraction. For the ESP sampler, recovery using a pipette, a swab, and an automated EWOD (Electrowetting on Dielectric) system [3] was used. The recovered samples were stored in a -20 ºC freezer until DNA was extracted using Quick-DNA Fungal/Bacterial Microprep Kit (Zymo Research). The extracted DNA was quantified using a qPCR assay with specific *Botrytis Cinerea* primers [4]. Suitable statistical tests were conducted using RStudio.

## Results and discussions

### Aerosolisation of fungal spores in an 8 m$^3$ test chamber

The speaker demonstrated potential for controlled aerosolisation of *Botrytis Cinerea* spores, as nebulisation of bigger particles, like fungal spores, is harder to achieve. Even though airborne concentration varied across the trials, the mean spore counts were broadly comparable between 4 OPCs and the APS (fig 2.). The coefficient of variation (CV) was similar across the particle counters, ranging from 41.7% to 43.3%. This indicates that the OPCs and APS showed comparable trial-to-trial variability in response to changes in airborne spore concentration. This supports the suitability of the aerosol chamber setup (arrangement of OPCs, APS, speaker and sampling devices) for comparative assessment, particularly as the devices operated simultaneously and rotated between positions across the trials.

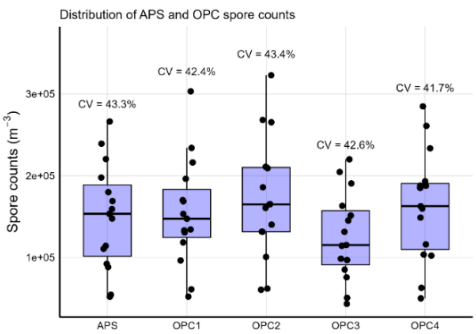



*Fig. 2. Distribution of spore counts in OPCs and APS.*

### DNA recovery by sampling devices

The DNA samples recovered from devices (ng per m$^3$) was tested with a linear mixed effects model, after accounting for trial-to-trial variation in airborne spore concentration, and showed that sampling devices had a significant effect on DNA concentration recovered, $F(5, 35.05) = 84.04$, $p < 0.001$.From the estimated marginal means, Rotorod had the highest recovered DNA amounts (fig 3), followed by the Burkard, while SKC Biosampler and ESP showed comparatively lower DNA collection. In addition, the difference in DNA recovery between Rotorod and Burkard was not statistically significant ($p$=0.160).

https://doi.org/10.18745/pb.27307

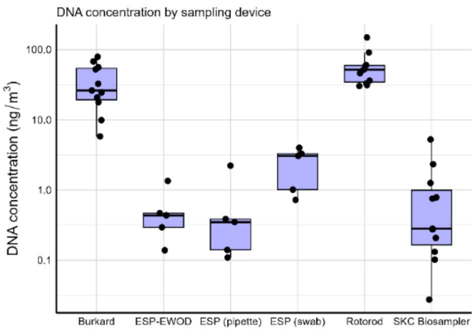



*Fig. 3. Botrytis cinerea DNA concentration (ng $m^{-3}$) recovered by each sampling device.*

The higher DNA recovery of the Rotorod and Burkard may be associated with fewer intermediate recovery steps of spores to DNA extraction tube. In contrast, the ESP and SKC Biosamplers require additional steps to recover spores to DNA extraction tube, which might have introduced additional sample losses [5]. The sample recovered from the ESP by cotton swab (ESP-swab) had a higher median DNA concentration. However, ESP (swab), ESP (pipette) and ESP-EWOD-based recovery showed no significant difference, indicating the ESP-EWOD system is comparable with manual methods and can be used to concentrate the collected spores onto 3-6µl droplet. This highlights its potential for integration with on-chip DNA extraction, a key future objective of the project.

## Conclusion

The Burkard and Rotorod collected higher amounts of DNA compared with both the ESP and SKC Biosampler. In addition, the fewer recovery steps and higher flowrate might have helped them achieve a higher capture efficiency than other samplers. The collection performance also depends on the spore size [6]. Future work will include the use of *Calvatia gigantea* spores, 3.1µm and 1.1µm polystyrene microspheres[7]. *Calvatia gigantea* spore recovery will be assessed by qPCR assay and the collection efficiency of polystyrene microspheres will be evaluated by counts. This will extend the comparison enabling sampler performance to be assessed across different particle sizes and characteristics. ESP (pipette), ESP (swab) and ESP-EWOD based spore recovery were generally comparable. An additional merit of ESP-EWOD-based recovery is its ability to highly concentrate spores into a microscale droplet, volume-matched for on-chip DNA detection within an automated EWOD system. This work is supported by the Biodetection Technologies Hub funded by Research England.

## References

[1] D. S. Meredith, "Significance of Spore Release and Dispersal Mechanisms in Plant Disease Epidemiology," *Annu. Rev. Phytopathol.*, vol. 11, no. 1, pp. 313–342, Sep. 1973, doi: 10.1146/annurev.py.11.090173.001525.

[2] E. Scheermeyer and I. E. Agranovski, "Design and evaluation of a new device for fungal spore aerosolization for laboratory applications," *J. Aerosol Sci.*, vol. 40, no. 10, pp. 879–889, 2009, doi: 10.1016/j.jaerosci.2009.06.003.

[3] T. G. Foat *et al.*, "A prototype personal aerosol sampler based on electrostatic precipitation and electrowetting-on-dielectric actuation of droplets," *J. Aerosol Sci.*, vol. 95, pp. 43–53, May 2016, doi: 10.1016/j.jaerosci.2016.01.007.

[4] C. F. Diguta *et al.*, "Development of a qPCR assay for specific quantification of Botrytis cinerea on grapes," Dec. 2010. doi: 10.1111/j.1574-6968.2010.02127.x.

[5] G. Mainelis, "Bioaerosol sampling: Classical approaches, advances, and perspectives," May 03, 2020, *Taylor and Francis Inc.* doi: 10.1080/02786826.2019.1671950.

[6] S. A. Grinshpun *et al.*, "Collection of airborne spores by circular single-stage impactors with small jet-to-plate distance," *Aerosol Science*, vol. 36, pp. 575–591, 2005, doi: 10.1016/j.aerosci.2004.06.078.

[7] A. Shirkhani *et al.*, "Development of a platform for in-situ airborne virus detection by interfacing RT-LAMP assay with a condensational growth collector," *Aerosol Science and Technology*, 2025, doi: 10.1080/02786826.2025.2568690.

https://doi.org/10.18745/pb.27307

## 4. Modelling the Universe: Astrophysics, Simulations, and Observational Systems

Chair: Dr Sophie Koudmani

Judges: Dr Krupa Kuruvila, Dr Ragesh Thelakkadan Puthiyaveettil

https://doi.org/10.18745/pb.27307

## Invited Speaker - Dr Lily Whitler

Bio: Dr Lily Whitler is a Gavin Boyle/Kavli Institute Fellow at the Kavli Institute for Cosmology, University of Cambridge, and a Fellow of Selwyn College. She is an observational astrophysicist, and her research focuses on understanding how the earliest generations of galaxies formed and evolved, and how these early galaxies impacted the overall evolution of the Universe

**Title: New Eyes on the Sky: Insights into the Early Universe from the James Webb Space Telescope**

The James Webb Space Telescope (JWST), NASA's current flagship observatory, was launched in December 2021 and has revolutionised the field of extragalactic astronomy in only a few short years. JWST's unprecedented sensitivity and infrared observational capabilities have opened a new window into the earliest luminous sources that existed in the Universe – stars, active black holes, and the galaxies they live in– which were completely impossible to observe with previous telescopes. JWST observations have revealed that the early Universe was unexpectedly rich with activity, with a large number of galaxies formed only a few hundred million years after the Big Bang, surprisingly massive black holes, and huge amounts of star formation. In this talk, I will discuss some of the surprising findings from JWST and their implications for our understanding of the Universe.

https://doi.org/10.18745/pb.27307

## Invited Speaker - Dr Will Coulton

Bio: Dr Will Coulton is an Ernest Rutherford Fellow in the astrophysics subdepartment at the University of Oxford. His research focuses on using novel analysis techniques to explore the physics of both the early Universe and the astrophysics governing the evolution of galaxy groups and clusters. He is very excited to be working with data from state-of-the-art telescopes, such as the Simons Observatory.

**Title: Glimpses into the Universe's youth: a search for new information on the primordial Universe.**

What was the Universe like in its first instances? What processes shaped it? These are some of the key questions motivating cosmologists working today. To answer these questions, cosmologists analyse the distribution of galaxies as observed with our ever-increasingly precise telescopes and attempt to tease out new information on the primordial Universe. The challenge is that these signals are buried under the complex physics that governs the formation and evolution of galaxies - an as-of-yet poorly understood problem. In this talk, I will discuss how to address this problem by identifying signals that can only be produced by primordial physics and thereby allowing new and unobscured views into the early Universe.

https://doi.org/10.18745/pb.27307

# From stardust to us: Predicting planet habitability in the Universe from computational simulations of galaxy chemical evolution

**Rob Yates**[1*]
[1]*University of Hertfordshire*
*Corresponding author: r.yates3@herts.ac.uk

**One of the most fundamental questions we can ask is "when and where does life form in the Universe?". However, relatively little research has been done on tackling this question from a theoretical perspective. At UH, we are pioneering the use of computational simulations to address this, by modelling the chemical elements, dust, and molecules in galaxies – the building blocks of life. In this talk, I will present the current state of this burgeoning field and signpost our ambitious plans to link simulations using Bayesian Inference to coherently track these building blocks and habitable planets across multiple scales.**

Keywords: astrophysics; computational simulations; galaxies; astrochemistry; habitability

## Introduction

Computational simulations of galaxy evolution are now able to confidently predict a wide range of chemical element abundances both in nearby galaxies [1-3] and back to high redshift (z) [4,5]. These predictions are in promising agreement with the latest precise observations at $z \sim 0-4$ [6,7] and at $z \sim 7-8$ [8,9]. Similar successes have also been found with dust modelling [5,10]. However, higher-order galaxy chemical evolution (GCE) phenomena such as planet formation and habitability remain largely unprobed from a theoretical perspective.

At the Centre for Astrophysics Research (CAR) in the University of Hertfordshire, we have been pioneering the use of computational simulations to address this for many years [5,11,12], by modelling the formation of chemical elements, dust, and molecules – the building blocks of life – in galaxies. The natural next step is to extend these GCE studies to planet formation and habitability. However, to do this we require simulations that are able to coherently span the huge range of scales that GCE physics operates on. Therefore, in a major STFC-funded project that is due to start later this year, we will be harnessing the power of a synchronised multi-scale simulation suite to predict GCE and habitability in our Milky Way and other galaxies with unrivalled detail and statistical significance.

## Experimental/Simulation

The bulk of the research to be presented has been carried out using the L-Galaxies simulation [5]. L-Galaxies is a cosmological-scale semi-analytic simulation containing millions of galaxies, each radially resolved down to a minimum of 30 pc (an average of ~835 pc within **the solar radius**). L-Galaxies runs x10,000 quicker than full hydrodynamical simulations of a similar scale, allowing it to (a) run on very large volumes ($0.5-1$ Gpc)$^3$ providing galaxy samples of unrivaled size and diversity, and (b) incorporate a wide range of physics, including a cutting-edge GCE model containing individual chemical elements, dust, and (soon) complex molecules [1,5].

We also plan to use the super-high-resolution EDGE hydrodynamical simulation [13] to study GCE on the smallest galactic scales. EDGE models dwarf galaxies at resolutions down to $\sim 20\ M_{sun}$ and $\sim 3$ pc, allowing it to resolve individual massive stars and SNe [14]. This makes EDGE ideal for tracing small-scale GCE effects [12]. It is also run from cosmological initial conditions, providing realistic merger histories that are lacking in isolated galaxy simulations.

Regarding planet habitability in galaxies, traditionally, 'empirical models' have been used [15,16]. These models rely on scaling relations from observations to infer the link between generic galaxy properties and planet formation. This means empirical models cannot provide us with detailed insights into the physics driving these processes, and only apply to 'typical galaxies' which follow the average observational scaling relations. Therefore, we plan to use L-Galaxies and EDGE in a synchronized way to help model GCE and planet formation in a coherent and physically motivated way.

## Results and discussion

The L-Galaxies has been successful in helping to establish the prevailing GCE paradigm in the field. For example, it has demonstrated that the low metallicities observed in low-mass star-forming galaxies are only reproducible by simulations with very metal-rich galactic winds [3]. This also helps explain the high metal abundances observed in the intracluster medium outside of galaxies [2].

The GCE model in L-Galaxies has also been applied to other galaxy types, from dwarf galaxies to massive ellipticals [1,4], showing that the chemistry in these diverse systems can be simultaneously explained by a consistent formalism. I will present how a recent extension of this modelling which includes binary stars and dust [5] – a first for cosmological-scale simulations – has successfully pushed this paradigm back to higher redshift and more detailed chemistry (see Fig. 1).

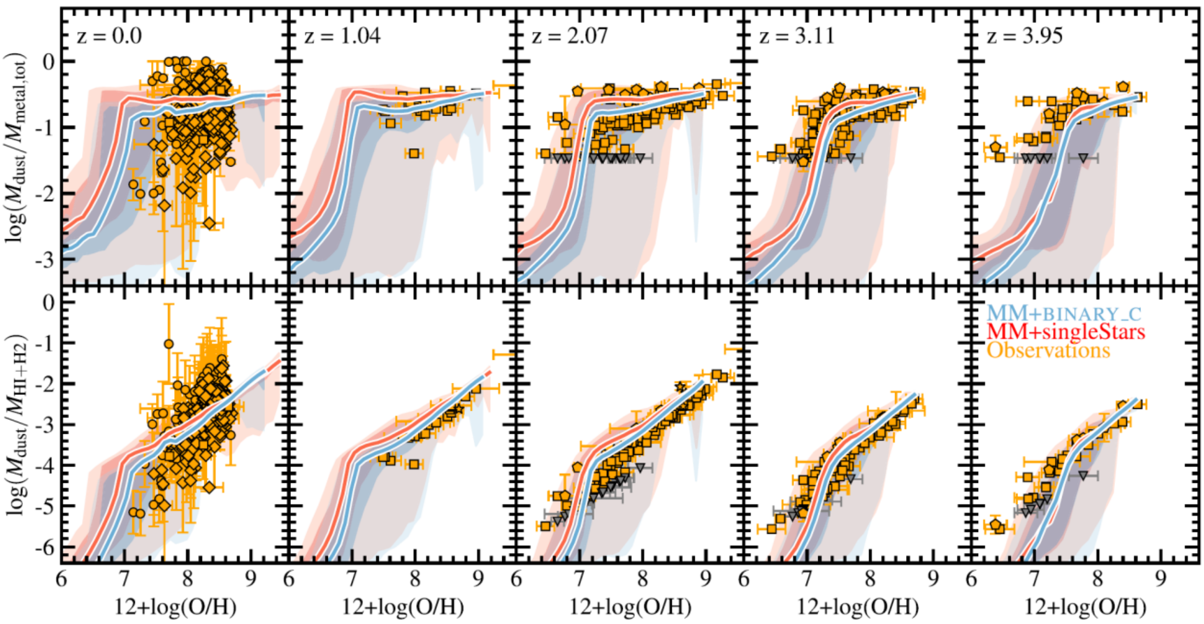



*Fig. 1. The dust-to-metal ratios (top row) and dust-to-gas ratios (bottom row) for galaxies from two versions of the L-Galaxies simulation (with binary stars [blue] and without binary stars [red]). Both versions of the model are in good agreement with a collection of observational data (orange symbols) back to z ~4. Figure adapted from Yates et al. (2024).*

The natural next step in this methodology is to include molecules and planets into the mix. Therefore, I will present on-going work by my team on incorporating complex molecule formation into L-Galaxies and EDGE, and subsequent plans to include a planet formation model in post processing which will use the detailed chemical information generated by these simulations to predict the number and composition of planets within galaxies.

In order for this to work coherently across the two simulations, we need to synchronise them by tuning the free parameters in L-Galaxies to replicate the physics modelled at high resolution in EDGE. I will demonstrate how we plan to do this using direct gradient-based Bayesian Inference (a statistical method similar to some machine learning techniques), to create a novel simulation suite called L-EDGE (see Fig. 2).

https://doi.org/10.18745/pb.27307

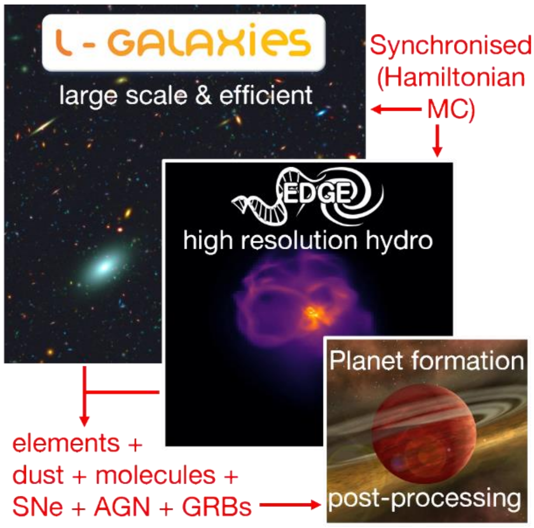



*Fig. 2. A schematic illustrating the L-EDGE simulation suite, combining the L-Galaxies and EDGE simulations, along with a planet formation model applied in post-processing.*

## Conclusion

I will present some of the world-leading research being carried out at UH on galaxy chemical evolution (GCE) and habitability using computational simulations of galaxy evolution. The results we've obtained thus far have helped shape the prevailing GCE paradigm in this field, and our planned future work will allow us to understand the frequency and location of habitable planets within our galaxy and across the cosmos better than ever before.

## References

[1] Yates R. M., Henriques B., Thomas P. A., Kauffmann G., Johansson J., White S. D. M., 2013, MNRAS, 435, 3500
[2] Yates R. M., Thomas P. A., Henriques B. M. B., 2017, MNRAS, 464, 3169
[3] Yates R. M., Henriques B. M. B., Fu J., Kauffmann G., Thomas P. A., Guo Q., White S. D. M., Schady P., 2021a, MNRAS, 503, 4474
[4] Yates R. M., Péroux C., Nelson D., 2021b, MNRAS, 508, 3535
[5] Yates R. M., Hendriks D., Vijayan A. P., Izzard R. G., Thomas P. A., Das P., 2024, MNRAS, 527, 6292
[6] Yates R. M., Schady P., Chen T. W., Schweyer T., Wiseman P., 2020, A&A, 634, A107
[7] Sanders R. L., et al., 2021, ApJ, 914, 19
[8] Jones T., Sanders R., Roberts-Borsani G., Ellis R. S., Laporte N., Treu T., Harikane Y., 2020, ApJ, 903, 150
[9] Curti M., et al., 2023, MNRAS, 518, 425
[10] Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
[11] Kobayashi C., Karakas A. I., Lugaro M., 2020, ApJ, 900, 179
[12] Rey M., et al., 2025, MNRAS, 541, 1195
[13] Agertz O., et al., 2020, MNRAS, 491, 1656
[14] Andersson E. P., Rey M. P., Pontzen A., Cadiou C., Agertz O., Read J. I., Martin N. F., 2025, ApJ, 978, 129
[15] Behroozi P., Peeples M. S., 2015, MNRAS, 454, 1811
[16] Dayal P., Ward M., Cockell C., 2016, arXiv e-prints, p. arXiv:1606.09224

https://doi.org/10.18745/pb.27307

# Modelling Extreme Nitrogen Emitting Galaxies from Cosmic Dawn to the Present

**Alessandro Angrilli Muglia[1,*], Chiaki Kobayashi[1], and Souradeep Bhattacharya[1]**
[1]*University of Hertfordshire, Centre for Astrophysics Research*
*Corresponding author: a.angrilli-muglia@herts.ac.uk

**We study how certain chemical elements evolve inside galaxies over cosmic time using numerical simulations of dwarf galaxies. In particular, we focus on nitrogen, which in some very distant galaxies appears unusually abundant compared to oxygen. We show that these extreme chemical patterns can emerge naturally from standard physical processes when the evolution of gas, stars, and feedback is properly modelled, without requiring rare or exotic explanations.**

Keywords: galaxy formation; hydrodynamical simulations; chemical evolution; high-redshift galaxies; computational astrophysics

## Introduction

Galaxies are made of gas, stars, and elements produced inside stars over time. By measuring the relative amounts of different elements, scientists can reconstruct how galaxies formed and evolved. Surprisingly, some of the most distant (and therefore earliest) galaxies show unusually high amounts of nitrogen compared to oxygen, which is difficult to explain with standard models of chemical evolution.

Observations of extremely distant systems suggest that this nitrogen enrichment appears very early in the history of the Universe. For example, the galaxy GN-z11 shows unusually strong nitrogen signatures, and proposed explanations include very rapid star formation in dense environments or rare energetic events [1]. Similar chemical patterns have also been observed in more nearby galaxies, suggesting that this is not unique to the earliest cosmic times [2].

From a modelling perspective, it is still unclear how such high nitrogen levels can develop so quickly. Some studies suggest that bursts of star formation can rapidly change the chemical composition of galaxies [3], while large-scale simulations show that gas flows, stellar feedback, and star formation history all play important roles in shaping elemental abundances [4]. In this work, we use numerical simulations to test whether these standard physical processes are sufficient to explain the observations, or whether additional mechanisms are needed.

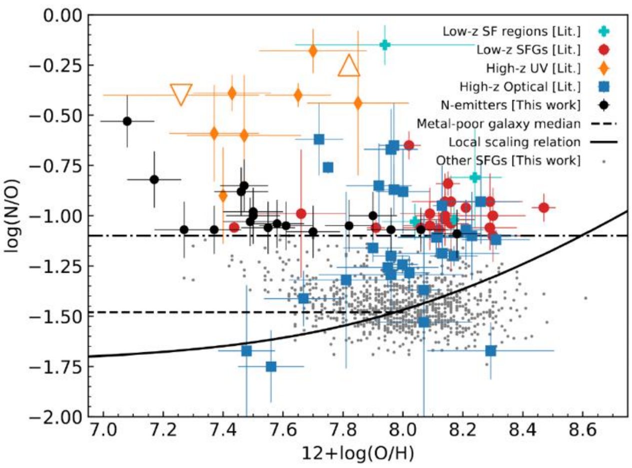



*Fig. 1. Adapted Figure from [2], the plot shows some of the latest observational measurements of extreme nitrogen-emitting galaxies compared to the local nitrogen-to-oxygen scaling relation (solid line). The highlighted systems lie significantly above the local relation, indicating enhanced nitrogen abundance relative to expectations from nearby galaxies.*

https://doi.org/10.18745/pb.27307

## Experimental/Simulation

We perform hydrodynamical simulations of dwarf galaxies in order to follow the coupled evolution of gas, stars, and chemical enrichment over cosmic time. Each simulation consists of gas particles with masses of $6250\ M_{\odot}$, embedded within a fixed dark matter potential and evolved self-consistently under gravity and hydrodynamics. The equations of motion, thermodynamics, and entropy are solved for each particle using adaptive timesteps.

After a galaxy is free to evolve, and under physically motivated star formation criteria, gas particles are converted into star particles representing stellar populations with shared ages and chemical properties. As stars evolve, they return newly synthesised elements to the surrounding interstellar medium through stellar feedback processes, enriching nearby gas particles with metals such as nitrogen and oxygen.

Our simulations are based on and calibrated against the framework developed by Kobayashi et al. [4] in earlier chemical enrichment studies, while continuously incorporating updated stellar yields and feedback prescriptions as our understanding of galaxy evolution improves (for the latest review see [5]). In this work, we specifically track the evolution of nitrogen and oxygen abundances across different evolutionary stages of the simulated galaxies.

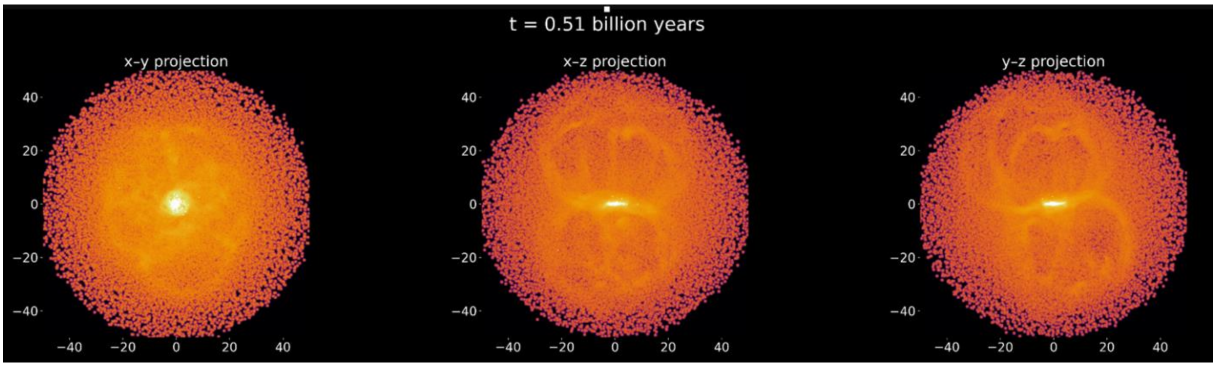



*Fig. 2. Snapshot of a simulated galaxy from our model, showing the spatial distribution of gas and its evolution during galaxy formation.*

## Results and discussion

To compare our simulations with observations, we compute star-formation-rate-weighted averages of the nitrogen and oxygen abundances in the gas phase. This approach reflects the fact that observations of distant galaxies primarily probe gas ionised by young, massive stars responsible for strong nebular emission lines.

We evaluate these abundances over time intervals of 10 Myr and 50 Myr, corresponding to characteristic timescales associated with observable star-forming regions. Our simulations show that all galaxies spend at least two-thirds of their lifetimes evolving close to the local nitrogen-to-oxygen scaling relation shown in Fig. 1. This indicates that the observed relation represents the long-term equilibrium state of chemical evolution in dwarf galaxies.

However, during specific evolutionary phases, galaxies can temporarily exhibit significant nitrogen overabundances or underabundances. We find that the systems displaying the strongest nitrogen enhancement are typically those with the largest standard deviation in star formation rate, corresponding to highly bursty star formation histories. These bursts rapidly alter the balance between metal production, gas dilution, and feedback, producing short-lived but extreme departures from the local relation.

The galaxies exhibiting the largest scatter are also characterised by strong galactic outflows combined with inflows of pristine gas. In our simulations, dwarf galaxies lose more than 60% of their metals through feedback-driven outflows over their lifetimes. Because of their shallow gravitational potentials, dwarf galaxies are especially efficient at ejecting enriched gas, making them ideal environments for generating extreme abundance fluctuations.

This behaviour motivates the interpretation of dwarf galaxies as an amplified representation of the physical processes that may also occur in more massive systems. Although massive galaxies generally retain a larger fraction of their metals, sufficiently strong outflow episodes can still temporarily drive them into nitrogen-enhanced phases similar to those observed in extreme nitrogen emitters.

Importantly, while nitrogen-enhanced phases are observable due to their strong emission features, the simulations predict the existence of equally extreme nitrogen-poor phases that remain largely undetectable because they produce weak or absent emission lines. Our results therefore suggest that currently observed extreme nitrogen emitters may represent only the visible half of a broader population of chemically fluctuating galaxies.

## Conclusion

Our results show that unusually high nitrogen-to-oxygen ratios observed in distant galaxies can be explained using standard physical processes when galaxy evolution is modelled in sufficient detail. This includes accurately tracking gas dynamics, star formation, and observational biases. We conclude that extreme nitrogen enrichment does not necessarily require exotic physics, but could naturally arise from the complex internal evolution of galaxies over time.

## References

[1] A. Cameron, H. Katz, M. P. Rey, and A. Saxena, “Nitrogen enhancements 440 Myr after the Big Bang: super-solar N/O, a tidal disruption event or a dense stellar cluster in GN-z11?,” *Monthly Notices of the Royal Astronomical Society*, vol. 523, no. 3, pp. 3516–3525, May 2023, doi: https://doi.org/10.1093/mnras/stad1579.

[2] S. Bhattacharya and C. Kobayashi, “The origin of extreme N-emitters in star-forming galaxies at $z < 0.5$ with DESI DR1,” *Monthly Notices of the Royal Astronomical Society*, vol. 548, no. 1, Mar. 2026, doi: https://doi.org/10.1093/mnras/stag545.

[3] C. Kobayashi and A. Ferrara, “Rapid Chemical Enrichment by Intermittent Star Formation in GN-z11,” *The Astrophysical Journal Letters*, vol. 962, no. 1, pp. L6–L6, Feb. 2024, doi: https://doi.org/10.3847/2041-8213/ad1de1.

[4] C. Kobayashi, Volker Springel, and Simon, “Simulations of Cosmic Chemical Enrichment,” *Monthly Notices of the Royal Astronomical Society*, vol. 376, no. 4, pp. 1465–1479, Mar. 2007, doi: https://doi.org/10.1111/j.1365-2966.2007.11555.x.

[5] C. Kobayashi, “Nucleosynthesis and the chemical enrichment of galaxies,” *arXiv.org*, 2025. http://arxiv.org/abs/2506.20436.

https://doi.org/10.18745/pb.27307

## 5. Interactive Intelligent Systems: Robotics, XR, and Adaptive Control

Chair: Prof. Salvatore Livatino

Judges: Dr Javed Khan, Anna Ojinnaka

https://doi.org/10.18745/pb.27307

## Invited Speaker – Dr Caroline Whyatt

Bio: Dr Caroline Whyatt is a senior lecturer in Psychology at the University of Hertfordshire. Her research interests focus on perception-action coupling: the relationship between sensory information and movement. She is particularly interested in how this process develops, adapts and deteriorates across the lifespan, and the role of this within clinical populations. Prior to joining the University of Hertfordshire, she held roles at Queen's University Belfast and Rutgers University, New Jersey, investigating the perception-action process in Autistic individuals and people with Parkinson's Disease. Using technology, such as motion capture and virtual reality, this work explored the role of perception-action coupling in applications such as diagnostics, intervention and rehabilitation.

**Title: From virtual worlds to real-world behaviours: Using VR to study Perception, Action and Decision Making.**

Advances in technology, including virtual reality, motion-capture and eye-tracking are transforming the research questions that psychologists can ask, and the precision with which they can answer them. The presentation explores how immersive technologies can be used to probe perception-action coupling: the recursive process of sensory information guiding action and behaviour. Ongoing research using a distance replication paradigm is presented to demonstrate how this technology can be harnessed to manipulate levels of sensory input and quantify how this systematically alters behaviour and judgements. How these technologies provide new levels of insight into the dynamic interaction of perception and action is discussed, moving beyond simple behavioural observations to examine the processes that generate them. Future applications are considered, including the study of affordance-based decision making, sensory integration in clinical populations, and the broader challenge of understanding how humans construct internal models of themselves and their environment to guide real-world behaviours.

https://doi.org/10.18745/pb.27307

# Attention-Based Soft Supervisory for Controlling Multiple Systems

**Klinsmann Agyei, Pouria Sarhadi**

School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK

Corresponding author: k.agyei@herts.ac.uk

## Abstract

**We present a learning-based approach that recognises dynamical systems in real time by observing closed-loop state and control-action trajectories. We train a neural classifier on trajectories generated by five diverse systems spanning linear (stable differential drive, non-minimum phase, Boeing 747 model), nonlinear (Nomoto ship model), and unstable dynamics. The architecture combines state and action encoders with LSTM-based temporal processing and an attention mechanism, trained on expert trajectories produced by an LQI controller tracking step references. Simulation results under previously unseen conditions demonstrate classification accuracy across all five systems. These results could enable label-free generalist control, allowing the controller to autonomously identify the system under control and eliminating the need for external labelling, thereby supporting fully autonomous operation across heterogeneous platforms. The approach can also be employed as a soft supervisory algorithm in classical switching controllers.**

**Keywords —** *Generalist control, system identification, AI, adaptive control, soft switching*

## I. Introduction

In recent years, the computer science and robotics communities have introduced so-called *generalist controllers*: single models trained to perform multiple tasks or operate across diverse environments within a single training framework. Notable examples include DeepMind's Gato [4], Octo [5], and NVIDIA's GR00T N1 [2]. These systems, typically based on transformer architectures, are trained via imitation learning and behavioural cloning on large expert datasets. Classical control has long pursued similar ambitions. Adaptive and robust control frameworks (more specifically, simultaneous stabilisation theory [6]) aimed to design single-entity controllers capable of handling uncertainty or multiple plants. However, despite their success and sustained interest, these approaches remain constrained by restrictive structural assumptions. Consequently, most controllers remain tailored to a single plant or a narrow class of dynamics. Motivated by advances in attention-based networks, we recently demonstrated a single AI controller trained on Linear Quadratic Integral (LQI) expert data that successfully controls 25 heterogeneous systems (linear and nonlinear, stable and unstable, minimum and non-minimum phase) across standard control engineering benchmarks [1]. An outstanding limitation is the need for a system label before operation. Removing this requirement would bring the controller closer to true generalisability.

In this paper, we propose an attention-based AI algorithm as a system recognition network. The framework can also be viewed as a soft supervisory switching layer [3], extending attention-based generalist control toward label-free operation across heterogeneous systems.

## II. Problem Formulation

Consider $N$ dynamical systems $\{\mathcal{S}_i\}$:

$$\dot{x}_{pi} = f_i(x_{pi}, u_{pi}), \qquad y_{pi} = h_i(x_{pi})$$

where $x_{pi} \in \mathbb{R}^{n_i}$, $u_{pi} \in \mathbb{R}$, and $y_{pi} \in \mathbb{R}$. The systems may differ in order, stability, phase characteristics and linearity. We propose a learning-based recognition network that recognises the active system from observed trajectories, enabling autonomous supervisory selection for generalist [1] and switching controllers. We demonstrate the results on $N = 5$ Single input single output systems:

**S1 — Differential Drive:** A stable first-order system representing simplified differential-drive vehicle dynamics:

$$\dot{x}_p = \begin{bmatrix} 0 & 1 \\ 0 & -0.5 \end{bmatrix} x_p + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_p, \qquad y_p = \begin{bmatrix} 1 & 0 \end{bmatrix} x_p$$

https://doi.org/10.18745/pb.27307

**S2 — CSTR Chemical Process (Non-Minimum Phase):** A second-order system exhibiting inverse response behaviour:

$$G_p(s) = \frac{-1.135s + 3.119}{s^2 + 4.719s + 5.47}$$

**S3 — Unstable System:** A second-order unstable system:

$$G_p(s) = \frac{1}{s(s - 1)}$$

**S4 — Boeing 747 Short Period:** The longitudinal short-period dynamics of a Boeing 747 aircraft:

$$\dot{x}_p = \begin{bmatrix} -0.32 & 0.86 \\ -0.93 & -0.43 \end{bmatrix} x_p + \begin{bmatrix} -0.02 \\ -1.16 \end{bmatrix} u_p, \qquad y_p = [1 \quad 0] x_p$$

**S5 — Nonlinear Nomoto Ship Model:** A first-order nonlinear model describing ship yaw dynamics:

$$T\dot{r} + a_1 r + a_2 r^3 = Ku, \qquad \dot{\psi} = r, \qquad y_p = \psi$$

## III. The Proposed Approach

The proposed recogniser builds on the generalist controller architecture of [1]. The overall architecture is illustrated in Fig. 1. Two parallel state encoders map the masked state history (where a validity mask $M \in \{0,1\}^{T_h \times n_{max}}$ handles variable-order systems) to latent representations via feedforward networks. A separate action encoder processes the control history; all outputs are concatenated. Two parallel LSTMs process the features at different timescales, followed by a multi-head attention where each head computes:

$$a^{(h)} = softmax\left(\frac{Q^{(h)} K^{(h)^{\top}}}{\sqrt{d_k}}\right) V^{(h)}$$

with $Q^{(h)}$, $K^{(h)}$, $V^{(h)}$ being learned query, key and value projections and $d_k$ the dimension per head. The classifier head maps the attended features to a probability distribution over the N system classes via softmax. The predicted identity is $\hat{S} = \mathrm{argmax}_i\, p(S_i)$ and confidence is $\max_i p(S_i)$.

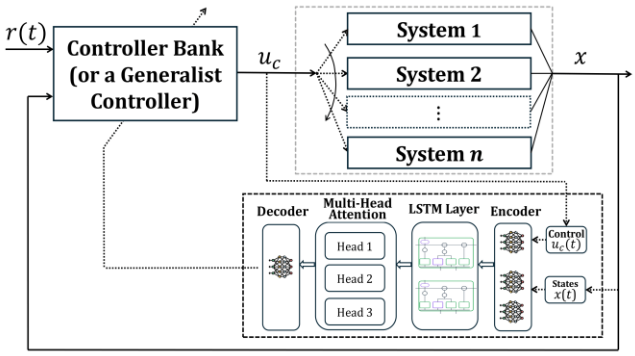



Fig. 1. Architecture of the proposed algorithm.

## V. Simulation Results

Fig. 2 presents the abrupt switching test with a step reference r = 1.0 and measurement noise = 0.01, with noise not present during training. The classifier correctly identifies all five system transitions, maintaining high confidence throughout, confirming that the network extracts dynamical features resilient to measurement noise. Fig. 3 presents switching under a sinusoidal reference at f = 0.2 Hz and measurement noise = 0.01, for S5, S3 and S2, a signal type never encountered during training.

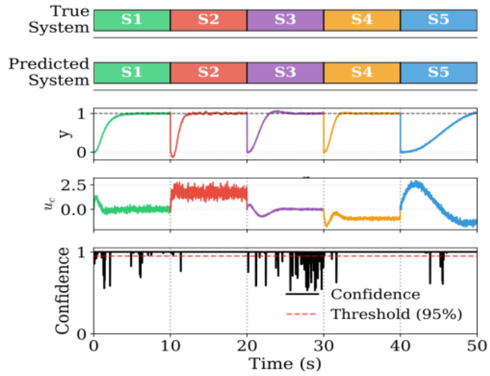



Fig. 2. System switching test with step reference r = 1.0 and measurement noise = 0.01

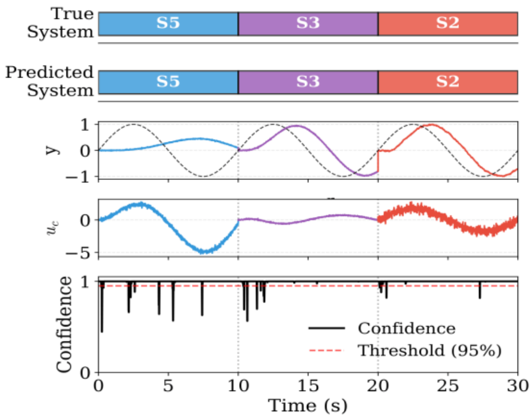



Fig. 3. System switching test with sinusoidal reference at f = 0.2 Hz and measurement noise = 0.01.

## VI. Conclusion

This paper presented an attention-based system recognition network that identifies the active dynamical system from state and control-action trajectories, eliminating manual system labelling in generalist and switching control architectures. Future work includes integrating the recogniser with the generalist controller as a unified end-to-end architecture, and extending beyond five systems to test scalability.

### References

[1] Agyei, Klinsmann and Sarhadi, Pouria. "Generalist AI control: Towards Multi-Purpose Adaptive Algorithms." *Control Engineering Practice* (2026): 106915.

[2] Bjorck, Johan and others. "GR00T N1: An Open Foundation Model for Generalist Humanoid Robots." arXiv preprint arXiv:2503.14734, 2025.

[3] Morse, A Stephen. "Supervisory Control of Families of Linear Set-Point Controllers." *IEEE Transactions on Automatic Control* (2002).

[4] Reed, Scott and others. "A generalist agent." *Transactions on Machine Learning Research* (2022).

[5] Team, Octo Model and others. "Octo: An Open-Source Generalist Robot Policy." *Robotics: Science and Systems (RSS) conference*. 2024.

[6] Vidyasagar, Mathukumalli. "Control Systems Synthesis: A Factorization Approach, Part II." *Springer Nature* (2022).

https://doi.org/10.18745/pb.27307

# Manipulating Perceived Agency in Human-Robot Interaction: A Dialogue and Behaviour Design Framework for Pepper and Fetch Robots

**Ali Fallahi[1*], Patrick Holthaus[1], Farshid Amirabdollahian[1], and Gabriella Lakatos[1]**

*[1]Robotics Research Group, University of Hertfordshire, Hatfield, United Kingdom*

Corresponding author: a.fallahi@herts.ac.uk

**Understanding how perceived agency shapes trust in robots is a central challenge in human-robot interaction (HRI). Despite growing interest, no established framework exists for systematically manipulating perceived agency in live interactions. This paper presents a structured dialogue and non-verbal behaviour design framework for two physically distinct robots, Pepper and Fetch that operationalises five theoretically grounded agency factors: Goal Directness, Social Awareness, Accountability, Shared Decision-Making, and Embodiment. The framework underpins an ongoing mixed-design experiment comparing High and Low Agency conditions in a between-subjects design, while examining embodiment within subjects across three scenarios, with perceived agency and trust as primary outcomes.**

Keywords: perceived agency; human-robot interaction; trust; embodiment; social robots

## Introduction

Trust is a key determinant of whether people accept and collaborate with robots [1]. An important factor shaping that trust is perceived agency, the extent to which a robot's actions are seen as driven by its own internal intentions rather than external commands [2]. Despite growing theoretical interest, no established method exists for systematically manipulating perceived agency in live HRI. Our prior work forms the empirical foundation for the present study. Fallahi et al. [4] showed that autonomy framing, describing the robot as autonomous versus remotely controlled, increased sincerity ratings but had limited impact on other trust dimensions or behavioural compliance, with pre-existing negative attitudes toward robots proving stronger predictors of trust outcomes. Building on this, Fallahi et al. [5] argued that perceived agency is the key mediator between robot autonomy and user trust, and proposed a behaviour-level manipulation as more effective than framing alone. The present paper directly realises that proposal through scripted verbal dialogue and non-verbal behaviour across two physically distinct robots.

## Background and Related Work

Perceived agency is defined as the extent to which an outside observer assumes an entity's actions are driven by its own internal thoughts and feelings rather than by external forces [2]. This definition underpins the Perceived Agency (PA) scale a psychometrically validated, Rasch-derived instrument covering thoughts (goal directness, decision-making), feelings (social awareness), and environmental responsiveness (accountability, embodiment) [2, 3] which outperforms existing measures across a wide range of entities [3]. The literature distinguishes autonomy a technical property of the robot from agency, which is a human-attributed quality: the sense that the robot has a mind of its own [6, 7]. Social agency theory argues that users infer agency from interactivity, adaptability, and autonomous behaviour [8], and research consistently shows that more agentic robots elicit greater trust [1], moral concern [2], and positive social engagement [8]. Key cues include goal-directed behaviour, emotional responsiveness, and transparent decision-making [9, 10]. Embodiment further moderates these effects: Pepper's screen, expressive LEDs, and arm gestures, and Fetch's physical grasping capability, are not merely hardware differences but theoretically meaningful sources of agency cues [13]. Research on perceived agency and trust across different age groups has also highlighted that individual differences including pre-existing attitudes toward robots significantly shape how robot behaviour is interpreted [14], reinforcing the importance of including attitude measures as covariates and of embedding agency signals in live behaviour rather than relying on framing alone.

https://doi.org/10.18745/pb.27307

## Experimental Design and Agency Manipulation

**Design.** Forty participants are randomly assigned to a High Agency (HA) or Low Agency (LA) condition (n = 20). All interact with both Pepper and Fetch in counterbalanced order. Three scenarios foreground different factor subsets: Scenario 1 (Security and Password Check) targets Goal Directness, Social Awareness, and Accountability; Scenario 2 (Rock–Paper–Scissors) targets Shared Decision-Making, Embodiment, and Social Awareness; Scenario 3 (Table Cleanup) targets all five. Perceived agency and trust are measured after each robot interaction using the PA scale [2], RoSAS [15], and MDMT [16]. NARS [17] is administered at baseline as a covariate, given its predictive power in our prior study [4]. Behavioural compliance provides an objective trust measure [18]. **Agency Factors and Verbal Manipulation:** Table 1 defines the five factors and contrasts HA and LA verbal manipulations. In the HA condition the robot states its purpose before acting, justifies decisions with explicit reasoning, invites the participant to jointly evaluate choices, and uses warm expressive language. In the LA condition the robot completes the same tasks with bare, context-free instructions that omit rationale, delegate decisions to the participant, and avoid emotional responsiveness.

*Table 1. The five agency factors operationalised in this experiment, with theoretical definitions and verbal manipulation contrasts.*

| *Factor* | *Definition* | *High Agency Manipulation* | *Low Agency Contrast* |
|---|---|---|---|
| *Goal Directness / Intentionality* | Actions perceived as purposeful and directed toward self-chosen goals rather than reactive [3, 7]. | States purpose before acting; identifies problems autonomously (e.g., "To check the sensors, I will need the master password"; "I see the table needs cleaning"). | Bare requests with no rationale (e.g., "I need a five-digit password"; "The table needs cleaning"). |
| *Social Awareness* | Perceives and responds adaptively to social cues [8, 9]. | Expressive, personalised responses (e.g., "That was fun! You won this time!"); adaptive recovery when participant input is missed. | Functional, neutral language; no emotional responsiveness or round-result feedback. |
| *Accountability / Responsibility* | Justifies or takes responsibility for actions, signalling transparency [8, 10]. | Explicit ethical reasoning (e.g., "I will terminate the password for safety reasons. It is not ethical to keep it active"). | Action stated without explanation (e.g., "Ok, the password is no longer needed"). |
| *Shared Decision-Making* | Human and robot jointly participate in choices [11, 12]. | Invites participant evaluation (e.g., "Do you agree the recycling bin is not ok for the bread?"). | Directive only; no consultation (e.g., "Please throw the bread in a proper bin"). |
| *Embodiment* | Physical form and movement shape how agency is experienced [9, 13]. | Each participant interacts with two robots: Fetch physically receives objects and places them in the bin/box in S3; Pepper visually displays its choice on its chest screen in S2. The Embodiment factor varies within-subjects (robot type) rather than between-subjects (HA/LA) | |

**Non-Verbal Manipulation.** Table 2 specifies the non-verbal behaviour differences. The HA robot actively orients toward the participant and relevant objects, pauses before decisions to convey deliberation, and leads physical navigation. Between scenarios it performs purposeful autonomous actions (e.g., checking room temperature after S1, monitoring the network after S2). In the LA condition the robot maintains neutral posture, acts deterministically, and remains idle between scenarios. Robot-specific HA cues include Pepper's chest screen display during the game, LED chest pulse during deliberation, and Fetch's varied trajectory and independent object placement.

*Table 2. Non-verbal behaviour differences between High and Low Agency conditions.*

| *Factor* | *High Agency* | *Low Agency* | *Applies To* |
|---|---|---|---|
| *Gaze & Head Orientation* | Actively orients toward participant; shifts gaze to relevant objects before initiating dialogue. | Neutral or fixed orientation; no proactive gaze shifts. | Both |
| *Approach & Spatial Behaviour* | Moves toward participant at scenario start; leads physical navigation (e.g., moves to kitchen first in S3). | Remains stationary at scenario start; follows participant rather than leading. | Both |
| *Deliberation & Timing* | 1–2 s pause before key decisions; motion timing varied across executions. | Immediate action after input; deterministic timing; identical motion sequences. | Both |
| *Inter-scenario Behaviour* | Performs autonomous purposeful task between scenarios (S1.5: temperature check; S2.5: network monitoring). | Remains idle between scenarios; no self-initiated behaviour. | Both |

https://doi.org/10.18745/pb.27307

| Factor | High Agency | Low Agency | Applies To |
|---|---|---|---|
| *Fetch Object Handling (S3)* | Orients gaze to item and bin before receiving; places items independently; leads to kitchen. | Asks participant to show the bin; follows participant to kitchen; deterministic placement. | Fetch only |

## Results and Discussion

Data collection is ongoing. We hypothesise that HA participants will rate both robots significantly higher on the PA scale, RoSAS warmth and competence subscales, and MDMT trust dimensions. The effect is expected to be moderated by embodiment, with Fetch's physical manipulation capability amplifying the HA manipulation specifically in Scenario 3. Unlike framing-based approaches [4] where only sincerity differed significantly between conditions, the present framework embeds agency signals across multiple channels simultaneously. This is motivated by PA scale evidence that the most predictive items concern emotional responsiveness and integrative behaviours requiring both thoughts and feelings [2, 3], cues that framing alone cannot convey. NARS will be included as a covariate following its demonstrated predictive role in our prior work [4], and a manipulation check using the PA-R Rasch scale [2] will confirm condition separation before main analyses.

## Conclusion

This paper presents a replicable, theoretically grounded framework for manipulating perceived agency in live HRI, implemented for Pepper and Fetch across three scenarios. Building on the empirical finding that autonomy framing is insufficient [4] and the theoretical position that perceived agency mediates the autonomy–trust relationship [5], the five-factor framework addresses a recognised gap in the field. Empirical results will be reported in a follow-up paper.

## References

[1] P. A. Hancock et al., "A meta-analysis of factors influencing the development of human-robot trust," Human Factors, vol. 53, no. 5, pp. 517–527, 2011.

[2] J. G. Trafton et al., "The perception of agency," ACM Transactions on Human-Robot Interaction, vol. 13, no. 1, pp. 1–23, 2024.

[3] J. G. Trafton, C. R. Frazier, K. Zish, B. J. Bio, and J. M. McCurry, "The perception of agency: Scale reduction and construct validity," in 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Busan, Republic of Korea, pp. 936–942, 2023

[4] A. Fallahi et al., "Effects of perceived robot autonomy and personal differences on trust in human-robot interactions," in International Conference on Social Robotics, Singapore: Springer Nature Singapore, 2025.

[5] A. Fallahi, P. Holthaus, F. Amirabdollahian, and G. Lakatos, "Autonomy, agency, and trust: Towards integrated calibration in human-robot interaction," presented at the TRUST: Building Trust in HRI Workshop (SCRITA 2025), in conjunction with the IEEE RO-MAN 2025 Conference, 29 Aug. 2025.

[6] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," Human Factors, vol. 46, no. 1, pp. 50–80, 2004.

[7] A. Waytz, J. Heafner, and N. Epley, "The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle," *Journal of Experimental Social Psychology*, vol. 52, pp. 113–117, 2014.

[8] R. B. Jackson and T. Williams, "A theory of social agency for human–robot interaction," *Frontiers in Robotics and AI*, vol. 8, 2021, Art. no. 687726.

[9] E. Short et al., "No fair!! An interaction with a cheating robot," in *2010 5th ACM/IEEE International Conference on Human–Robot Interaction (HRI)*, IEEE, 2010.

[10] P. Langley, J. E. Laird, and S. Rogers, "Cognitive architectures: Research issues and challenges," Cognitive Systems Research, vol. 10, no. 2, pp. 141–160, 2009

[11] M. Q. Azhar and E. I. Sklar, "A study measuring the impact of shared decision making in a human-robot team," The International Journal of Robotics Research, vol. 36, no. 5–7, pp. 461–482, 2017.

[12] A. Tausch et al., "Human–robot interaction: how worker influence in task allocation improves autonomy," Ergonomics, vol. 65, no. 9, pp. 1230–1244, 2022.

[13] F. Ciardo et al., "Attribution of intentional agency towards robots reduces one's own sense of agency," Cognition, vol. 194, 2020, Art. no. 104109.

https://doi.org/10.18745/pb.27307

[14] P. Holthaus et al., "Agency effects on robot trust in different age groups," in *Proceedings of the 12th International Conference on Human–Agent Interaction*, 2024.

[15] C. M. Carpinella et al., "The robotic social attributes scale (RoSAS): Development and validation," in Proceedings of the 2017 ACM/IEEE International Conference on Human–Robot Interaction, 2017

[16] D. Ullman and B. F. Malle, "What does it mean to trust a robot? Steps toward a multidimensional measure of trust," in Companion of the 2018 ACM/IEEE International Conference on Human–Robot Interaction, 2018.

[17] T. Nomura et al., "Measurement of negative attitudes toward robots," Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, vol. 7, no. 3, pp. 437–454, 2006.

[18] M. Salem et al., "Would you trust a (faulty) robot? Effects of error, task type and personality on human–robot cooperation and trust," in Proceedings of the Tenth Annual ACM/IEEE International Conference on Human–Robot Interaction, 2015.

https://doi.org/10.18745/pb.27307

# Data Driven Virtual Reality Tensile Testing Laboratory Development

**Eugen Guritanu[1*], Israt Kabir[2]**
*Corresponding author: e.guritanu@herts.ac.uk

## Abstract

This paper presents a mixed reality tensile testing laboratory built in Unity for the Meta Quest 3 HMDs. The virtual machine replays real experimental data rather than simulated behaviour, preserving the output format of the physical lab and requiring no changes to the downstream assignment structure. Five flexible stations guide learners through the full testing workflow at their own pace.
Keywords: Augmented Reality; Virtual Reality; Unity; Blender; Tensile Testing Lab.

## Introduction

The standard Tensile Testing of materials is a fundamental practical-based teaching activity of engineering materials related modules across all major engineering programmes in the higher education institutes (Llanos-Ruiz et al., 2025). Tensile testing provides key information on the mechanical behaviors under axial pulling load, concepts of stress-strain behavior (Tech, n.d.), to design parameters complying mechanical performance requirements of an engineered part (The Welding Institute, n.d.). However, access to the universal testing machines (UTMs) to conduct tensile test is constrained by equipment cost, throughput, and safety supervision (Bourguet et al., 2020; Vergara et al., 2017). In a typical session, single student of 100 cohort size, observes a half-an-hour single demonstration or receive a ten-minute brief turn at the machine, limiting hands-on engagement with specimen preparation, software operation, and data interpretation. Virtual reality (VR) and mixed reality (MR) tensile testing laboratory offer an immersive route to enhanced access for learners (McCloskey et al., 2024; Vergara et al., 2017, 2019). The existing Virtual Reality Tensile Testing (VRTT) laboratories often rely on simulated material behavior which significantly varies from the physical lab results (Rivai Wardhani et al., 2023; Srinivasa et al., 2021).

This paper presents a mixed reality tensile testing laboratory, driven by real experimental data. Rather than simulating material response, the virtual UTM uses previously recorded data. At the end of the session, the format of the data gathered is identical to that of the physical experiment.

The contributions of this work are threefold. First, a data-driven mixed reality laboratory sidesteps the fidelity problem of simulated mechanical behaviour by replaying validated experimental data. Second, hybrid animation architecture for running the machine provides more realistic experience. Third, a flexible Virtual Lab layout scaffolds the full procedural workflow from specimen measurement to post-test data interpretation enhancing learning process for students with variable learning needs.

## Developing VRTT lab

#### Hardware and Software Platform

The VRTT application is developed using Unity Game Development Software and deployed on the Meta Quest 3 Head Mounted Display (HMD), creating passthrough mixed reality where the virtual UTM is anchored to a physical space in the room. Students retain sight of their surroundings, peers, and instructor throughout the session, which reduces the disorientation common to fully immersive VR.

#### Animation Approach

The design of the VRTT lab based on driving the virtual test from real experimental data rather than from a simulation model. Three reference specimens — MDF, Aluminium, and Stainless Steel — were tested on a Tinius Olsen 50 kN UTM, producing CSV files with three columns: time, force, and extension. These files form the data banks of the virtual tests. This data driven approach can extend the data bank with other advanced and costly materials, like Titanium (Ti-6Al-4V).

When a student selects a specimen, the corresponding CSV is loaded, and its rows are played back in real time. Time and extension values drive the position of the upper jaw and the deformation of the specimen. The blendshape value controlling specimen extension is computed as: blendshape = ($extension_{mm}$ / $extension_{max,mm}$);
For example, an 80 mm extension against a 200 mm blendshape range maps to a value of 0.4. The Force and extension values are simultaneously plotted on a virtual display. At test completion, the original CSV is written to the student's

https://doi.org/10.18745/pb.27307

session output in the same format as that produced by the Tinius Olsen's Horizon software, preserving compatibility with the existing downstream assignment of data post processing and writeup. A comparison of the VRTT and a traditional Physical laboratory pathway can be found below in Figure 1:

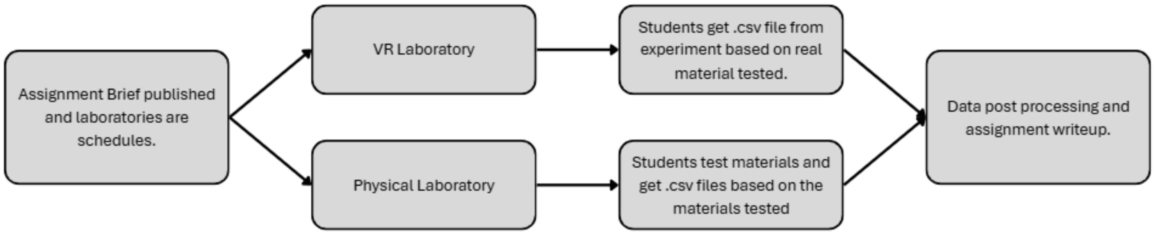



*Figure 7 VR and Physical Laboratory pathway comparison*

**Animation Architecture**

Animations operate via the hybrid animation architecture suited to a different class of motion, highlighted in *Table 1 Animation Mechanisms*.

| Layer | Purpose | Driven by |
|---|---|---|
| Unity Transforms | Upper jaw position during test | C# script reading corresponding .CSV file |
| Keyframed Animations | Jaw open/close actions | Button press events |
| Blendshapes | Specimen extension, necking, fracture | CSV extension value and fracture flag |

*Table 1 Animation Mechanisms*

The Jaw position is operated via a C# script reading the .csv file in real time. The current position is calculated as the original upper jaw position adding to the extension from the csv file. Jaw operation is button-triggered event, a deliberate departure from the physical machine rather than physically tightening the jaws. This is due to the lack of perceivable feedback when the jaw is tightened within the virtual environment. However, it allows animation events to be gated — for instance, the test cannot begin until both jaws are closed — and ensures that procedural steps are completed in the correct order. The specimen itself is represented by a single Blender mesh with three blendshape channels: *extension*, *necking*, and *fracture*. Fracture is binary, represented as a discrete state rather than an animated event because the physical failure occurs too rapidly for meaningful interpolation at typical test speeds.

**VR Lab Layout** The laboratory is organised as five stations distributed across the virtual space, through which students can flexibly progress and revisit stations:

*Figure 2 Video of Narrator for Conceptual Overviews*

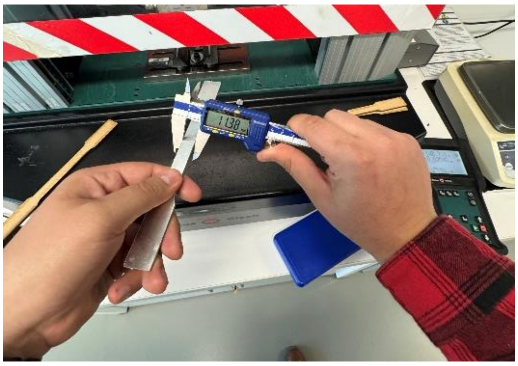


*Figure 3 First Person Video of Manual Operations*

**Introduction** — a video overview of tensile testing and session objectives as shown in Figure 2.

**Measuring** — a first-person tutorial demonstrating vernier calliper use and measuring specimens as shown in Figure 3. **Guided walkthrough** — a video how-to guide explaining machine operation, including the button-operated grip convention. **Working machine** — the interactive test itself: specimen selection, mounting, jaw closure, test execution, inspection, and CSV export. This is aided by a written how-to guide that mirrors the steps in the Guided walkthrough video. **Post-processing** — a closing video covering calculations, including yield stress, ultimate tensile stress, elastic modulus, interpretation of the stress–strain curve, and how the exported data feeds the assignment.

## Results and discussion.

The VRTT lab uses recorded CSV data from the physical Tinius Olsen machine that avoids the simulation fidelity problems also reported in similar studies (Rivai Wardhani et al., 2023; Srinivasa et al., 2021). The exported CSV file matches the experimental output so that downstream assignment stays alike regardless of how the data was collected. This approach is compatible to add new materials like Ti-6Al-4V, simply by including a vetted dataset, helping overcome the cost and access issues others have flagged (Bourguet et al., 2020).

Hybrid animation architecture separates data-driven motion from event-triggered actions, to add specimens without reworking the underlying logic. Gated events direct students through the correct sequence reinforcing good lab practice that is highlighted as important for effective VR learning (Vergara et al., 2019).

The station layout gives students control over their own pacing, letting them revisit content or repeat tests as needed (McCloskey et al., 2024). In a physical session with a large cohort this kind of flexibility is rarely possible. Each station uses the media format that fits best — first-person video for measurement, Video of Narrator for conceptual overviews and written guides for the test itself.

These contributions will be validated through a student survey run in Semester A alongside the physical lab in an engineering materials module.

## Conclusion

This work successfully embedded hybrid animation architecture for developing the VRTT which demonstrates the deformation of the materials more realistically during the virtual tensile test. Unlike the existing VRTT lab, a data-driven mixed reality laboratory was built to overcome the fidelity issue of simulated mechanical behaviour by replaying validated experimental data. Anchoring a VRTT lab to real experimental data in this work preserves output compatibility with existing assessments while removing the access and cost constraints of the physical session. The modular layout and hybrid animation architecture of this VRTT allow the system to grow as new materials are added.

## Future work

VRTT will be added as a supporting tool alongside physical laboratories as a starting point. Student experience will be captured using feedback from the students after undertaking both the real and the virtual Tensile Testing Labs. Based on that data, future works will be written to assess the pedagogical impact.

## References

Bourguet, M. L., Wang, X., Ran, Y., Zhou, Z., Zhang, Y., & Romero-Gonzalez, M. (2020). Virtual and augmented reality for teaching materials science: A students as partners and as producers project. Proceedings of 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2020, (November), 452–459. https://doi.org/10.1109/TALE48869.2020.9368381

Llanos-Ruiz, D., Abella-García, V., & Ausín-Villaverde, V. (2025). Virtual Reality in Higher Education: A Systematic Review Aligned with the Sustainable Development Goals. Societies, 15(9), 1–60. https://doi.org/10.3390/soc15090251

McCloskey, D. W., McAllister, E., Gilbert, R., O'Higgins, C., Lydon, D., Lydon, M., & McPolin, D. (2024). Embedding Civil Engineering Understanding through the Use of Interactive Virtual Reality. Education Sciences, 14(1). https://doi.org/10.3390/educsci14010006

Rivai Wardhani, Naufal Alfandra Setyawan, & Moh. Rafi Bima Adi Saputra. (2023). Development of Interactive Virtual Reality Application for Simulating Experimental Tensile Test. International Conference on Digital Advance Tourism, Management and Technology, 1(2), 570–580. https://doi.org/10.56910/ictmt.v1i2.121

Srinivasa, A. R., Jha, R., Ozkan, T., & Wang, Z. (2021). Virtual reality and its role in improving student knowledge, self-efficacy, and attitude in the materials testing laboratory. International Journal of Mechanical Engineering Education, 49(4), 382–409. https://doi.org/10.1177/0306419019898824

Tech, M. (n.d.). Tensile Test Experiment. Materials Science and Engineering. Retrieved April 20, 2026, from https://www.mtu.edu/materials/k12/experiments/tensile/

The Welding Institute, T. (n.d.). Mechanical testing - Tensile testing, Part 1. The Welding Institute. Retrieved April 20, 2026, from https://www.twi-global.com/technical-knowledge/job-knowledge/mechanical-testing-tensile-testing-part-1-069

Vergara, D., Extremera, J., Rubio, M. P., & Dávila, L. P. (2019). Meaningful learning through virtual reality learning environments: A case study in materials engineering. Applied Sciences (Switzerland), 9(21). https://doi.org/10.3390/app9214625

Vergara, D., Rubio, M. P., & Lorenzo, M. (2017). On the design of virtual reality learning environments in engineering. Multimodal Technologies and Interaction, 1(2). https://doi.org/10.3390/mti1020011

https://doi.org/10.18745/pb.27307

# Bold Hearts: Developing new generations of humanoid robots

**Alessandra Rossi[1*], Daniel Barry[1], Bente Riegler [1], Oliver Samuel [1] and Daniel Polani[1]**

*[1]RoboCup Bold Hearts, University of Hertfordshire*

*Corresponding author: a.rossi@herts.ac.uk

*RoboCup is an international championship whose purpose is to advance state-of-the-art humanoid robots, with the objective of winning against the human FIFA world champion soccer team. This ambitious goal implies the development of fully autonomous robots that can adapt to the fast dynamics of a soccer game – spanning from physical to cognitive architecture advancements. Dwelling on such a scenario, we discuss the significant technical and technological challenges expected in human-robot soccer games. Finally, we provide an overview of future development of Bold Hearts robots needed to meet the 2050 RoboCup goal[1].*

Keywords: RoboCup; Humanoid Robots; Bold Hearts

## Introduction

RoboCup (RC) was founded in 1997 in Nagoya (Japan) where over 40 university teams participated in real and simulation challenges, and over 5000 spectators attended. Since 1997, the attention to robotics and development of robotic technologies rapidly evolved, from hardware to software, and in terms of different challenges. Now, RoboCup is among the largest annual robotics competitions, aimed at advancing robotic technologies by combining high-performance hardware, innovative physical designs enabled by cutting-age materials, and state-of-the-art perceptual and cognitive intelligence [1]. In line with these objectives, RoboCup serves as a benchmark for the development of autonomous robots capable of understanding their contextual environment – including their surroundings and other agents – adapting to the dynamics of the environment, and interacting safely with humans. To support this mission, RC is organised in several leagues. For example, the @Home League focuses on robots performing household tasks while interacting with people in a simulated domestic setting. The Rescue League involves robots operating in realistic search-and-rescue scenarios. However, the most prominent leagues are the Soccer Leagues, and particularly the Humanoid Soccer League, where teams of robots with varying sizes, hardware and software compete in soccer matches. The goal is to create a robotic team capable of defeating the best human soccer team by 20250. To reach such ambitious goal, the soccer competitions address a wide range of scientific challenges by introducing progressively more complex tasks each year and encouraging collaboration across leagues and academic institutions.

Despite the remarkable advances of robotic systems, such as the use of Large Language Models and humanoid hardware capabilities for running, jumping, and stair climbing, robotic platforms are still not ready for competing and winning against human soccer players. This gap highlights a significant scientific and research challenge for the Bold Hearts team this means that its robots need substantial adaptation and continued development to achieve the level of performance required for effective participation in such competitions. Addressing this need is not only essential for competitiveness but also for advancing the broader understanding of integrated perception, cognition and control in dynamic, human-centred environments. In this work, we discuss the key modifications and research directions necessary to move towards this goal.

## Bold Hearts

The Bold Hearts is the RoboCup team from the University of Hertfordshire. We are the longest continuously active RoboCup teams in the United Kingdom, and one of the most successful. The team was founded in 2002 by Prof. Daniel Polani, and consists of computer science and engineering BSc and MSc students, and staff members at UH who have a keen interest in AI and Robotics. The team started competing in Soccer 2D Simulation League in 2003, continued in the 3D Simulation in 2004, humanoid simulation in 2007 and finally moved into the

[1] RoboCup Federation: Objective, The Dream: https://robocup.org/objective

https://doi.org/10.18745/pb.27307

Humanoid League in 2013 playing with DARwin OP robots[2]. In the past years, RoboCup started to redefine its vision to incorporate new generations of humanoid robots, advanced AI methods and technologies, and human–robot interaction to bridge the substantial gap between the technical roadmap prioritized by the RoboCup community and the broader challenges involved in a true human–robot match. As a consequence, these rapid advancements in humanoid robotics have driven substantial development of the Bold Hearts robots' hardware. In the next section, we will present the main changes undertaken to meet the increasing demand of the Humanoid League.

## Building new generations of robots

The DARwin OP initially owned by the UH Bold Hearts, is a small humanoid robot with approximately 20 actuated joints, standing about 45 cm tall and weighing around 2.8 kg, that were widely used for research and RoboCup competitions. To address the scientific and research challenges of RoboCup, the team has, in recent years, begun replacing the original DARwin OP platform by designing and developing new components - ranging from redesigned feet that enable walking on turf instead of flat carpet, to improved arms, legs, torso, and head - aimed not only at increasing the robots' height but also enhancing their overall performance and adaptability. However, in the past year, there has been a boom in humanoid robotic platforms with rapid advancements across the field, in particular, the physical achievement. Figure 1 shows the development of Bold Hearts robots moving from DARwin OP to the Bold Bot platform. To accommodate the increased height target, the platform gained approximately 16 centimeters in height, requiring proportional elongation of the arms, shins, legs, and torso. Each structural component is 3D printed using a combination of PLA and ABS materials, with aluminum reserved for the torso to provide additional rigidity and durability. The feet underwent significant redesign, both in shape and functionality. Inspired by human soccer cleats, spikes were added to the sole to improve grip and stability on turf surfaces — which present a notably different challenge compared to carpet, as turf patches are rarely seamlessly connected and the surface itself is considerably softer and more uneven. Finally, the head was redesigned to evoke the visor of a knight's helmet, paying homage to the team's origins. The Bold Hearts name draws inspiration from the legend of King Arthur — his knights and the Lady of the Lake. The robots' names are: Oberon, Gareth, Nimue, Ywain (who is on tour around the world at the "AI: More than Human tour" organised by the Barbican[3]), Bors, Dagonet, Tor.

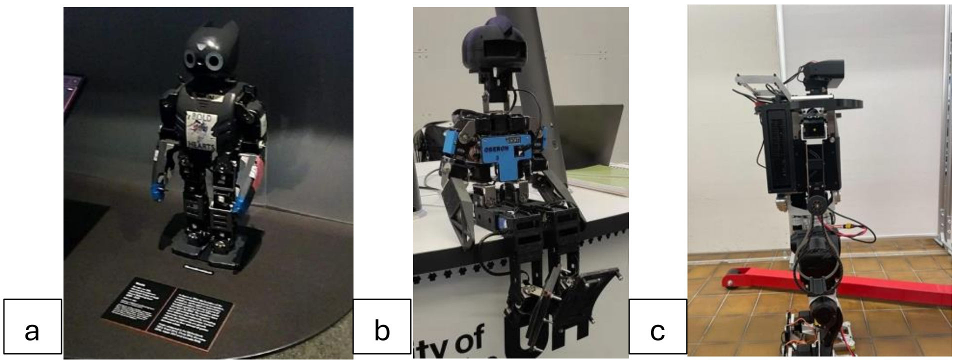



*Fig. 1. The Figure on the left (a) shows the original Darwin OP on display at the Barbican, the Figure* in the centre *(b) shows the most recent version of the Bold Bot robots, and Figure on the right (c) shows the latest developed platform BoldRbot.*

Modern humanoid robots, such as Booster T1, Unitree G1, Tiangong Ultra and Themis V2, can walk at high pace, run over uneven terrain, and participate in half-marathons (e.g., in April 2025 in Beijing). Moreover, high-speed and high-acceleration movements demand more precise control—for instance, when a robot needs to come to a stop. Additionally, current motors are still far less efficient than human muscles in terms of power density, and a substantial gap persists between the fastest humanoid robots (ASIMO at 9 km/h, Themis V2 at 10 km/h) and

[2] https://emanual.robotis.com/docs/en/platform/op/getting_started/
[3] https://www.barbican.org.uk/hire/exhibition-hire-barbican-immersive/ai-more-than-human

the fastest human, Usain Bolt, at 44.72 km/h. These advances, combined with the cumulative changes made to the platform over the years, necessitated a further round of redesign. On one hand, the robot required additional height gains to remain competitive within the evolving demands of the Humanoid League. On the other hand, the existing electronics and onboard computational hardware had become outdated, no longer sufficient to support the increased physical demands or the complexity of modern software pipelines. As a result, the team is currently undertaking a full redesign of the robotic platform. The major extension of the existing design, called the BoldRbot, funded by CapEX, is being developed to be taller and more physically capable, equipped with more powerful motors to handle the increased load and range of motion. The platform will also feature upgraded control boards capable of running computationally intensive software, including multiple machine learning and AI algorithms, enabling more sophisticated perception, decision-making, and real-time autonomy during competition.

*Table. 1. Comparison of humanoid robots for football at the University of Hertfordshire.*

| | Darwin OP1 | Bold Bot | Darwin OP3 | BoldRbot |
|---|---|---|---|---|
| Height | 45.5cm | 61.5cm | 51.0cm | 90.0cm |
| Weight | 2.90kg | 3.05kg | 3.50kg | 20.0kg |
| Walking speed | - | 20cm/s | 24cm/s | >100cm/s (est) |

The new BoldRbot represents a significant improvement on previous humanoid platforms as demonstrated in Table 1. The platform is expected to be technically capable of running and jumping with suitable control software. The team is planning to test the robotic platform to the next RC competition.

## Conclusion & Future Works

The Bold Hearts team at the University of Hertfordshire has made significant strides in advancing their humanoid robotic platforms, moving from the commercial DARwin OP1 to the substantially more advanced BoldRbot. While the gap between robotic and human athletic performance remains, ongoing hardware innovation and progress in AI, perception and control position the team well for continued progress towards RoboCup's ambitious log-term vision of developing a robot team capable of competing – and winning – against the best human soccer players.

## References

[1] Rossi, A., Paetzel-Prüsmann, M., Keijsers, M. et al. The human in the loop Perspectives and challenges for RoboCup 2050. Auton Robot 48, 8 (2024). https://doi.org/10.1007/s10514-024-10159-3

https://doi.org/10.18745/pb.27307

## 6. Resilient Infrastructure: Materials, Modelling, and Decision Systems

Chair: Dr Asal Pournaghshband

Judges: Dr Antonis Kanellopoulos, Dr Mike Otieno

https://doi.org/10.18745/pb.27307

## Invited Speaker – Dr Antonis Kanellopoulos

Bio: Antonis is a Reader in Innovative Construction Materials at the School of Physics, Engineering and Computer Science at the University of Hertfordshire. He is currently the Director of the Centre for Engineering Research and the REF-2029 coordinator for Engineering. Previously, he held the position of research fellow at the University of Cambridge (2013-2017), working on self-healing cement-based materials at the Division of Civil Engineering (2013-2016) and on graphene-modified construction materials at the Cambridge Graphene Centre (2016-2017). Prior to this, Antonis was a lecturer in Civil Engineering at Frederick University, Nicosia (2011-2013) and a senior research fellow at the University of Cyprus (2007-2011). Antonis is a chartered engineer (CEng), chartered scientist (CSci), a member of the Institute of Materials, Minerals and Mining, a member of the Institute of Concrete Technology and a Fellow of the Higher Education Academy.

**Title: Upcycling Industrial Byproducts for Sustainable Built Environment – Opportunities, Barriers and Case Studies**

Rapid urbanisation is driving an unprecedented demand for construction materials such as cement, aggregates, steel, and water, while simultaneously generating enormous quantities of construction, demolition, and mining waste. Although recycling and recovery rates for these waste streams are often reported as high, much of the recovered material is downcycled into lower-value applications that do not significantly reduce the need for virgin raw materials. As cities continue to expand and global infrastructure requirements increase, the extraction of natural resources, energy consumption, carbon emissions, and waste generation are expected to rise substantially, creating major environmental and sustainability challenges.

A key opportunity lies in shifting from traditional waste management practices toward a circular, zero-waste economy in which industrial byproducts and mineral wastes are treated as valuable resources rather than disposal burdens. Many of these materials possess chemical and mineralogical properties that make them suitable for high-value applications in construction, enabling the conservation of natural resources, reduction of landfill use, and lowering of environmental impacts. Through innovative reuse and upcycling strategies, supported by policy initiatives and industry collaboration, waste materials can be transformed into sustainable construction products that contribute to both economic value creation and the development of more resilient and resource-efficient infrastructure.

https://doi.org/10.18745/pb.27307

# Reactivity of Calcined CDW Fines and Mine Tailings in Blended Cements through Integrated Experimental and Modelling Approaches

**Bernasconi D.[1*], Maruthupandian S.[1], Chrysanthou A.[1], and Kanellopoulos A.[1]**
[1]*Centre for Engineering Research, University of Hertfordshire, Hatfield, UK*
*Corresponding author: d.bernasconi@herts.ac.uk

**The upcycling of construction and demolition waste fines (CDW fines) and graphite mine tailings as supplementary cementitious materials offers a viable pathway to reduce clinker demand while valorising mineral waste. Their reactivity was quantified in Ordinary Portland Cement blends using an mass balance approach combined with thermodynamic modelling. Degree of reaction was derived from clinker hydration and validated through phase assemblage evolution. Calcinated CDW exhibits moderate reactivity with significant secondary hydrate formation, while mine tailings shows limited reactivity dominated by filler effects. Thermodynamic predictions reproduce experimental trends in portlandite consumption and AFm development, enabling a consistent interpretation of calorimetry and mechanical strength.**

Keywords: SCM reactivity; thermodynamic modelling; construction and demolition waste; mine tailings; low-carbon cement

## Introduction

The use of supplementary cementitious materials (SCMs) is essential to reduce the environmental impact of Portland cement (OPC) systems, yet their reactivity remains difficult to quantify in complex, heterogeneous waste-derived materials. Traditional approaches based solely on calorimetry or strength development are insufficient, as they do not isolate the intrinsic contribution of SCM reaction from dilution and filler effects. Recent developments in degree of reaction (DoR) quantification, particularly the mass balance framework proposed by [1], provide a robust methodology to evaluate SCM reactivity by linking reactive oxide composition, hydrate assemblage, and portlandite consumption. When coupled with thermodynamic modelling (e.g. GEMS), this approach enables predictive interpretation of phase evolution and microstructural development. This study aims to quantify and compare the reactivity of two calcinated mineral waste SCMs, namely graphite mine tailings (MT) and construction and demolition waste fines (ST) blended with OPC, and to link their reaction kinetics to mechanical and durability behaviors [2,3].

## Experimental

An OPC 52.5N system was partially replaced (10-40 wt%) with two SCMs: ST (around 40%wt of reactive amorphous Ca-Si-Al phase) and MT (around 19%wt amorphous content, mostly aluminosilicate in nature). Quantitative phase composition was determined by XRD and TGA. Clinker hydration degrees ($C_3S$, $C_2S$, $C_3A$, $C_4AF$) were quantified over time (7-700 days), and used as input for DoR estimation. The SCM degree of reaction was calculated following a mass balance approach, where the redistribution of reactive $SiO_2$, $Al_2O_3$, CaO and MgO into hydrates is constrained by chemical equation and experimental portlandite and AFt/AFm contents. Thermodynamic modelling was performed using GEMS (Cemdata database), incorporating measured C-A-S-H composition (SEM-EDS-derived Ca/Si and Al/Si ratios) and SCM oxide release. Model outputs included phase assemblages (C-A-S-H, AFt, AFm, hydrotalcite, hydrogarnet) and water partitioning.

## Results and discussion

The integration of experimental data with thermodynamic modelling provides a consistent description of SCM reactivity and its impact on hydration mechanisms. The $^{27}Al$ SSNMR spectra of the ST system (Fig. 1a) show consumption of the raw SCM resonance near 60 ppm and the development of AFt and AFm environments at

https://doi.org/10.18745/pb.27307

lower chemical shift, confirming active participation of reactive Al in hydrate formation. This is consistent with the higher amorphous content of ST, around 41.5 wt%, compared with 19.0 wt% for MT. Portlandite evolution (Fig. 1b) further distinguishes the two systems. ST blends show a clear deviation from the dilution reference, particularly at later ages, reflecting ongoing pozzolanic reactions and consumption of CH. In contrast, MT blends follow near-dilution behaviour, indicating limited chemical interaction with the pore solution. This is consistent with their lower amorphous content and the presence of slowly dissolving phases. The degree of reaction derived from mass balance (Fig. 1c) confirms these trends. ST reaches values up to 60% at long term, while MT remains below 45%, with negligible early-age contribution. These results are in agreement with complementary R3 systems and indicate that standard accelerated tests may underestimate SCM behaviour under realistic cement conditions. Thermodynamic modelling (Fig. 1d) successfully reproduces the evolution of hydrate assemblages. In ST systems, increased formation of AFm and C-A-S-H with lower Ca/(Si+Al) ratios reflects the incorporation of reactive aluminosilicates into the binding phase. MT systems instead show phase assemblages largely controlled by clinker hydration, with limited modification of C-A-S-H chemistry. The model also captures the sustained AFt content in ST blends due to sulfate availability, and the progressive development of secondary phases such as hydrotalcite and Fe-bearing hydrates. These mechanistic differences are reflected in macroscopic performance. At 7 days, OPC reaches 52 MPa, while MT blends range from 36 to 49 MPa and ST blends from 47.2 to 52.8 MPa. At 90 days, OPC reaches 70 MPa, MT blends 51 to 63 MPa, and ST blends 63 to 69 MPa. In both waste families, the optimum replacement is 20 wt%, with MT20 showing a strength activity index of 82 to 93% and ST20 92 to 101%, indicating that ST20 can match or slightly exceed OPC at later age.

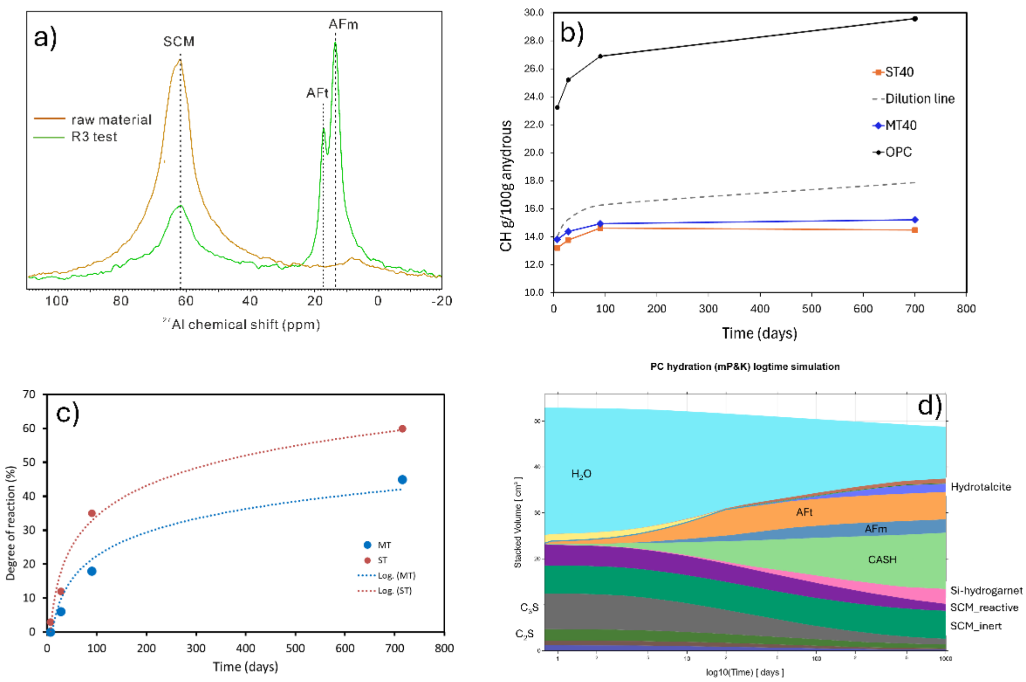



*Fig. 1. (a) $^{27}$Al SSNMR spectra of ST raw materials and R$^3$ test highlighting the formation of AFt and AFm and the consumption of reactive SCM phases; (b) evolution of portlandite content compared with the dilution reference, evidencing secondary reactivity in ST and MT blends; (c) degree of reaction of ST and MT over time derived from mass balance; (d) ST predicted phase assemblage evolution showing the development of C-A-S-H, AFt, AFm and secondary phases as a function of hydration time.*

https://doi.org/10.18745/pb.27307

## Conclusion

The quantitative results demonstrate that calcined CDW fines are distinctly more reactive than calcined mine tailings in OPC blends. For the 40 wt% systems, the degree of reaction reaches about 60% for ST40 and 55% for MT40 after 715 days, but the difference is much larger at practical ages, with 30% versus 18% at 90 days . This higher reactivity is reflected in stronger AFm formation, lower portlandite relative to dilution, and greater Al incorporation into C-A-S-H. From a performance perspective, both wastes are technically promising at moderate replacement. The optimal formulation is around 20 wt% replacement in terms of mechanical strength and durability. Overall, calcined CDW fines behave as a moderately reactive SCM with measurable secondary hydrate formation, whereas calcined mine tailings act mainly as a low-reactivity filler with limited late pozzolanic contribution. The combined mass balance and thermodynamic modelling successfully link degree of reaction, phase assemblage evolution, and engineering performance, and therefore provides a robust basis for designing low-carbon cements from heterogeneous mineral wastes.

## References

[1] F. Avet, X. Li, K. Scrivener. “Determination of the amount of reacted metakaolin in calcined clay blends” *Cement and Concrete Research*, vol. 106, pp. 40–48, 2018.
[2] S. Maruthupandian, N. A. Chaliasou, S. Akram, A. Chrysanthou, A. Kanellopoulos. “Characterisation and activation of graphite mine tailings for use in cementitious binders” *Construction and Building Materials*, vol. 473, p. 140923, Mar. 2025.
[3] S. Maruthupandian, A. Chrysanthou, A. Kanellopoulos. “Assessing the upcycling potential of construction and demolition waste derived silt: Activation, physicochemical and mineralogical characterisation” *Construction and Building Materials*, vol. 493, p. 143163, Aug. 2025.

https://doi.org/10.18745/pb.27307

# Non-Destructive Evaluation of Bond Integrity in Adhesively Bonded Joints in Composite Structures Using Acousto-Ultrasonics

**Immanuel Onwochei[1*], Gbanaibolou Jombo[1]**

[1]*Centre for Engineering Research, School of Physics, Engineering and Computer Science, University of Hertfordshire, UK.*

*Corresponding author: immanuelonwochei@gmail.com

**Adhesively bonded joints are integral to composite structures in aerospace, automotive, and marine applications due to their superior strength-to-weight ratio and fatigue resistance. However, assessing bond integrity non-destructively remains challenging. This study investigates the use of Acousto-Ultrasonic (AU) techniques for characterising bond behaviour in adhesively bonded carbon fibre reinforced polymer (CFRP) composite joints under defect-free conditions. Specimens with bonded overlap lengths of 5 cm, 10 cm, and 15 cm were tested using an offset inline transmission configuration. AU waveform features analysed include peak amplitude, signal-to-noise ratio, temporal characteristics, frequency parameters and energy distribution metrics. Results demonstrated systematic variation in AU waveform behaviour with increasing overlap length, reflecting changes in wave propagation and load transfer mechanisms. The findings support the capability of AU testing as a reliable non-destructive approach for evaluating bond behaviour and structural consistency in composite joints.**

Keywords: Structural Health Monitoring; Acousto-Ultrasonic Technique (AUT); CFRP; Adhesive Bond Integrity; Non-Destructive Testing

## Introduction

Adhesively bonded composite joints are widely used in engineering applications due to their ability to provide efficient load transfer while maintaining a high strength-to-weight ratio. However, assessing the structural performance of bonded joints remains challenging, particularly where non-destructive evaluation methods are required [1].

Acousto-Ultrasonic (AU) testing provides a means of evaluating internal structural behaviour by analysing stress wave propagation through materials [2]. Unlike conventional ultrasonic testing, AU focuses on diffuse waveform characteristics and energy transmission, making it particularly suitable for assessing bonded interfaces [3].

This study aims to characterise bond behaviour in adhesively bonded CFRP joints using AU techniques. The work is carried out under the assumption of defect-free bonding conditions, with emphasis on understanding how variations in bond overlap length influence AU waveform features and corresponding structural behaviour.

## Experimental/Simulation

Rectangular CFRP specimens of 23 cm x 5 cm x 0.2 cm are cut from Easy Composites High Strength Carbon Fibre Sheet. To form the test coupons, two specimens are bonded with 3M Scotch-Weld DP8805 adhesive. Three overlap bond configurations were prepared: 5 cm, 10 cm, and 15 cm.

Acousto-Ultrasonic testing is conducted using an offset inline transmission setup. A transmitting transducer (i.e., resonant ultrasonic transducer with 1 MHz centre frequency) introduced stress waves at one end of the bonded region, while a receiving transducer (i.e., acoustic emission transducer with a flat frequency response up to 750 kHz and resonant peak at 150 kHz) captured the transmitted signal at the opposite end. Transducer alignment and acquisition parameters were maintained consistently to ensure repeatability.

Signal processing was applied to extract key AU waveform features, including peak amplitude, signal-to-noise ratio (SNR), rise time, signal duration, energy distribution between defined gated time gates (arrival and decay), centroid frequency, decay time and damping proxy.

https://doi.org/10.18745/pb.27307

Tensile testing was also performed to evaluate the mechanical performance of the adhesively bonded joints, enabling comparison between AU-derived features and bond strength.

## Results and discussion

As shown in Fig. 1, the AU results reveal systematic changes in waveform characteristics across bonded overlap classes, indicating evolving wave propagation behaviour and energy distribution with increasing bonded area. The observed trends suggest interdependencies between key AU parameters, reflecting changes in load transfer mechanisms within adhesively bonded joints.

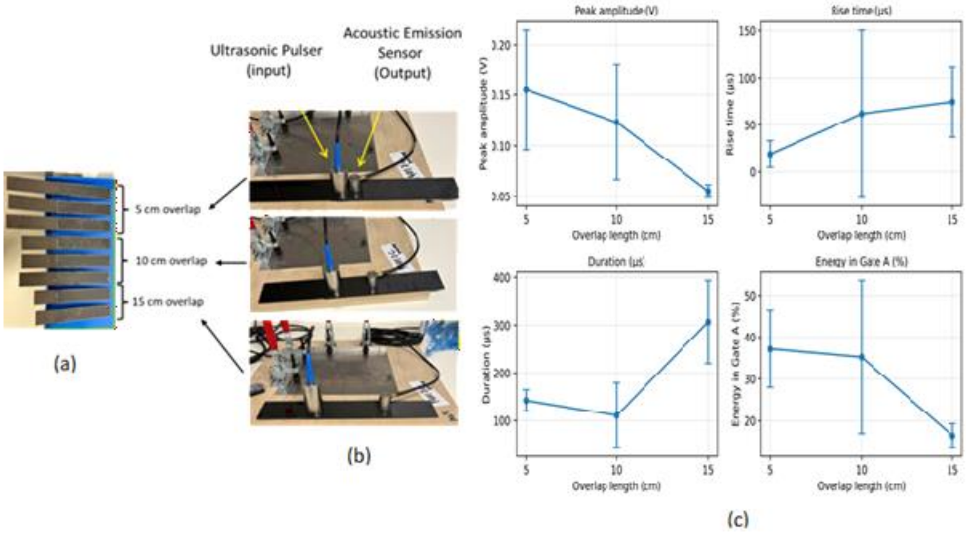



*Fig. 1. (a) Adhesive Bonding Coupon Classes; (b) AUT offset inline transmission setup; (c) Relationships between overlap length and key AU parameters,*

Peak amplitude decreases consistently with overlap length, from 0.155 V (5 cm) to 0.055 V (15 cm). Rather than indicating signal degradation, this reduction suggests progressive redistribution of transmitted wave energy across longer bonded paths. The substantially lower variability observed in the 15 cm class further indicates increased signal stability within larger bonded regions.

Signal-to-noise ratio follows a similar declining trend, reducing from 26.63 dB (5 cm) to 13.67 dB (15 cm). This behaviour likely reflects increasing temporal dispersion of wave energy, producing broader and less concentrated waveform responses as overlap length increases.

Temporal characteristics exhibit the opposite trend. Rise time increases significantly from 18.93 µs (5 cm) to 74 µs (15 cm), while signal duration expands from 142.33 µs to 306.67 µs. These changes indicate prolonged wave interaction and extended transmission pathways within larger bonded interfaces, consistent with increasingly distributed load transfer mechanisms.

Energy distribution further supports this interpretation. The proportion of energy contained within Gate A decreases from 37.13% (5 cm) to 16.4% (15 cm), suggesting that early arriving wave energy becomes progressively less dominant as overlap length increases.

Collectively, the evolution of peak amplitude, rise time, duration and energy distribution reveals interconnected changes in wave propagation behaviour with increasing overlap length. While peak amplitude decreases, rise time and duration increase substantially, indicating a transition from concentrated early-wave transmission toward more distributed energy propagation within larger bonded regions. These linked changes are consistent with improved load sharing and reduced local stress concentration.

The tensile results support these observations, with longer overlap configurations demonstrating improved mechanical performance relative to shorter bonded regions. The agreement between evolving AU waveform

https://doi.org/10.18745/pb.27307

characteristics and tensile behaviour suggests that parameters such as rise time, signal duration and energy distribution may act as surrogate indicators of bond performance and load transfer efficiency.

Overall, the results demonstrate that AU techniques are highly sensitive to changes in joint geometry, with measurable waveform variations reflecting underlying structural responses. The observed relationships between AU features and tensile behaviour further support the potential application of AU as a non-destructive approach for evaluating bond behaviour under defect-free conditions.

*Table. 1. Comparative Cross-Class Summary of Selected AU Waveform Parameters (Mean ± Standard Deviation)*

| AU Waveform Parameters | Unit | Bond Overlap 5cm | Bond Overlap 10cm | Bond Overlap 15cm |
|---|---|---|---|---|
| **Peak Amplitude (Mean)** | **V** | 0.155 | 0.123 | 0.055 |
| **Peak Amplitude (Standard Deviation)** | **V** | 0.059093 | 0.056719 | 0.006245 |
| **Rise Time (Mean)** | **µs** | 18.93333 | 61.66667 | 74 |
| **Rise Time (Standard Deviation)** | **µs** | 14.35874 | 88.86132 | 37.24245 |

## Conclusion

This study demonstrates that Acousto-Ultrasonic testing is effective for characterising bond behaviour in adhesively bonded CFRP joints, with overlap length significantly influencing AU waveform features, particularly peak amplitude and temporal parameters. The observed relationship between AU features and tensile performance highlights the potential of AU techniques as non-destructive indicators of bond performance in composite structures. The method shows practical value for quality assurance, structural health monitoring, and maintenance decision-making in aerospace, automotive, and other composite-intensive industries.

However, the study is limited by defect-free assumptions, small sample size, restricted overlap configurations, and controlled laboratory conditions. Future work should investigate defective bonded interfaces and broader loading conditions to further validate AU applicability

## References

[1] M. J. L. Cheraghi and A. S. L. Chen. “Defects and uncertainties of adhesively bonded composite joints” SN Applied Sciences, vol. 3, no. 9, pp. 1-17, 2021.
[2] E. D. Sarantinos, T. Tomic, and V. K. Srivastava. “Acousto-ultrasonic analysis of defects in composite specimens used in transportation domain” Journal of Measurements in Engineering, vol. 9, no. 2, pp. 117-124, 2021.
[3] J. Guo, A. Doitrand, C. Sarr, S. Chataigner, L. Gaillet, and E. Sarantinos. “Numerical voids detection in bonded metal/composite assemblies using acousto-ultrasonic method” Applied Sciences, vol. 12, no. 9, pp. 4153, 2022

https://doi.org/10.18745/pb.27307

# Debonding-on-demand for structural silicone joints in façade applications

**Sajjad Karimi[1*], Yiding Liu[1], Antonios Kanellopoulos[1], and Kyriaki Corinna Datsiou[1]**

*[1] Centre for Engineering Research, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, Hertfordshire AL10 9AB, United Kingdom*
*Corresponding author: s.karimi2@herts.ac.uk

**Adhesive bonding has gained popularity for its capacity to bond dissimilar materials, however, hinders end-of-life circularity. This study investigates the debonding-on-demand potential of structural silicone, typically used in façade applications, modified with 10wt.% Thermally Expandable Particles to enable circularity. Experimental tests characterized joint strength through single lap shear tests. Synchronized thermal-mechanical tests at 180°C quantified debonding potential. Results revealed dual-functional behaviour: the 10wt.% TEP formulation enhanced joint load-bearing capacity by approximately 19% compared to the neat adhesive, while achieving complete on-demand separation under thermal activation. This approach facilitates high-value facade recycling and reuse, directly contributing to the circular economy.**

Keywords: Adhesive joints; glass-aluminium joints; debonding on demand; Thermally Expandable Particles; single lap shear tests

## 1. Introduction

The construction industry is increasingly driven by circular economy goals and environmental regulations focusing on material recyclability [1]. This is particularly challenging for façades applications relying on glass and aluminium assemblies. While adhesive bonding offers superior stress distribution and design flexibility over mechanical fasteners [2], its permanent nature complicates end-of-life separation and hinders high-value recycling [3]. Conventional disassembly via manual cutting or bulk heating is problematic; structural silicones require 350–400 °C for degradation [4], exceeding safety limits for glass coatings and substrates while risking toxic gas release [5]. Consequently, research has shifted toward Debonding-on-Demand (DoD) technologies. Thermally Expandable Particles (TEPs) offer a promising physical solution. TEPs consist of a thermoplastic shell encapsulating a hydrocarbon blowing agent; upon reaching activation temperatures (80–190 °C), internal pressure induces significant volumetric expansion (Fig. 1). This expansion creates micro-cracking and porosity, reducing cohesive strength to facilitate separation [7]. While TEPs can lower the thermal barrier for disassembly [8], their impact on baseline mechanical strength remains inconsistent in existing literature [9]. This study investigates the mechanical performance and debonding efficiency of glass-aluminium joints using a structural silicone modified with 10 wt.% TEP to support façade systems that maintain structural integrity while enabling efficient end-of-life disassembly.

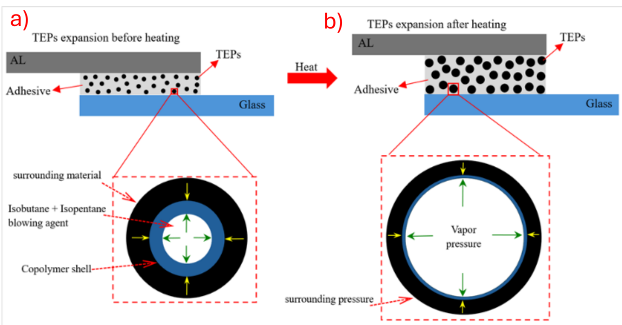



*Fig. 1. TEP expansion mechanism: (a) unexpanded state with encapsulated hydrocarbon, (b) expanded state at activation temperature causing volumetric growth.*

## 2.Experimental details

### 2.1. Materials and Sample Preparation

The study employed heat-strengthened soda-lime silica glass and anodised aluminium 6082-T6 (Al) substrates. A two-part structural silicone adhesive was used as the bonding agent, modified with 10 wt.% Thermally Expandable Particles (TEP, grade 930 DU 120). This specific TEP grade was selected for its thermal stability below

https://doi.org/10.18745/pb.27307

120 °C to ensure structural integrity under service conditions while enabling activation near 180 °C (Table 1). To ensure a homogeneous matrix and prevent particle agglomeration, the adhesive and TEPs were processed using a FlackTek 330-100 PRO high-shear centrifugal mixer (2000 rpm), followed by a 10-minute vacuum degassing stage to eliminate air entrapment. Al-Glass-Al single lap joints were then fabricated using a bespoke jig to maintain a precise 6 mm bondline thickness. Before bonding, substrates were cleaned with isopropyl alcohol and treated with 3M 111 primer to enhance interfacial adhesion. All specimens were cured at 21 ±2 °C and 50% relative humidity for seven days.

*Table. 1. Physical and thermal properties of TEPs.*

| Grade | $T_{start}$ (°C) | $T_{max}$ (°C) | Density (kg/m³) | Particle Size (µm) |
|---|---|---|---|---|
| 930 DU 120 | 117–127 | 189–204 | ≤ 6.5 | 28–38 |

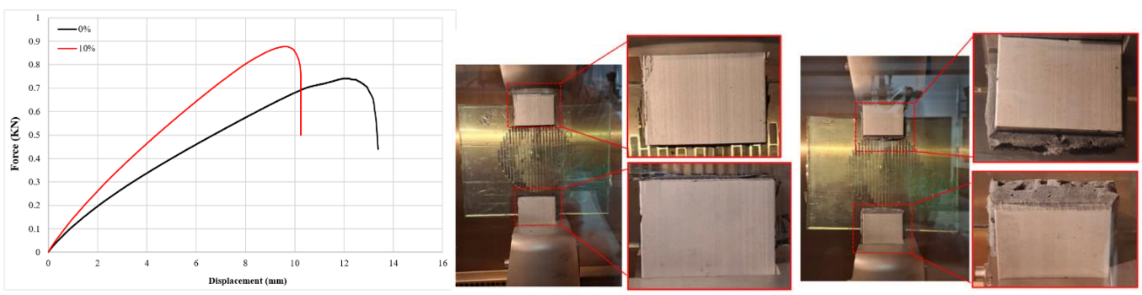
## 2.2. Mechanical and Debonding Testing

Mechanical performance was evaluated through single lap shear tests at ambient temperature using an Instron 68FM-100 machine, following ASTM D1002 at a crosshead speed of 1.3 mm/min.

To assess the debonding-on-demand potential, a synchronized thermal-mechanical procedure was conducted at 180 °C. The test involved four phases: (1) application of a 20 N preload to eliminate structural slack; (2) linear heating at 3°C/min from ambient to 180°C; (3) a 10-minute thermal soaking period to ensure uniform heat distribution; and; (4) a constant loading phase of 90 N (representing ~10% of the ultimate joint capacity). The "debonding time" ($t_d$) was recorded as the duration required for the internal expansion pressure of the TEPs to overcome the interfacial bond strength, leading to complete separation. Each condition was tested in triplicate to ensure statistical validity.

## 3.Results and discussion

The average mechanical performance of the single lap joints is illustrated in Fig. 2a. The neat adhesive exhibited a ductile response with a peak force of 0.75 kN. Upon incorporating 10 wt.% TEPs, a significant stiffening effect and a 19% increase in load-bearing capacity (reaching 0.88 kN) were observed. This enhancement suggests that TEPs effectively redistribute stress within the low-modulus silicone matrix, mitigating concentrations at the overlap ends. However, a trade-off in ductility was noted, as the failure displacement decreased from 13.4 mm to 10.2 mm, indicating that the filler loading restricts large-strain deformations of the elastomeric system.
The on-demand debonding performance was triggered by the volumetric expansion of TEPs. As shown in Fig. 2b-c, the TEP-modified formulation exhibits substantial volumetric expansion. Microscopy (Fig. 2d-e) confirmed that microspheres expanded from an average diameter of 30 µm to 60–145 µm, generating sufficient internal pressure to overcome cohesive strength. The debonding potential was quantified through force-time profiles (Fig.2f), where the 10 wt.% formulation achieved complete separation within 32 minutes at 180 °C under a minimal static load (0.09 kN). In contrast, the baseline specimens maintained structural integrity throughout the test. The expansion kinetics, shown in Fig. 2g, revealed a sharp increase in displacement rate for the modified joints after thermal activation, directly attributed to the TEP-induced volumetric growth. This confirms that




disassembly is driven by internal mechanical pressure, enabling non-destructive separation at temperatures significantly below the silicone's thermal degradation threshold.

a)

https://doi.org/10.18745/pb.27307

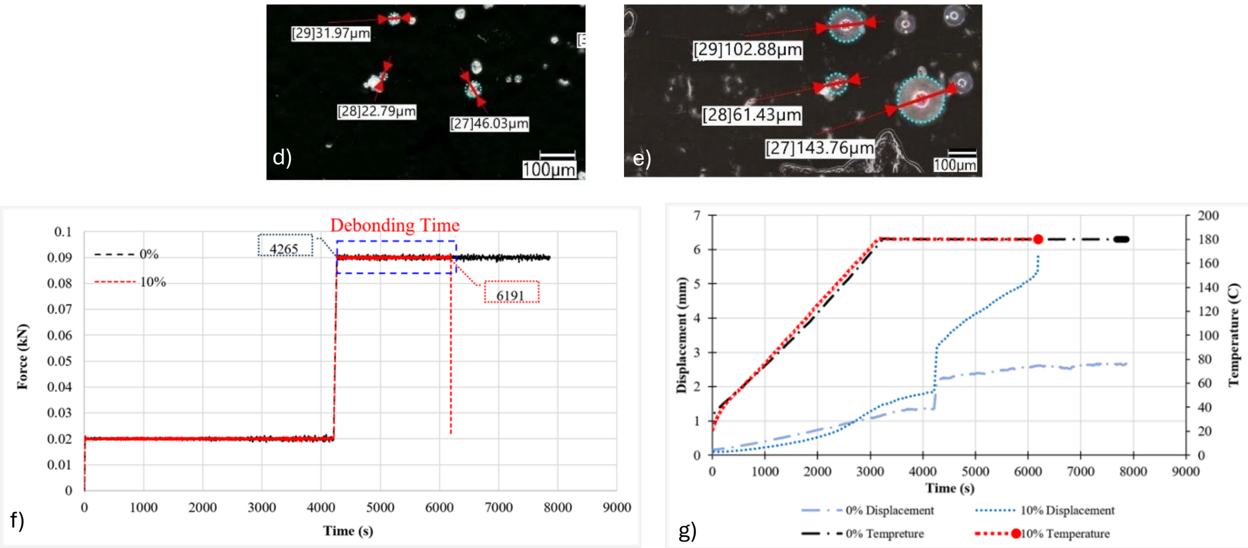



*Fig. 2. Single lap shear and debonding results: (a) force-displacement curves at ambient temperature; (b,c) macro-scale joint expansion at 20°C and 180°C; (d,e) micro-scale TEP expansion at 20°C and 180°C; (f) force-time profiles and separation behaviour during debonding phase, and; (g) Evolution of displacement and temperature over time.*

## Conclusions

This study illustrates the feasibility of integrating thermally expandable particles (TEPs) into structural silicone adhesives to achieve debonding-on-demand for glass-aluminium systems. Results indicate that a 10 wt.% TEP concentration increases the joint load-bearing capacity by approximately 19% compared to the neat adhesive. Microstructural analysis confirms that separation is driven by TEP volumetric expansion at 180 °C, facilitating a controlled failure mode while maintaining structural integrity at ambient temperatures. This methodology provides a robust, non-destructive solution for disassembling glass-aluminium facades, aligning with circular economy principles and demonstrating a viable pathway for high-value component reuse in the construction sector.

## References

[1] J.F.G. Timm, V.G. Maciel, and A. Passuello. “Towards Sustainable Construction: A Systematic Review of Circular Economy Strategies and Ecodesign in the Built Environment” Buildings, vol. 13, no. 2059, 2023.

[2] X. He. “Effect of mechanical properties of adhesives on stress distributions in structural bonded joints” Proc. World Congr. Eng., vol. 2, 2010.

[3] S. Tsoka and K. Tsikaloudaki. “Design for Circularity, Design for Adaptability, Design for Disassembly” Circular Economy Design and Management in the Built Environment, pp. 257–272, 2024.

[4] J.P. Lewicki, B.P. Mayer, and C.T. Alviso. “Thermal Degradation Behavior and Product Speciation in Model Poly(dimethylsiloxane) Networks” J. Inorg. Organomet. Polym., vol. 22, pp. 636–645, 2012.

[5] K.R. Mulcahy, A.F.R. Kilpatrick, G.D.J. Harper, A. Walton, and A.P. Abbott. “Debondable adhesives and their use in recycling” Green Chem., vol. 24, pp. 36–61, 2022.

[6] A. Hutchinson, Y. Liu, and Y. Lu. “Overview of disbonding technologies for adhesive bonded joints” J. Adhes., vol. 93, pp. 737–755, 2017.

[7] R.H. McCurdy, A.R. Hutchinson, and P.H. Winfield. “The mechanical performance of adhesive joints containing active disbonding agents” Int. J. Adhes. Adhes., vol. 46, pp. 100–113, 2013.

[8] M.D. Banea, L.F.M. da Silva, R.J.C. Carbas, and R.D.S.G. Campilho. “Structural Adhesives Modified with Thermally Expandable Particles” J. Adhes., vol. 91, pp. 823–840, 2015.

[9] G. Piazza, M. Burczyk, M. Gerini-Romagnoli, G. Belingardi, and S.A. Nassar. “Effect of thermally expandable particle additives on the mechanical and reversibility performance of adhesive joints” J. Adv. Joining Process., vol. 5, no. 100088, 2022.

https://doi.org/10.18745/pb.27307

# Optimising Green Building Design Using BREEAM and Artificial Intelligence Modelling (BREE-AIM)

**Mina Taheri, Farzad Piadeh***

*University of Hertfordshire*

*Corresponding author:f.piadeh@herts.ac.uk

## Short Abstract

This study addresses the limited integration of sustainability certification into early-stage building design decision-making. Although BREEAM provides structured performance benchmarks, it is often applied retrospectively, reducing its influence on strategic engineering decisions. In response, this research presents BREE-AIM, a feasibility-driven and certification-aware optimisation framework that integrates BREEAM UK New Construction (v6.1) data, expert-informed feasibility modelling, and the Shuffled Frog-Leaping Algorithm (SFLA). Applied to a pilot office building case, the framework generates and evaluates certification-compliant pathways while minimising implementation burden. A Python-based decision-support application was also developed to collect user inputs, check BREEAM rules, and export optimised scenarios. Preliminary validation provides initial evidence that embedding certification logic into optimisation can strengthen early-stage decision support and align sustainability targets with practical implementation.

**Keywords**: BREEAM; optimisation; artificial intelligence; feasibility modelling; decision support

## Introduction

Early-stage design decisions play a critical role in determining the environmental performance of buildings across their whole life cycle. Decisions made during concept and schematic design stages, such as material selection, structural systems, and building services, strongly influence embodied carbon, operational energy use, and long-term resource consumption [1], [2]. Because design flexibility is highest at these early stages, sustainability strategies need to be considered before key project decisions become fixed.

BREEAM is one of the most widely used sustainability certification systems in the UK and provides a structured framework for assessing building performance across categories such as energy, water, materials, management, health and wellbeing, transport, land use, and pollution [3]. However, in practice, BREEAM is often applied after major design decisions have already been made, which can reduce its ability to actively shape design strategies [4]. In this situation, certification may become a late-stage compliance exercise rather than an early-stage decision-support tool.

At the same time, digital tools, Building Information Modelling (BIM), and Artificial Intelligence (AI)-based optimisation have created new opportunities for improving sustainability decision-making in the built environment. Optimisation methods can explore many design alternatives and identify efficient solutions under competing objectives [5]. However, existing AI-based approaches often focus on technical performance indicators such as energy, carbon, or cost, while formal certification rules, mandatory credits, scoring thresholds, and implementation feasibility are rarely embedded directly into the optimisation model [6].

This creates a gap between sustainability certification, feasibility assessment, and optimisation-based decision-making. This study addresses this gap by developing BREE-AIM, a feasibility-driven and certification-aware optimisation framework that embeds BREEAM UK New Construction logic into early-stage design optimisation. The framework aims to support more practical certification planning by linking BREEAM rules, expert-informed feasibility scores, and AI-based search within a single decision-support workflow. This paper presents the methodology, implementation, and validation of the proposed framework.

## Experimental/Simulation

The proposed BREE-AIM framework integrates three main components: BREEAM certification data, expert-based feasibility modelling, and AI-based optimisation. First, BREEAM UK New Construction v6.1 data were

https://doi.org/10.18745/pb.27307

structured into a machine-readable dataset. This included assessment credits, category weightings, mandatory requirements, rating thresholds, and scoring rules [3]. Second, feasibility modelling was developed using expert judgement. Feedback was collected from 15 experts from government, industry, and academia. Each BREEAM credit was assessed using five criteria: cost, complexity, expertise, service life, and social value. The responses were weighted and normalised to create a composite feasibility score. Third, the optimisation problem was formulated as a binary selection model. Each BREEAM credit was represented as either selected or not selected. The objective was to minimise total feasibility burden while satisfying BREEAM mandatory requirements and the selected rating threshold. The optimisation process was implemented using the Shuffled Frog-Leaping Algorithm (SFLA), a population-based metaheuristic algorithm used for solving complex combinatorial optimisation problems [7]. A Python-based decision-support application was also developed. The application allows users to define project parameters, check BREEAM rules and mandatory credits, generate optimised certification scenarios, and export results to Excel.

## Results and Discussion

The framework is applied to a pilot case study of a small-scale office building targeting a BREEAM “Good” rating (≥45%). The results show that multiple certification-compliant scenarios can achieve the same target score while using different strategies. Scenario 1 provides a balanced distribution across categories but requires a higher number of optional credits. Scenario 2 focuses more on energy-related credits and achieves the target with fewer credits but with minimal margin above the threshold. Scenario 3 is the most efficient solution, achieving the highest score with the lowest number of optional credits. These findings demonstrate that certification outcomes are not unique and that multiple feasible pathways exist, each with different trade-offs among efficiency, feasibility, and coverage. This confirms that embedding certification logic within optimisation enables structured decision support rather than a single deterministic solution. This indicates that optimisation does not produce a single optimal solution but a set of feasible alternatives, supporting flexible decision-making in early-stage design.

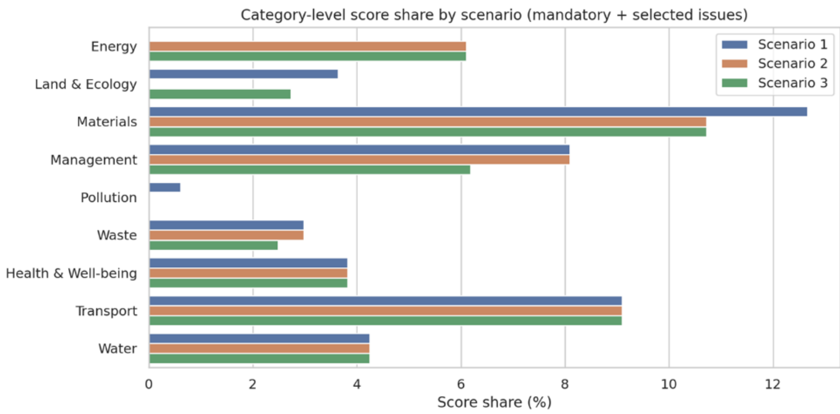



**Fig. 1. Category-level score distribution across optimisation scenarios.**

**Table 1. Comparison of optimisation scenarios**

| Scenario | Score (%) | Margin above 45% | Optional issues |
|---|---|---|---|
| Scenario 1 | 45.13 | +0.13 | 15 |
| Scenario 2 | 45.04 | +0.04 | 11 |
| Scenario 3 | 45.36 | +0.36 | 10 |

https://doi.org/10.18745/pb.27307

## Conclusion

This study presented BREE-AIM, a feasibility-driven and certification-aware optimisation framework for early-stage BREEAM planning. The framework integrates BREEAM certification logic, expert-informed feasibility modelling, and AI-based optimisation within a single decision-support workflow. The results show that embedding BREEAM constraints directly into the optimisation model enables the systematic identification of certification-compliant pathways. The pilot office case study demonstrated that multiple valid credit combinations can achieve the same certification target, while reflecting different trade-offs in feasibility, complexity, and performance. Overall, the framework helps shift BREEAM from a retrospective compliance tool to a proactive early-stage decision-support mechanism. Future work should focus on using real project data, testing the framework across more building types and project scales, integrating it with BIM workflows, and adapting the method to other certification systems. The work is currently being developed into a journal publication to support further validation and academic dissemination.

## References


[1] J. Basbagill, F. Flager, M. Lepech, and M. Fischer, “Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts,” *Building and Environment*, vol. 60, pp. 81–92, 2013.

[2] M. Röck *et al*., “Embodied GHG emissions of buildings—The hidden challenge for effective climate change mitigation,” *Applied Energy*, vol. 258, article 114107, 2020.

[3] BRE Global, *BREEAM UK New Construction Version 6.1 Technical Manual SD5079*. Building Research Establishment, 2023.

[4] A. N. Nielsen, R. L. Jensen, T. S. Larsen, and S. B. Nissen, “Early stage decision support for sustainable building renovation—A review,” *Building and Environment*, vol. 103, pp. 165–181, 2016.

[5] Y. Pan and L. Zhang, “Roles of artificial intelligence in construction engineering and management: A critical review and future trends,” *Automation in Construction*, vol. 122, article 103517, 2021.

[6] I. Elwy and A. Hagishima, “The artificial intelligence reformation of sustainable building design approach: A systematic review on building design optimization methods using surrogate models,” *Energy and Buildings*, vol. 323, article 114769, 2024.

[7] M. Eusuff and K. Lansey, “Optimization of water distribution network design using the shuffled frog leaping algorithm,” Journal of Water Resources Planning and Management, vol. 129, no. 3, pp. 210–225, 2003.

https://doi.org/10.18745/pb.27307

## 7. Applied Science and Industrial Innovation

Chair: Denis Smith

Judges: Ricky Linforth, Anju Viswanath

https://doi.org/10.18745/pb.27307

## Invited Speaker - Mel Disher

Bio: Mel Disher is a product specialist for a portfolio of particle characterisation instruments at Analytik Ltd, and a member of the RSC's particle characterisation interest group. His background is in vitro DMPK drug discovery. He spent 10 years working to support early-stage drug development, and prior to this, spent time working in regenerative medicine, culturing cells for autologous chondrocyte implantations. His current role allows me to match researchers' needs to characterisation techniques to best support their work.

**Title: The Less-Travelled Paths of Particle Characterisation.**

Highlighted some lesser-known particle characterisation techniques, including differential centrifugal sedimentation, single particle extinction and scattering, and the fluorescent capabilities for nanoparticle tracking analysis. Particle characterisation is a complex field, and there are often complementary techniques to the industry standards that can provide much deeper insights into your particles. However, these techniques are generally less well-known.

https://doi.org/10.18745/pb.27307

## ESG & Sustainable Product Framework - Andy Bonney (RS Pro)

RS Group has embedded sustainability at the centre of its business strategy through a comprehensive ESG framework aligned with long-term environmental, social, and governance goals. The company has achieved significant progress, including major reductions in carbon emissions and transport-related emissions, high recycling and recyclable packaging rates, and the expansion of its Better World product portfolio to around 33,000 products sourced from more than 170 suppliers across 30 countries. At the same time, RS continues to strengthen supplier engagement through science-based carbon reduction targets, ethical trading commitments, and sustainability performance assessments.

A key focus is the Better World Product Framework, which helps customers make trusted and sustainable purchasing decisions. Developed with sustainability experts and supported by rigorous governance processes, the framework ensures that all sustainability claims are evidence-based and transparent. Products qualify either through recognised third-party certifications and eco-labels or through verified sustainability benefits demonstrated across the product lifecycle. Looking ahead, RS plans to enhance customer experience with initiatives such as product carbon footprint labelling, lifecycle data transparency, Better World solutions, and advanced sales enablement tools to further support sustainable business practices.

https://doi.org/10.18745/pb.27307

# Combining Two Analytical Techniques to Provide Complementary Molecular Information - Clement Eliott (Shimadzu)

Clement Elliott, Spectroscopy Technical Specialist at Shimadzu UK, explores the latest developments in analytical instrumentation and how they are transforming materials characterisation across research, quality control and failure analysis applications. A key focus of the presentation is the AIRsight, an innovative microscope that uniquely integrates both Fourier Transform Infrared (FTIR) and Raman spectroscopy into a single platform. By combining these complementary analytical techniques, AIRsight enables users to rapidly identify and characterise a wide range of materials, contaminants and defects without the need to transfer samples between instruments.

The presentation demonstrates how this integrated approach simplifies workflows, improves analytical efficiency and provides richer chemical and structural information than either technique alone. Real-world application examples will highlight the benefits of combining FTIR and Raman microscopy for challenging analytical tasks, including particle identification, contaminant analysis, mapping, polymer characterisation and pharmaceutical investigations.

Attendees will gain an overview of the latest innovations in molecular spectroscopy and discover how advances in instrumentation are enabling laboratories to obtain faster, more reliable and more comprehensive analytical insight, supporting better decision-making across both industrial and academic environments.

https://doi.org/10.18745/pb.27307

## Evident Industrial (Materials) Microscopy - Jane Bratherton (Evident Scientific)

Evident Scientific, formerly known as Olympus, builds on over 100 years of optical excellence to deliver advanced industrial materials microscopy and digital imaging solutions. The technology integrates traditional materials microscopy with modern image analysis tools to streamline workflows for particle distribution, phase analysis, and pore fraction assessments. Utilising advanced digital and focus variation microscopy, these systems feature automated features like deep learning classification, live contrast enhancement for grain boundaries, and high-speed "best image" generation to dramatically improve efficiency and analytical precision in quality assurance and manufacturing processes.

Central to these innovations is the LEXT™ OLS5500 3D Optical Profilometer, a hybrid system that combines focus variation, laser scanning confocal microscopy, and white light interferometry into a single instrument. By leveraging these three distinct modalities, the microscope achieves exceptional X/Y resolution and precise height measurements spanning from sub-manometer scales to hundreds of microns. Whether capturing steep slopes or highly polished flat surfaces, the optical design eliminates blur, captures highly accurate 3D surface topographies, and delivers highly repeatable height data across a broad range of industrial materials.

https://doi.org/10.18745/pb.27307

https://doi.org/10.18745/pb.27307

# Poster Sessions

https://doi.org/10.18745/pb.27307

# Temperature-Dependent Loss Analysis of a PMSM for Electric Vehicle Traction Under WLTP Drive Cycle Using MATLAB/Simulink

**Mohisin Syed, Christos Kalyvas and Chandra Vendra**

*School of Physics, Engineering and Computer Science, University of Hertfordshire, United Kingdom*

*Corresponding author: m.syed3@herts.ac.uk

**Abstract** — This paper presents a coupled electro-thermal PMSM simulation model under the WLTP drive cycle in MATLAB/Simulink, incorporating vehicle dynamics for physically realistic torque demand generation. The Siemens SIMOTICS S 1PH8137 (57.5–66.7 kW, 104 A) is the reference motor. The model integrates temperature-dependent copper losses, iron losses, inverter losses, and mechanical losses within a first-order Lumped Parameter Thermal Network. A vehicle dynamics module — incorporating vehicle inertia (1800 kg), rolling resistance, and aerodynamic drag — generates real-time q-axis current demand from the WLTP speed profile, replacing simplified speed-proportional mapping. Under the 1800-second WLTP cycle, peak motor temperature reached 109.1°C (thermal margin 70.9°C below Class H limit), driving efficiency was 80.69%, and inverter losses were identified as the dominant loss mechanism (mean 951 W). These results provide a validated framework for EV thermal management and motor control design.

**Keywords:** PMSM; Electric Vehicles; WLTP; Vehicle Dynamics; Temperature-Dependent Losses; MATLAB/Simulink; FOC; Inverter Losses

## Introduction

Permanent Magnet Synchronous Motors (PMSMs) are the dominant traction technology in modern EVs due to high power density and efficiency [1]. However, motor performance is significantly affected by temperature: stator resistance increases with winding temperature, raising $I^2R$ losses, while NdFeB permanent magnet remanence decreases, reducing torque capability and risking demagnetisation [2, 3]. Conventional isothermal PMSM models introduce efficiency prediction errors of up to 10–15% under real driving conditions [4].

A critical but often neglected limitation of published models is the use of simplified current excitation — mapping motor speed linearly to a fixed rated current. Real motor current demand is governed by vehicle dynamics: the torque required to accelerate vehicle mass, overcome rolling resistance, and overcome aerodynamic drag determines the instantaneous q-axis current demand. This work addresses both limitations by implementing a temperature-dependent PMSM model with full vehicle dynamics under the 1800-second WLTP drive cycle, generating physically realistic loss and efficiency predictions for the Siemens SIMOTICS S 1PH8137 traction motor.

## Simulation Model

The MATLAB/Simulink model comprises five subsystems: (1) PMSM Constants — temperature-dependent flux linkage $\psi f(T)$ and electrical parameters; (2) FOC Controller with id = 0 strategy; (3) dq Current Dynamics solving voltage-current differential equations; (4) PMSM Thermal Model with LPTN and full loss computation; (5) Vehicle Dynamics generating real-time Iq_ref. Table 1 summarises all model parameters.

| Parameter | Symbol | Value | Source |
|---|---|---|---|
| Rated power | Pn | 57.5–66.7 kW | Siemens 1PH8137 nameplate |
| Rated current | Iq | 104 A | Siemens 1PH8137 nameplate |
| Pole pairs (4-pole) | p | 2 | T=(3/2)pψfIq=181.4 Nm matches rated [7] |
| Stator resistance at 25°C | $Rs_0$ | 0.035 Ω | Estimated from nameplate [6] |

| Parameter | Symbol | Value | Source |
|---|---|---|---|
| PM flux linkage at 25°C | $\psi f_0$ | 0.58 Wb | Derived from torque equation [7] |
| Copper temp. coefficient | $\alpha$ | 0.00393 /°C | IEEE Std 399-1997 [8] |
| NdFeB remanence coeff. | k | 0.0011 /°C | Kermanizadeh et al. 2024 [9] |
| Thermal resistance | Rth | 0.01 K/W | Calibrated LPTN [10] |
| Thermal capacitance | Cth | 10,000 J/K | Calibrated LPTN [10] |
| IGBT on-state voltage | $Vce_0$ | 1.7 V | IGBT datasheet [12] |
| Switching energy | Esw | 2 mJ, 10 kHz | IGBT datasheet [12] |
| Vehicle mass | M | 1800 kg | Typical EV (VW ID.4 class) |
| Gear ratio | G | 4.0 | Calibrated for WLTP max speed |
| Rolling resistance | Cr | 0.012 | Standard EV tyre [13] |
| Drag coefficient | Cd | 0.28, A=2.4 $m^2$ | Typical EV body [13] |

*Table 1. PMSM and vehicle model parameters with sources.*

## Loss Models

**Copper:** $Rs(T)=Rs_0[1+\alpha(T-25)]$, $Pcu=1.5\times Rs(T)\times iq^2$

**Inverter:** $Vce(T)=1.7\times[1+0.003(T-25)]$, $Pinv=6\times Vce(T)\times(iq/\sqrt{2})+120$ W

**Iron:** $Pfe=kh\times f\times B^2+ke\times f^2\times B^2$ ($kh=2.5\times10^{-3}$, $ke=5\times10^{-7}$)

**Thermal:** $dT/dt=[Ptotal-(T-25)/Rth]/Cth$, $\tau=Rth\times Cth=100$ s

**Efficiency:** $\eta=Pout/(Pout+Ptotal)\times100\%$, $Pout=Te\times\omega m$

## Vehicle Dynamics Model

Real-time motor torque demand is computed from Newton's second law applied to the full vehicle, replacing the previous simplified speed-proportional current mapping:

$$F\_total = M\times(dv/dt) + Cr\times M\times g + 0.5\times\rho\times Cd\times A\times v^2$$

$$T\_motor = F\_total \times r\_wheel / G \qquad Iq\_ref = T\_motor / (1.5 \times p \times \psi f(T))$$

Vehicle acceleration dv/dt is obtained via filtered derivative of the WLTP speed timeseries. Negative torque (braking) is clipped to zero; Iq_ref is saturated at 0–104 A (rated current).

## Results and Discussion

| Metric | Value |
|---|---|
| WLTP Duration | 1800 s |
| Temperature — start / peak / end | 25.0 / 109.1 / 102.1 °C |
| Thermal headroom (Class H = 180°C) | 70.9 °C |
| Overall / Driving mean efficiency | 80.60 / 80.69 % |
| Maximum efficiency | 91.06 % |
| Mean / Peak copper loss | 649.0 / 755.5 W |
| Mean / Peak inverter loss | 951.4 / 1059.3 W |
| Peak / Mean driving torque | 180.0 / 172.5 Nm |
| Inverter-to-copper loss ratio | 1.46× |

*Table 2. WLTP simulation summary — vehicle dynamics model, gear ratio = 4.*

https://doi.org/10.18745/pb.27307

Temperature rose steadily from 25.0°C to a peak of 109.1°C, stabilising at 102.1°C by cycle end — 70.9°C below the Class H thermal limit of 180°C. The thermal time constant $\tau$ = 100 s produces smooth thermal evolution tracking the WLTP load profile. Temperature rise is most rapid during the Extra-High sub-phase (t = 1477–1800 s) where sustained high-speed operation maximises loss generation.

Driving efficiency of 80.69% with a maximum of 91.06% is consistent with published traction PMSM performance under dynamic drive cycles [1, 5]. The small gap between overall (80.60%) and driving (80.69%) efficiency confirms that WLTP standstill phases are brief. Efficiency is lowest during urban Low-speed phases where inverter switching losses dominate at partial load.

Inverter losses (mean 951.4 W) were the dominant loss component, exceeding copper losses (649.0 W) by 1.46×. The fixed switching loss term (P_sw = 120 W) is present throughout the entire cycle, and conduction losses increase with both current and temperature. At 109.1°C, Rs(T) = 0.0466 Ω — a 33% increase from ambient — amplifying copper losses accordingly. Flux linkage at peak temperature reduces to ψf(109.1°C) = 0.525 Wb (−9.5%), reducing the torque constant and confirming the importance of temperature-dependent modelling for accurate torque prediction.

Comparison with the previous temperature-ramp model (25–120°C, fixed rated current) confirms that vehicle dynamics significantly increases predicted copper losses (649 W mean vs 570 W at 25°C fixed) and produces physically representative transient loss profiles, validating the supervisor's recommendation to implement vehicle-level physics for current excitation.

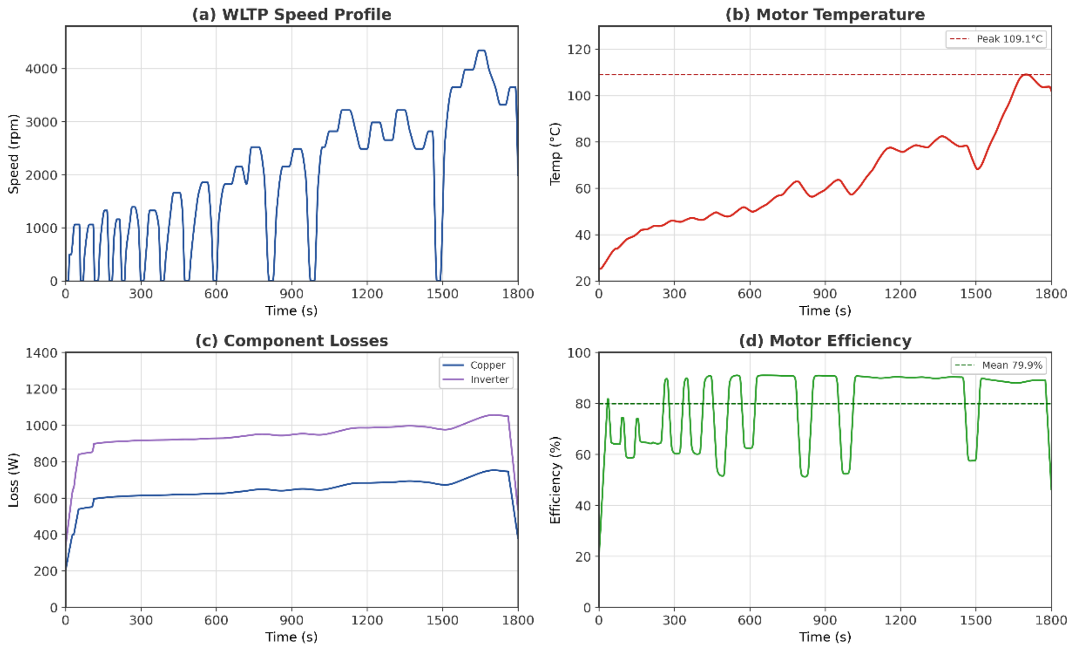



*Figure 1. PMSM performance overview under WLTP: (a) speed profile, (b) motor temperature, (c) component losses, (d) efficiency*

## Conclusion

A temperature-dependent PMSM loss model with vehicle dynamics under the WLTP drive cycle was developed in MATLAB/Simulink for the Siemens SIMOTICS S 1PH8137. Vehicle dynamics (M=1800 kg, G=4, Cr=0.012, Cd=0.28) generates physically realistic real-time Iq_ref, replacing simplified speed-proportional mapping. Key results: peak temperature 109.1°C (70.9°C thermal margin); driving efficiency 80.69%; maximum efficiency 91.06%; inverter losses dominant at 951.4 W mean (1.46× copper losses). The 93°C critical crossover from the temperature-ramp analysis — where resistive losses begin to dominate over flux-weakening effects — remains actionable for thermal management strategy design. Future work will implement regenerative braking and validate against dynamometer test-bench data.

https://doi.org/10.18745/pb.27307

## References

**[1]** A. Shandilya and V. Kumar, SAE Technical Paper 2024-28-0080, 2024. DOI: 10.4271/2024-28-0080

**[2]** S. Sirimanna, T. Balachandran, and K. Haran, Energies, vol. 15, no. 17, p. 6116, 2022. DOI: 10.3390/en15176116

**[3]** Y. Yang et al., Energy, vol. 213, p. 118859, 2020. DOI: 10.1016/j.energy.2020.118859

**[4]** M. R. Dilshad et al., Proc. AEEICB, Chennai, India, 2016. DOI: 10.1109/AEEICB.2016.7538266

**[5]** B. Chen et al., Proc. IMechE Part D, vol. 239, no. 8, pp. 3164–3179, 2024. DOI: 10.1177/09544070241254557

**[6]** M. S. Rafaq and J.-W. Jung, IEEE Trans. Ind. Informat., vol. 16, no. 7, pp. 4747–4758, 2020. DOI: 10.1109/TII.2019.2944413

**[7]** J. F. Gieras, Permanent Magnet Motor Technology, 3rd ed., CRC Press, 2010.

**[8]** IEEE Std 399-1997, IEEE Recommended Practice for Industrial and Commercial Power Systems Analysis, 1997.

**[9]** A. Kermanizadeh, M. T. Mercan, and P. Pillay, Proc. IEEE ITEC, Chicago, IL, 2024. DOI: 10.1109/ITEC60657.2024.10598879

**[10]** A. Kacenka et al., Proc. IEEE IEMDC, San Diego, CA, 2019. DOI: 10.1109/IEMDC.2019.8785176

**[11]** M. A. Azom, M. S. Hossain, and M. Y. A. Khan, Control Syst. Optim. Lett., vol. 3, no. 1, pp. 75–83, 2025. DOI: 10.59247/csol.v3i1.172

**[12]** N. Mohan, T. M. Undeland, and W. P. Robbins, Power Electronics, 3rd ed., Wiley, 2002.

**[13]** M. Ehsani, Y. Gao, S. Longo, and K. Ebrahimi, Modern Electric, Hybrid Electric, and Fuel Cell Vehicles, 3rd ed., CRC Press, 2018.

https://doi.org/10.18745/pb.27307

# Rheological Behaviour of Chitosan Solutions for Controlled Membrane Fabrication in $CO_2$ Separation Applications

**Hugh W. Shaw[1], Anna Ojinnaka[1] and Yuen-Ki Cheong[1*]**

*[1]University of Hertfordshire, School of Physics, Engineering and Computer Science*

*[*]Corresponding author: y.cheong2@herts.ac.uk*

**Membrane-based $CO_2$ separation requires controlled polymer rheology to ensure reliable film formation. Chitosan is a promising sustainable material, but its viscosity varies significantly with physicochemical conditions. This study investigates the effects of concentration, molecular weight and ionic strength on chitosan solution rheology. Results show strong concentration dependence, higher viscosity for high molecular weight samples and viscosity reduction with NaCl due to electrostatic screening. These findings demonstrate that chitosan rheology can be tuned to enable controlled and reproducible membrane fabrication for $CO_2$ separation applications.**

Keywords: Chitosan; Rheology; Membrane fabrication; $CO_2$ separation; Polymer viscosity

## Introduction

The increasing concentration of atmospheric carbon dioxide ($CO_2$) is a major contributor to climate change, driving the need for efficient carbon capture technologies. Membrane-based gas separation has emerged as a promising approach due to its energy efficiency, scalability and operational simplicity compared to conventional methods such as amine absorption [1].

Chitosan, a bio-based polymer derived from chitin, has gained attention for membrane applications due to its film-forming ability and affinity for $CO_2$ through amine functional groups [2]. However, the performance and reproducibility of chitosan membranes are strongly dependent on the rheological behaviour of the polymer solution during fabrication.

Viscosity governs film formation, thickness and structural integrity. However, chitosan solutions exhibit highly variable rheological behaviour influenced by parameters such as molecular weight, concentration and ionic environment [4]-[6]. This variability presents a challenge for controlled membrane fabrication.

This study aims to investigate how key physicochemical parameters affect the rheology of chitosan solutions and to establish relationships between solution behaviour and membrane fabrication performance.

## Experimental

Chitosan solutions with concentrations ranging from 0.5 to 2.0 wt% were prepared by dissolving the high-molecular-weight (HMW 21,370 Da) and the low-molecular-weight (LMW 6,280 Da) samples in 0.3 M acetic acid. To evaluate the effect of the ionic strength, 2 g sodium chloride (NaCl) was added to chitosan solutions.

Rheological measurements were conducted using a rotational rheometer (Anton Paar MCR 703) across a shear rate range of 0.1-100 $s^{-1}$. Shear viscosity was recorded to assess flow behaviour and determine the influence of concentration, molecular weight and ionic strength on solution properties.

## Results and discussion

All chitosan solutions exhibited shear-thinning behaviour, characteristic of polymer systems where viscosity decreases with increasing shear rate [8]. Viscosity increased significantly with concentration, following power-law scaling with exponents of approximately 2.14 for the HMW samples and 2.0 for the LMW samples as shown in **Figure 1**, which is consistent with the reported rheological behaviour of chitosan solutions [5], [6].

https://doi.org/10.18745/pb.27307

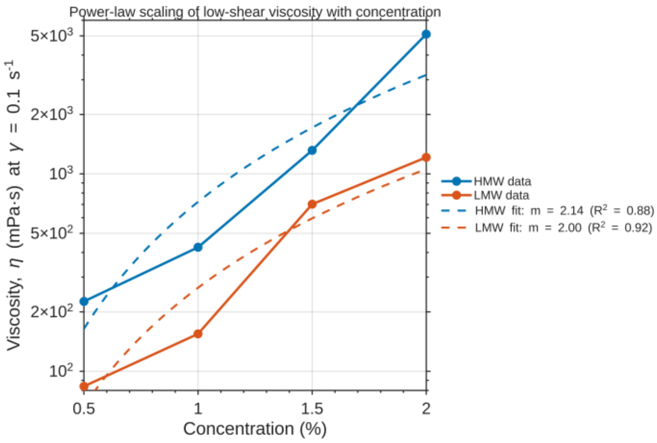



*Fig. 1.* Power-law scaling of low-shear viscosity with polymer concentration for high molecular weight (HMW) and low molecular weight (LMW) chitosan solutions. Viscosity values at $\gamma \approx 0.1\ s^{-1}$ increase strongly with concentration, following power-law relationships with exponents m ≈ 2.14 for HMW chitosan and m ≈ 2.00 for LMW, indicating concentration-dependent enhancement of intermolecular interactions

As shown in **Figure 2**, the HMW chitosan solutions showed consistently higher viscosity compared to LMW samples due to increased chain entanglement and larger hydrodynamic volume. These effects enhance resistance to flow and contribute to improved structural integrity during film formation, as described in polymer physics and intrinsic viscosity relationships [4], [8].

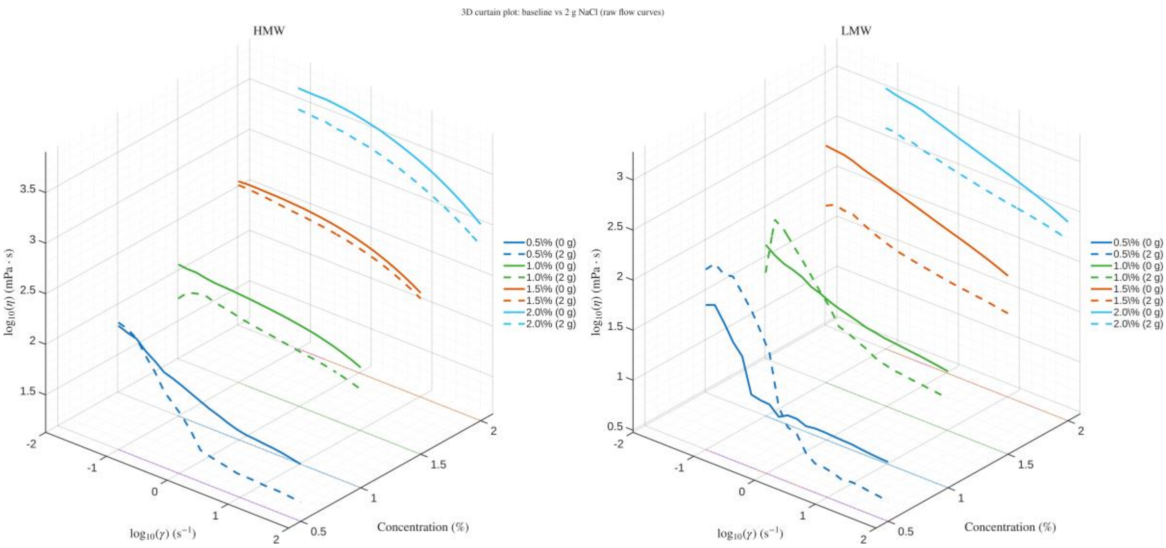



*Fig. 2. Three-dimensional visualisation of viscosity behaviour for high molecular weight (HMW) and low molecular weight (LMW) chitosan solutions as a function of shear-rate and polymer concentration. Baseline solutions and those containing 2 g NaCl are compared, illustrating the combined influence of concentration and ionic strength on chitosan rheology across the investigated shear-rate range.*

The addition of NaCl resulted in a reduction in viscosity across all samples. This behaviour is attributed to electrostatic screening, where ionic interactions screen electrostatic repulsion between polymer chains, resulting in reduced chain expansion and hydrodynamic volume, thereby decreasing resistance to flow and lowering viscosity [6].

These findings demonstrate that chitosan solution viscosity can be systematically tuned through control of concentration, molecular weight and ionic environment, enabling improved control over membrane fabrication processes and ensuring uniform film formation while maintaining processability.

https://doi.org/10.18745/pb.27307

## Conclusion

This study demonstrates that chitosan solution rheology is strongly influenced by concentration, molecular weight and ionic strength. Shear-thinning behavior was observed across all samples, with higher molecular weight and concentration leading to increased viscosity, while NaCl reduced viscosity through electrostatic screening.

This study establishes a framework for controlling chitosan solution rheology, enabling tunable, reproducible and scalable membrane fabrication for $CO_2$ separation applications in industrial carbon capture.

## References

[1] R. Borgohain, "A review on chitosan-based membranes for sustainable $CO_2$ separation applications: Mechanism, issues, and the way forward," Carbohydrate Polymers, vol. 267, 2021, Art. no. 118177.
[2] A.T Celeizabal and C. Coterillo, "Chitosan-based mixed matrix composite membranes for CO2/CH4 mixed gas separation: Experimental characterization and performance validation," Separation and Purification Technology, vol. 324, 2023, Art. no. 124524.
[3] L. A. El-Azzami, "Carbon dioxide separation from hydrogen and nitrogen: Facilitated transport in arginine salt-chitosan membranes," Journal of Membrane Science, vol. 323, no. 2, pp. 225-234, 2008.
[4] M. R. Kasaai, G. Charlet, and J. Arul, "Intrinsic viscosity-molecular weight relationship for chitosan," Biomacromolecules, vol. 1, no.2, pp. 259-263, 2000.
[5] J.K. Hwang and H. H. Shin, "Rheological properties of chitosan solutions," Journal of Food Engineering, vol. 48, no. 2, pp. 153-158, 2001.
[6] P. A. Sobral et al., "Rheological and viscoelastic properties of chitosan solutions prepared with different concentrations," Polymers, vol. 14, no. 3, pp. 1-15, 2022.
[7] M. Rinaudo, "Chitin and chitosan: Properties and applications," Progress in Polymer Science, vol. 31, no. 7, pp. 603-632, 2006.
[8] M. Rubinstein and R. H. Colby, Polymer Physics. Oxford: Oxford University Press, 2003.

https://doi.org/10.18745/pb.27307

# A Study on Key performance impact factors of ALPR

**Jiabao Li[1*], Li Meng[1+], Na Helian[1]**
[1]*School of Physics, Engineering and Computer Science, University of Hertfordshire*
*jl25abp@herts.ac.uk
[+]Corresponding author: l.1.meng@herts.ac.uk

**Automatic License Plate Recognition (ALPR) systems often experience severe performance bottlenecks in unconstrained environments. This study investigates the key factors affecting the Strict Recognition Rate (SRR) of ALPR systems through experimental evaluation on the CCPD dataset using a lightweight OCR baseline. Specifically, three primary factors are examined, which include character complexity, spatial positioning, and environmental conditions. Our experimental results show that character deletion (missed detections) is the dominate source of performance drop, with characters located at the plate edges exhibiting the highest deletion rates. Furthermore, adverse weather and challenging lighting conditions substantially reduce recognition accuracy. These findings expose the vulnerabilities in single-modal OCR systems and emphasise the necessity for more robust spatial rectification frameworks.**

Keywords: Automatic License Plate Recognition; Machine Vision; Character Complexity; Spatial Truncation; OCR Robustness

## Introduction

Automatic License Plate Recognition (ALPR) technology is a cornerstone of modern intelligent transportation systems. Despite significant advances in deep learning, the performance of ALPR systems continues to degrade in unconstrained environments. Adverse factors such as motion blur, extreme illumination, and severe weather compromise the quality of the input plate images, leading to a loss of fine-grained alphanumeric details.
This study investigates the key performance impact factors that contribute to recognition bottlenecks in ALPR systems. Specifically, we present a granular analysis of recognition failures by quantifying the impact of character complexity, the spatial position of characters on the plate, and varying environmental conditions.

## Experiments

To rigorously evaluate model robustness under adverse conditions, experiments were conducted on a curated subset of 5,200 images from the CCPD (Chinese City Parking Dataset) [1]. This subset was split into CCPD-Base (1,200 high-quality images) and CCPD-Challenge (4,000 degraded images exhibiting extreme tilt, motion blur, and adverse weather). The PP-OCRv3 architecture [2] was adopted as the baseline model. Our experimental evaluation utilized a sequence matching algorithm to align predicted strings with ground truth labels, enabling precise measurement of spatial error rates and detailed character-level confusion metrics.

## Results and discussion

Our empirical results demonstrate a severe performance dichotomy under varying imaging conditions, highlighting three key impact factors:
1. Environmental Conditions
The baseline accuracy is highly dependent on environmental conditions. As shown in Table 1, the model achieves an exceptional 97% accuracy on the high-quality CCPD-Base subset. However, the Strict Recognition Rate drops sharply to 13.80% on the CCPD-Challenge subset, which is characterized by motion blur, adverse weather (e.g., rain/snow), and extreme tilt. This confirms that single-modal recognition remains fundamentally vulnerable to severe optical distortions.

*Table 1. Baseline Recognition Performance across Environmental Conditions*

| Dataset Subset | Environmental Condition | Total Samples | Original Baseline Accuracy |
|---|---|---|---|
| CCPD-Base | High-Quality (Clear, Normal Angle) | 1,200 | 97.00% |
| CCPD-Challenge | Degraded (Blur, Rain/Snow, Extreme Tilt) | 4,000 | 13.80% |
| Overall (Total) | Mixed Scenarios | 5,200 | 32.65% |

https://doi.org/10.18745/pb.27307

2. Complexity of the Individual Characters

Further analysis reveals that character complexity directly influences the dominant failure mode. As illustrated in Fig. 1, character deletion ("Missed" detections) rather than conventional OCR substitution is the primary source of performance failure with degraded images. Such character deletion often occur when characters with dense strokes (e.g., "W") merge into unrecognizable blobs, or when simpler geometries (e.g., "J", "L") blend into plate borders or environmental artefacts such as rain streaks, thereby failing to meet detection thresholds.

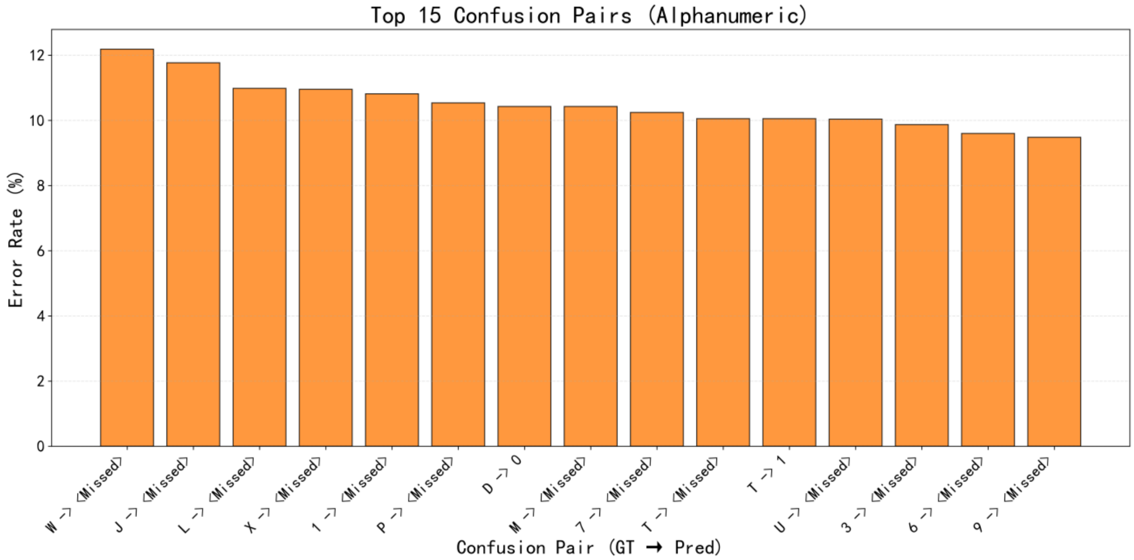



*Fig. 1. Top 15 Most Frequent Alphanumeric Confusion Pairs highlighting character deletion*

3. Position of the Character

Spatial position emerges as a critical vulnerability factor. A granular spatial analysis of character-level error rates reveals a pronounced spike in errors at the first and last positions. Characters in these locations are particularly susceptible to distortion from motion blur and physical cropping, causing their features to blend with the background noise and significantly increasing the likelihood of recognition failure.

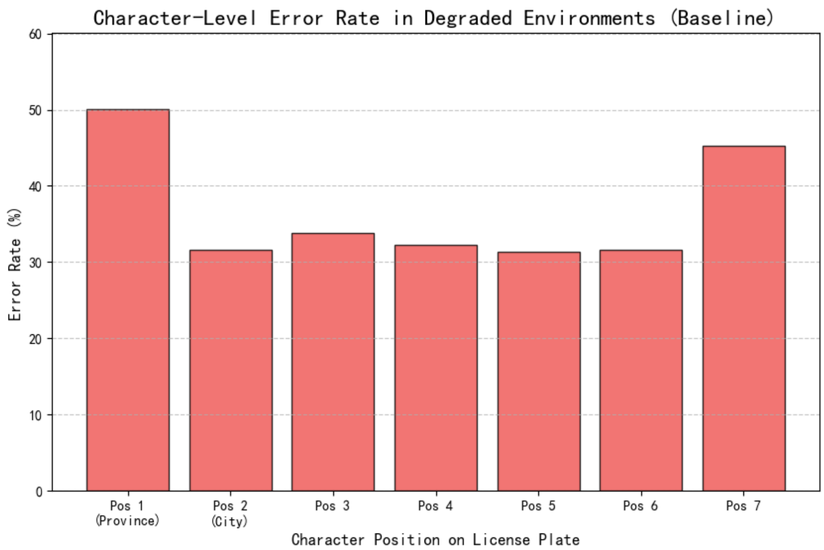



*Fig. 2. Spatial distribution of character-level error rates in degraded environments*

## Conclusion

This study confirms that the performance of ALPR systems in unconstrained environments is severely affected by environmental degradation, character complexity, and character position. To address these bottlenecks, future work should explore adaptive multi-modal frameworks [3] that fuse multiple models, each tailor trained

https://doi.org/10.18745/pb.27307

for a specific performance challenge. In particular, attention-based fusion mechanisms can be employed to combine the strengths of these models, enabling them to operate collaboratively and improve robustness across a wide range of real-world conditions.

## References

[1] Z. Xu, W. Yang, A. Meng, N. Lu, H. Huang, C. Ying, and L. Huang, "Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline," Proceedings of the European Conference on Computer Vision (ECCV), 2018.
[2] C. Li, et al., "PP-OCRv3: More Attempts for the Improvement of Ultra Lightweight OCR System," arXiv preprint arXiv:2206.03001, 2022.
[3] Z. Lan, Z. Li, C. Yan, X. Xiang, D. Tang, H. Zhou, and J. Lai, "Adaptive knowledge distillation with attention-based multi-modal fusion for robust dim object detection," IEEE Trans. Multimedia, vol. 27, pp. 2083-2096, 2025.

https://doi.org/10.18745/pb.27307

# Parametric Study of Aluminium Single-Lap Adhesive Joint Using Cohesive Zone Modelling

**Muhammad Awais Awan[1]*, Bhishm Dewangan[1], Yiding Liu[1]**

[1] *School of Physics, Engineering and Computer Science, University of Hertfordshire, UK*

**Corresponding Author:* mawaisawan572@gmail.com

## Abstract

**The paper gives a parametric study through a simulation on aluminium single-lap adhesive joints with Cohesive Zone Modelling (CZM) in ABAQUS. The effect of adhesive thickness, the length of overlap, the thickness of adherend and length of adherend, and the external taper on the load and displacement of joint failure in adhering Aluminium 6082-T6 adherends were investigated using a validated 2D finite element model of adherends bonded with Araldite 2015 adhesive. All geometric parameters were tested individually while keeping other parameters constant to isolate their respective effects. The model used traction separation laws including cohesive elements to simulate the initiation and propagation of damages in adhesive layer. Findings indicate that the most important strength increase (21.02% ), adherend thickness changes (5.64%), external taper optimisation (3.22%), and adhesive thickness increase (1.88%), but the adherend length did not have a significant effect on the same (2.13%), beyond a certain length. The work presents geometry configurations that enhance the stress distribution and postpone edge crack propagation to provide a viable design rules to lightweight bonded structures with high performance in both aerospace and automotive.**

**Keywords:** Cohesive Zone Modelling; Adhesive joint; ABAQUS; Fracture Mechanics; Failure load

## Introduction

Adhesive bonding is becoming popular in lightweight structural applications as it allows effective transfer of loads, minimizes stress concentrations that occur in mechanical fasteners, and permits bonding of thin or dissimilar materials[1]. The behaviour of bonded joints is, however, very sensitive to geometry of the joints and local stress concentrations especially in single-lap designs where peel and shear stresses control failure. This work aims at aluminium single-lap adhesive joints and examines the influence of geometric changes on the strength of the joints and deformation behaviour. The objective of the work was to derive and test an ABAQUS Cohesive Zone Modelling (CZM) 2D finite element model, and conduct a parametric study in a systematic manner in order to determine the effect of adhesive thickness[2], overlap length[3], adherend thickness[4], [5], adherend length[1], and external taper[6], [7]. The aim was to find out which design parameters were the most influential and offer practical suggestions on how to optimise the joint performance and strike a balance between strength and flexibility.

## Experimental/Simulation

It was a numerical simulation that was carried out in ABAQUS/CAE 2024. An Aluminium 6082-T6 adherends and Araldite 2015 adhesive were used to develop a 2D finite element model of a single-lap joint. The model with the base was made of a total joint length of 100 mm, literature-based material properties and overlap length of 25 mm. Continuum plane strain elements (CPE4R) were used to model the adherends of aluminium and cohesive elements (COH2D4) governed by a traction separation law were used to model the adhesive layer. The initiation of adhesive damage was characterized by nominal stress criteria and the development of damage was monitored by fracture energy parameters in Mode I and Mode II. The mesh was a structured quadrilateral with a global seed of 0.2 mm, and in the tipping region, the mesh was refined to resolve stress gradient and crack formation. A tensile case of lap-shear loading was replicated using the boundary conditions with one end fixed and the other end displaced. Five parametric variations were tested by varying adhesive thickness (0.20-0.30

https://doi.org/10.18745/pb.27307

mm), overlap length (20-30 mm), adherend thickness (1-3 mm), adherend length (90-110 mm) and external taper angle (60° - 30°). Load-Displacement behaviour, peak failure load, stress contours and evolution of cohesive damage were determined in each instance.

## Results and discussion

The published literature was found to align well with the validated baseline model, which validated the applicability of the CZM-based model to predict damage initiation and failure in bonded joints. Figure 1 shows the stress contour chart for the base configuration. The length of overlap was the most influential of all parameters examined: making the overlap longer (25 mm to 30 mm) led to the greatest change in failure load (maximum 21.02 per cent), owing to the increased bonded area and more even load distribution as shown in Figure 2. Thickening adhesive layer to 0.30 mm gave a 1.88% rise in failure load and a decrease in the displacement at failure, which demonstrates that the adhesive increases the stiffness and less localised stress concentration. Adherend thickness was also a significant factor in the study; it was found that the study experienced a 5.64% variation in the failure load due to the change in adherend thickness, which revealed that adherend stiffness is a strong factor in bending resistance and distribution of loads. A further increase in the adherend length to 110 mm did not significantly improve the number, by just 2.13 percent, suggesting that the effect becomes secondary above a certain point. External taper optimisation enhanced the behaviour of joints by minimising edge stress concentration: a taper angle of 30° was found to improve the failure load by 3.22% and delay premature crack initiation, but at the cost of a slightly higher displacement. The percentage changes in failure load for all tested parameters are as follows: adhesive thickness (0.30 mm: +1.88%), overlap length (30 mm: +21.02%), adherend thickness (3 mm: +5.64%), adherend length (110 mm: +2.13%), and external taper angle (30°: +3.22%) relative to their respective baseline configurations. Altogether, the findings indicate that geometry can play a significant role in determining the strength and the flexibility and that the best way to achieve performance improvement is to focus on the length of the overlap and the local management of stress at the adhesive edges.

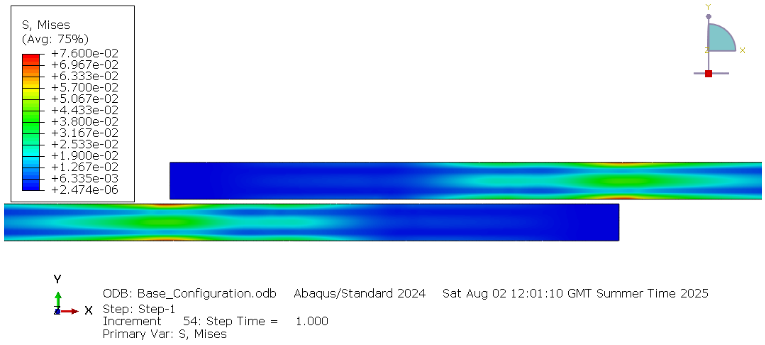



*Figure 8 Stress Contour Chart for Base Configuration*

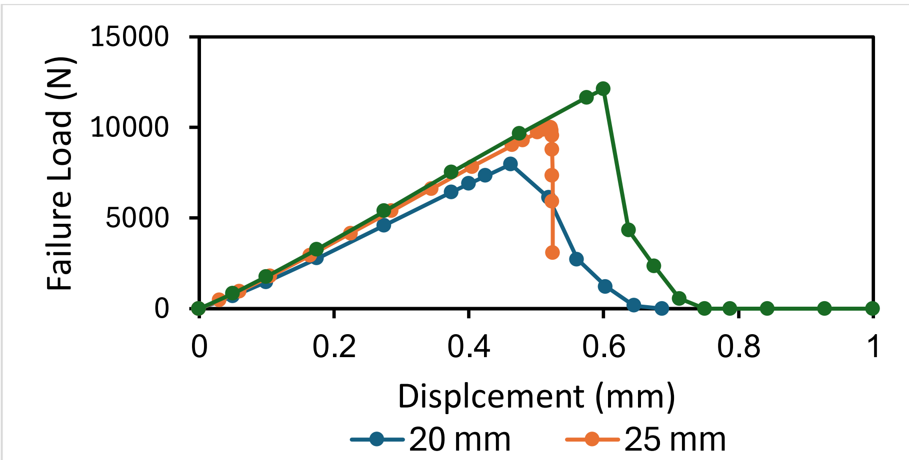



*Figure 9 Failure Load of Overlap Length (overall best performing)*

https://doi.org/10.18745/pb.27307

## Conclusion

This research has shown that Cohesive Zone Modelling in ABAQUS is a useful modelling tool to understand the failure mechanisms of aluminium single-lap adhesive joints. The parametric experiment proves that the joint geometry does directly and measurably affect the failure load and displacement. Overlap length was the most significant design variable with adherend thickness followed by external taper, adhesive thickness, and adherend length was of limited use beyond a practical level. The analysis provides concise design recommendations to enhance bonded joint performance through expansion of bonded area, maximization of stiffness ratio and minimization of end-of-overlap stress concentrations. These results are directly applicable to the design of lightweight bonded structures in aerospace, automotive and general engineering applications, and a solid basis on future research that includes experimental validation, fatigue behaviour, or extension to composite and hybrid adherend structures.

## References


[1] C. G. Dávila, P. P. Camanho, and A. Turon, "Cohesive Elements for Shells," 2007. [Online]. Available: http://www.sti.nasa.gov

[2] A. K. Salve and S. N. Jalwadi, "Implementation of Cohesive Zone in ABAQUS to Investigate Fracture Problems."

[3] "26.5.3 Modeling with cohesive elements." [Online]. Available: https://classes.engineering.wustl.edu/2009/spring/mase5513/abaqus/docs/v6.6/books/usb/default.htm?startat=pt06ch26s05alm40.html?utm_source=

[4] Y. Liu, S. Lemanski, X. Zhang, D. Ayre, and H. Y. Nezhad, "A finite element study of fatigue crack propagation in single lap bonded joint with process-induced disbond," *Int. J. Adhes. Adhes.*, vol. 87, pp. 164–172, Dec. 2018, doi: 10.1016/j.ijadhadh.2018.10.005.

[5] H. Quan and R. Alderliesten, "On the effect of plastic model on simulation of adhesive bonded joints with FM94," *Int. J. Adhes. Adhes.*, vol. 110, Oct. 2021, doi: 10.1016/j.ijadhadh.2021.102916.

[6] R. D. S. G. Campilho, M. D. Banea, J. A. B. P. Neto, and L. F. M. Da Silva, "Modelling adhesive joints with cohesive zone models: Effect of the cohesive law shape of the adhesive layer," *Int. J. Adhes. Adhes.*, vol. 44, pp. 48–56, 2013, doi: 10.1016/j.ijadhadh.2013.02.006.

[7] Z. Gu, Y. Liu, D. J. Hughes, J. Ye, and X. Hou, "A parametric study of adhesive bonded joints with composite material using black-box and grey-box machine learning methods: Deep neuron networks and genetic programming," *Compos. B Eng.*, vol. 217, Jul. 2021, doi: 10.1016/j.compositesb.2021.108894.

https://doi.org/10.18745/pb.27307

# Effects of Slots on Frame on the Performance of Millimeter-Wave Dual-Polarized Antenna

**Preethy Sethuraman*, Dr. Qi Luo[2]**
[1]*University of Hertfordshire*
[2]*University of Hertfordshire*
*Corresponding author: p.sethuraman@herts.ac.uk

**This study investigates the impact of slot configurations on the performance of a millimeter-wave dual-polarized stacked patch antenna integrated within a smartphone platform. The presence of metallic frames degrades antenna gain and radiation characteristics. To mitigate these effects, horizontal and vertical slots with varying positions and dimensions were introduced. Simulation results show that slot height and placement significantly influence bandwidth, impedance matching, and radiation behaviour. Middle slot configurations combined with ±45° polarization demonstrate improved gain, reduced reflections, and dual-band characteristics. The findings provide design insights for optimizing antenna performance in compact mobile devices.**

Keywords: Millimeter-wave antenna; slot antenna; dual polarization; smartphone integration; antenna optimization

## Introduction

The development of millimeter-wave (mmWave) antennas for 5G mobile devices has introduced significant design challenges due to compact device constraints. The presence of metallic components such as frames, ground planes, and back covers leads to reduced antenna gain, impedance mismatch, and distorted radiation patterns [1].

Previous studies have demonstrated that antenna performance degrades significantly when integrated into realistic mobile platforms, primarily due to electromagnetic coupling with surrounding structures [1]. One effective technique to improve performance is the introduction of slots in the metallic frame, which can act as secondary radiators and enhance radiation efficiency [2].

This study investigates the influence of slot configurations on antenna performance, focusing on slot orientation, position, and polarization. The objective is to identify optimal configurations that improve gain, bandwidth, and radiation characteristics in practical smartphone environments.

## Simulation Setup

A stacked patch antenna designed for mmWave operation was analysed under two polarization configurations:

- Horizontal/Vertical (H/V) polarization
- ±45° polarization

Slots were introduced in the metallic frame in two orientations:

- Horizontal slots
- Vertical slots

Each configuration was further divided into:

- **Half Slot:** Extending from the centre towards the edge
- **Middle Slot:** Located at the centre of the frame

The Slot placement on Honor mobile is shown in Fig.1.

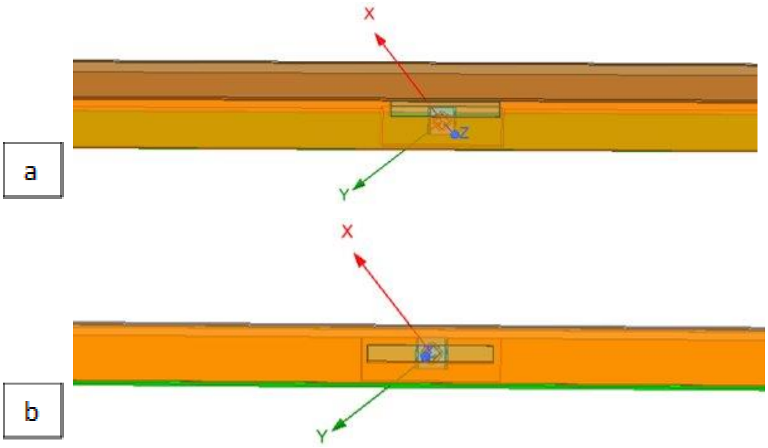



*Fig. 1. Types of Horizontal Slots (a) Half Slot (b)Middle Slot*

Slot height was varied while maintaining a constant slot width. Full-wave electromagnetic simulations were performed to evaluate reflection coefficients (S-parameters), gain, radiation patterns, and electric field (E-field) distribution. The simulations included realistic mobile device components to ensure practical relevance.

## Results and Discussion

### 3.1 Effect of Slot Height

Slot height significantly influences antenna performance. At lower slot heights, strong reflections from the metallic frame reduce gain and limit bandwidth. As slot height increases, coupling between the antenna and slot improves, resulting in enhanced radiation and reduced reflections. Effective radiation is observed beyond a threshold height, where the slot begins to behave as a secondary radiator [2]. The effect on 3D radiation pattern on 45/-45 polarised antenna is shown in Fig.2.

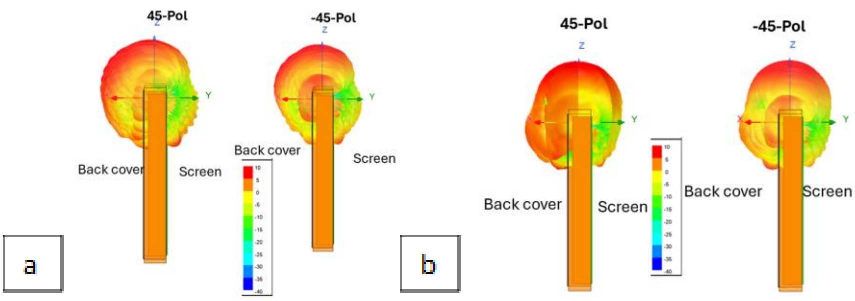



*Fig. 2. 3D Radiation Pattern of 45°/-45°polarised antenna with Slot Height=4.5mm at 30GHz (a)With Half Slot (b)Middle Slot.*

### 3.2 Slot Configuration Comparison

A comparison between half slot and middle slot configurations shows that:

- Middle slots provide higher bandwidth and improved gain
- Dual-band behaviour is achievable in middle slot configurations
- Enhanced coupling occurs due to exposure of the central radiating patch

The central patch plays a critical role in radiation, and its interaction with the slot significantly improves overall antenna performance.

### 3.3 Effect of Polarization

The ±45° polarization demonstrates superior performance compared to H/V polarization. It shows:

- Reduced impedance mismatch

- Lower reflections from the frame
- More stable radiation characteristics

This behaviour aligns with previous findings that ±45° polarization provides more balanced interaction with surrounding metallic structures [3]. In contrast, H/V polarization is more sensitive to slot dimensions, particularly slot height [4].

**3.4 Radiation and Field Behaviour**

Electric field analysis indicates that increasing slot height strengthens field coupling between the antenna and the slot. Middle slot configurations produce stronger and more uniform field distributions, leading to improved radiation efficiency. Additionally, optimized slot placement reduces reflections from the frame and improves radiation pattern stability.

## Conclusion

This study demonstrates that slot integration in metallic frames is an effective approach to improving mmWave antenna performance in smartphone platforms. Among the configurations analysed, the middle slot combined with ±45° polarization provides the best overall performance.

Key findings include:

- Slot height is critical for achieving efficient radiation
- Middle slot configurations outperform half slots in bandwidth and gain
- ±45° polarization reduces reflections and improves stability
- Slot placement directly influences radiation characteristics

These results provide practical guidelines for designing high-performance antennas for next-generation mobile devices.

## References

[1] R. Rodriguez-Cano, S. Zhang, K. Zhao, and G. F. Pedersen, “mm-wave beam-steerable endfire array embedded in a slotted metal-frame lte antenna,” IEEE Transactions on Antennas and Prop agation, vol. 68, no. 5, pp. 3685–3694, 2020.
[2] W.-C. Hsiao, H.-S. Huang, C.-Y. Ho, S.-C. Hsieh, and C.-C. Wang, “A 2x2 broadband dual polarized antenna array using aip techniques for 5g mmwave beamforming systems,” in 2022 17th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IM PACT), 2022, pp. 1–4.
[3] X. Xia, F. Wu, C. Yu, Z. H. Jiang, R. Lu, Y. Yao, and W. Hong, “Millimeter-wave ± 45°dual lin early polarized end-fire phased array antenna for 5g\b5g mobile terminals,” "IEEE Transactions on Antennas and Propagation", vol. 70, no. 11, pp. 10391–10404, 2022.
[4] R. Montoya Moreno, J. Ala-Laurinaho, A. Khripkov, J. Ilvonen, and V. Viikari, “Dual-polarized mm-wave endfire antenna for mobile device

https://doi.org/10.18745/pb.27307

# A Combined-Indicator Model for Detecting Workflow Bottlenecks and Delivery Delays in Agile Software Development.

**Hanifat Owuri[1]*, Mr. Kelechi Emerole[1]**
[1]*University of Hertfordshire, United Kingdom*
*Corresponding author: ho24aas@herts.ac.uk

## Abstract

**This study investigates bottleneck detection in Agile software development workflows using Jira-based process data. While Agile methodologies emphasise continuous delivery and flow efficiency, teams often experience delays caused by hidden bottlenecks within workflow stages. Existing approaches typically rely on either queue-based indicators, which capture work accumulation, or waiting-time indicators, which measure delays, but rarely combine both.**

**To address this gap, this study proposes a lightweight, reproducible Python-based analysis pipeline that operates on standard Jira exports without specialised tools. The findings demonstrate that integrating queue-based and waiting-time indicators improves bottleneck detection, enabling more reliable analysis and supporting proactive workflow optimisation in Agile teams.**

**The approach was evaluated on Apache Kafka Jira workflow data and successfully identified Code Review as the primary delivery bottleneck.**

<u>Keywords</u>: Agile Software Development; Bottleneck Detection; Workflow Analysis; Jira Data; Delivery Performance

## Introduction

Agile software development emphasises continuous delivery and rapid workflow progression; however, teams frequently experience hidden workflow bottlenecks that reduce delivery efficiency and predictability [3]. Existing bottleneck detection approaches often rely on either queue-based indicators, which capture work accumulation, or waiting-time indicators, which measure delay, while rarely integrating both within a unified analytical framework [1], [7], [8].

To address this limitation, this study proposes a lightweight combined-indicator bottleneck detection model using Jira workflow data from the Apache Kafka open-source project [2]. Apache Kafka was selected because it provides a large-scale, publicly accessible Jira dataset with detailed workflow transition histories, making it suitable for reproducible bottleneck analysis in a realistic Agile development environment.

The approach analyses workflow stages using waiting time, throughput, work-in-progress, lead time, and cycle time metrics computed from Jira status transition histories. Workflow bottlenecks are identified through comparative stage analysis and normalized severity scoring, while cumulative flow analysis is used to distinguish capacity constraints from temporary queue accumulation.

The model was evaluated using a filtered 12-month Jira dataset comprising 2,148 issues and 526 workflow transition records. Results identified Code Review as the primary bottleneck, characterised by low throughput, prolonged waiting time, and strong correlation with overall lead time. Correlation analysis demonstrated that prolonged time spent in Code Review was strongly associated with increased delivery delay, validating its impact on workflow performance.

https://doi.org/10.18745/pb.27307

Overall, the findings demonstrate that integrating queue-based and waiting-time indicators improves bottleneck detection reliability while supporting reproducible and practically applicable workflow analysis for Agile software teams.

## Experimental/Simulation

This study employs a quantitative case study design using Jira workflow data from the Apache Kafka project. The dataset was filtered to a 12-month period (November 2023 – November 2024), resulting in 2,148 issues and 526 status transitions, ensuring both temporal relevance and sufficient sample size for analysis.

Workflow stages were defined by mapping Jira statuses into Backlog, Development, Code Review, and Done based on observed transition behaviour.

The analysis computes workflow metrics including waiting time, throughput, work-in-progress, lead time, and cycle time across mapped workflow stages. The combined-indicator model mathematically represents bottleneck severity by normalising workflow metrics onto a common 0–1 scale and comparatively ranking workflow stages according to relative bottleneck intensity. Bottlenecks are identified by combining queue-based and waiting-time indicators, while correlation analysis is used to evaluate the relationship between detected bottlenecks and delivery performance outcomes.

All analysis was conducted in Python, and the complete analytical pipeline is available on GitHub to support reproducibility.

## Results and Discussion

Results identified Code Review as the primary workflow bottleneck, characterised by approximately 45% lower throughput than Development and a mean waiting time exceeding 15 days. Severity scoring consistently ranked Code Review as the most constrained active stage. Correlation analysis further demonstrated a strong positive relationship between Code Review duration and overall lead time, indicating that issues spending longer in review experienced significantly delayed delivery. Cumulative flow analysis additionally suggested that the bottleneck reflected a capacity constraint rather than temporary queue accumulation, due to sustained low throughput and limited work progression.

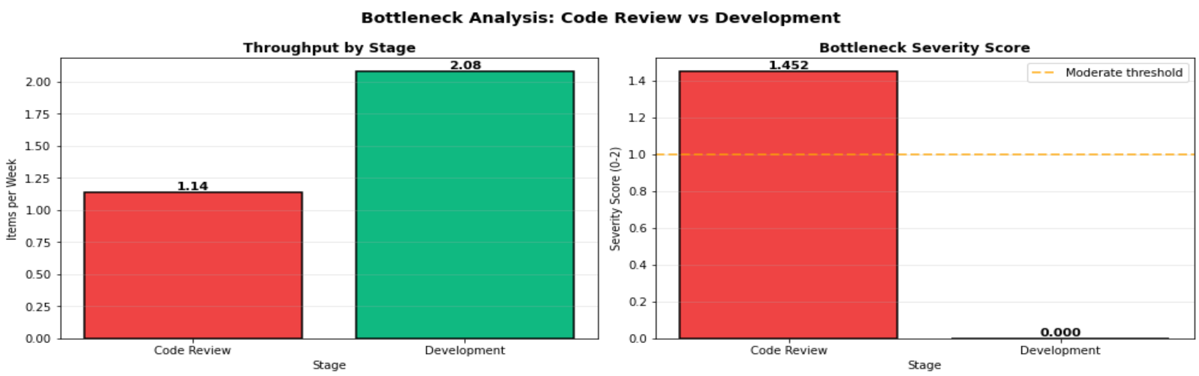



*Fig.1. Comparative bottleneck severity analysis across active workflow stages using normalized throughput, waiting time, and work-in-progress indicators.*

https://doi.org/10.18745/pb.27307

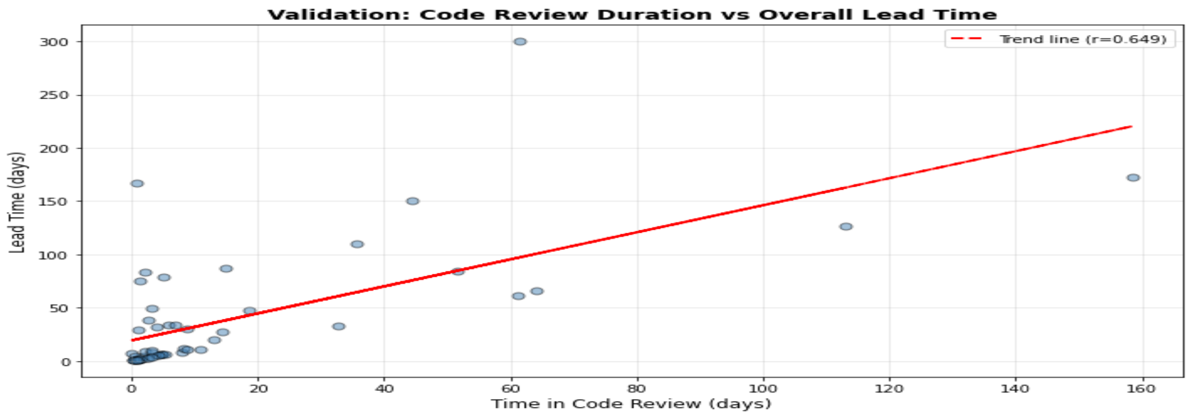



*Fig.2. Correlation between Code Review waiting time and overall lead time, demonstrating the impact of workflow bottlenecks on delivery performance*

## Conclusion

This study presented a combined-indicator approach for identifying workflow bottlenecks in Agile software development using Jira event log data. By integrating waiting-time and throughput-based indicators within a unified detection algorithm, the approach enables systematic and objective identification of workflow inefficiencies.

Applied to the Apache Kafka dataset, the model identified the Code Review stage as the primary bottleneck, characterised by low throughput (1.14 items/week), high mean waiting time (15.15 days), and sustained work-in-progress accumulation. In contrast, the Development stage exhibited stable flow characteristics and was classified as a normal stage. These findings were supported by converging evidence from throughput and waiting-time analysis.

The results demonstrate that bottleneck conditions can be quantitatively detected and linked to delivery performance, while also highlighting the potential of early indicators, such as rising queue size and waiting time, to signal emerging workflow constraints. Practically, these findings suggest that Agile teams should prioritise optimising code review processes, for example by improving reviewer availability to enhance overall delivery flow.

## References

[1] F. S. Abdullah, S. Riyanarto, and R. S. Kelly. “Identifying bottlenecks of hierarchical process model by using discrete event simulation and process mining,” *2022 IEEE 8th International Conference on Computing, Engineering and Design (ICCED)*, Sukabumi, Indonesia, pp. 1-6, 2022.

[2] Apache Software Foundation, “Apache Kafka JIRA Issue Tracking Data,” 2024. [Online]. Available: https://issues.apache.org/jira/projects/KAFKA

[3] T. Dingsøyr, S. Nerur, V. Balijepally, and N. B. Moe, “A decade of agile methodologies: Towards explaining agile software development,” Journal of Systems and Software,Volume 85, Issue 6, 2012, pp 1213-1221, 2012.

[4] K. Elvan, E. G. Eric, A. van der Hoek, and G. Gousios, “Factors affecting on-time delivery in large-scale agile software development,” in *IEEE Transactions on Software Engineering*, vol. 48, no. 9, pp. 3573-3592, 2022,

[5] P. Jirat, J. Fan, T. Kemp, F. Arik, N. Fenton, and I. Kwan, “Improving agile planning for reliable software delivery,” *2023 IEEE/ACM 20th International Conference on Mining Software Repositories (MSR)*, Melbourne, Australia, pp. 25-26, 2023.

https://doi.org/10.18745/pb.27307

[6] P. Kai and C. Wohlin, “Measuring the flow in lean software development,” Pract. Exper. 00, pp. 1–7, 2011.

[7] A. R. Muhammad, R. S. Kelly, S. Riyanarto, and S. W. Cahyaningtyas, “Detecting bottleneck and fraud in agile development by using Petri net performance and trace clustering,” *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, Indonesia, pp. 214-218, 2019.

[8] R. Nemesio, R. S. Kelly, S. Riyanarto, and S. W. Cahyaningtyas, “Detection of bottleneck and social network in business process of agile development,” *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, Indonesia, pp. 208-213, 2019.

https://doi.org/10.18745/pb.27307

# Characterisation of Chitinous Polymers for Antibiofilm Applications

**Holly Goodall[1*], Yuen-Ki Cheong[1*] and Anna Ojinnaka[1]**
[1]*University of Hertfordshire, School of Physics, Engineering and Computer Science*
*Corresponding author: h.goodall@herts.ac.uk and y.cheong2@herts.ac.uk

**Systematic review focused on the novel approach to extracting chitinous polymers from black soldier flies *(Hermetia illucens)* using a deep eutectic solvent (DES) method, and the antibiofilm properties of chitosan for various applications. This project also involves the characterisation of chitinous polymers, highlighting the differences between industry standard polymers extracted from crustaceans using harsh chemicals, and polymers extracted from black soldier fly. Analytical studies include morphological studies of chitin and chitosan micrographs and Ubbelohde viscosity experiments to calculate the molecular weight of chitosan. Furthermore, preliminary studies showed the feasibility of forming microspheres using chitosan for potential antibiofilm applications.**

Keywords: Chitin; Chitosan, Black Soldier Fly (BSF), Deep eutectic solvent (DES); Antibiofilm properties

## Systematic review

Applications of chitinous polymers has been demanding increasing attention in various fields including water treatment and biomedical industries due to their biodegradable, biocompatible and non-toxic properties[1]. Using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)[2] approach, screening was carried out using results from the Elsevier database from the last 15 years (2011-2026). The review firstly addresses the green method of using deep eutectic solvents (DES) to extract chitin from black soldier fly (BSF), as an alternative to the traditional method which uses harsh chemicals to extract chitin from crustaceans. 109 papers were screened to reveal 46 relevant papers and analysis showed that the most successful DES in terms of chitin purity and yield is the combination of choline chloride as the hydrogen bond acceptor (HBA) and lactic acid as the hydrogen bond donor (HBD). For BSF extraction from the pupae stage of the life cycle, a successful paper reported that the use of this solvent produced a chitin purity of 93.10% and yield of 93%, in addition to a clear porous, honey-comb structure observed by scanning electron microscope (SEM) [3]. The review also discusses the antibiofilm properties of chitosan, and its applications to treat or prevent biofilms in various contexts. This Elsevier search gave 631 papers, of which 243 were deemed relevant. Papers discussed chitosan's antibiofilm properties is a variety of contexts including applications in biomedical industries such as drug delivery [4] and treatment for chronic wounds [5]. Studies also discussed chitosan's antibiofilm properties in dentistry, to eliminate oral biofilms [6] and for water treatment applications [7].

## Methodology

Industry standard chitin and chitosan (purchased from Sigma Aldrich) is extracted from crustaceans using the traditional harsh acid and base method of extraction, with hydrochloric acid, sodium hydroxide and hydrogen peroxide. Experimental BSF samples include chitin extracted from BSF pupae using both industrial and DES methods. The subsequent chitosan samples are obtained from refluxing the chitin in aqueous sodium hydroxide (NaOH) at various conditions, depending on the required deacetylation degree.

*Morphological characterisation of polymer micrographs:*

A Keyence VHX-7000 Digital Microscope, and a JEOL JSM-IT210 Scanning Electron Microscope are used to characterise and compare the surface morphology of the chitinous polymers; crustacean derived chitin, BSF derived chitin, and BSF derived chitosan.

*Viscosity measurements and molecular weight calculation of Chitosan:*

Viscosity measurements are carried out using an Ubbelohde capillary viscometer on four concentrations of each industry standard chitosan sample: low molecular weight (LMW) and high molecular weight (HMW). To prepare the chitosan samples for measurement, 0.4g of each chitosan is dissolved in separate tubes containing 200mL of buffer (0.3M acetic acid and 0.2M sodium acetate[8]) giving a final concentration of 0.20g/dL. Following complete dissolution, serial dilutions are carried out to give the following concentrations for each sample;

https://doi.org/10.18745/pb.27307

0.05g/dL, 0.10g/dL. 0.15g/dL, 0.20g/dL. For each concentration, 20mL is transferred into the filling tube of the viscometer. A syringe connected to silicone tubing is then attached to the capillary tube of the viscometer and used to gently draw the liquid to sit above the upper calibration mark. Following detachment of the syringe, the flow time is measured as the time taken for the liquid to pass from the upper line to the lower line. This is repeated 3 times for accuracy, and a mean is calculated and recorded as the solvent flow time ($t_0$). The solution flow time (t) is measured in the same way using the blank buffer with no chitosan. Both parameters are then used to calculate the intrinsic viscosity of the samples. The Mark Houwink equation is rearranged, using the intrinsic viscosity (η) and Mark Houwink constants (K and a) appropriate to the specific system to calculate the relative molecular weight (M) of each chitosan sample in g/mol (Da): $M = \left(\frac{[\eta]}{K}\right)^{1/a}$[9]

*Formation and characterisation of chitosan microspheres*:
Chitosan microspheres are formed by ionic cross-linking. This is carried out by chitosan being dissolved in either acetic acid or salicylic acid, then adding the solutions drop-by drop to a slow stirring mixture solution containing tripolyphosphate (TPP) and sodium hydroxide (NaOH) in a 15:5 ratio [10]. Microsphere samples are then washed and air dried to be analysed.

## Results

*Morphological characterisation of polymer micrographs:*

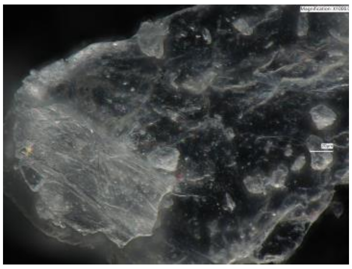
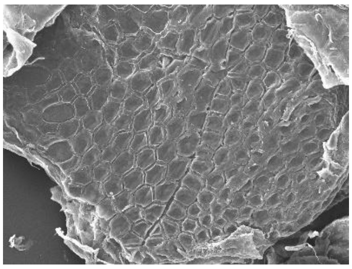
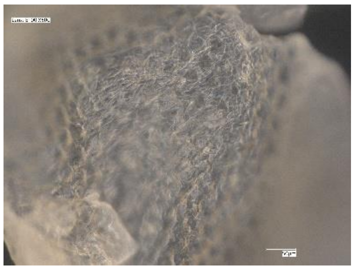
A)  B)  C) 

**Figure 1:** *A) Industry standard chitin extracted from crustaceans using chemical method. x1000 magnification Keyence digital microscope. B) BSF chitin extracted from pupae using chemical method. x200 magnification scanning electron microscope (SEM). C) BSF chitosan obtained from NaOH reactions. High degree of deacetylation, low molecular weight. x500 magnification Keyence digital microscope.*

Micrographs of the different sources of chitinous polymer show distinct morphological differences. The industry standard, crustacean derived chitin exhibits a fragmented morphology with irregular compressed flakes whereas BSF chitin possesses a distinct porous, honey-comb like structure. This honeycomb struture remains in the BSF chitosan sample following the NaOH reactions used to produce chitosan from chitin.

*Viscosity measurements and molecular weight calculation of Chitosan:*

**Industry standard crustacean chitosan, low molecular weight:**

*Intrinsic viscosity (η) = 0.57 dL/g      Mark Houwink constants for this system[2]: K = 0.74x10$^{-3}$dL/g    a= 0.76*

Calculation: $\frac{0.57}{0.74\times10^{-3}} = (770.27)^{\frac{1}{0.76}} = 6283.69 g/mol$      M(LMW Chitosan) = 6.28kDa

**Industry standard crustacean chitosan, high molecular weight:**

*Intrinsic viscosity (η) = 1.445 dL/g      Mark Houwink constants for this system[2]: K = 0.74x10$^{-3}$dL/g    a= 0.76*

Calculation: $\frac{1.445}{0.74\times10^{-3}} = (1952.702)^{\frac{1}{0.76}} = 21{,}369.01 g/mol$    M(HMW Chitosan) = 21.37kDa

*Formation and characterisation of chitosan microspheres*:
The formation of chitosan microspheres from crustacean derived chitosan was successfully performed using acetic acid and salicylic acid. Micrographs show differentiation in sphere properties between the two solvents including differences in surface texture, shape and colour.

https://doi.org/10.18745/pb.27307

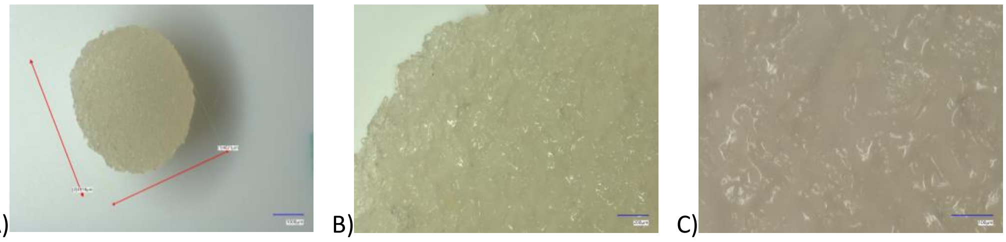



**Figure 2.1:** Micrographs of chitosan microspheres formed using salicylic acid taken on Keyence digital microscope. *A) x30 magnification with measurement bars. B) x150 magnification C) x400 magnification*


**Figure 2.2:** Micrographs of chitosan microspheres formed using acetic acid taken on Keyence digital microscope. *A) x30 magnification with measurement bars. B) x150 magnification C) x400 magnification*

There are distinct differences between the microspheres created using the two different solvents. The salicylic acid microspheres exhibit a more spherical, uniform shape, whereas the acetic acid microspheres showed a spherical body with small tails. Surface morphology also differed between the two solvents, with salicylic acid microspheres showing a rough, uneven surface and the acetic acid microspheres demonstrating a very smooth surface. Additionally, there was a difference in colour as the salicylic acid microspheres exhibited a yellow tone and the acetic acid microspheres appeared more white in colour. This project will go on to explore the formation of microspheres using BSF chitosan in the same solvents, and the use of a 3D printed microfluidic channel to standardise the size and shape of the microspheres to later functionalise for antibiofilm studies and investigate the potential for various applications, including biomedical and environmental applications.

## **References**:

[1]D. Chicea and A. Nicolae-Maranciuc, 'A Review of Chitosan-Based Materials for Biomedical, Food, and Water Treatment Applications', Dec. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*.

[2]M. J. Page *et al.*, 'The PRISMA 2020 statement: An updated guideline for reporting systematic reviews', Mar. 29, 2021, *BMJ Publishing Group*.

[3]R. D. Sharma and P. Thangadurai, 'Sustainable bioprocessing of Hermetia illucens exuviae into structurally defined chitooligosaccharides within a circular biorefinery framework', *Bioresour. Technol. Rep.*, vol. 32, p. 102400, Dec. 2025

[4]J. B. de Souza *et al.*, 'Ceftazidime and Usnic Acid Encapsulated in Chitosan-Coated Liposomes for Oral Administration against Colorectal Cancer-Inducing Escherichia coli', *Pharmaceuticals*, vol. 17, no. 6, Jun. 2024

[5]B.-Y. Lim, F. Azmi, and S.-F. Ng, 'LL37 Microspheres Loaded on Activated Carbon-chitosan Hydrogel: Anti-bacterial and Anti-toxin Wound Dressing for Chronic Wound Infections', *AAPS PharmSciTech*, vol. 25, no. 5, p. 110, May 2024

[6]E. M. Costa, S. Silva, A. R. Madureira, A. Cardelle-Cobas, F. K. Tavaria, and M. M. Pintado, 'A comprehensive study into the impact of a chitosan mouthwash upon oral microorganism's biofilm formation in vitro', *Carbohydr. Polym.*, vol. 101, pp. 1081–1086, Jan. 2014

[7]A. M. Youssef Moustafa, M. M. Fawzy, M. S. Kelany, Y. A. Hassan, R. F. M. Elsharaawy, and F. H. A. Mustafa, 'Synthesis of new quaternized chitosan Schiff bases and their N-alkyl derivatives as antimicrobial and anti-biofilm retardants in membrane technology', *Int. J. Biol. Macromol.*, vol. 267, p. 131635, May 2024

[8]N. D. Sanandiya, C. Ottenheim, J. W. Phua, A. Caligiani, S. Dritsas, and J. G. Fernandez, 'Circular manufacturing of chitinous bio-composites via bioconversion of urban refuse', *Sci. Rep.*, vol. 10, no. 1, Dec. 2020

[9]S. Koltzenburg, M. Maskos, and O. Nuyken, *Polymer Chemistry* , 1st ed. 2017.

[10]H. N. Cuong, N. D. Kien, and N. Van Hoa, 'Chitosan-based aerogel microspheres for wastewater treatment', *Vietnam Journal of Chemistry*, vol. 62, no. 5, pp. 615–623, Oct. 2024

https://doi.org/10.18745/pb.27307

# Advanced functionalisation of Borophene/ GCN composites as photocatalyst for textile wastewater treatment application

**SeyedehSadrieh Emadian[1,2*], Ajay Kumar[1], James Bowen[1], Irwing Ramirez[1], Silvia Varagnolo[1], Sudhagar Pitchaimuthu[3] and Satheesh Krishnamurthy[1*]**

[1]*School of Engineering and Innovation, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK*
[2]*School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK*
[3]*School of Engineering & Physical Sciences, Institute of Mechanical, Process & Energy Engineering, Heriot-Watt University, Riccarton, UK*
*Corresponding author: s.emadian@herts.ac.uk

**Keywords: 2D materials, photocatalysis, surface engineering**

## Introduction

Water connects every aspect of life and plays a pivotal role in human health and well-being. Industrial wastewater pollution presents a significant challenge, not only in access to clean water but also through contamination of terrestrial ecosystems. Approximately 20% of global industrial water pollution originates from textile manufacturing [1]. Therefore, the development of low-cost, environmentally friendly materials and efficient treatment systems is essential to address this issue.

Photocatalysis is a promising technology that utilizes the UV–visible spectrum of solar radiation to break down organic pollutants. Two-dimensional (2D) materials offer unique properties that allow researchers to tune layered structures for enhanced performance. Among photocatalytic materials, graphitic carbon nitride (GCN) has gained significant attention due to its high stability, suitable electronic properties, atomically thin structure, and mechanical strength [2]. However, its performance is limited by rapid charge recombination, low visible-light absorption, limited active sites, and slow surface reaction kinetics.

To overcome these limitations, heterojunction systems have been widely explored to improve charge separation and enhance photocatalytic efficiency [2,3]. Borophene, a newly emerging 2D material, has attracted considerable interest as a potential co-catalyst due to its unique physical and chemical properties [4]. In this study, borophene was successfully synthesized and plasma-functionalized, and its electronic, optical, and catalytic properties were investigated using borophene/GCN composites for wastewater remediation.

## Experimental Method

Borophene was synthesised via liquid exfoliation using different solvents to optimise layer morphology. Graphitic carbon nitride (GCN) was prepared and subsequently combined with borophene to form composite materials. The structural and morphological properties of borophene were analysed using Scanning Electron Microscopy (SEM) and Transmission Electron Microscopy (TEM), revealing variations in layer thickness and atomic arrangement. In addition, a custom-built, low-cost LED-based photocatalytic reactor was developed to evaluate the photocatalytic performance of the composites. Indigo carmine wastewater was used as the model pollutant to assess degradation efficiency under visible-light irradiation.

## Results

SEM and TEM observations further confirmed the layered morphology and provided insight into the atomic arrangement of the material (Figure 1). Morphological analysis confirmed successful exfoliation of borophene

https://doi.org/10.18745/pb.27307

with solvent-dependent structural variations. Atomic Force Microscopy (AFM) analysis indicated that the exfoliated sheets possess thicknesses in the few-nanometre range. The borophene/GCN composites exhibited significantly enhanced photocatalytic performance compared to pure GCN, attributed to improved charge separation and increased active sites due to heterojunction formation.

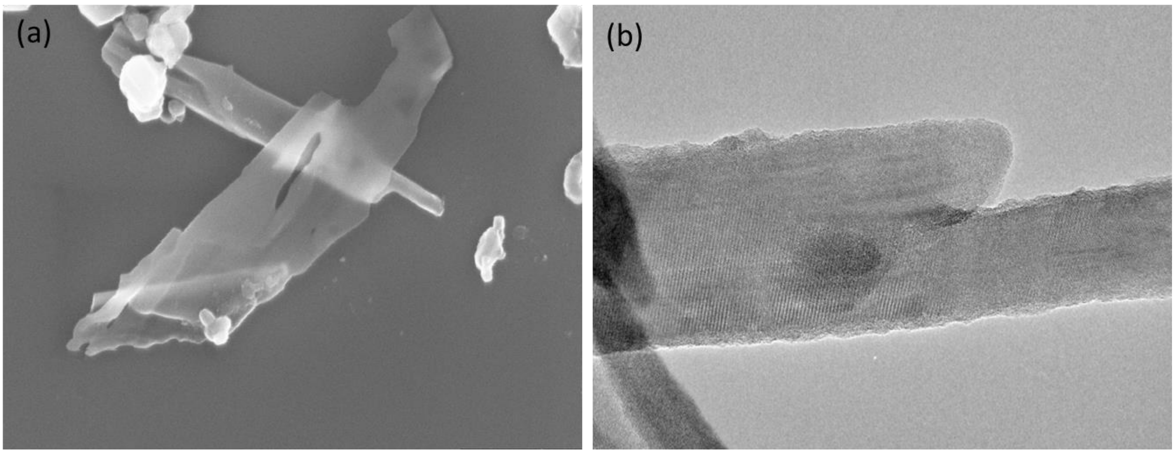



***Figure 10.*** *a) SEM b) TEM images of borophene sheets*

## Conclusion

Borophene/GCN composites show significant potential as efficient photocatalysts for wastewater treatment. The improved performance is due to heterojunction formation and defect engineering, offering a promising pathway for sustainable environmental remediation.

## References

[1] Yue, C., Chen, L., Zhang, H., Huang, J., Jiang, H., Li, H. and Yang, S. Metal-organic frameworks-based materials: Emerging catalysts of high-efficiency for heterogeneous photocatalytic degradation of pollutants in water. Environmental Science: Water Research & Technology, 2023.
[2] Kumar, A., Majithia, P., Choudhary, P., Mabbett, I., Kuehnel, M.F., Pitchaimuthu, S. and Krishnan, V.. MXene coupled graphitic carbon nitride nanosheets based plasmonic photocatalysts for removal of pharmaceutical pollutants. Chemosphere, 308, p.136297, 2022.
[3] Wang, L., Wang, K., He, T., Zhao, Y., Song, H. and Wang, H., 2020. Graphitic carbon nitride-based photocatalytic materials: preparation strategy and application. ACS Sustainable Chemistry & Engineering, 8(43), pp.16048-16085, 2020.
[4] Chen, Y., Cheng, M., Lai, C., Wei, Z., Zhang, G., Li, L., Tang, C., Du, L., Wang, G. and Liu, H. The Collision between g-C3N4 and QDs in the Fields of Energy and Environment: Synergistic Effects for Efficient Photocatalysis. Small, p.2205902, 2023.

University of Hertfordshire UH
https://doi.org/10.18745/pb.27307

# Methanol's Hidden Memory: Reading the Thermal History of Interstellar Ices from a Single Spectrum

**Marta Podgórny[1*]**
[1]*SPECS, University of Hertfordshire, Hatfield, United Kingdom*
*Corresponding author: m.podgorny@herts.ac.uk

**Methanol ice is abundant in space, but its temperature-dependent structure remains poorly understood. This work systematically investigates solid methanol and water/methanol ices (20–130 K) using IR spectroscopy. Results show that deposition temperature uniquely determines the crystalline structure of pure methanol ice, with phase transitions at 102 K. Presence of water molecules in the ice system further modifies the structure and brings new spectral signatures. Presented Principal Component Analysis of IR spectra enable detailed analysis, including determination of thermal history with ~1.2 K accuracy, based on single spectrum, providing a novel tool for interpreting observational data.**

Keywords: solid methanol; temperature-dependent structure; infrared spectroscopy; Principal Component Analysis; astrophysical ices

## Introduction

Methanol ($CH_3OH$) is a key organic molecule in star-forming regions, serving as a precursor to more complex species. Despite its astrochemical importance, the structural and thermal properties of solid methanol remain poorly characterised. Reported phase transition temperatures are inconsistent across the literature, and although previous studies suggest that deposition temperature influences methanol's spectral properties [1], this effect has not been systematically explored. These limitations directly impact the interpretation of astronomical observations, including those from the James Webb Space Telescope, where spectral fitting often relies on initial assumptions and existing databases, leading to degeneracy in the analysis [2]. This work systematically investigates the effects of deposition temperature and its interactions with water on the structure and spectral properties of methanol ice under astrophysically relevant conditions. Additionally, Principal Component Analysis (PCA) is evaluated as a tool to reduce degeneracy and enable more detailed interpretation of spectral data.

## Experimental/Simulation

Five ice systems were investigated: pure methanol, binary mixtures with water (1:1 and 10:1 $H_2O$:$CH_3OH$), and two layered configurations in which the components were deposited sequentially. Ices were prepared under ultra-high vacuum (UHV, ~$10^{-6}$ mbar) using the Portable Astrochemistry Chamber (PAC) [3,4], with deposition carried out over a temperature range of 20–130 K. Precise control of substrate temperature (±1 K) enabled systematic variation of initial growth conditions. The structural evolution of the ice samples was monitored in situ using infrared (IR) spectroscopy, allowing changes in vibrational band positions, shapes, and relative intensities to be tracked as a function of temperature and thermal processing.

A novel experimental protocol was implemented, combining high-temperature deposition with subsequent isothermal holding, enabling direct observation of structural relaxation and phase evolution without continuous temperature ramping. To support interpretation of the large spectral dataset, Principal Component Analysis (PCA) was applied as an automated, data-driven approach. This enabled retrospective reconstruction of the thermal history of the samples and identification of subtle spectral signatures associated with phase transitions and molecular reorganisation.

## Results and discussion

The phase transition of pure methanol was identified at 103 K during annealing and 102 K during direct deposition. Notably, direct deposition above the transition temperature results in the formation of distinct crystalline structures that differ from those obtained via thermal annealing. These high-temperature phases

https://doi.org/10.18745/pb.27307

exhibit clear temperature-dependent spectral variations, particularly in the O–H and C–O stretching regions, indicating differences in hydrogen-bonding environments and molecular ordering.

The presence of water significantly modifies methanol's structural behaviour. In systems containing amorphous water, the methanol crystallisation temperature is elevated (e.g., up to 110 K in layered configurations), suggesting inhibition of molecular mobility and delayed reorganisation. In contrast, crystalline water promotes earlier methanol crystallisation, lowering the transition temperature to ~99 K, likely due to templating effects or enhanced surface diffusion. Moreover, in mixed ices, high-temperature deposition promotes partial segregation of methanol and water, whereas in layered systems it induces interfacial mixing, resulting in a stable structure that retains both segregated and mixed characteristics.

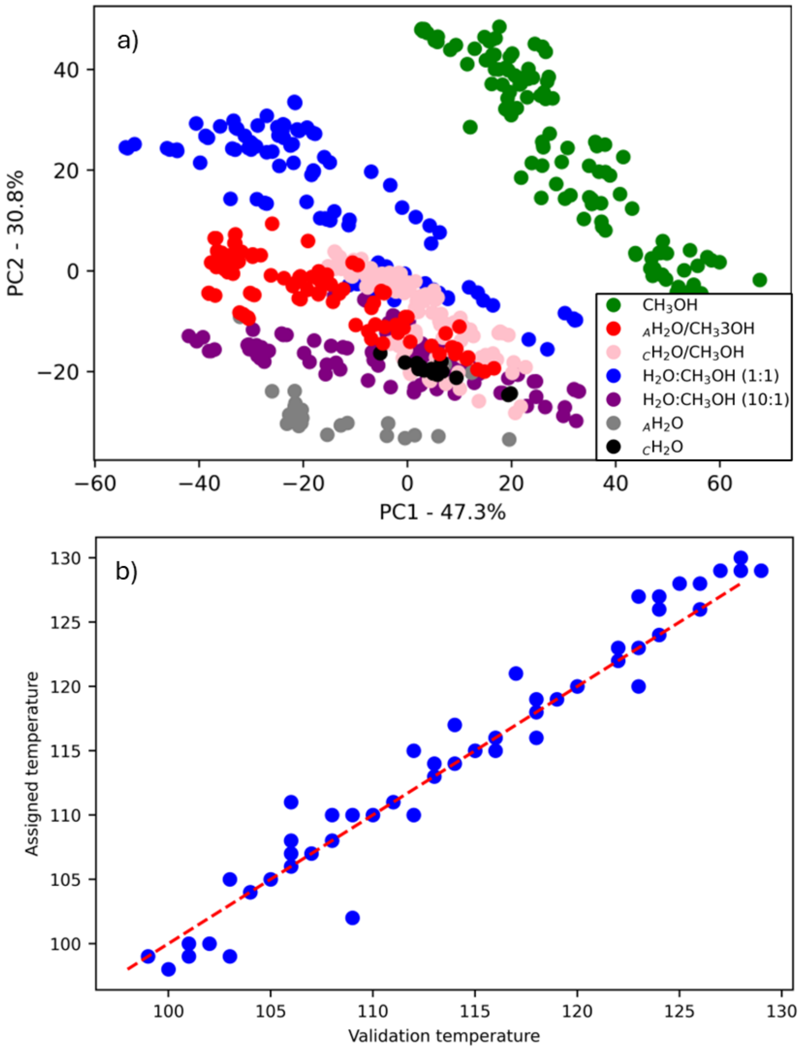



*Fig. 1. PCA model of all investigated ice systems based solely on the O–H stretching region (2930–3520 cm$^{-1}$): a) 2D scores plot (PC1 vs. PC2) showing clear separation of pure methanol, mixed (1:1 and 10:1 $H_2O$:$CH_3OH$), layered (amorphous and crystalline water substrates), and reference water ices. The first two principal components capture 78.1% of the total variance; b) validation plot comparing actual deposition temperatures (x-axis) with model-predicted temperatures (y-axis) for independent test spectra. The dashed line (y = x) represents perfect prediction.*

The $CH_3$ rocking mode was confirmed as a robust and sensitive spectral marker for distinguishing methanol phases. Application of Principal Component Analysis (PCA) enabled effective classification of ice structures and reconstruction of their thermal history. Notably, models based solely on the O–H stretching region achieved high predictive accuracy, with a mean temperature error of ~1.2 K, demonstrating that even spectrally congested regions contain sufficient information to reliably decode thermal processing pathways (Fig. 1.).

https://doi.org/10.18745/pb.27307

## Conclusion

This work establishes a systematic, high-resolution framework for understanding the structural evolution of methanol ice, directly addressing longstanding inconsistencies in reported phase behaviour. It demonstrates that both deposition temperature and interactions with water critically influence the resulting ice structure, leading to distinct and spectroscopically identifiable phases. These findings highlight that structurally different forms of methanol ice can arise under conditions relevant to interstellar environments, even within a narrow temperature range.

The successful application of Principal Component Analysis (PCA), including models based solely on the broad and traditionally ambiguous O–H stretching region, demonstrates a powerful new approach for extracting physical information from complex spectra. This approach has strong potential for astronomical applications, where overlapping features often limit direct interpretation, and may enable the resolution of spectral degeneracies as well as retrospective reconstruction of the thermal history of interstellar ices.

## References

[1] R. Luna, G. Molpeceres, J. Ortigoso, M. A. Satorre, M. Domingo, and B. Maté, “Densities, infrared band strengths, and optical constants of solid methanol,” *Astronomy & Astrophysics*, vol. 617, p. A116, 2018.

[2] M. K. McClure, W. Rocha, K. Pontoppidan, N. Crouzet, L. E. Chu, E. Dartois, T. Lamberts, J. Noble, Y. Pendleton, and G. Perotti, “An Ice Age JWST inventory of dense molecular cloud ices,” *Nature Astronomy*, vol. 7, no. 4, pp. 431–443, 2023.

[3] M. Podgórny and A. Dawes, “A temperature study of methanol interstellar ice analogues,” *Progress*, no. 12, pp. 88–102, 2023.

[4] M. Podgórny, “Investigating the role of solid phase methanol in space environments,” Ph.D. dissertation, The Open University, 2025.

https://doi.org/10.18745/pb.27307

# Classifying the Variability of Young Stellar Objects

**Anushka Shinde[1*], Philip Lucas[1], and Michael Kuhn[1]**
[1]*University of Hertfordshire, Centre for Astrophysics Research*
*Corresponding author: a.shinde2@herts.ac.uk

**Brightness variations of Young Stellar Objects (YSOs) yield valuable information on the structure of newborn solar systems on size scales too small to image directly. Star spots and accretion and extinction changes create complex variability patterns. Our project analyses 10-years of infrared time series data on 30,000 YSOs from the VISTA Variables in the Via Lactea (VVV) survey, probing the very early "Class I" evolutionary stage. By applying advanced statistical methods to a larger sample with a longer time baseline than previous infrared studies, we aim to gain a deeper understanding of the physics of solar system assembly.**

Keywords: Young stellar objects; Variability; Time domain surveys; Infrared (IR) observations; Star formation

## Introduction

Young stellar objects (YSOs) are stars in the earliest stages of a star's evolution. Star formation starts from the gravitational collapse of dense regions in molecular clouds. To conserve angular momentum, the collapsed clump rotates faster and flattens into a disc, at the centre of which a protostar is assembled. YSOs are categorised via the stage of evolution. Class 0 is the earliest stage, in which the protostar accumulates most of its mass. Class I corresponds to protostars with circumstellar discs and a surrounding infalling envelope. In Class II systems the envelope has dispersed and in Class III systems the disc becomes optically thin as the remaining matter is accreted, dispersed or bound into planets. A "flat spectrum" (FS) category refers to sources between Class I and II, i.e. more evolved than Class I but retaining a residual envelope [1].

Variability in YSOs can be caused by diverse phenomena that include stellar brightness changes from the star spots, inhomogeneities in the circumstellar disk, variable extinction, accretion events from the disk onto the star and more. These appear as changes in stellar magnitudes ranging from as low as 0.05 magnitudes to 2+ magnitudes and over hours to decades [2].Due to the nature of these occurring over a large time frame, large time domain surveys are required that take repeat observations of these stars over years to study all possible changes both on short- and long- term timescales. Majority of past work on variability has focused on optical observations of short duration. We aim to advance on this by analysing over a decade of infrared observations of a sample of 30,000 YSO candidates. The data are drawn from the VVV survey [2] which observed the Milky Way bulge and southern disc in the near infrared (0.8 – 2.5 µm). The 30,000 YSOs were selected by cross matching the VVV star catalogue with the Spitzer/IRAC Candidate YSO (SPICY) catalogue of 118,000 YSO candidates [3].

Previous literature has classified variability into categories such as periodic, quasi-periodic and aperiodic, where aperiodic includes stochastic, bursts, dips, outbursts and fades, all corresponding to different physical processes [4]. For example, periodic variability is thought to arise from cool magnetic or hot accretion spots like sunspots. Aperiodic on the other hand is thought to come from phenomena such as flares, irregular accretion of new material onto stellar surface and variable extinction. Shorter timescale variability generally corresponds to processes occurring from the stellar photosphere or close to the inner circumstellar disk. Longer timescales are more likely associated with the accretion disc physics [4]. Some larger accretion events might arise from the disc caused by instabilities that could include gravitational or thermal instability, turbulence due to magnetic fields and interactions with companions. These different processes present themselves as complex light curve data. Currently, most of the work focused on smaller samples, enabling a manual checking of categories as confirmation. However, with this sample of 30,000 sources the project is working on different methods to try and classify these in a more automated way.

https://doi.org/10.18745/pb.27307

## Methods

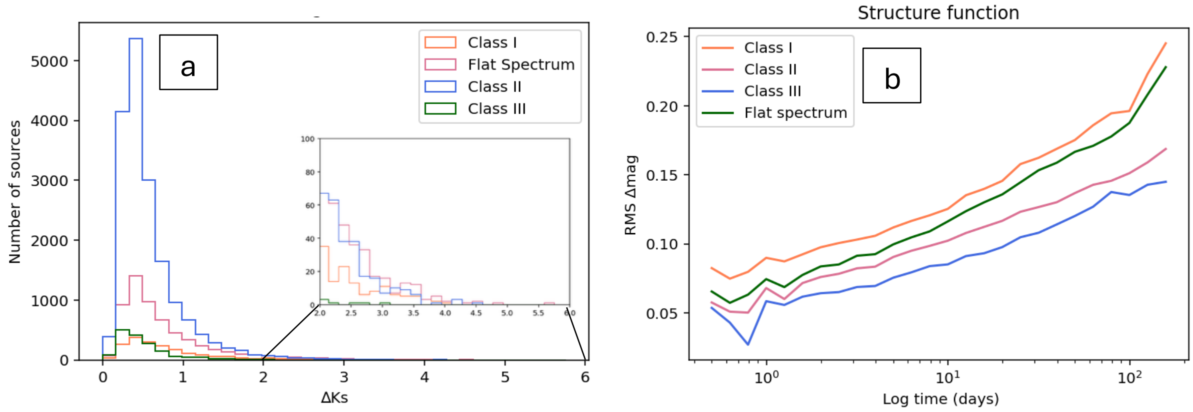
To study these light curves there is high importance for ensuring well cleaned light curves for better classification. The VVV survey was matched within 0.5 arcseconds of the SPICY sources. Light curves of the data were cleaned to keep only Ks band (2.14μm) observations. To remove anomalies and artefacts, the light curve was binned in 0.5-magnitudes interval and then into 0.1-day bins to remove points with large errors or magnitudes using the method of inverse variance weighting and sigma and median absolute clipping.

As the sample contains both non-variable sources and variable sources, the Stetson J parameter was used to distinguish variability from measurement error. This is used to measure the correlation between the brightness changes between pairs of observations by calculating normalised residuals from the mean, where a higher value indicates more fluctuation hence more variable. To set the parameter threshold, VVV data for non-variable sources, selected using data from the Gaia satellite, were used to define the 90th percentile value in the distribution. To examine the broad variability scale for the sources in different classes, the structure function was used, which measures how the brightness changes over different time intervals $(\Delta t)$ by measuring the mean square variation in $\log(\Delta t)$ bins.

To identify periodic sources that are and find periods, a Fourier transform designed for unevenly sampled data, the Lomb Scargle method, is currently being used. The period is then used to compute the Q index parameter [5] which measures the ratio of the mean square residuals to the fit to the total variance of the light curve. The aim is to use parameters like these and others such as the M index, amplitude changes, etc. and to apply them to a clustering algorithm to create a more automated way to classify the different sources. This will then give a sample of sources that are of interest to follow up for comparison with mid-infrared light curves from the WISE satellite to study further the physics behind what's happening in these sources.

## Results and discussion

Figure 1 shows the earlier classes overall have a larger amplitude and ranges of variability timescales shown through subfigure (a) where the histogram distribution extends beyond 3.5 mag for Class I – II. Panel (b) of the structure function shows a higher rms offset in the Class I and FS curves, indicating larger variability and the steepness shows a larger proportion of variability on long timescales. This is expected as earlier stages with protostars being embedded in its envelope leads to more complex interactions between the protostar and the environment, including infalling material from the envelope and circumstellar disks. Whereas, with later stages most of the variability is primarily factors like star spots or stochastic accretion instabilities. A variety of stellar variabilities on different scales and times are shown in Figure 2. This gives an insight into how much the variability can differ between classes, with all the samples either from FS or Class II stage. The variability often shows complex behaviour that is not trivial to classify. A sophisticated automated method will be needed to characterise and classify the behaviours in our large sample. In certain light curves, variability is shown for example on a long timescale but potentially also on shorter time periods, which will be investigated in the future to find ways of separating both.



https://doi.org/10.18745/pb.27307

*Fig. 1. (a) Histogram showing the distribution of the total K band variability for all stars in the different YSO classes. The inset panel zooms into the large amplitude tail of the distribution. (b) The structure function (see text above) indicating the variability on timescales for the different YSO classes.*

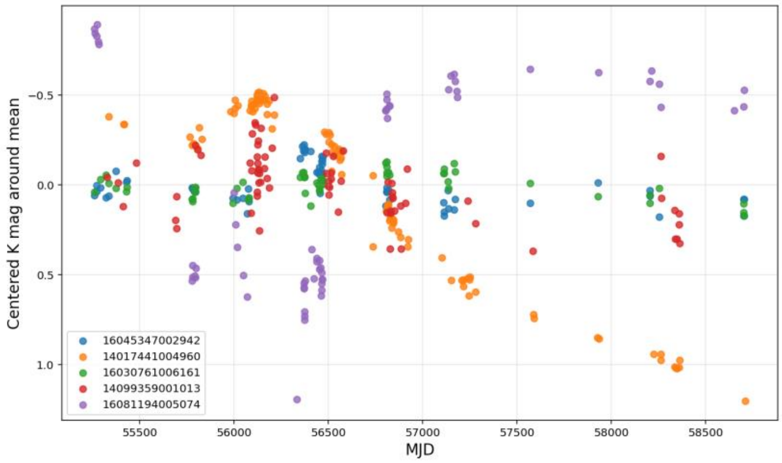



*Fig.2. Light curves of the data for a sample of 5 sources, where the K mag differences are normalised around the mean. The figure shows the large variety in both scale and magnitudes variability just across 5 different random samples.*

## Conclusion

The first stages of the project involving cleaning and selection of data have given insights into the different scales of the variability in the data. Ongoing work is currently focusing on identifying and defining the parameters that can be used to characterise the variability and then classify the large sample of sources using automated clustering methods. This will be followed by further studies with other surveys at different wavelengths, to deepen our understanding on the physics behind this variability for different timescales by using this large sample.

## References

[1] I. Gezer, G.Marton, J. Roquette. "SED Modelling of Young Stellar Objects in the Orion Star Formation Complex" *Astronomy & Astrophysics* 696, A196. issn: 0004-6361, 1432-0746, 2025

[2] L. C. Smith, P. W. Lucas, S. Koposov., 'VIRAC2: NIR astrometry and time series photometry for 500M+ stars from the VVV and VVVX surveys', *Monthly Notices of the Royal Astronomical Society*, vol. 536, no. 4, pp. 3707–3738, 2025

[3] M. A. Kuhn , R. de Souza, A. Krone-Martins, 'SPICY: The Spitzer/IRAC Candidate YSO Catalog for the Inner Galactic Midplane', *ApJS*, vol. 254, no. 2, p. 33, 2021

[4] W. Fischer , L. A Hillenbrand, G . J Herczeg "Accretion Variability as a Guide to stellar mass assembly" *PP VII ASP conference series,* vol 534, p. 355, 2023.

[5] A. M. Cody, J. Stauffer, A. Bagglin, "CSI 2264 : Simultaneous Optical and Infrared Light curves for young disk-bearing stars in NGC2264 with CoRoT and Spritzer – Evidence for multiple origins of variability" *AJ,* vol 147, 4, no. 4, p. 82, 2014.

https://doi.org/10.18745/pb.27307

# Simulation of Recrystallization Using MICRESS: Effect of Shielding Time in a Phase-Field Framework

**Ayaz Hussain[1*], Nusrat Tamanna[1]**

*[1]University of Hertfordshire, Hatfield, UK*

*Corresponding author: ayazh806@gmail.com

**Short Abstract: Static recrystallisation was investigated using MICRESS multiphase-field simulations, with focus on shielding time as a nucleation-control parameter. A two-dimensional isothermal parameter study was performed in which shielding time was varied while other model conditions were kept fixed. Shorter shielding times promoted earlier nucleation, faster early-stage recrystallisation and finer grain structures. Increasing shielding time delayed nucleation and increased the mean grain radius. Beyond about 50 s, further increases had little additional effect, indicating a saturated shielding regime. The results show that shielding time strongly influences predicted recrystallisation kinetics and microstructural evolution within the phase-field framework.**

Keywords: recrystallisation; phase-field modelling; MICRESS; shielding time; microstructure evolution

## Introduction

Recrystallisation is an important restoration process in plastically deformed metals because it replaces a highly strained microstructure with new strain-free grains and therefore strongly affects properties such as strength, ductility and formability [1], [2]. The process is driven by the reduction of stored energy associated with deformation, which is commonly related to dislocation density, while the resulting grain structure depends on the balance between nucleation and grain boundary migration [1], [2].

Predicting recrystallisation remains challenging because the transformation depends on local heterogeneity in stored energy, nucleation sites and boundary behaviour. For this reason, spatially resolved modelling approaches are widely used to investigate both recrystallisation kinetics and microstructural evolution. Cellular automata, Monte Carlo and phase-field methods have all been applied, but phase-field modelling is particularly attractive because it provides a thermodynamically consistent description of interface motion and allows grain growth and microstructural evolution to be represented in a full-field framework [3], [4].

MICRESS implements a multiphase-field formulation that is well suited to recrystallisation and grain-growth simulations in polycrystalline materials. In this framework, stored energy can be introduced as the driving force for boundary migration, while nucleation and mobility laws can be prescribed within the simulation framework [4], [5]. In the present study, attention is focused on shielding time, a nucleation-control parameter that defines the minimum interval between successive nucleation events in a local region. The aim was to determine how shielding time influences recrystallisation kinetics, grain structure and the emergence of a saturated regime at large shielding times.

## Experimental/Simulation

The simulations were carried out in MICRESS and performed in an effectively two-dimensional, isothermal domain. They were designed as a parameter study so that the influence of shielding time could be isolated without adding unnecessary complexity. Shielding time was the main study variable, while the remaining model conditions were kept fixed. The shielding times examined were 5 s, 10 s, 20 s, 50 s, 1000 s, 1500 s and 2000 s.

Within the multiphase-field framework, grains are represented by continuous order parameters and grain boundaries evolve as diffuse interfaces [4]. Recrystallisation is driven by stored-energy differences associated with dislocation density, so that highly deformed regions are progressively replaced by recrystallised grains of lower stored energy [1], [2]. Nucleation is restricted to grain interfaces and controlled by the MICRESS nucleation formulation, with shielding time governing how frequently new nuclei can form locally. This makes shielding

https://doi.org/10.18745/pb.27307

time a useful parameter for testing the sensitivity of recrystallisation predictions to nucleation assumptions [5], [6].

## Results and discussion

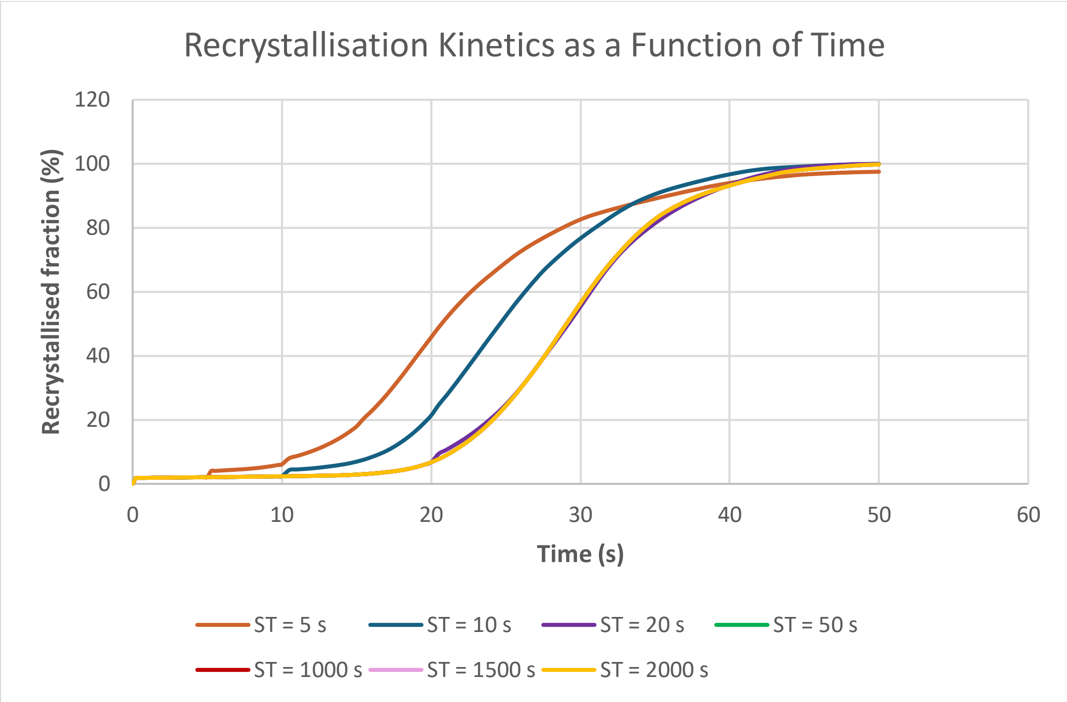



*Fig. 11. Recrystallised fraction as a function of simulation time for different shielding times (ST = 5 s, 10 s, 20 s, 50 s, 1000 s, 1500 s and 2000 s). Shorter shielding times promote earlier nucleation and earlier recrystallisation, whereas large shielding times produce very similar behaviour, indicating a saturated shielding regime.*

As shown in Fig. 1, shielding time strongly affects the early stages of recrystallisation. The recrystallised-fraction curves follow the expected sigmoidal pattern, but their position changes with shielding time. Shorter shielding times lead to earlier nucleation and therefore an earlier onset of recrystallisation, whereas increasing shielding time delays the transformation. At large shielding times, the curves become very similar, showing that the influence of shielding time weakens once the parameter becomes large relative to the simulated transformation timescale. This trend is one of the clearest results in this study and is also consistent with the broader recrystallisation literature, where kinetics are highly sensitive to nucleation treatment [2], [6].

The grain-structure results support the same interpretation. Short shielding times promote a higher density of nuclei, producing finer final microstructures, while longer shielding times suppress nucleation and allow existing grains to grow further before impingement. This shift is captured particularly clearly in Fig. 2, where the mean grain radius increases with shielding time before reaching a plateau. As shielding time increases from 5 s to 20 s, the mean radius rises gradually from about 6.13 µm to 6.89 µm, then increases more sharply to about 9.68 µm at 50 s. For 1000 s, 1500 s and 2000 s, the mean radius remains essentially unchanged, showing that the grain-size response has already reached a plateau.

This behaviour indicates the emergence of a saturated shielding regime. In practical terms, once shielding time becomes sufficiently large, further increases have little additional influence on either recrystallisation kinetics or final grain size within the simulated time window. That is an important modelling result, because it suggests that very large shielding times do not continue to change the predicted microstructure in a meaningful way under the present assumptions. Taken together, these results suggest that shielding time strongly affects kinetics and morphology, but beyond approximately 50 s the response changes very little.

https://doi.org/10.18745/pb.27307

More broadly, the results reinforce a point already made in the literature: recrystallisation simulations are highly sensitive to how nucleation is represented. In that sense, shielding time should be understood as a phenomenological but influential nucleation-control parameter rather than as a direct physical material constant [5], [6]. Within the present MICRESS framework, it is precisely this parameter that governs the differences in transformation timing, grain density and final mean grain size.

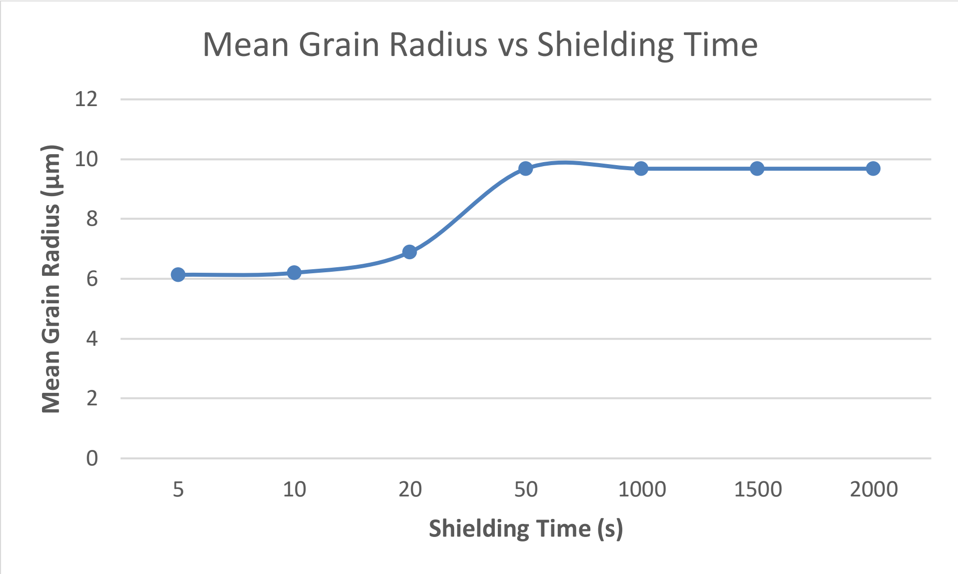



*Fig. 12. Mean grain radius as a function of shielding time. Shorter shielding times produce smaller average grain radii, while larger shielding times increase the mean grain radius before the values stabilise at high shielding times.*

## Conclusion

This study investigated the effect of shielding time on static recrystallisation using MICRESS multiphase-field simulations. Shorter shielding times promoted earlier nucleation, faster early-stage recrystallisation and finer grain structures, while longer shielding times delayed nucleation and increased the mean grain radius. The simulations also showed that, beyond about 50 s, further increases in shielding time produced little additional change in the predicted behaviour, indicating a saturated shielding regime. Overall, the results suggest that shielding time is a highly influential nucleation-control parameter in this modelling framework and that careful parameter selection is important when interpreting phase-field recrystallisation simulations.

## References


[1] D. Raabe, "Recovery and Recrystallization: Phenomena, Physics, Models, Simulation," in *Physical Metallurgy*, 5 ed., D. E. Laughlin and K. Hono, Eds., Oxford, Elsevier, 2014, p. 2291–2397.

[2] R. D. Doherty, D. A. Hughes, F. J. Humphreys, J. J. Jonas, D. Juul Jensen, M. E. Kassner, W. E. King, T. R. McNelley, H. J. McQueen and A. D. Rollett, "Current issues in recrystallization: A review," *Materials Science and Engineering A,* vol. 238, p. 219–274, 1997.

[3] Y. Liu, M. Militzer and M. Perez, "Phase Field Modelling of Abnormal Grain Growth," *Materials,* vol. 12, no. 24, p. 4048, 2019.

[4] I. Steinbach and F. Pezzolla, "A generalized field method for multiphase transformations using interface fields," *Physica D,* vol. 134, p. 385–393, 1999.

[5] M. Militzer, "Phase field modeling of microstructure evolution in steels," *Current Opinion in Solid State and Materials Science,* vol. 15, no. 3, p. 106–115, 2011.

[6] S. P. Gentry and K. Thornton, "Sensitivity analysis of a phase field model for static recrystallization of deformed microstructures," *Modelling and Simulation in Materials Science and Engineering,* vol. 28, no. 6, p. 065002, 2020.

https://doi.org/10.18745/pb.27307

# Integrated Agent-Based Framework with Hydrodynamic Modelling for Urban Flood Risk Assessment

**Saeid Najjarghabel[1], Farzad Piadeh[2], and Kourosh Behzadian[1*]**

*[1]Smart Infrastructure and Green Technologies Research Group, School of Computing and Engineering, University of West London, St Mary's Rd, London, W5 5RF, UK*

*[2]Centre for Engineering Research, School of Physics, Engineering, Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK*

*Corresponding author: Kourosh.Behzadian@uwl.ac.uk

## Abstract

**This study develops an integrated framework that couples a hydrodynamic flood model with an activity-based agent-based model (ABM) to analyse how urban flooding affects transport systems and traveller behaviour. The hydrodynamic model was calibrated using historical flood events and linked to an ABM that simulates daily mobility and adaptive responses to flood hazards. A dynamic Risk Priority Index (RPI) was introduced to capture agents' risk awareness based on flood depth, demographic vulnerability, institutional communication, and social information diffusion using a Watts-Strogatz network. Results show that flooding closed 9% of roads, causing major disruption with 54% failed trips, highlighting vulnerable traveller groups and critical transport network locations.**

Keywords: Agent-based model; Demographic attributes; Flood simulation model; Risk priority index; Road networks vulnerability

## Introduction

Floods are among the most damaging natural hazards and are becoming more frequent due to climate change and urbanisation, particularly in densely populated urban areas [1]. Traditional flood risk assessments mainly rely on hydraulic modelling and engineering measures but often neglect human behaviour, mobility patterns, and social interactions that influence flood impacts. Agent-based models (ABMs) have increasingly been used to represent behavioural responses and mobility during flood events; however, existing frameworks still inadequately capture dynamic risk perception, social information diffusion, and uncertainty in hydrodynamic modelling [2]. This study addresses these gaps by proposing an integrated framework that couples an hydrodynamic model with a risk-aware ABM to simulate flood dynamics, human behaviour, and transport disruptions simultaneously.

## Methodology

Figure 1 presents an integrated framework that combines a hydrodynamic model with an ABM for adaptive flood risk assessment. The ABM incorporates stakeholder behaviour and dynamic Risk Priority Index (RPI) using a Watts-Strogatz social network. The coupled system evaluates impacts on mobility and population vulnerability. The performance of the model was tested for the 100-year flood event at the part of River Thames shown in Figure 2, where agents are categorized by age (children, adults, seniors) and travel modes (walking, cycling, public transport, and driving).

As shown in Equation (1), the RPI combines the perceived flood depth at the agent's location $f(D_i(t))$, institutional flood warnings $C(t)$, and a demographic vulnerability modifier $M_i$. The parameters $\alpha$ and $\beta$ control the relative importance of environmental exposure and official warnings. Also, the sigmoid function (Equation 2) models how agents perceive floodwater depth. In this equation, $T$ and $k$ show the depth threshold and sensitivity of the response.The institutional warning component is represented by Equation (3), where the warning level $C(t)$ depends on the observed water level at the monitoring station. When the water level is below the alert threshold, no warning is issued; when it exceeds predefined thresholds, alert and severe warning levels are triggered. Equation (4) introduces demographic heterogeneity through a vulnerability modifier $M_i$,

https://doi.org/10.18745/pb.27307

which adjusts the perceived risk based on the agent's characteristics. Agents in vulnerable age groups (i.e., children and seniors) or those experiencing financial insecurity (i.e., unemployment agents) have a higher vulnerability. For the vulnerable age groups, an additional value of 0.1 was assigned to reflect increased physical vulnerability during flood events. Similarly, agents who were unemployed were assigned an additional value of 0.1 (i.e., $EmploymentFactor = 0.1$) to represent heightened sensitivity to potential economic and safety-related losses. Moreover, the dynamic evolution of risk is modeled through Watts-Strogatz model shown in Equation (5). Each agent updates their RPI based on a weighted combination of their own previous priority and the priority of neighboring agents in their social network. The parameter $\lambda$ controls the strength of social influence, while $W_{ij}$ represents the influence weight between connected agents.

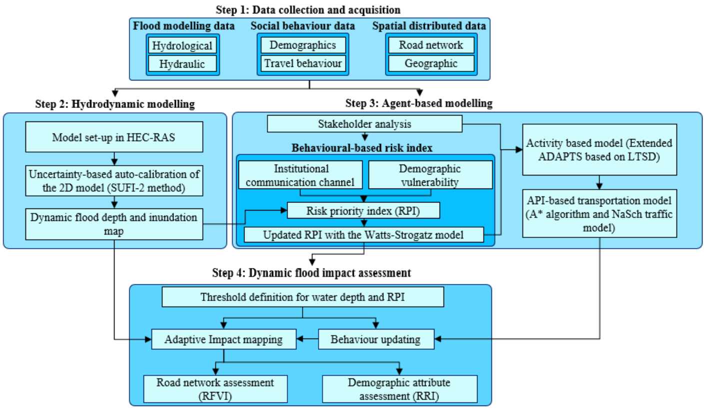



*Figure 13. Workflow of the proposed method for dynamic flood risk assessment using coupled agent-based and hydrodynamic models*

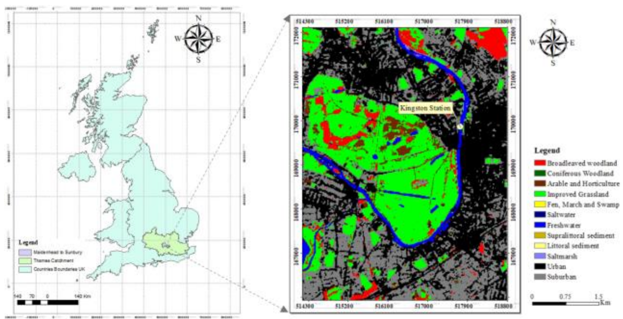



*Figure 14. The River Thames from Maidenhead to Sunbury, and the land cover map*

$$RPI_i(t) = \min(1, \max(0, (\alpha \times f(D_i(t)) + \beta \times C(t)) \times M_i)) \tag{1}$$

$$f(D_i(t)) = \frac{1}{1 + e^{-k(D_i(t)-T)}} \tag{2}$$

$$C(t) = \begin{cases} 0, & H_t < 5.15 \\ 0.5, & 5.15 \leq H_t < 5.77 \\ 1, & H_t \geq 5.77 \end{cases} \tag{3}$$

$$M_i = 1 + (AgeFactor_i + EmploymentFactor_i) \tag{4}$$

$$RPI_i(t+1) = \lambda * RPI_i(t) + (1-\lambda) \sum_{j \epsilon S_i} W_{ij}\, RPI_j(t) \tag{5}$$

https://doi.org/10.18745/pb.27307

## Results and discussion

The hydrodynamic model developed in HEC-RAS,  was calibrated against historical flood events at Kingston Station using the SUFI-2 algorithm to optimise Manning's n values across different land-cover classes, ensuring accurate and reliable flood-dynamics simulations. The 2020 calibration (Figure 3) achieved good agreement with observations, achieving an $R^2$ value of 0.91. The p-factor = 0.897 indicates that 89.7% of the observed data fell within the 95PPU band. Also, the d-factor = 1.047 demonstrates that the uncertainty-band width was well-proportioned relative to the observed variability.

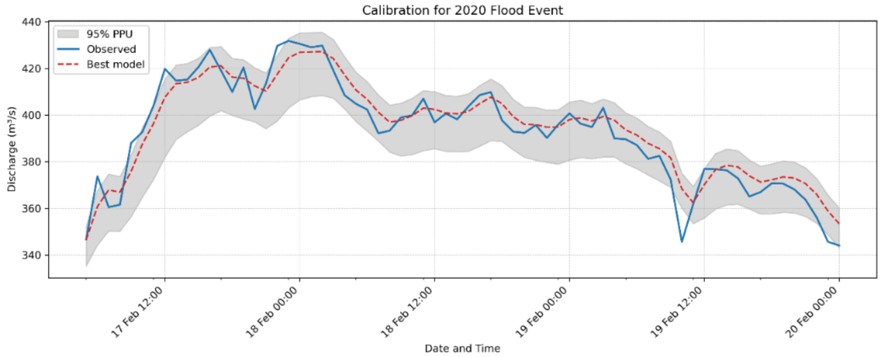
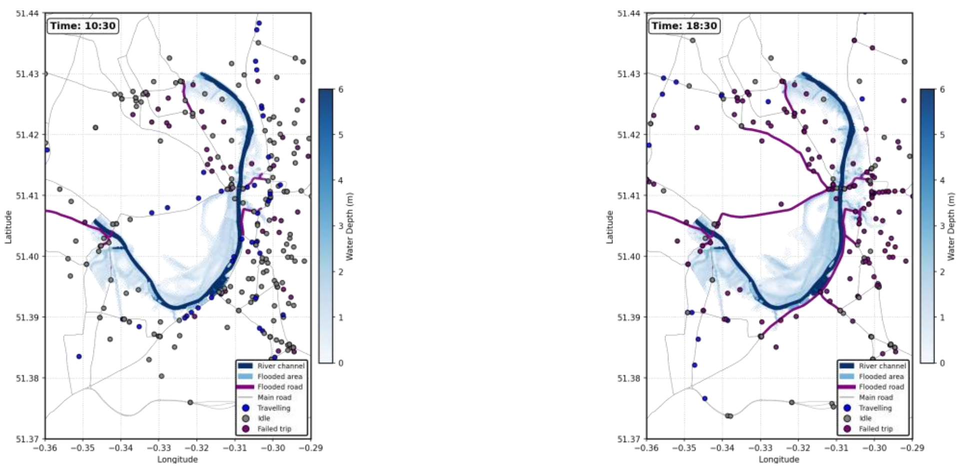



*Figure 15. Observed vs Simulated discharge at Kingstone Station for  calibration (flood event in 2020)*

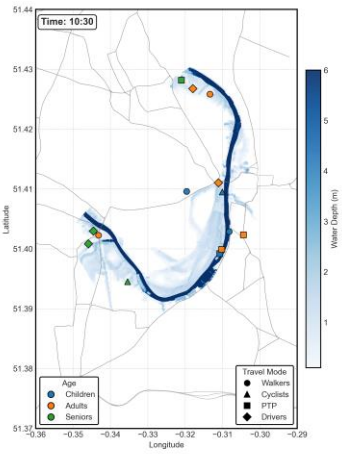
The spatial results (Figures 4a and 4b) show that flood propagation along the River Thames progressively disrupts the transport network and agent mobility during weekdays. Flooding initially impacts riverside corridors and bridges during the late morning period (approximately 10:00-12:00) and expands by evening, increasing flooded road segments and causing network fragmentation. Results indicate that approximately 9% of roads are closed due to flooding, leading to major disruption with 54% of trips failing. As inundation intensifies, clusters of failed trips form near disconnected links and bridge access points, while the number of travelling agents declines due to reduced route availability, particularly around critical commuting corridors.

Furthermore, Figures 4a and 4b show the spatial distribution of agents exposed to high risk (RPI > 0.8) or hazardous flood depths (≥ 0.3 m) across age groups and travel modes. During the morning (10:30), high-risk agents are mainly concentrated near river-adjacent routes and bridges, with adults (walkers and drivers) most exposed along commuting paths. By the evening period (18:30 UK local time), risk intensifies around flooded corridors due to return trips, with more vulnerable groups such as children and seniors appearing closer to deeper flood zones, reflecting reduced mobility and limited alternative routes.

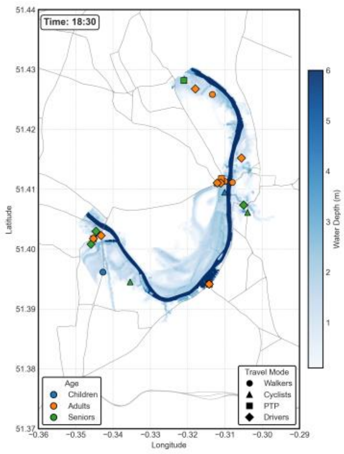



**(a)** **(b)**

https://doi.org/10.18745/pb.27307




**(c)**




**(d)**

*Figure 16. Flooded roads and agent behaviour during a 100-year flood at (a) 10:30 and (b) 18:30, and the spatial distribution of high-risk agents (RPI > 0.8) or those exposed to water depths ≥ 0.3 m across different age groups and travel modes at (c) 10:30 and (d) 18:30.*

## Conclusion

This study demonstrates that integrating hydrodynamic modelling with an agent-based framework provides a more comprehensive understanding of urban flood impacts by capturing both physical processes and human behaviour. The results highlight how dynamic risk perception, social interactions, and travel patterns significantly influence mobility disruption, beyond flood severity alone. The proposed framework effectively identifies vulnerable populations and critical network hotspots, offering valuable insights for improving urban resilience and supporting informed decision-making in flood risk management and emergency planning.

Despite the advantages of the proposed framework, several limitations remain. The current model simplifies certain behavioural and mobility assumptions and relies on synthetic mobility rather than observed real-time mobility trajectories. In addition, social interaction parameters and demographic vulnerability coefficients were generalized for the study area. Future work will focus on validating the framework using empirical mobility datasets, transport disruption records, and real-world behavioural observations collected during historical flood events.

## References

[1] F. Piadeh, K. Behzadian, and A. M. Alani, “A critical review of real-time modelling of flood forecasting in urban drainage systems,” *J. Hydrol. (Amst).*, vol. 607, p. 127476, Apr. 2022

[2] A. Templeton, H. Xie, S. Gwynne, A. Hunt, P. Thompson, and G. Köster, “Agent-based models of social behaviour and communication in evacuations: A systematic review,” *Saf. Sci.*, vol. 176, p. 106520, Aug. 2024

https://doi.org/10.18745/pb.27307

# Simulation of Satellite Attitude Control using Deep Reinforcement Learning

**Dharshini Subramani[1*], Pouria Sarhadi[1]**
[1]*Department of Engineering, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK*
*Corresponding author: dharshinishanmugham@gmail.com

**This study proposes Deep Reinforcement Learning (DRL) as an alternative to the traditional Proportional – Integral-Derivative (PID) controllers for single-axis satellite attitude control. Four controllers (PID + Relay, PID + Dead zone + Relay, DDPG and DQN) were evaluated across realistic satellite scenarios including tracking, large inertia variation and detumbling. The results show that while PID is superior in eliminating steady-state error using integral action, the DRL-based controllers, especially the DQN shows higher level of robustness in detumbling and parameter uncertainty cases. The DQN aligns with the real – world on-off thruster actuation, and helps in maintaining a balance between fuel efficiency and stability.**

Keywords: Satellite Attitude Control; Deep Reinforcement Learning; DQN; DDPG; PID Control

## Introduction

Satellites are becoming increasingly a vital part of life, as they play important roles not only in research but also in other applications such as communications and weather forecasting. For any satellite mission, the main objective is to be able to orient itself precisely and maintain the desired orientation even when faced with disturbances or changing operation conditions. This is why the Attitude Control System is said to be one of the most important subsystems as it observes the satellite's position and applies corrective forces to help the satellite maintain the intended position and trajectory.

Traditional control methods such as the PID controller have been the industry standard due to its flight-proven heritage.[1] However, they use rigid mathematical models which makes them more vulnerable in scenarios involving non-linear dynamics, parameter variations and external disturbances.[2] This project aims to propose DRL as a flexible alternative to PID controllers, as the control agent in DRL learns to control and stabilize the satellite through direct interactions with the environment instead of relying on rigid mathematical models.[3] The main objective of this study is to provide a direct comparative evaluation of DRL algorithms against traditional baseline controllers, with a specific focus on metrics such as robustness and fuel-efficiency.

## Experimental/Simulation

The dynamics of the single-axis satellite were modeled using Euler's rotational equations of motion, with a baseline moment of inertia $J = 11.4\ kg.m^2$. The plant model was deployed in Simulink as a double integrator system. The controller models are:

- PID with relay: Relay converts the continuous signals into on-off pulses.
- PID with relay and dead zone: Inclusion of a dead zone threshold of $\pm\ 0.3$ to reduce chattering.
- Deep Deterministic Policy Gradient (DDPG): A DRL- based algorithm, which has an actor-critic network whose output is a continuous torque output
- Deep Q-Network (DQN): Another DRL-based algorithm which uses discrete action spaces (-1, 0, +1 Nm), matching the on-off thruster pulses of real satellites.

The reward function used to train the DRL agents penalized attitude error, angular velocity and the control effort (torque). Conditional rewards were included so that the agent can learn to fine-tune the outputs according to the scenario. Techniques such as domain randomization and inclusion of random disturbance during training were employed to increase to make the agents more robust.

https://doi.org/10.18745/pb.27307

## Results and discussion

Performance metrics such as settling time, steady state error, and overshoot were evaluated for the scenarios.

| Metric | PID + Relay | PID + Relay + Dead zone | DDPG | DQN |
|---|---|---|---|---|
| C1 Settling time (sec) (nominal condition) | 7.10 sec | ~6.0 sec | 3.2 | 5.9 |
| C2 Overshoot (Large angle recovery) (%) | 46.56 | ~40.0 | oscillations | negligible |
| C3 Steady state error (external disturbance case) (rad) | 0 | ~0 | 0.10 | 0.10 |
| C4 Parameter uncertainty robustness | Retuning required | Retuning required | Failed | Partially successful |
| C5 Detumbling scenario (settling time) | 35s | ~35 seconds | ~14 seconds | ~14.9 secs |

*Table 1: Results*

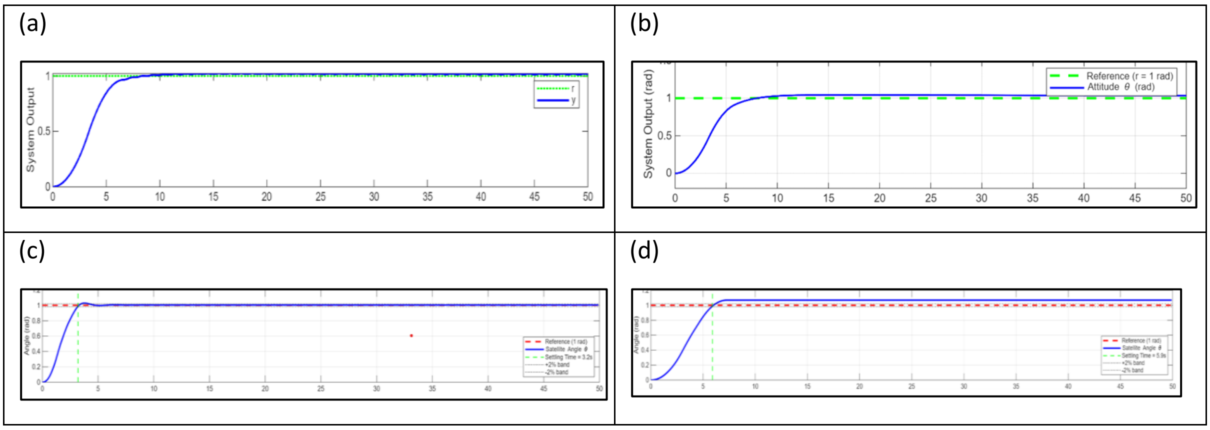



*Fig.1. (a) PID + Relay Case 1 Nominal Step Response; (b) PID + Relay + Dead zone Case 1 Nominal Step Response (c) DDPG Case 1 Nominal Step Response; (d) DQN Case 1 Nominal Step Response*

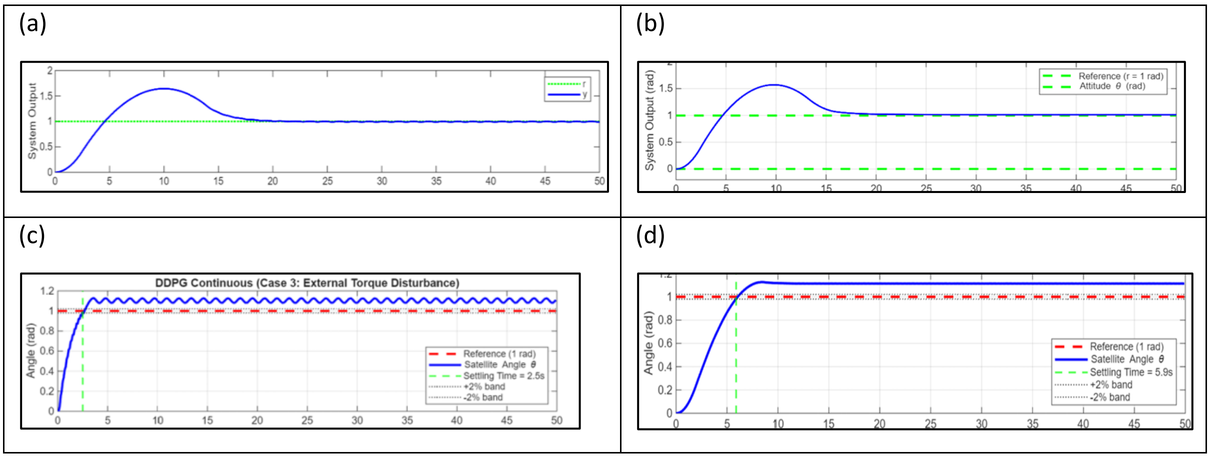



*Fig.2. (a) PID + Relay Case 3 External Disturbance Response; (b) PID + Relay + Dead Zone Case 3 External Disturbance Response; (c) DDPG Case 3 External Disturbance Response; (d) DQN Case 3 External Disturbance Response*

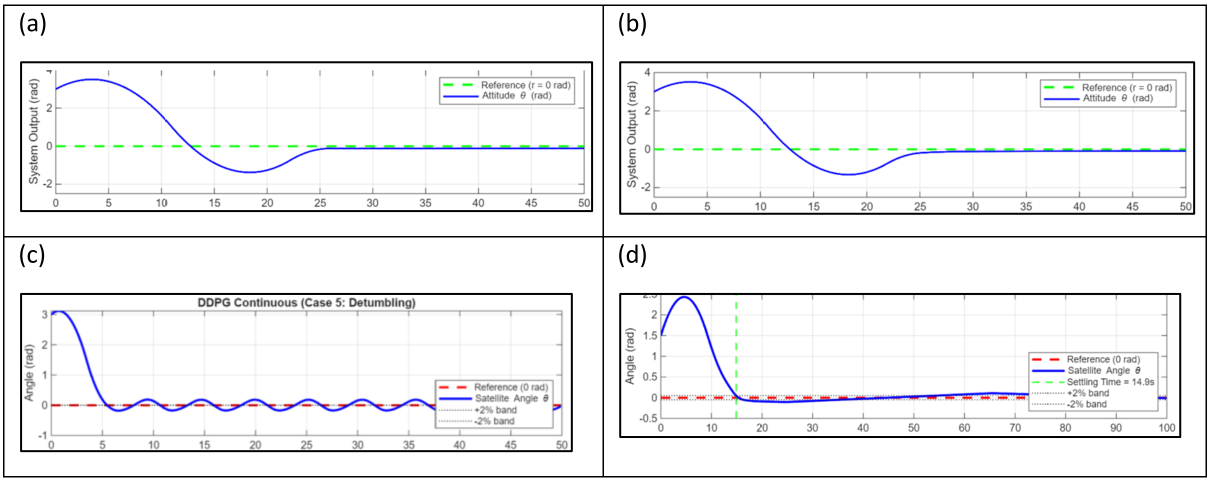



*Fig.3. (a) PID + Relay Case 5 Detumbling Response; (b) PID + Relay + Dead Zone Case 5 Detumbling Response; (c) DDPG Case 5 Detumbling Response; (d) DQN Case 5 Detumbling Response*

Case 1 is used to evaluate if the controller is able to follow a simple tracking command. Figure 1 shows that the DDPG controller settles the system within 3.2 seconds, demonstrating most aggressive stabilization technique. The discrete DQN controller brings the system to its desired angle within 5.9 seconds, using only three types of pulses (-1, 0, +1Nm) as shown in Table 1. This is more suitable for real-satellite missions as satellite actuators and thrusters follow the on-off control logic.

Case 3 evaluates the system's disturbance rejection ability. In this case, the system is subjected to a constant disturbance torque of 0.5 Nm. Figure 2 shows that, due to the integral action of PID controllers, they were able to achieve zero steady state error but as the DRL controllers do not have the memory of past errors, they struggle to eliminate the error completely. This is a key limitation of DRL.

The objective of the controllers in Case 5 is to bring the satellite from an extreme initial displacement of 5 radian and a high angular rate. This is where the DRL algorithm really shines, as the DQN agent is not subjected to integrator windup like the PID agents. Therefore, Table 1 shows that the DRL agent achieves steady state in 14.9 seconds, less than half of the time required by PID models. Another key advantage of DRL is that the agents offered more resistance to inertia changes in Case 4, while PID models required manual returning.

## Conclusion

DQN on-off control provides the best balance for practical spacecraft attitude control as it is compatible with discrete thruster (similar to real actuators), which is robust to parameter variations and other disturbances if adequately trained. Its ability to handle non-linear dynamics and parameter uncertainty without the need for manual tuning makes it the best choice for the primary control system while PID controllers can be used as the backup system if the DRL agent diverges.

## References

[1] A. Khosravi and P. Sarhadi, "Tuning of pulse-width pulse-frequency modulator using PSO: an engineering approach to spacecraft attitude controller design," Automatika, vol. 57, no. 1, pp. 212–220, 2016, doi: 10.7305/automatika.2016.07.618.
[2] W. De Sa Marques, "Intelligent attitude control of satellites via deep reinforcement learning," Master's dissertation, Instituto Nacional de Pesquisas Espaciais (INPE), 2021. [Online]. Available: https://www.researchgate.net/publication/377396432_Intelligent_Attitude_Control_of_Satellites_via_Deep_Reinforcement_Learning. [Accessed: April 14, 2026].
[3] Á. G. Pérez-Muñoz, G. López-García, I. García-Villoria, A. Alonso, A. Porras-Hermoso, and M. S. Pérez, "Feasibility of deep reinforcement learning for the real-time attitude control of satellite system," J. Syst. Archit., vol. 167, p. 103513, 2025, doi: 10.1016/j.sysarc.2025.103513.

https://doi.org/10.18745/pb.27307

# Design and Implementation of a Low-Cost Real-Time Gesture-Controlled Smart Wheelchair Using ESP32 and IMU Sensors

**Zhouyu Wang[1*], and K. Oo Htet[1]**

*[1]Department of Engineering, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK*

**Corresponding author: zw25aat@herts.ac.uk*

**This paper presents a low-cost real-time gesture-controlled smart wheelchair based on ESP32 controllers, a QMI8658 IMU, and WiFi SoftAP UDP communication. Wrist tilt is filtered and interpreted using thresholding, dead-zone compensation, and axis-priority arbitration, then mapped proportionally to differential wheel speeds. Safety locking, timeout detection, and fail-safe stopping are included. Experiments showed 95–100% gesture recognition, about 10.1 ms communication latency, about 40 ms emergency stopping, and faster task completion than joystick control.**

Keywords: Gesture recognition; ESP32 embedded systems; IMU-based control; Wireless real-time communication; Smart wheelchair

## Introduction

Gesture-based human–machine interaction can provide a more intuitive control method than joysticks for users with limited finger dexterity or reduced hand mobility [1], [2]. Low-cost embedded platforms such as the ESP32 also make it possible to integrate sensing, wireless communication, and real-time control in a compact assistive system.


*Fig. 1. System architecture of the gesture-controlled smart wheelchair.*

The proposed wheelchair uses an ESP32-S3 wearable controller with a QMI8658 IMU, an ESP32-based wireless receiver, and an Arduino Uno with an L298N driver for differential-drive motion. Wrist pitch and roll are converted into motion commands, while WiFi SoftAP UDP enables infrastructure-independent low-latency control.

## Experimental/Simulation

Accelerometer data from the wearable IMU are used to estimate wrist pitch and roll relative to gravity [1], [2]. A neutral posture is calibrated at start-up, then low-pass filtering, dead-zone compensation, and thresholding suppress noise and unintended motion. When both axes are active, the dominant tilt is selected and mapped proportionally to left- and right-wheel speeds for smooth control.

https://doi.org/10.18745/pb.27307

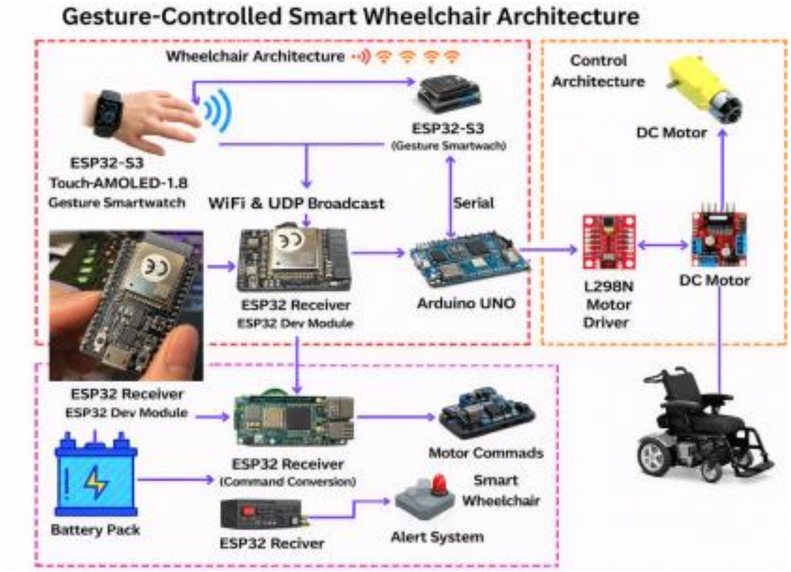



*Fig. 2. Functional block diagram of the distributed wheelchair control system.*

The ESP32-S3 broadcasts UDP packets at about 50 Hz, each containing normalised wheel commands [5]. The receiver validates the packets and forwards the commands to the Arduino Uno, which generates PWM outputs for the L298N motor driver [4], [5]. Safety is implemented through a 10° dead zone, gesture limiting to about 60°, a >65° safety lock, and timeout-based emergency stopping.

## Results and discussion

Tests confirmed stable real-time navigation performance: gesture recognition reached about 95–100%, average communication latency was about 10.1 ms, and emergency stop response was about 40 ms. Navigation time decreased from 43.4 s with joystick control to 39.4 s with gesture control, suggesting better convenience after brief familiarisation. Based on the project report, the proposed ESP32–IMU platform is also markedly cheaper than typical commercial smart wheelchair control hardware (about £546–£936) because it avoids proprietary controllers, specialist sensors, and external camera systems.

**Table 1. Summary of key experimental results**

| Metric | Measured value | Interpretation |
|---|---|---|
| Gesture recognition accuracy | ≈95-100% | High recognition reliability for the tested gestures |
| Average communication latency | ≈10.1 ms | Near real-time wireless response |
| Emergency stop time | ≈40 ms | Fast reaction to unsafe gestures |
| Control update frequency | 50 Hz | Smooth and frequent command refresh |
| Estimated hardware cost | < £62 | Much lower than typical commercial control hardware (£546–£936) |

**Table 2. Cost comparison between the proposed system and typical commercial smart wheelchair control solutions**

| Component / Aspect | Typical commercial solution | Proposed system |
|---|---|---|
| Main controller | £156–£390 | ESP32: £6–£12 |
| Sensor module | £117–£312 | QMI8658 IMU: £4–£8 |
| Communication module | £62–£117 | Built-in WiFi |
| Motor control unit | £234+ | Arduino + L298N: ~£16 |

https://doi.org/10.18745/pb.27307

| Movement | Illustration | Function |
|---|---|---|
| Moving the hand downwards |  | Forward movement |
| Moving the hand upwards |  | Backward movement |
| Tilting the hand to the right |  | Turn right |
| Tilting the hand to the left |  | Turn left |
| Straightening the hands |  | Stop |
| Exceeding safety tilt entering lock mode |  | Safety degree（10° -65° ） |

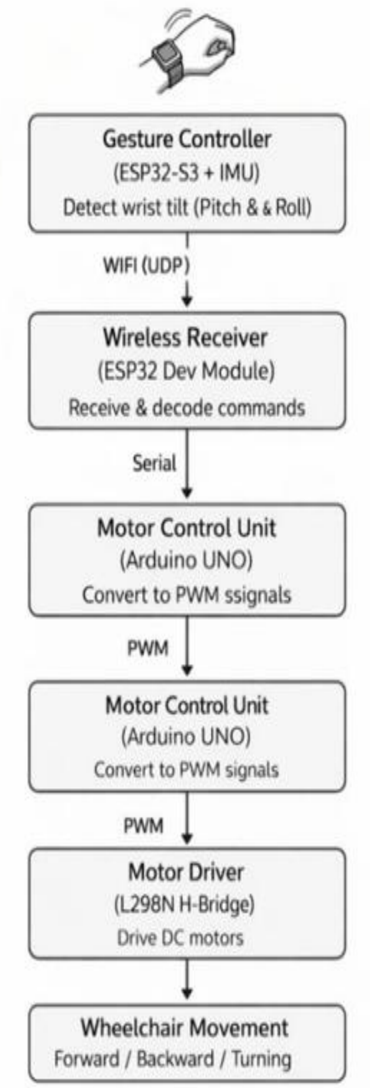



*Fig. 3. Experimental prototype and command-mapping views of the gesture-controlled smart wheelchair system.*

## Conclusion

A low-cost ESP32- and IMU-based gesture-controlled wheelchair was successfully implemented and evaluated. Compared with higher-cost commercial alternatives, the system offers a more convenient wearable interface and substantially lower control-system cost while retaining responsive wireless control and layered safety.

## Testing methodology paragraph

During the testing phase, the prototype was tested using repeated forward, backward, left, right, stop, and safety-lock gestures. The response of each command was recorded to check recognition accuracy, communication latency, and emergency stop time. The gesture-control performance was then compared with joystick control using the same navigation task.

## References

[1] F. Al-Wesabi, M. Alamgeer, F. Al-Yarimi, and A. Albaadani, “A smart-hand movement-based system to control a wheelchair wirelessly,” Sensors and Materials, vol. 31, no. 9, 2019.
[2] M. Huda, M. Ali, and M. Sadi, “Developing a real-time hand-gesture recognition technique for wheelchair control,” PLoS ONE, vol. 20, 2025.
[4] T. Kaffale and J. Liu, “Speed and direction control of DC motor using Arduino Uno microcontroller,” OALib, vol. 12, 2025.
[5] M. Anshar et al., “Control system design for smart wheelchair robot,” in AIP Conference Proceedings.

https://doi.org/10.18745/pb.27307

# Comparative Study of the Effect of Ply Orientation and Stacking Sequence on the Tensile Strength of Natural Fibre Composite for Motorsport Applications

**Ashton Smith-Thorn[1*], and Gbanaibolou Jombo[1]**

[1]*Centre for Engineering Research, School of Physics, Engineering and Computer Science, University of Hertfordshire, UK.*

*Corresponding author: as23adx@herts.ac.uk

**Motorsport is facing increasing pressure to reduce the environmental impact, especially from carbon fibre composites. This study evaluates the viability of natural fibre composites such as flax as a sustainable alternative for structural applications within Formula 1. Finite element analysis (FEA) is performed to compare the tensile strength performance between flax, glass and carbon fibre-based composites across multiple stacking sequences and ply orientations. Results show that although carbon fibre-based composites maintained superior stiffness and tensile strength properties, with glass fibre acting as an intermediary between performance and environmental impact. For flax composite testing, the angle-ply sequence provided the strongest result, with a predicted failure displacement of 5.99mm and a failure load of 44.8kN. This force is 91% higher than the quasi-isotropic flax fibre configuration. These results are based on a linear elastic model and do account for a failure criteria or progressive model damage.**

Keywords: Composite Material; Tensile Strength; Finite Element Analysis (FEA); Flax

## Introduction

Formula 1 relies heavily on carbon fibre-based composites for structural applications such as monoque chassis, space frames, etc. due to their high stiffness and superior strength to weight ratio. Carbon fibre makes up 75% of a Formula 1 whilst only accounting for 25% of its mass and requires an incredible energy-intense production, costing 234 MJ/Kg to produce a carbon chassis[1][2]. With increasing pressure to reduce the environmental impact caused by Formula 1, this creates the need to analyse sustainable composite materials as alternatives for structural applications within the sport.

Natural fibre composites, such as flax, offer a reduced energy-intense production process meaning improved sustainability; however, they are limited by their mechanical properties and further research into their use for high-performance structural applications. As of recent developments, flax is largely restricted to non-structural components in motorsport such as racing seats[3].

This study presents a numerical analysis of a flax fibre composite in comparison to glass and carbon fibre composite under tensile loading conditions. Finite element analysis (FEA) is used to assess the tensile strength behaviour across varied stacking sequences such as: cross-ply, angle-ply, uni-directional (UD) dominated ply, quasi-isotropic ply and non-quasi isotropic ply, to analyse the effect of fibre orientation and stacking sequence on the mechanical performance of natural fibre-based composites.

The objective of this study is to determine the performant stacking sequence for flax natural fibre composite in tensile loading for structural application within motorsport. This will be achieved by analysing the stress, strain and load bearing capabilities using a load-controlled displacement within the ANSYS workbench.

## Experimental/Simulation

For this study three composite materials were selected, such as: carbon fibre, glass fibre and flax fibre. Carbon fibre is selected as it is the industry standard for high-performance structural material in motorsports, glass fibre as an intermediary material between performance and environmental impact, and flax fibre as the optimal sustainable alternative as compared to hemp, jute, and sisal. All three materials were modelled using unidirectional laminates to ensure consistency throughout the comparison.

https://doi.org/10.18745/pb.27307

A standardised dogbone specimen geometry was used for the tensile testing as per standard ASTM D3039. The dogbone is used to promote stresses towards the gauge section as opposed to the location of the grips. The geometry was developed using CATIA v5 and imported into the ANSYS workbench for structural analysis. Composite layups were defined within the ANSYS ACP workbench, allowing for control of the fibre orientations and stacking sequence.

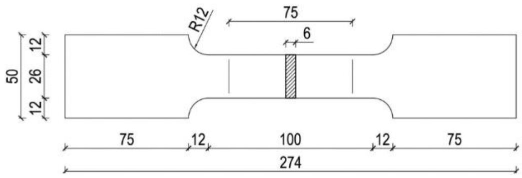



*Figure 17 Dogbone Specimen Geometry [5]*

Five stacking sequence variations were used for the analysis, they are as follows: cross-ply $[(0/90)_{10}]$, angle-ply $[(\pm45)_5]_s$, uni-directionally (UD) dominated $[0_3/+45/0_2/90/0_3]_s$, quasi-isotropic $[0/\pm45/90/0/\pm45/90/0/+45]_s$ and non-quasi-isotropic $[0/\pm45/0/90/0/\pm45/0_2]_s$. This allowed for the evaluation of a wide scope of tensile mechanical behaviour that was determined by the range of stacking sequences due to the anisotropic characteristics of composite materials.

The FEA model was developed using a linear elastic assumption as opposed to a designed failure criterion such as Tsai-Wu. Inclusion of the composite failure criterion would be considered in further study to improve model failure mechanism. A mesh convergence study was conducted to ensure that the results were independent from the mesh size. The results found that running simulations at 1 mm mesh size reduced the computational time by 96% whilst only effecting the results accuracy by 0.04%. Boundary conditions represented tensile testing by using a fixed support at one end of the specimen and applying a uniaxial load-controlled displacement at the other end, generating stresses along the length of the specimen.

Using the ANSYS Static Structural workbench, normal stress, normal strain and reaction force were recorded taking data readings from the gauge section to evaluate stiffness and load-bearing capabilities. The use of a controlled displacement allowed for a consistent analysis of force-displacement across all materials and stacking sequences.

## Results and discussion

The tensile results demonstrate clear differences in the mechanical behaviour between the three materials and highlight the significance of stacking sequences on the structural integrity of composite structures.

Across all possible configurations, as expected carbon fibre exhibited the highest stiffness and load-bearing capabilities confirmed by the steepest force-displacement gradients. Glass fibre provided an intermediate performance, maintaining a relatively high stiffness but a reduced failure displacement. Flax fibre performed the worst out of the three materials demonstrating a lower stiffness and reduced load-bearing capability under identical loading conditions.

Stacking sequence was found to a be significant factor in the influence of mechanical behaviour of the test specimens. For flax, the angle-ply sequence provided the highest predicted displacement of 5.99mm and a failure load of 44.8kN. The predicted failure displacement is 209% higher then quasi-isotropic, the worst performing stacking sequence. The other sequences had predicted failure loads of cross ply at 27.5kN, UD dominated at 38.5kN and non-quasi-isotropic at 30.4kN. This data shows that the angle-ply sequence performed the best, having the greatest resistance to failure.

The linear relationships seen in Figure 2, validates the use of a linear elastic model and supports the comparison between the materials. However, due to the model using a linear elastic behaviour assumption, the simulations do not include a failure criterion such as Tsai-Wu. This means that the results should be used a predictive analysis to determine mechanical behaviour during elastic deformation and not a true failure analysis of ultimate tensile strength.

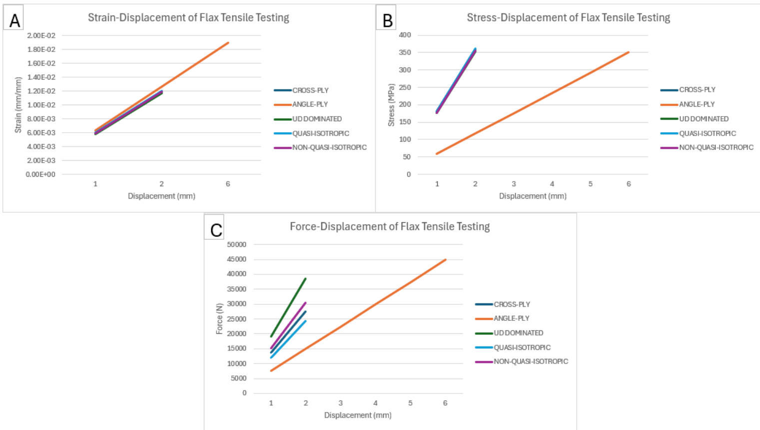



*Figure 18 Flax Testing Results a) Strain-Displacement b) Stress-Displacement c) Force-Displacement*

## Conclusion

The result from this study shows, although the tensile mechanical performance of flax-fibre composites is not equivalent to carbon-fibre composites, improved performance can be achieved using stacking sequence optimisation. Based on a consideration of five stacking sequences, such as: cross-ply, angle-ply, UD dominated, quasi-isotropic and non-quasi-isotropic, it is found that the angle-ply stacking sequence provided the best tensile mechanical performance considering stress, strain and load-displacement relationships. As such in motorsport, natural fibres can offer a pathway for organisations to reduce their environmental impact and work towards Formula 1's target of being net zero by 2030. [6]

## References

[1 ]G. Hatton, "How Are Composites Used in Motorsport?," *Racecar Engineering*, Jul. 24, 2019. https://www.racecar-engineering.com/tech-explained/how-are-composites-used-in-motorsport/ (accessed Mar. 24, 2026).
[2] S. Das, "Life Cycle Assessment of Carbon fiber-reinforced Polymer Composites," *The International Journal of Life Cycle Assessment*, vol. 16, no. 3, pp. 268–282, Feb. 2011, doi: https://doi.org/10.1007/s11367-011-0264-z.
[3] "McLaren Racing - Revealed: How McLaren Is Pioneering the Use of Sustainable Composites in F1," *www.mclaren.com*, Aug. 18, 2020. https://www.mclaren.com/racing/sustainability/natural-fibre-sustainable-composite-racing-seat/
[4] A. International, *Standard Test Method for Tensile Properties of Polymer Matrix Composite Materials*. ASTM International, 2025. Accessed: Mar. 26, 2026. [Online]. Available: https://store.astm.org/d3039_d3039m-17.html
[5] T. Dokšanović, H. Draganić, I. Radić, and D. Damjanović, "Stress–strain Relationships and Influence of Testing Parameters on Coupon Test Results," *Elektronički Časopis Građevinskog Fakulteta Osijek*, pp. 50–63, Jul. 2018, doi: https://doi.org/10.13167/2018.16.5.
[6] F. 1, "Formula 1 reduced carbon footprint by 26% by end of 2024," *Formula1.com*, 2025. https://www.formula1.com/en/latest/article/formula-1-on-track-to-be-net-zero-by-2030-with-26-reduction-in-carbon.4p2vZMDrgDu6lVHLvOcLPy

https://doi.org/10.18745/pb.27307

# Video-Based Human Fall Detection Using a Mixture of Experts Framework

**Dileepa Joseph Jayamanne[1*], Na Helian[1], Yi Sun[1], Volker Steuber[1], Say Meng Toh[2], and Kudiwa Pasipamire[2]**
[1]Biocomputation Research Group, Centre for AI and Robotics Research, University of Hertfordshire, Hatfield AL10 9AB, UK
[2]Delight Supported Living, Letchworth SG6 3EG, UK
*Corresponding author: dj22aay@herts.ac.uk

**Accurate fall detection remains a central challenge in human activity analysis, particularly in elderly care. The application of optical-flow-based experts provides high sensitivity, while keypoint-based experts offer stronger specificity. We propose a Mixture of Experts (MoE) model that combines optical-flow and keypoint-based fall-detection experts to exploit their complementary strengths. Evaluated on the UP-Fall dataset, the MoE model preserves the optical-flow expert's ability to detect true falls (96.62% sensitivity) while substantially reducing false positives, improving specificity to 98.21% and achieving the highest F1-score of all models. These results show that the gated fusion of experts yields a more accurate and balanced fall-detection system than either unimodal approach.**

Keywords: Video-based fall detection; Mixture of Experts; Optical flow; UP-Fall dataset; keypoint-driven rule-based detection

## Introduction

Video-based human fall detection remains challenging due to the need for both high accuracy and real-time efficiency. Existing methods can broadly be categorised into pixel-based and keypoint-based approaches. Pixel-based models, most commonly optical-flow CNNs [1,2] learn motion patterns directly from temporal pixel dynamics and typically achieve high sensitivity by capturing rapid changes associated with falls. However, these models often suffer from high false-positive rates, indicating limited discriminative power when used alone. In contrast, keypoint-based methods [3,4] extract 2D human pose landmarks using detectors such as MediaPipe [5] and apply lightweight, rule-driven reasoning based on geometric ratios, velocities, and orientation changes. While computationally efficient and interpretable, their reliance on fixed thresholds makes them sensitive to noise and less capable of modelling the variability of human motion.

These two families of methods exhibit complementary error profiles: optical-flow models reduce false negatives by capturing rapid motion, while keypoint-rule systems suppress false positives through geometric constraints. This complementarity suggests a natural opportunity for ensemble fusion. In this work, we explore a Mixture of Experts (MoE) framework that combines an optical-flow 3D-CNN expert [2] with a rule-based keypoint expert [3] to achieve more balanced and reliable fall-detection performance. We evaluate this approach on the UP-Fall dataset [6], demonstrating that gated fusion can enhance the accuracy of a high-sensitivity based optical-flow expert by substantially improving its specificity, thereby reducing false positives while preserving its fall-detection sensitivity.

## Experiments/Simulations

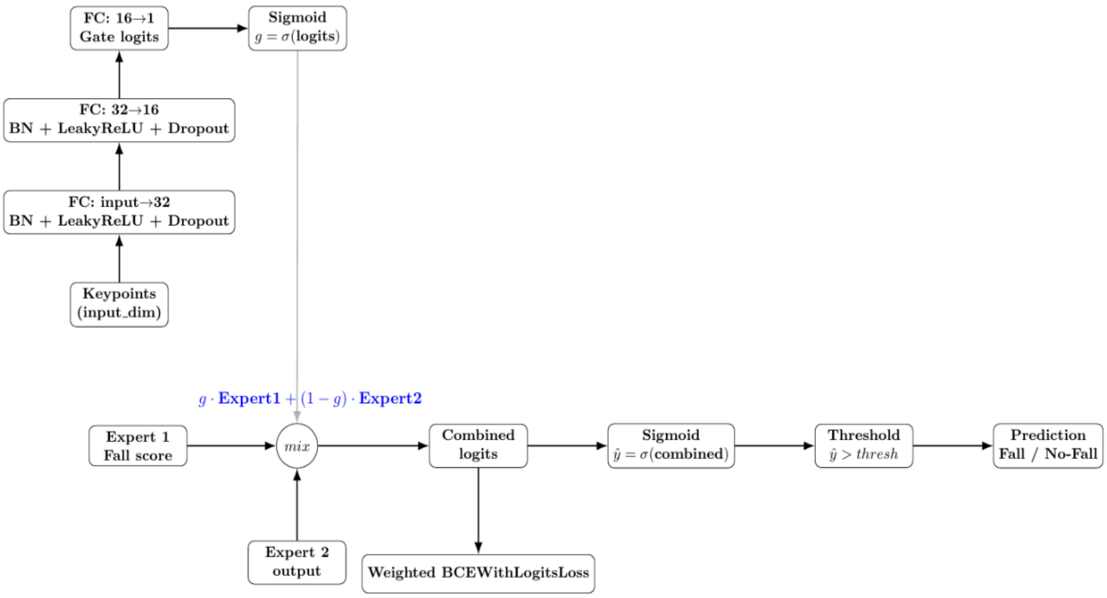
The UP-Fall dataset [1] is used for conducting our experiments. The dataset contains 1122 videos recorded from 17 participants using two cameras across 11 activities with three trials each; it is a multimodal dataset providing wearable and ambient sensor data, with actors aged 18–24 (nine males, eight females; mean weight 66.8 kg), and most clips feature a single person with no occlusions. We apply sliding windows of length 18 frames with 50% overlap to the UP-Fall dataset [6]. Windows are retained only when both valid optical-flow fields and MediaPipe keypoints are detected [5]; missing keypoints are interpolated whenever possible to avoid discarding usable motion segments. The resulting window population is highly imbalanced, with far fewer fall windows than non-fall windows. To address this, we identify all fall windows and sample an equal number of non-fall windows to construct a balanced subset. From this balanced set, 80% (2,987 windows) is used for training and

https://doi.org/10.18745/pb.27307

20% for testing, while all remaining unused windows from the full population are also included in the test set, yielding a total test size of 55,785 windows.

The Mixture-of-Experts (MoE) framework shown in Fig. 1 is applied to two complementary fall-detection experts: the 3D-CNN model of [2] evaluated using the best fold from five-fold cross-validation, which outputs a continuous fall-likelihood score, and the rule-based model of [3], which produces a binary decision. A gating network is trained to weight these experts on a per-window basis. As input, the gate received the four handcrafted motion features proposed by Pasipamire et al., namely h-ratio, w-ratio, average velocity, and average acceleration, together with window-averaged, nine normalized 2D human-body keypoint coordinates (nose, left shoulder, right shoulder, left hip, right hip, left knee, right knee, left ankle, right ankle) [3].

*Fig. 1. The MOE model. FC: Fully connected layers, BN: Batch normalization.*




For training, we use a decision threshold of 0.7. The model is optimized with Adam (learning rate = 0.001, weight decay = 0.001), and binary cross-entropy with logits is used as the loss function, incorporating a positive-class weight of 5.5 to compensate for the rarity of fall events. Training is performed in full-batch mode for 300 epochs, ensuring stable convergence under the balanced window sampling described earlier. A five-fold cross-validation procedure is used, and the test results reported correspond to the best-performing fold.

## Results and discussion

In fall-detection systems, sensitivity and specificity are as critical as they determine overall accuracy. High sensitivity corresponds to low false-negative rates, ensuring that true falls are not missed, an essential requirement in healthcare and assisted-living settings where undetected falls can have severe consequences. Conversely, high specificity reflects low false-positive rates, reducing unnecessary alerts and minimising the burden on caregivers and monitoring staff. A reliable fall-detection model must therefore achieve strong performance on both metrics, balancing the need to detect genuine falls while avoiding excessive false alarms. Table 1 reports the performance of the two unimodal experts and the proposed Mixture of Experts model, comparing their F1-score, accuracy, sensitivity, and specificity on the fall-detection task.

https://doi.org/10.18745/pb.27307

*Table. 1. Performance of the optical-flow expert, keypoint expert, and MoE model on fall-detection using the UP-Fall dataset*

| | TP | FN | FP | TN | F1-score | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|---|---|---|
| Optical-flow-based expert [2] | 343 | 12 | 1393 | 53037 | 0.3280 | 97.44% | 96.62% | 97.44% |
| Keypoint-based expert [3] | 263 | 92 | 875 | 53555 | 0.3520 | 98.23% | 74.08% | 98.39% |
| Our MOE model | 343 | 12 | 975 | 53455 | 0.4100 | 98.20% | 96.62% | 98.21% |

Across the two unimodal experts, the optical-flow-based model demonstrates the stronger baseline performance for fall detection. It achieves 96.62% sensitivity, meaning it rarely misses true falls, an essential property in safety-critical healthcare applications. However, this comes at the cost of a high false-positive rate, reflected in its lower specificity of 97.44% and a modest F1-score of 32.80%. The keypoint-based expert shows the opposite behaviour: it achieves a higher specificity of 98.39%, indicating fewer false alarms, but its sensitivity drops to 74.08%, resulting in many missed falls. This trade-off highlights the complementary strengths and weaknesses of the two approaches: the optical-flow expert excels at detecting falls (low FN), while the keypoint expert excels at avoiding false alerts (low FP).

The proposed Mixture of Experts (MoE) model successfully combines these complementary behaviours. It preserves the high sensitivity of the optical-flow expert (96.62%), ensuring that true falls continue to be detected reliably. At the same time, it substantially improves specificity, increasing it from 97.44% to 98.20%, which corresponds to a large reduction in false positives (1393 to 975). This improvement directly boosts the F1-score to 41.00%, the highest among all models, and raises overall accuracy to 98.20%. These results demonstrate that the gated fusion of experts enabled the MoE model to retain the strengths of the optical-flow expert while addressing its limitation, producing a more balanced and reliable fall-detection system.

## Conclusion

This work demonstrates that combining complementary fall-detection experts within a Mixture of Experts (MoE) framework yields a more balanced and reliable system than either unimodal approach alone. While the optical-flow expert provides high sensitivity and the keypoint-based expert offers stronger specificity, the MoE model effectively integrates these strengths, improving accuracy and substantially reducing false positives while preserving the optical-flow expert's ability to detect true falls. Evaluated on the UP-Fall dataset, the model achieves the highest F1-score and provides a more dependable balance between sensitivity and specificity, showing the benefit of the MoE framework.

## References

[1] R. Espinosa, H. Ponce, S. Gutiérrez, L. Martínez-Villaseñor, J. Brieva, and E. Moya-Albor, “A vision-based approach for fall detection using multiple cameras and convolutional neural networks: A case study using the UP-Fall detection dataset,” *Computers in Biology and Medicine*, vol. 115, p. 103520, 2019.

[2] S. M. Toh, N. Helian, K. Pasipamire, and Y. Sun, “Vision-Based Human Fall Detection Using 3D Neural Networks,” in *Artificial Intelligence XLI*, Springer Nature Switzerland AG, 2025.

[3] K. Pasipamire, S. M. Toh, N. Helian, Y. Sun, and A. Pasipamire, “Human fall detection based on gait speed and body pose,” presented at the SPECS Research Conference, University of Hertfordshire, Hatfield, U.K., Jun. 12, 2024.

[4] S. Saraswat and G. Malathi, “Pose estimation based fall detection system using MediaPipe,” in *Proc. 10th Int. Conf. Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, Apr. 12–14, 2024.

[5] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann, “MediaPipe: A framework for perceiving and processing reality,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, CA, USA, Jun. 2019, pp. 1-8.

[6] L. Martínez-Villaseñor, H. Ponce, J. Brieva, E. Moya-Albor, J. Núñez-Martínez, and C. Peñafort-Asturiano, “UP-Fall detection dataset: A multimodal approach,” *Sensors*, vol. 19, no. 9, p. 1988, 2019.

https://doi.org/10.18745/pb.27307

# Efficient Vectorization for Sentiment Analysis

**Chiedozie Onyearugbulem[1*] and Deepak Panday[2]**
*Corresponding author: doziesixtus@gmail.com

**Online platforms have normalized how people express opinions about products and services. Sentiment analysis enables the extraction of these opinions to understand public perception. However, model performance depends heavily on preprocessing steps, including vectorization. This study evaluates the impact of different vectorization methods on sentiment analysis performance and proposes a hybrid approach combining Term Frequency-Inverse Document Frequency (TF-IDF) and TinyBERT vectors through horizontal stacking. Support Vector Machine (SVM) and Artificial Neural Network (ANN) models are trained and tested using IMDB, Amazon Electronics, and Cherwell datasets. Results show the hybrid method consistently outperforms existing vectorization techniques across all three datasets.**

Keywords: machine learning; sentiment analysis; vectorization; natural language processing; neural networks

## Introduction

Opinions have long served as a channel for providing feedback on the quality of services or products. Individuals express their levels of satisfaction or dissatisfaction by sharing their thoughts and experiences through various platforms, including forums, social media, emails, and other digital communication channels. The increasing prevalence of technology has expanded the avenues available for consumers to effortlessly share their opinions with other users, enabling customer understanding and driving product development and improvement.

Sentiment analysis is a natural language processing (NLP) task that evaluates the sentiment polarity (e.g. happy, sad, neutral) of a text input. Human-generated texts often reflect the emotions of the writer at the time of writing. This is why sentiment analysis is an effective method of understanding how people feel through text. In machine learning, extracting sentiments from unstructured text inputs involves converting text into a computer-readable format, such as vectors or numbers, and training these vectors with a machine learning algorithm. The most common way of converting a text input into a computer-readable format is through vectorization.

Vectorization can be obtained through different processes. Term Frequency Inverse Document Frequency (TFIDF) vectorizer is one of the most common vectorization methods with high performance when used in combination with machine learning algorithms [1,2]. Vectorizers that perform well with traditional machine learning algorithms can be used in combination with more robust sentiment analysis algorithms like neural networks for increased performance. Neural networks have seen increased performances in recent years due to recent advancements in models, availability of data and compute power. Deep neural networks have been used in various tasks like image classification and NLP tasks. These models contain deeper layers of computation that enable them to understand textual data better than traditional machine learning algorithms like Support Vector Machine (SVM), Naïve Bayes, etc.

Other vectorization techniques use the semantic meanings of words to create a feature vector representation for each word. Word embedding is a vectorization process that creates dense vector representations of words by capturing semantic meanings, context, and syntactic relationships in texts. Some examples of word embedding techniques are Word2Vec, Bidirectional Encoder Representations from Transformers (BERT), etc. [3]. TinyBERT, a distilled variant of BERT, was used in this study because it is considerably smaller in size when compared to BERT, and its performance is quite comparable to the parent BERT model.

This study evaluates the performance of multiple vectorizers, such as TF-IDF, Count, Hashing and TinyBERT, when trained on Artificial Neural Networks (ANN) and SVM models, and proposes a hybrid vectorizer to compare the effectiveness against other vectorization techniques. The datasets used in this study are publicly available data (Amazon reviews and IMDB movie reviews) and dataset from the industry (service desk data - Cherwell). The service desk data contains texts from customers who are expressing their challenges with a service.

https://doi.org/10.18745/pb.27307

## Methodology

The type of sentiment analysis performed in this study is a binary classification problem. Binary classification occurs when only two classes are the target variables. In this study, the classes are positive and negative classes. Three labelled datasets from different domains were used in the study: Amazon Electronics reviews dataset, the IMDB movie reviews dataset, and Cherwell service desk data from a database management company. Each dataset contains reviews and the polarities (positive or negative) of the reviews.

Data in their raw format are not always suitable for use in machine learning models. Text data usually contain lots of noise, poorly structured sentences, irregular expressions, and other issues [4]. It is a rule of thumb to pre-process raw data before using it for sentiment analysis. The type of pre-processing performed depends on the data used, the task involved, and the machine learning models employed in the sentiment analysis. The data cleaning process used in this study includes stop words removal, special characters cleaning, case folding and stemming. The cleaned texts are vectorized to convert texts input into vector representations. Five vectorization techniques were evaluated in this study: TF-IDF, Count Vectorization, Hashing Vectorization, TinyBERT embeddings, and a proposed hybrid vectorizer. The hybrid approach was developed by combining TF-IDF vectors with TinyBERT embeddings through horizontal stacking, enabling the model to leverage both traditional frequency-based textual features and contextual semantic representations.

The datasets were partitioned into training and testing subsets using an 80:20 train-test split ratio. The training data were used to train both the SVM and ANN models, while the unseen test data were used for evaluation. The models were optimized to reduce overfitting on the training dataset. The predictions from the models are evaluated to determine the level of correctness between the predicted outputs and the ground truth outputs. The closer the predicted values are to the ground truth values, the better the models are generalizing for the task. For sentiment analysis, the outputs can be evaluated using accuracy, precision, recall, and F1 score.

## Results and Discussion

As shown in Fig 1 and Table 1, the proposed hybrid vectorizer (TFIDF + Tiny Bert) achieved consistently high accuracy across all three datasets. The highest performance was observed on the Cherwell dataset, where SVM achieved 96.89% and ANN achieved 96.67% using the hybrid approach. This indicates that the hybrid method is effective on helpdesk service ticket data, likely due to the domain-specific terminology and structured nature of customer support requests. On the Amazon Electronics reviews dataset, the hybrid vectorizer outperformed the rest, achieving accuracy as high as 91.29% using SVM. On the IMDB movie reviews dataset, the hybrid vectorizer achieved 89.42% accuracy using ANN and 89.08% using SVM. Overall, SVM and ANN produced comparable results. SVM performed marginally better on the Cherwell and Amazon datasets, while ANN showed a better performance on the IMDB dataset. These findings suggest that the hybrid vectorizer is effective regardless of the classifier used, making it a reliable and generalizable approach for sentiment analysis tasks.

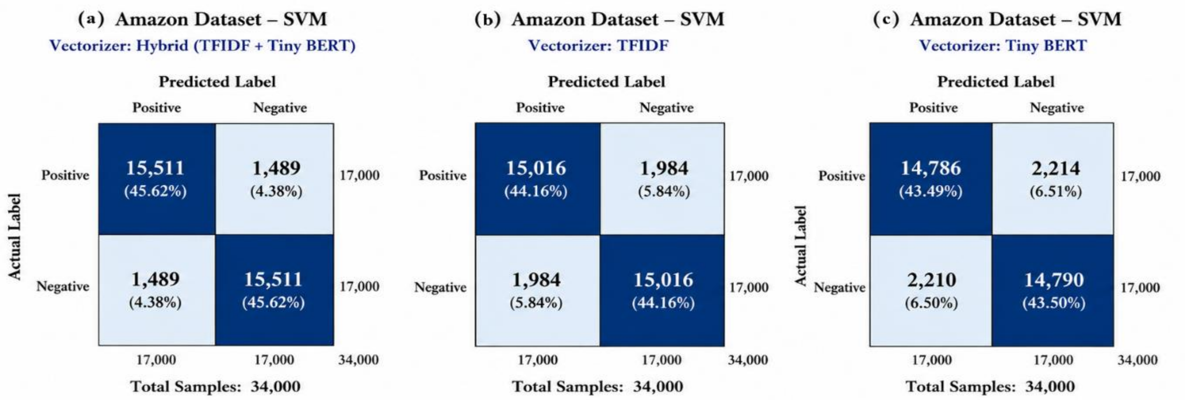



*Fig. 1. Confusion matrices of SVM algorithm with different vectorizers on the Amazon dataset left (a), hybrid vectorizer (b), TFIDF vectorizer (c) TinyBERT vectorizer*

| Dataset | Model | Vectorizer | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|---|
| Cherwell | ANN | Hybrid | 96.67% | 91.99% | 93.65% | 92.80% |
| Cherwell | SVM | Hybrid | 96.89% | 93.26% | 92.95% | 93.11% |
| Amazon | ANN | Hybrid | 91.03% | 91.03% | 91.03% | 91.03% |
| Amazon | SVM | Hybrid | 91.29% | 91.30% | 91.29% | 91.29% |
| IMDB | ANN | Hybrid | 89.42% | 89.48% | 89.41% | 89.41% |
| IMDB | SVM | Hybrid | 89.08% | 89.09% | 89.08% | 89.08% |
| Cherwell | ANN | TFIDF | 96.56% | 92.84% | 91.83% | 92.33% |
| Cherwell | SVM | TFIDF | 96.81% | 94.43% | 91.16% | 92.71% |
| Amazon | ANN | TFIDF | 88.13% | 88.17% | 88.13% | 88.13% |
| Amazon | SVM | TFIDF | 88.32% | 88.32% | 88.32% | 88.32% |
| IMDB | ANN | TFIDF | 88.04% | 88.04% | 88.04% | 88.04% |
| IMDB | SVM | TFIDF | 89.26% | 89.28% | 89.25% | 89.26% |
| Cherwell | ANN | Tiny Bert | 96.85% | 94.05% | 91.77% | 92.67% |
| Cherwell | SVM | Tiny Bert | 96.13% | 91.71% | 91.12% | 91.41% |
| Amazon | ANN | Tiny Bert | 86.55% | 86.96% | 86.84% | 86.84% |
| Amazon | SVM | Tiny Bert | 86.99% | 87.00% | 86.98% | 86.98% |
| IMDB | ANN | Tiny Bert | 78.54% | 78.57% | 78.55% | 78.54% |
| IMDB | SVM | Tiny Bert | 79.30% | 79.33% | 79.31% | 79.3% |

*Table. 1. Experimental results for all the datasets, models and vectorizers (TFIDF, Tiny Bert and hybrid)*

## Conclusion

This study investigated the effects of different vectorization methods on the performance of sentiment analysis models. Traditional vectorization methods, such as TF-IDF, Count and Hashing vectorizers were used alongside a more advanced vectorizer, TinyBERT embedding and a hybrid vectorizer (TF-IDF + TinyBERT). These vectorizers were trained on the SVM and ANN models for sentiment analysis. The hybrid vectorizer outperformed the other vectorizers on the three datasets presented, achieving an accuracy as high as 96.89%. Furthermore, the hybrid vectorizer demonstrated strong performance well with both the SVM and ANN models.

## References

[1] I. Irawaty, R. Andreswari, D. Pramesti. Vectorizer Comparison for Sentiment Analysis on Social Media Youtube *3rd International Conference on Computer and Informatics Engineering (IC2IE)*, pp. 69–74, 2020.
[2] V. Kumar, B. Subba. A Tfidf Vectorizer and SVM based sentiment analysis framework for text data corpus *National Conference on Communications (NCC)*, pp. 1–6, 2020.
[3] X. Jiao, Y. Yin, L. Shang, X. Jiang, X. Chen, L. Li, F. Wang, Q. Liu. TinyBERT: Distilling BERT for Natural Language Understanding, 2019.
[4] Z. Jianqiang, G. Xiaolin. Comparison research on text pre-processing methods on twitter sentiment analysis *IEEE Access* 5, 2870–2879, 2017.

https://doi.org/10.18745/pb.27307

# A Bayesian approach for classifying radio sources fitting photometry and spectroscopy

**Dr Soumyadeep Das[1*] and Prof. Daniel J. B. Smith[1]**
[1]Centre for Astrophysics Research, University of Hertfordshire, Hatfield, UK
*Corresponding author: s.das4@herts.ac.uk

**Radio synchrotron emission associated with distant galaxies is derived from two main processes: (i) star formation, through cosmic-ray electrons accelerated by supernova remnants, and (ii) jets driven by the central supermassive black hole in the subset with an Active Galactic Nucleus (an AGN). Understanding the nature of the radio source population requires distinguishing the physical processes powering the radio emission. Here, we present a novel Bayesian approach that brings together photometric observations spanning UV to far-IR wavelengths with optical spectroscopy, jointly modelling both with a new-generation Spectral Energy Distribution (SED) modelling code, *Prospector*. By incorporating three complementary diagnostics into this probabilistic scheme, we identify nearly twice as many AGN compared to photometry-only approaches. The framework is particularly effective at recovering dust-obscured AGN that may be missed by spectroscopy-based methods. This methodology is well suited to ongoing and upcoming massively multiplexed spectroscopic surveys (e.g., the Herts-led WEAVE-LOFAR and ORCHIDSS), where robust, scalable AGN identification is essential.**

Keywords: galaxies, active galaxies, star formation, galaxy classification, radio continuum surveys

## Introduction

Radio observations provide a unique window to study the extragalactic sky, unobscured by the ubiquitous astrophysical dust which plagues other wavelengths. Low-frequency radio emission is dominated by synchrotron radiation, powered by both star formation (via supernova-driven cosmic ray acceleration) and Active Galactic Nuclei (AGN) through bipolar jets, making radio surveys a key tool for studying the co-evolution of star formation and black hole accretion across cosmic time[1]. The faint radio source population is a heterogeneous mixture of star-forming galaxies (SFGs), radio-quiet AGN (RQ AGN), and high- and low-excitation radio galaxies (HERGs/LERGs). Identification of sources in each class is a critical step towards understanding their nature and the cosmic evolution of the processes that power them[2]. For most sources, radio observations lack the angular resolution to morphologically separate synchrotron emission powered by a galaxy's star formation from that produced by AGN jets, and as a result, attempts to distinguish rely on complementary data. Optical spectroscopy has long been regarded as the gold standard for doing this, especially with the recent development of probabilistic classification algorithms which allow the end user to identify samples based on completeness or purity[3][4][5]. However, spectroscopic samples remain highly incomplete due to selection biases (a majority of sources do not have measured spectra), redshift limitations (since the distances are large, and the sources are faint), and dust obscuration. Photometric data, by contrast, are more readily available, but perhaps provide weaker diagnostic power on their own[6][7]. A classification framework that models both together, leveraging the precision of spectroscopy and the wide wavelength coverage and sensitivity of photometry, is therefore essential for producing complete and reliable samples across the full radio source population. In this work, we present a Bayesian spectro-photometric classification framework that jointly models photometry and spectroscopy using the Prospector SED fitting code[8].

## Data and methods

We focus on 6,085 radio sources detected at 150 MHz by LOFAR in the ELAIS-N1 deep field, cross-matched with 26-band UV-to-far-IR photometry and optical spectra from the DESI Early Data Release[9][10]. The SEDs of these sources encode information about the underlying emission mechanisms, including the stars, dust, accretion disk, and jets. SED modelling is performed with Prospector, which uses a physically-motivated model to simultaneously fit photometry and spectra, estimating full Bayesian posteriors to measure a range of properties including stellar masses and star formation rates. We propagate the posterior distributions into a probabilistic classification scheme, that combines three complementary diagnostics (Fig. 1): (i) mid-IR excess tracing

https://doi.org/10.18745/pb.27307

radiative-mode AGN, (ii) emission-line classifications [e.g. 3,4,5], and (iii) identifying sources with a radio-excess relative to that expected on the basis of their star formation. By using Bayes' theorem in this way, we can produce general purpose probabilistic classifications – allowing the end user to identify samples for further study on the basis of e.g. confidence or completeness as their use case requires – which better reflect our understanding.

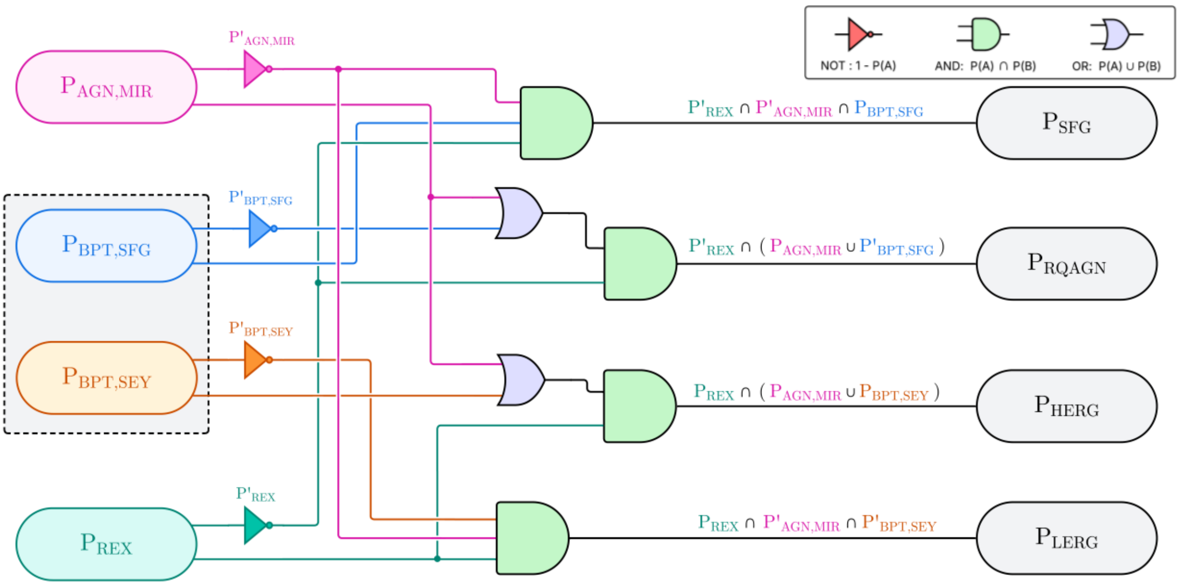



*Fig. 1. Flowchart describing the classification workflow, combining probabilities derived from three complementary diagnostics (left) to produce final probabilities of sources belonging to the SFG, RQ AGN, LERG, or HERG classes. The NOT, AND, and OR logic gates are represented by their usual symbols (see legend at the top-right corner).*

## Results and discussion

At a >90% confidence threshold, we identify 2,109 SFGs, 510 RQ AGN, 512 LERGs, and 121 HERGs from the 6,085-source sample (53% classified). By way of comparison, a maximum-likelihood scheme classifies 93% of sources, trading reliability for a 'best guess' for a larger fraction of the population. Comparison with a photometry-only classification of the same sources reveals agreement (in the sense that our 90% confidence classification agrees with the maximum-likelihood photometric classification) on 98.5% of SFGs but only 21% on RQ AGN; the discrepancy highlights the value of probabilistic classifications by revealing doubt that was previously not apparent. Similarly, comparison with a probabilistic spectroscopy-only classification reveals that 20% of our RQ AGN are classified as SFGs by the spectroscopic method (Fig. 2). A detailed analysis of the individual sources reveals that these sources have a strong mid-IR excess, but their emission-line ratios are consistent with star formation, identifying them as dust-obscured AGN whose hard ionising radiation is hidden from spectroscopic diagnostics. This directly demonstrates the value of the joint approach: the spectrophotometric approach resolves degeneracies by bringing together the information.

## Conclusion

We have demonstrated that classifications produced by jointly modelling photometric and spectroscopic data are able to classify sources that are mis-classified by photometry-only or spectroscopy-only approaches. This has important ramifications for e.g. the cosmic star formation rate density, or the role of accretion processes in shaping galaxies in the distant Universe. With upcoming massively multiplexed spectroscopic surveys such as the Herts-led WEAVE-LOFAR Survey[11] and ORCHIDSS[12] set to obtain millions of optical spectra of faint radio sources, this framework is scalable and directly applicable to next-generation datasets.

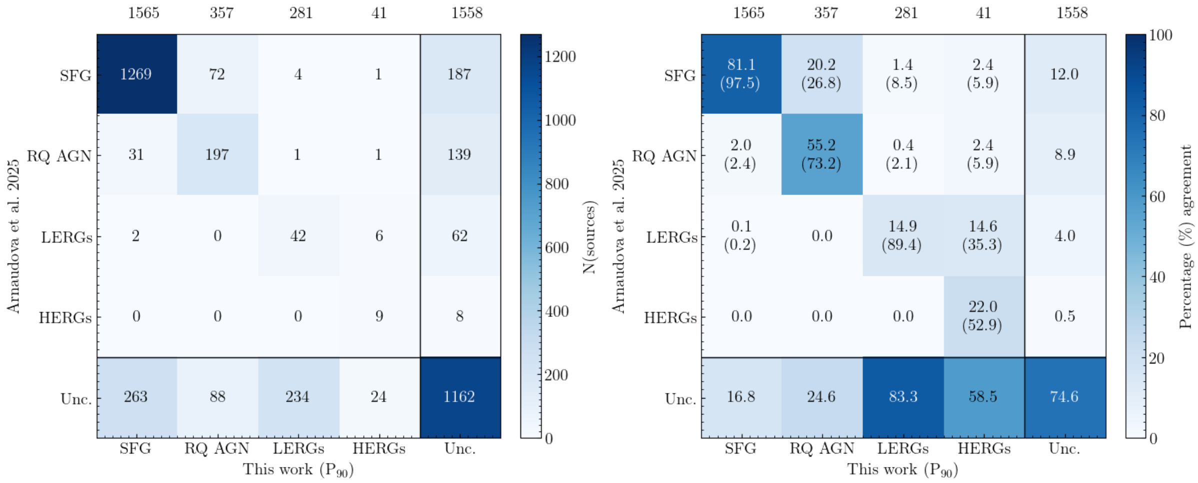



*Fig. 2. Confusion matrices comparing the spectro-photometric classifications with spectroscopy-only classifications from [5]. Both classifications are >90% purity limited. The left panel shows source counts; the right shows percentage agreement between classifications. In each right-hand cell, the top value includes unclassified sources, while the value in parentheses excludes them. The total number of sources in each class (from this work) is shown at the top.*

## References

[1] J. J. Condon and S. M. Ransom, Essential Radio Astronomy. Princeton University Press, 2016.
[2] P. Padovani, 'The faint radio sky: radio astronomy becomes mainstream', A&AR, 24, 1, 13, 2016.
[3] P. N. Best and T. M. Heckman, 'On the fundamental dichotomy in the local radio-AGN population: accretion, evolution and host galaxy properties', MNRAS, 421, 2, 1569–1582, 2012.
[4] A. B. Drake et al., 'The LOFAR two metre sky survey data release 2: probabilistic spectral source classifications and faint radio source demographics', MNRAS, 534, 2, 1107–1126, 2024.
[5] M. I. Arnaudova et al., 'The LOFAR Two-metre Sky Survey Deep Fields: new probabilistic spectroscopic classifications and the accretion rates of radio galaxies', MNRAS, 2025.
[6] P. N. Best et al., 'The LOFAR Two-metre Sky Survey Deep Fields Data Release 1: V. Survey description, source classifications and host galaxy properties', MNRAS, 523, 2, 1729–1755, 2023.
[7] S. Das et al., 'The LOFAR Two-metre Sky Survey: The nature of the faint source population and SFR-radio luminosity relation using PROSPECTOR', MNRAS, 531, 1, 977–996, 2024.
[8] B. D. Johnson, J. Leja, et al., 'Stellar Population Inference with Prospector', ApJS, 254, 2,22, 2021.
[9] J. Sabater et al., 'The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. II. The ELAIS-N1 LOFAR deep field', A&A, 648, A2, 2021.
[10] DESI Collaboration et al., 'Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument', ApJ, 164, 5, 207, 2022.
[11] D. J. B. Smith et al., 'The WEAVE-LOFAR Survey', SF2A-2016, 271-280, 2016.
[12] K. Duncan et al., 'Optical, Radio Continuum and HI Deep Spectroscopic Survey (ORCHIDSS)', The Messenger, 190, 25–27, 2023.

https://doi.org/10.18745/pb.27307

# Uncertainty-Aware Anomaly Detection in Satellite Telemetry using Self-Supervised Learning

**Karthikeyan Arivazhagan[1], and Okechi Onuoha[1*]**
[1]*University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK*
*Corresponding author: o.onuoha@herts.ac.uk

**Satellite telemetry anomaly detection is critical for mission safety, yet deep learning models often produce overconfident errors. This work proposes an uncertainty-aware framework using an LSTM autoencoder with Monte Carlo Dropout. A selective prediction mechanism abstains when uncertainty is high, trading coverage for reduced risk. Evaluated on NASA SMAP/MSL (channels P-1 and M-1) across five seeds, MC Dropout significantly reduces risk at 90% coverage ($p < 0.05$, Cohen's d = 0.8). The risk-based threshold offers a balanced trade-off (FAR = 0.051) without labelled validation data. The method generalises across spacecraft.**

**Keywords: Anomaly detection; satellite telemetry; uncertainty quantification; Monte Carlo Dropout; selective prediction.**

## Introduction

Modern satellites generate high-dimensional telemetry streams that are multivariate, temporally correlated, and unlabelled. Anomalies are rare but can be catastrophic, making reliable detection essential. Traditional rule-based monitors lack adaptability, while deep learning anomaly detectors typically output binary decisions without any confidence measure, leading to overconfident errors that erode operator trust [1].

This project addresses the research question: *"When should a satellite telemetry anomaly detection system abstain from making a prediction, and how does predictive uncertainty affect reliability?"* We develop an uncertainty-aware framework based on a self-supervised LSTM autoencoder augmented with MC Dropout [2] to estimate epistemic uncertainty. A selective prediction mechanism [3] is used to abstain when uncertainty is high, formalised through coverage–risk curves. The system is evaluated on two benchmark channels from the NASA SMAP and MSL datasets [4] (P-1 and M-1) using five random seeds and statistical significance testing.

## Experimental Study

**Dataset and Preprocessing:** The NASA SMAP/MSL datasets contain real telemetry from two spacecraft. We used channels P-1 (SMAP, contextual anomalies) and M-1 (MSL, contextual anomalies). A sliding window of width 50 and stride 1 was applied; windows were normalised using a ***StandardScaler*** fitted only on training data (no leakage). Window-wise labels were created (anomalous if any point inside is anomalous).

**Model Architecture**: LSTM autoencoder: encoder LSTM (1-layer, hidden size 64, dropout 0.2), linear bottleneck 64→16, decoder LSTM (1-layer, hidden size 64, dropout 0.2), output linear 64→input dimension. Trained only on normal windows (20 epochs, batch 64, Adam ***lr***=0.001) for 5 seeds.

**Uncertainty Estimation**: During inference, dropout is enabled (MC Dropout). For each test window, 50 stochastic forward passes give a distribution of reconstructions. Predictive uncertainty = mean variance across features/time; reconstruction error = MSE between original and mean reconstruction.

**Threshold Strategies**: Three thresholds compared: percentile (95th percentile of training errors), ROC-optimal (maximising Youden's J), risk-based (95th percentile of reconstruction error + α×uncertainty, α=0.95).

**Evaluation Metrics**: AUROC, PR-AUC, precision/recall/F1, false alarm rate (FAR), detection delay, coverage–risk curves, risk at 90% coverage. Statistical validation: mean ± 95% CI, paired t-test, Cohen's d.

https://doi.org/10.18745/pb.27307

## Results and discussion

**Figure 1** shows the **coverage-risk curve for P-1**. At 90% coverage (abstaining on the most uncertain 10% of windows), MC Dropout **risk drops from 0.149 to 0.138**, while deterministic only improves from **0.149 to 0.145**. The reduction (0.007 absolute, 5% relative) is significant ($p<0.05$, Cohen's d=0.8).

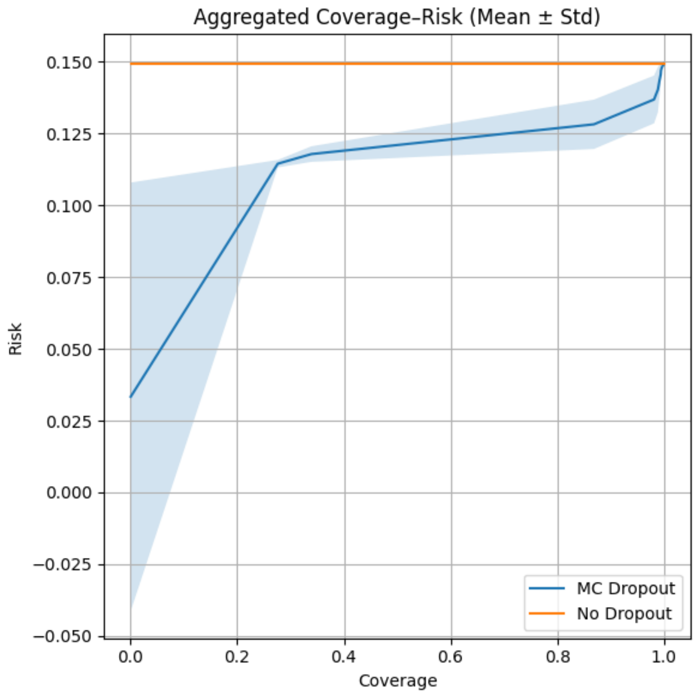



*Figure 19: Coverage Risk Curve for P-1 Channel (mean ± std)*

**Figure 1 - Aggregated coverage-risk curve** (mean ± std). MC Dropout (blue) achieves lower risk than deterministic (orange) at all coverages.

**Table 1** summarises **aggregated AUROC and PR-AUC** for both channels **(5 seeds, MC Dropout)**. The modest values reflect extreme class imbalance **(<5% anomalies)** and the difficulty of detecting contextual anomalies. The small confidence intervals confirm stability across seeds. The primary benefit of uncertainty awareness is not in raw AUROC but in the coverage–risk trade-off (Fig. 1).

*Table 1: Aggregated AUROC and PR-AUC (mean ± 95% CI)*

| Channel | AUROC (MC Dropout) | PR-AUC (MC Dropout) |
|---|---|---|
| **P-1** | 0.422 ± 0.006 | 0.088 ± 0.004 |
| **M-1** | 0.498 ± 0.012 | 0.121 ± 0.009 |

**Figure 2** compares the three thresholds. ROC-optimal gives the **lowest FAR (0.015)** but exceptionally low **recall (0.019)**. Percentile and risk-based thresholds have similar precision (~0.074) and recall (~0.035). The risk-based threshold for MC Dropout has **FAR=0.051 (vs. 0.052)** and **F1=0.045 (vs. 0.047)** – a balanced, unsupervised option.

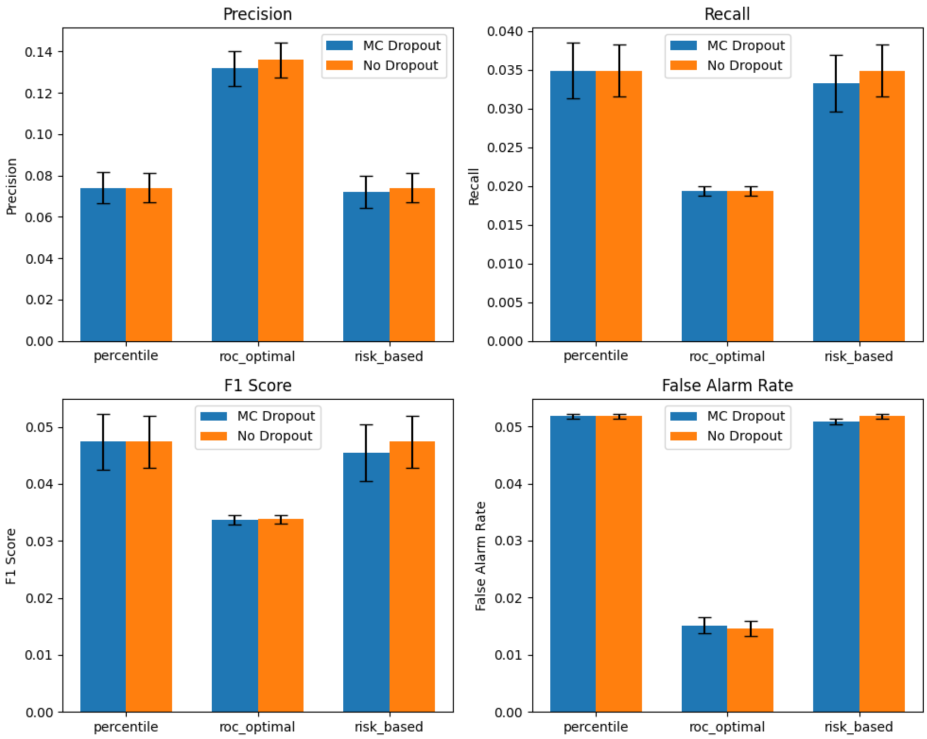



*Figure 20: Comparison of Threshold Metrics for P-1*

**Figure 2 – Threshold metrics comparison (2×2 grid).** ROC-optimal minimises FAR at recall cost; risk-based offers a trade-off without labels.

**Detection delay** for the percentile and risk-based thresholds is approximately **97 windows** (≈4,850 time steps). This high delay is inherent to contextual anomalies, which unfold slowly over hundreds of time steps. The model only raises an alert after a sustained period of elevated reconstruction error, which is acceptable given the reduction in false alarms. The method generalises across spacecraft (similar performance on M-1), indicating channel-agnostic applicability.

## Conclusion

We have developed and validated an uncertainty-aware anomaly detection system for satellite telemetry that can abstain when uncertain. Coverage–risk analysis provides a principled answer to the research question: abstain when predictive uncertainty exceeds a threshold that reduces risk to an acceptable level (e.g., **remove the most uncertain 10% of windows** to **lower risk from 0.149 to 0.138**). The risk-based threshold offers a practical decision rule without requiring labelled validation data. The framework is reproducible, statistically validated, and generalises across different telemetry channels. Future work will extend to more channels, explore other uncertainty methods (e.g., deep ensembles), and address real-time deployment.

## References

[1] D. Amodei et al., "Concrete problems in AI safety," *arXiv:1606.06565*, 2016.
[2] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation," in *ICML*, 2016, pp. 1050–1059.
[3] Y. Geifman and R. El-Yaniv, "Selective classification for deep neural networks," in *NeurIPS*, vol. 30, 2017, pp. 4878–4887.
[4] K. Hundman et al., "Detecting spacecraft anomalies using LSTMs," in *KDD*, 2018, pp. 387–395.

https://doi.org/10.18745/pb.27307

# Mechanical and Durability Performance of Low-carbon Concrete Utilizing Recycled Concrete Coarse Aggregates and Fly Ash for Road Pavement Construction

**Goda K. Baral and Mike Otieno**

*Department of Engineering, School of Physics, Engineering and Computer Science, Centre for Engineering research, University of Hertfordshire*

*Corresponding author: gb25aao@herts.ac.uk

## Abstract

**Concrete production using construction and demolition waste (CDW) together with low-carbon binders such as fly ash (FA) can significantly reduce the negative environmental impact associated with the construction, repair and maintenance of durable and sustainable road pavements. This study focused on the mechanical and durability performance of concretes made using recycled concrete coarse aggregates (RCCA) and FA. Specimens made using 70/30 CEM II/FA and 50/50 RCCA/NA (natural aggregates) at a 0.50 w/b ratio were tested for mechanical (compressive and flexural strength, and elastic modulus) and durability (chloride penetration and freeze-thaw damage) performance. Preliminary results indicate a potential to use RCCA and FA in pavement construction.**

Keywords: Construction and demolition waste; Recycled concrete coarse aggregates; Natural aggregates; Fly ash

## Introduction

Concrete is one of the most widely used construction material, but its production has significant environmental impacts – consuming large quantities natural resources [1]. In addition, the construction sector generates vast quantities of CDW estimated at over 450 million tons annually in the EU, with nearly 75% still ending up in the landfills despite its strong recycling potential [2]. These issues highlight the need for sustainable construction practices. This study aims to assess the feasibility of using RCCA and FA in rigid road pavement concretes and the outcomes are expected to promote sustainable construction practices.

## Method

Table 1 presents a summary of the concrete mix proportions used. Compressive strength (BS EN 12390-3) [3], flexural strength (BS EN 12390-5) [4] and modulus of elasticity (BS EN 12390-13:2021) [5] were determined to assess the respective mechanical properties of concrete. Durability performance was assessed by carrying out freeze-thaw resistance test (CEN/TS 12390-9:2016) [6], carbonation resistance test (BS EN 12390-12) [7] and rapid chloride penetration test (RCPT, ASTM C1202).

*Table.1. Concrete mix proportions (kg/m$^3$) [8]*

| Mix label | Binder | | Coarse aggregate | | Fine aggregate | Water |
|---|---|---|---|---|---|---|
| | CEM II/A-L 32.5 R (70%) | Fly ash (30%) | Natural (50%) | Recycled (50%) | | |
| Reference concrete | 475 | – | 1000 | – | 710 | 237.5 |
| RCCA-concrete | 332.5 | 142.5 | 500 | 500 | 710 | 237.5 |

https://doi.org/10.18745/pb.27307

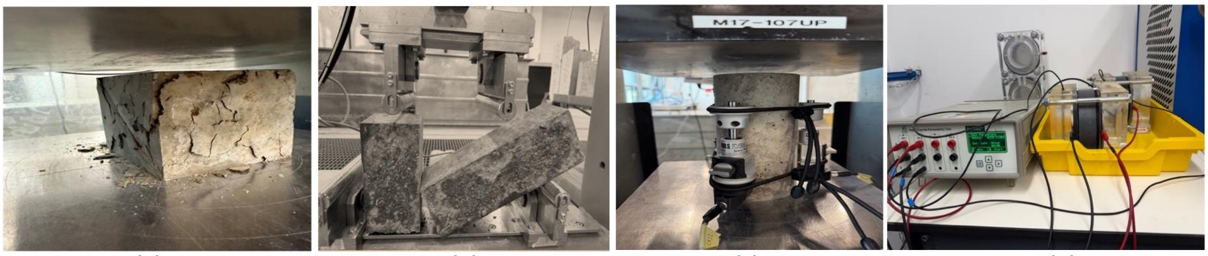



*Fig. 14. Tests performed – (a) compressive strength, (b) flexural strength, (c) modulus of elasticity, (d) chloride penetration*

## Preliminary Results

The compressive strength and flexural strength tests were evaluated to determine the load bearing capacity and ability of resistance to cracking under flexural stress of the pavement. Modulus of Elasticity was measured to understand the stiffness of the concrete pavement under wheel load. Durability of the concrete was assessed by carrying out accelerated 28-day and 56-day chloride penetration test (rapid chloride penetration test, RCPT) and by exposing the concrete to 56 freeze-thaw cycles (–15 to 20 °C). Preliminary results show that RCCA concrete improved mechanical properties. As expected, a steady increase in compressive strength was observed between 7 and 56 days. A similar trend was observed in the flexural strength and elastic modulus (E-mod) results – see Fig. 2 (a)-(c). Negligible (0.02%) mass loss was recorded after 56 freeze-thaw cycles relative to the reference concrete.

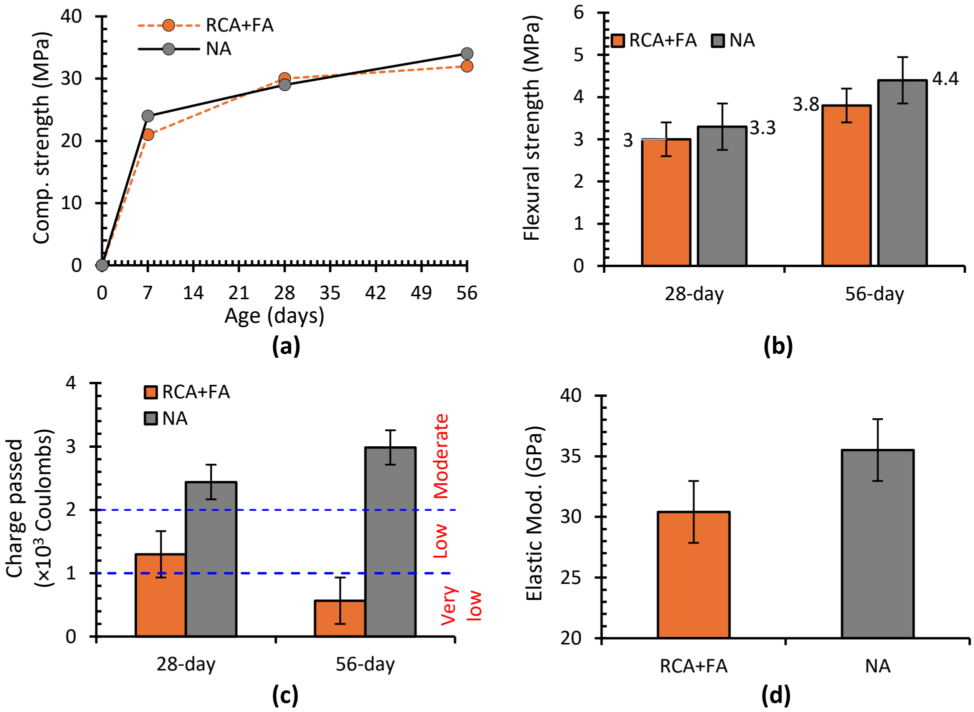



*Fig. 2. (a) Compressive strength, (b) flexural strength, (c) Chloride penetration (RCPT) and (d) Elastic modulus results*

## Conclusion and way forward

Based on the preliminary results, the following conclusions can be made: (i)The RCCA blended concrete (50% RCCA and 30% FA replacement) demonstrates a consistent strength development in terms of both compressive and flexural strength, with time meeting adequate structural requirement for concrete pavements; (ii)The mix exhibits good durability performance, specially high resistance to chloride and negligible mass loss in 56 freeze-thaw cycles, indicating superior surface integrity under environmental conditions; (iii) The replacement of natural aggregates with RCCA combined with FA provides a viable sustainable alternatives for road pavement construction without compromising the essential mechanical and durability properties. In general, the results show satisfactory mechanical and durability performance of FA-blended RCCA concrete and provide a basis for further tests to ascertain its suitability for use in concrete road pavements. Ultimately, a critical assessment will be made on its suitability to contribute to reducing the negative environmental impact associated with the construction, repair and maintenance of durable and sustainable road pavements. The study is ongoing.

## References

[1] M. Alexander, S. Mindness and M. Otieno, (2025), Aggregates in Concrete (2$^{nd}$ Edition), Taylor and Francis, ISBN: 9781032221304, doi.org/10.1201/9781003271253.

[2] A. Kanellopoulos, D. Nicolaides and M. Petrou, (2014), Mechanical and durability properties of concretes containing recycled lime powder and recycled aggregates, Construction and Building Materials, Vol. 53, pp. 253-259, 2014.

[3] BS EN 12390-3:2019, (2019), Testing hardened concrete – Part 3: Compressive strength of test specimens, British Standards Institution.

[4] BS EN 12390-5:2019, (2019), Testing hardened concrete – Part 5: Flexural strength of test specimens, British Standards Institution.

[5] BS EN 12390-13:2021, (2021), Testing hardened concrete – Part 13: Determination of secant modulus of elasticity in compression, British Standards Institution.

[6] PD CEN/TS 12390-9:2016, (2016), Testing hardened concrete – Part 9: Freeze-thaw resistance with de-icing salts- Scaling British Standards Institution, British Standards Institution.

[7] BS EN 12390-12:2020, (2020), Testing hardened concrete – Part 12: Determination of carbonation resistance of concrete – Accelerated carbonation method British Standards Institution, British Standards Institution.

[8] G. Owens, (2013), Fundamentals of Concrete, Midrand: The Concrete Institute, South Africa.

https://doi.org/10.18745/pb.27307

# Probabilistic Robotics Approach to EKF-SLAM for Autonomous Navigation in GPS-Denied Environments

**Don Joy[1], Frank Foerster[1], and Okechi Onuoha[1*]**
[1]*Department of Computer Science, University of Hertfordshire*
*Corresponding author: o.onuoha@herts.ac.uk

**This paper presents an Extended Kalman Filter Simultaneous Localisation and Mapping (EKF-SLAM) framework for an autonomous vehicles operating on a cone-defined closed circuit using on-board sensing only. The aim is to investigate the suitability of probabilistic filtering framework in maintaining accurate pose estimation while simultaneously constructing a landmark map without external positioning systems. Experimental evaluation over 211.4 s and 1,489 logged samples produced a mean position error of 0.694 m, a root mean square error (RMSE) of 0.756m and a mean yaw error of 0.483°. The results indicate EKF-SLAM remains effective for structured GPS-denied navigation.**

Keywords: Extended Kalman Filters; Simultaneous Localization and Mapping; Pose Estimation, Robotics; Control Systems.

## Introduction

Autonomous vehicles require accurate localisation to operate safely and efficiently. In many practical environments, GPS signals may be unavailable, degraded or unreliable. Simultaneous Localisation and Mapping (SLAM) estimates vehicle pose while concurrently constructing a map of unknown surroundings [1],[2]. This work investigates EKF-SLAM in a cone-defined racing-style environment where sparse coloured cones form track boundaries.

## EKF-SLAM Formulation and Motion Model

The system state contains vehicle pose, motion variables and landmark coordinates. Vehicle motion is propagated using wheel speed and inertial yaw-rate measurements. Cone observations are incorporated through the EKF correction stage [1]. Nearest-neighbour association with Mahalanobis gating rejects inconsistent matches, while unmatched observations initialise new landmarks [1],[2].

The filter maintains a joint Gaussian estimate of vehicle state and landmark coordinates [2],[3]. The state vector used in the project is

$$x = [p_x,\ p_y,\ \psi,\ v,\ b_g,\ l_{x1},\ l_{y1},\ \dots,\ l_{xN},\ l_{yN}]^T$$

Where, $p_x$ and $p_y$ are global position coordinates in the x- and y-axis respectively, $\psi$ is vehicle yaw angle, $v$ is forward velocity, $b_g$ is gyroscope bias, and each landmark is represented by Cartesian coordinates $(l_{x1}, l_{y1}, \dots, l_{xN}, l_{yN})$.

The EKF prediction step propagates the state mean through a discrete-time nonlinear motion model using odometry and IMU measurements [2]. The discretized vehicle kinematics used is

$$p_x^{k+1} = p_x^k + v^k cos\,\psi^k \Delta t + \frac{1}{2} a_x^k cos\,\psi^k \Delta t^2 p_y^{k+1} = p_y^k + v^k sin\,\psi^k \Delta t + \frac{1}{2} a_x^k sin\,\psi^k \Delta t^2 \psi^{k+1} = \psi^k + (\omega_z^k - b_g^k)\Delta t$$

$$v^{k+1} = v^k + a_x^k \Delta t$$

Where, $a_x$ is longitudinal acceleration, $\omega_z$ is yaw rate, and $\Delta t$ is the sampling interval.

## Prediction Step and Measurement Update

Cone landmark observations were incorporated using innovation-based correction [2],[4]. The innovation vector was defined as:

https://doi.org/10.18745/pb.27307

$$y_k = z_k - \hat{z}_k$$

where $z_k$ is the measured observation and $\hat{z}_k$ is the predicted observation.

The Kalman gain was computed as

$$K_k = P_{k|k-1}H_k^T(H_k P_{k|k-1}H_k^T + R_k)^{-1}$$

where $H_k$ is the observation Jacobian and $R_k$is the measurement noise covariance [2] .

The corrected state estimate gives

$$x_k = x_{k|k-1} + K_k y_k$$

To preserve covariance symmetry and improve robustness, the Joseph-form covariance update used is

$$P_k = (I - K_k H_k)P_{k|k-1}(I - K_k H_k)^T + K_k R_k K_k^T$$

This formulation is particularly useful in long-duration EKF runs where floating-point asymmetry may otherwise accumulate [2].

### Experimental Setup

Experiments were conducted using a simulated cone-track circuit representing a GPS-denied autonomous navigation scenario. Ground-truth pose and landmark coordinates were available for offline evaluation. A runtime dataset spanning 211.4 s with 1,489 samples was analysed.

### Results

The result of the study is presented in Figure 1, and the performance metric is presented in Table 1. The estimated trajectory closely followed the reference path over the full circuit. Small deviations were visible in higher-curvature regions. The peak RMSE was slightly above 1.0m and occurred near the mid-run phase, likely associated with reduced geometric observability or temporary landmark ambiguity. Following re-observation of landmarks, the filter recovered rapidly as suggested in [1],[5].

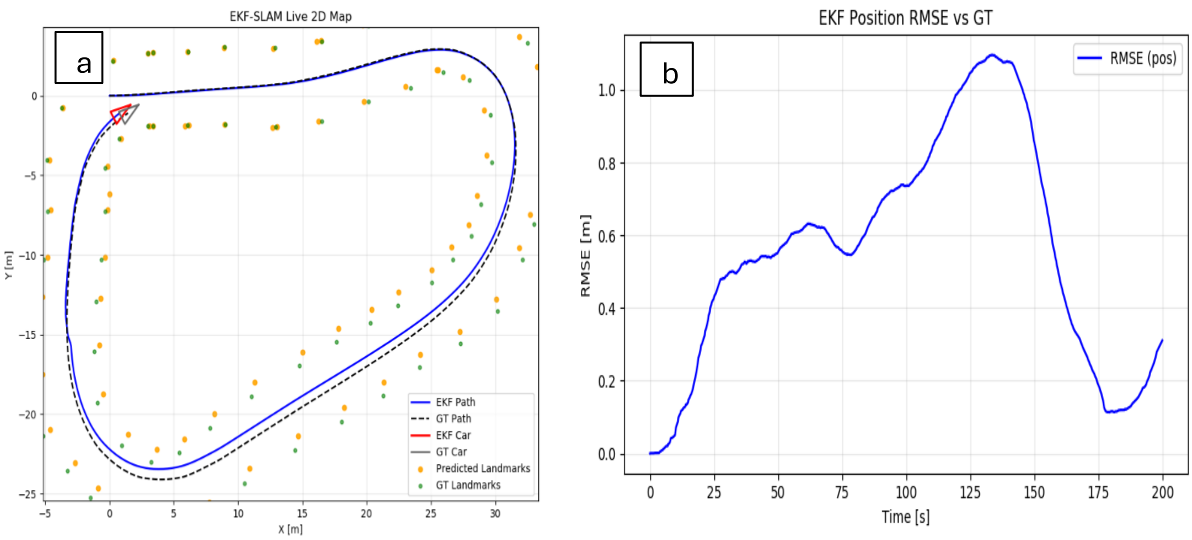



*Figure 1. The results of the experiments. (a) EKF predicted map (b) EKF predicted position vs ground truth (GT) position.*

https://doi.org/10.18745/pb.27307

*Table. 1. Performance Metrics*

| Metric | Result |
|---|---|
| Mean Position Error | 0.694 m |
| Position RMSE | 0.756 m |
| Maximum Position Error | 1.374 m |
| Mean Yaw Error | 0.483° |
| Yaw RMSE | 0.643° |
| Maximum Yaw Error | 3.004° |

## Conclusion

A probabilistic EKF-SLAM framework for autonomous cone-track navigation in GPS-denied environments was presented and validated. By fusing odometry, inertial sensing and cone observations, the system achieved sustained sub-metre localisation accuracy while simultaneously constructing an accurate landmark map.

## .References

[1] Dissanayake, M.W.M.G., Newman, P., Clark, S., Durrant-Whyte, H. and Csorba, M. "A solution to the simultaneous localisation and map building problem," *IEEE Transactions on Robotics and Automation*, vol. 17 no 3, pp. 229–241, 2021.

[2] Thrun, S., Burgard, W. and Fox, D. "Probabilistic Robotics*" Cambridge, MA: MIT Press*, 2005

[3] Durrant-Whyte, H. and Bailey, T. "Simultaneous localisation and mapping: Part I," *IEEE Robotics and Automation Magazine*, vol. 13 no. 2, pp. 99–110, 2006.

[4] Bailey, T. and Durrant-Whyte, H. "Simultaneous localisation and mapping (SLAM): Part II", *IEEE Robotics and Automation Magazine*, vol. 13 no. 3, pp. 108–117, 2006

[5] Castellanos, J.A., Martinez-Cantin, R., Tardós, J.D. and Neira, J. "Robocentric map joining: Improving the consistency of EKF-SLAM," *Robotics and Autonomous Systems*, vol. 55 no. 1, pp. 21–29, 2007

University of Hertfordshire UH
https://doi.org/10.18745/pb.27307

# A Multi-Platform Pipeline for Deriving Consumer Intelligence from User Generated Content

**Linda Mobuogwu 1[1]*, Dr. Wei Ji 2[2], and Dr. Barry Ip 3[3]**

[1]School of Physics, Engineering and Computer Science, University of Hertfordshire

[2]School of Physics, Engineering and Computer Science, University of Hertfordshire

[3]School of Physics, Engineering and Computer Science, University of Hertfordshire

*Corresponding Author: L.n.mobuogwu@herts.ac.uk

**Abstract**

**This study introduces a novel multi-platform pipeline for extracting valuable consumer insights from social media data. Unlike existing approaches that focus on mining social media data for business-oriented intelligence, the proposed approach supports the integration and transformation of large-scale unstructured social data from YouTube and Reddit into structured insights that support consumers in making informed decisions. The pipeline uses a combination of text preprocessing, consumer-aligned feature extraction, topic modeling, and sentiment analysis to identify the preferences, needs, concerns, and motivations of real users, thereby offering a scalable approach to transforming multi-platform social content into actionable consumer-oriented intelligence.**

**Keywords**: **Consumer Intelligence; Social Media Analytics; Text Mining; Topic Modeling; Sentiment Analysis**

## Introduction

Social media users rely on social platforms to interact with other users, seek information, relax, and learn from the real-life experiences of others for entertainment or to make informed decisions [1]. As a result, a massive amount of User-Generated Content (UGC) is available in the form of unstructured data that is useful in deriving actionable insights. Existing approaches focus on deriving business-oriented insights from the available UGC to support business goals and profitability [2],[3]. Also, useful information that users seek to make informed purchasing decisions is often fragmented across multiple platforms, requiring them to filter through large volumes of data to find relevant information. There is lack of a unified multi-platform pipeline with consumer-focused feature alignment across diverse social sources for consumer decision support. Unlike prior approaches, this study utilizes consumer-aligned feature extraction to map core product aspects across social media platforms, and text modeling techniques, including topic modeling and sentiment analysis, to derive intelligence for consumers. Through this process, latent features in fragmented user conversations can be identified and mapped to aspects relevant to the needs, concerns, and motivations of consumers, enabling structured consumer-aligned insights for decision support from large-scale unstructured data.

https://doi.org/10.18745/pb.27307

## Methods:

To demonstrate the proposed approach, the social media conversations for a smartphone brand, Oppo, were collected from YouTube and Reddit using the developer APIs for each platform. The methodology followed a multi-step process for deriving consumer intelligence from smartphone-related social media data, as illustrated in Figure 1. The study used publicly available social media data. No direct interaction with social media users was conducted, and user identities were anonymized during analysis. The multi-phase pipeline involved ingesting the datasets and applying preprocessing techniques to remove missing values and noise such as URLs, special characters, and HTML tags that do not add semantic meaning to the data. Also, lowercasing and normalization were applied to the text for data uniformity. Term Frequency-Inverse Document Frequency (TF-IDF) was then used to extract features in the data by highlighting the importance of words in the text data. Unsupervised machine learning techniques, Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF), were used to uncover hidden themes in the data. Following a similar feature lexicon used by [4], predefined keywords to categorize consumer-aligned aspects of smartphones were selected - 'battery', 'camera', 'software', 'price', 'display', 'screen', 'design', 'performance', 'value', 'storage', 'ui', 'audio', 'network', 'security', 'connectivity', 'durability', 'charger', 'brand', 'memory'. Sentiment analysis was then used to classify the smartphone aspects into positive, negative, or neutral sentiments. The pipeline standardizes text data processing to ensure that the feature-based analysis is comparable across diverse social sources.

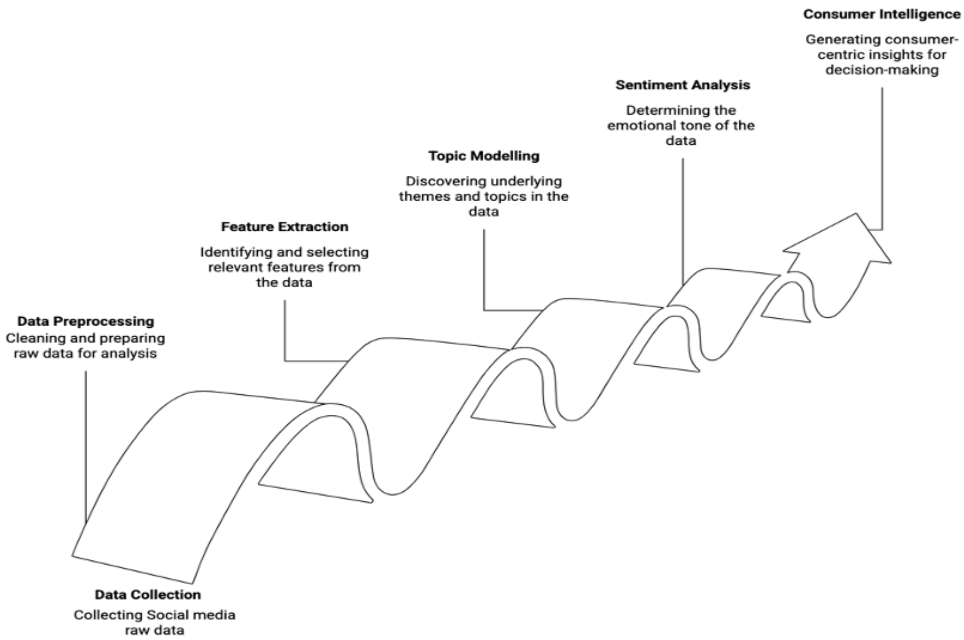



***Figure 1: Data-Driven Consumer Intelligence Framework***

## Results and Discussions:

Feature extraction using predefined smartphone keywords quantified smartphone users' priorities, and topic modeling showed the dominant user needs, including conversations about smartphone features, firmware upgrades, accessories, region/shipping, and general user experiences. Comparison of LDA and NMF topic results showed that NMF produced more coherent and interpretable topics, making it the preferred approach for deriving consumer intelligence from smartphone-related social media data. Sentiments varied across the smartphone features, suggesting more positive sentiments across most of the smartphone aspects for the two platforms (Figure 2). Subfigures 2a and 2b show the frequency distributions of smartphone features in YouTube and Reddit data, with camera being the most frequently mentioned feature at 19.24% and 25.1% respectively.

Price emerged as the second highest mentioned feature for YouTube at 18.84%, while for Reddit, the price of Oppo smartphone was less frequently mentioned (9.6%) than battery and ui. These indicate platform-specific priorities and feature importance that support informed purchase decision-making. In Subfigure 2c, smartphone price and screen were the dominant aspects for YouTube and showed the highest positive sentiment among all the features, while in Subfigure 2d, battery, camera, price, and screen were dominant for Reddit, also showing the highest positive sentiment. Across the two platforms, most of the features showed higher positive sentiments, indicating that those aspects of the smartphone products are viewed favorably. These results show that social media contains insights relevant to consumer needs and concerns consistent with deriving intelligence for consumers. The visualizations provide high-level feature-based comparisons that align with the goal of deriving consumer intelligence from UGC.

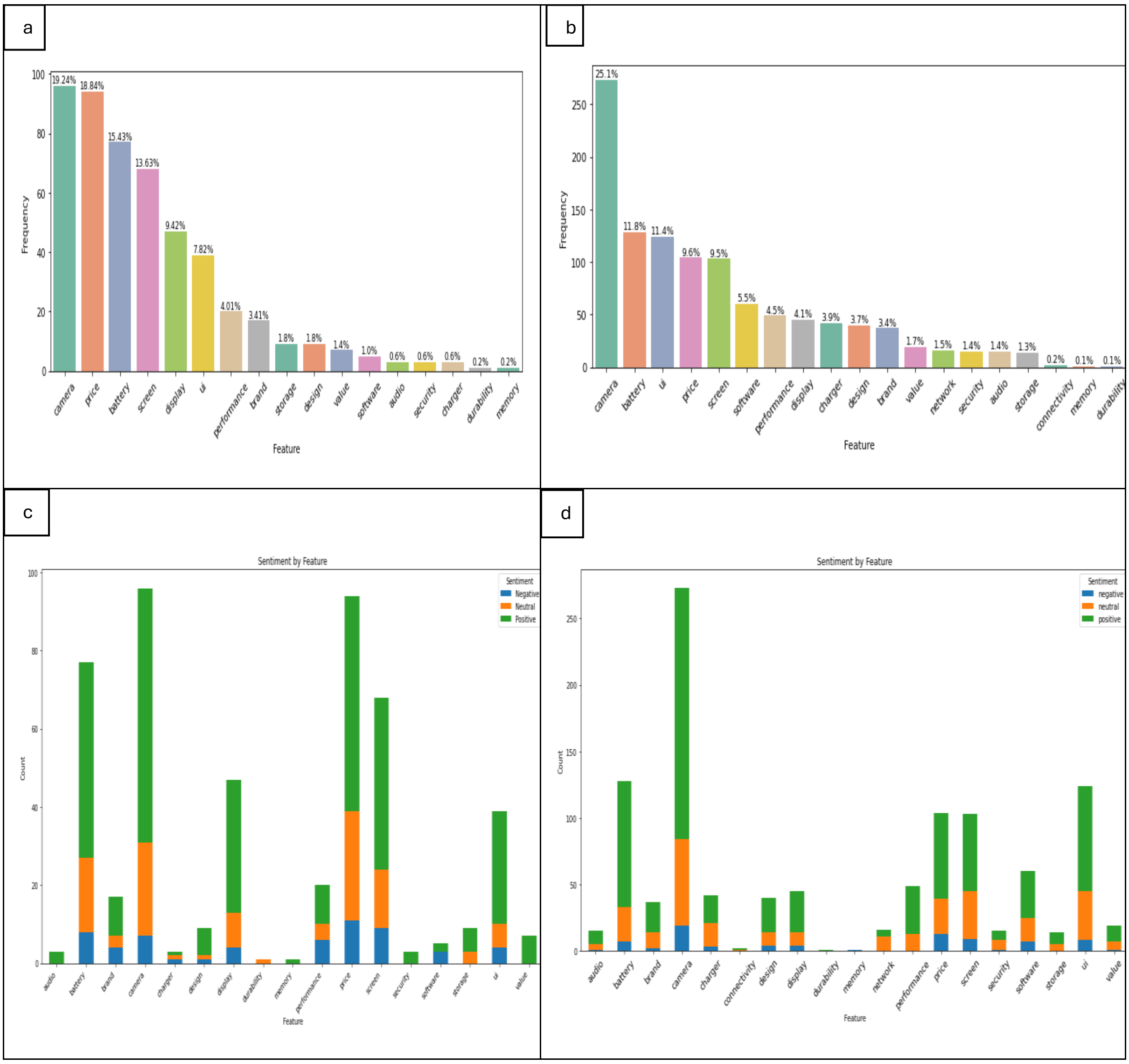



*Figure 2: Frequency and Sentiment Distributions for Oppo by Smartphone feature (a) Frequency distribution on YouTube (b) Frequency distribution on Reddit (c) Sentiment distribution on YouTube (d) Sentiment distribution on Reddit*

https://doi.org/10.18745/pb.27307

## Conclusions

As demonstrated, multi-platform pipeline was used to extract structured insights from unstructured social media data. This approach improves coverage by collecting useful data fragmented across different platforms to generate useful information to support consumers in their decision-making process. Feature-level sentiment analysis allows consumers to understand the opinions of real-users regarding different smartphone aspects. Using visualization tools, the results derived from topic modeling and sentiment analysis can be presented to users in a clear manner to help them make informed decisions.

## References

[1] M. Chen and A.Y. Peng, "Why do people choose different social media platforms? Linking use motives with social media affordances and personalities" *Social Science Computer Review*, vol. *41*, no. 2, pp.330-352, 2023.

[2] C. Jiang, A. Alqadhi and M. Almesbahi, "Conducting product comparative analysis to outperform competitor's product using Teardown" JST Model. Concurrent Engineering, vol 30, no.1, pp.21-31, 2022.

[3] O. Babafemi, D. Itakpe Jonathan and T. Afolabi Ibukun "Twitter Sentiment Based Mining for Decision Making using Text Classifiers with Learning by Induction" Journal of Physics: Conference Series, vol. 1299, no 1, 2019

[4] Y. Kim, R. Dwivedi, J. Zhang and S. R. Jeong, "Competitive intelligence in social media Twitter: iPhone 6 vs. Galaxy S5" Online Information Review, vol 40, no.1, pp.42–61, 2016.

https://doi.org/10.18745/pb.27307

# Analysis for the Implementation of a Manufacturing Enterprise Resource Planning System into a Formula Student Team

**Fernanda Herrejon[1*], Alix Vargas[1]**
[1]*School of Physics Engineering and Computer Science, University of Hertfordshire*
**Corresponding author: fh22aaq@herts.ac.uk*.com

**Formula Student (FS) teams operate as small-scale engineering organizations but face challenges due to high member turnover, leading to inconsistent processes and limited knowledge transfer. This project investigates the feasibility of implementing a custom Enterprise Resource Planning (ERP) system to improve standardisation and process efficiency. A PRISMA-based literature review and structured interviews identified key operations that supported the TO-BE workflow development, to be translated into a prototype using Oracle APEX, integrating 5 main activity modules. The findings demonstrate strong alignment between team needs and ERP capabilities, highlighting the potential to enhance coordination, transparency, and overall project performance.**

Keywords: ERP System; Digitalization; Project Planning; Manufacture; Formula Student;

## Introduction

The University of Hertfordshire aims to prepare its students for their professional life by giving them the opportunity to participate in extracurricular engineering projects, such as Formula Student. Joining a Formula Student team, requires students to develop competencies in project coordination, project planning, procurement management, process integration, among other cross-functional collaboration activities. UH Racing is well known as a team of innovation and success by building single-seat race cars since 1997. However, despite several successful years, the team faces recurring challenges in maintaining standardized processes across areas such as procurement, manufacturing tracking, cost analysis, and documentation. These inefficiencies frequently result in duplicated efforts, communication gaps, and inconsistent data management. The environment also develops practical competencies in engineering management, digital coordination, workflow integration, and operational decision-making.

Enterprise Resource Planning (ERP) systems are integrated software platforms that unify an organization's processes and daily operations into a centralized database [1]. By centralizing data and process flows with an ERP system, communications across departments and data-driven decisions can improve. In large-scale industries, ERP systems have proven successful for improving operational efficiency and resource coordination [2]. However, their implementation in a small-scale project remains limited, creating an opportunity to adapt ERP principles to the specific needs of an engineering project and student-lead organisation. This project aims to analyse the feasibility and requirements of implementing a Manufacturing ERP System tailored to an FS team. The proposed system would centralize the teams' operations (manufacturing processes, procurement, financials, design validation, etc) into a digital platform, enhancing process standardization, traceability, and productivity. Developing this ERP system looks to transform the team´s management structure into one that closely aligns with industrial standards, improving operational performance, process standardization and knowledge continuity for future generations of UH Racing.

## Methodology

This study adopts a hybrid research and development methodology, combining quantitative and qualitative data collection, allowing theoretical foundation and operational experiences to identify and analyse the organizational requirements for implementing an ERP system within an FS team. The qualitative literature review provides a theoretical foundation by identifying functional requirements with ERP systems, such as Industry 4.0 and manufacturing digitalization, in different organizations. Then, quantitative data is collected through interviews with key members of the FS team with the purpose of identifying operational challenges, workflow inefficiencies, and user expectations to then define the new ERP process workflows. These methods support the development of a prototype ERP system using Oracle APEX, showing the relevance, applicability,

https://doi.org/10.18745/pb.27307

and potential impact of the proposed ERP solution. Ethical approval for the interviews was obtained (Disclaimer 1595-ST-HSET-2025) in accordance with University of Hertfordshire research procedures.

## Results and discussion

Quantitative analysis of interview responses (Fig. 1) identified 4 recurrent operational challenges within the team such as fragmented data management, limited workflow visibility, lack of process traceability, and weak cross-functional coordination. The analysis also identified that procurement and manufacturing activities represented the highest-priority areas for ERP integration and process improvements.

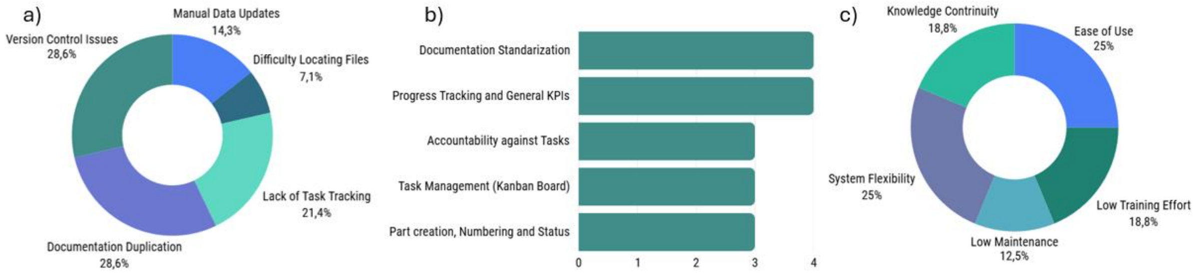



*Figure 22. Quantitative Interview analysis results and ERP implementation priorities. –(a) Current Tools and Systems Areas of Opportunity, (b) ERP Features Requirements, (c) ERP Adoption Considerations*

Before developing the Oracle APEX prototype, the main functional areas of the FS team (Project Management, Design & Engineering, Manufacturing, Procurement, Human Resources, and Marketing & Partnerships) were identified and their current processes defined. The AS-IS process mapping provided a good understanding of how activities flow throughout a complete season and established a foundation for defining the functional requirements of a centralised ERP system and the necessary modules. The TO-BE process model was then developed to address the limitations identified, introducing a digitally integrated operational workflow without changing the team's organizational structure and instead it introduces controlled interactions between departments through ERP logic. Compared to the AS-IS model, the TO-BE process flow introduces operational transparency, reduce coordination risk, and increase managerial control through adding automated workflow validation, real-time KPI tracking, and a centralized data management system. The proposed system supports operational monitoring through KPI visibility, workflow tracking, task status updates, and centralized reporting across functional areas. Fig. 2 illustrates the final ERP architecture structure around five core functional modules aligned with the TO-BE model.

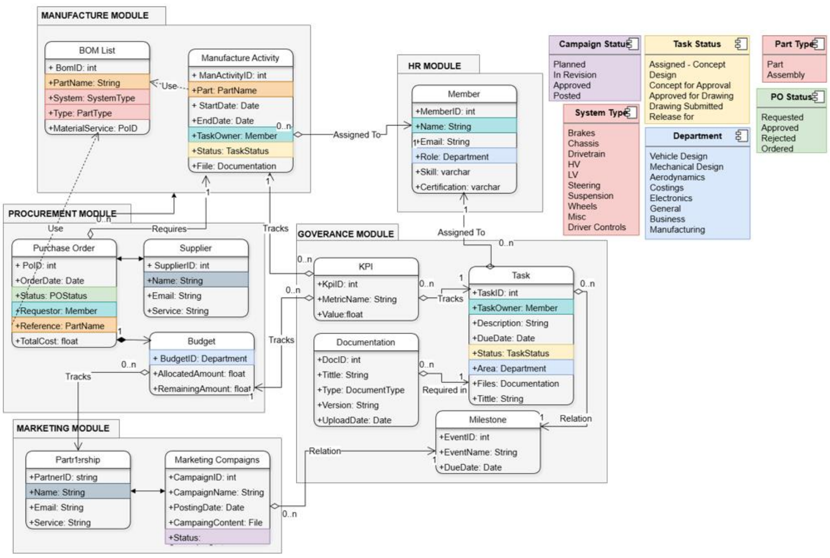



*Figure 23. ERP System Prototype Class Diagram*

The ERP system prototype aims to create a modular and integrated platform that maintains individual functions while supporting cross-functional coordination through shared data structures and workflow-based control. The

https://doi.org/10.18745/pb.27307

architecture is guided by key design principles including modular decomposition aligned with FS subsystems, centralised data management to ensure a single source, process standardization and scalability for future growth. These principles ensure the ERP system is both technically strong and customized to the operational needs of an FS team, supporting efficient collaboration, reliable data management and long-term continuity.

After concluding all the modules and process planning for the ERP system, the prototype was developed using Oracle Application Express (APEX) to demonstrate the feasibility of translating the TO-BE process model into a digital platform. The first development stage consisted of establishing the code variables that would support all the workflow for the design phase.  After creating the database structure and the initial application pages, the next stage focused on implementing the functional modules of the ERP system. These modules were translated into different individual pages that users can access through the system's main navigation page, as shown in Figure 3. The main navigation menu of this platform provides access to the main system module pages: Members, KPIs, Tasks, Manufacturing, Procurement, Social media plan, and the Team's calendar.

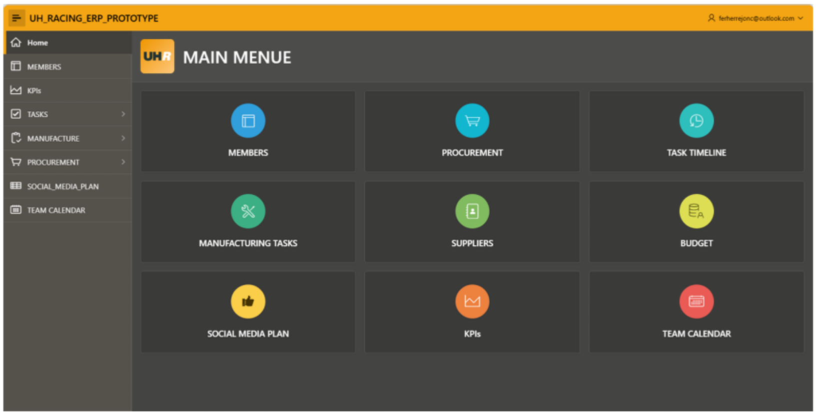



*Figure 24.  ERP Prototype Main Interface*

## Conclusion

This study examined how digital enterprise management principles improve a Formula Student team's operation. Findings from literature and interviews revealed that, despite operating in a highly technical environment, team management relies on fragmented tools and informal communication. Key challenges identified include limited task visibility, lack of traceability, and poor integration between departments, making it difficult to monitor progress and ensure completion of activities. These issues reflect those faced by small and medium-sized enterprises lacking structured digital systems.

The translation of the TO-BE model into an ERP prototype using Oracle APEX demonstrated how a low-code platform can address these challenges by centralising data and structuring workflows around a defined task lifecycle. This improves coordination, traceability, and visibility across departments while reducing reliance on manual processes. However, the effectiveness of such a system depends not only on the technology but also on organisational adoption, establishing clear data governance policies, providing ongoing user training, and continuously monitoring and refining data management practices. Overall, the study highlights the potential of low-code ERP solutions to enhance efficiency and coordination in small, project-based engineering teams.

## References

[1] Gattiker, T. F., & Goodhue, D. L. (2004). Understanding the local-level costs and benefits of ERP through organizational information processing theory. *Information and Management*, 41(4), 431-443. https://doi.org/10.1016/S0378-7206(03)00082-X
[2] Ellen F. Monk, & Bret J. Wagner. (2013). Business Functions and Business Processes. In Concepts in Enterprise Resource Planning (4th ed., pp. 21–25). *Course Technology*.

https://doi.org/10.18745/pb.27307

# Marginal Cost of Flexibility in CCS-Retrofitted UK Gas Plants: Comparing Plant-Integrated Thermal Energy Storage and Grid-Scale Battery Storage

**Nizam Uddin Payel[1*], Dr Mahmoud Eltaweel[1], and Prof. Hongwei Wu[1*]**

*1 School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK*

*Corresponding author: n.payel@herts.ac.uk

**CCS-retrofitted natural gas combined-cycle (NGCC) plants lose flexibility because solvent regeneration couples reboiler heat demand to the steam cycle. This study compares plant-integrated sensible-hot thermal energy storage (TES) with grid-scale lithium-ion battery energy storage (BESS) using a net marginal cost of flexibility metric (£/MWh-flex), 2024 GB hourly prices and selected UK cost/revenue assumptions. HOMER Pro models the baseline and BESS cases; Python represents TES as a reduced-order reboiler heat-shifting proxy. TES gives £43.56/MWh-flex, versus £57.03 and £60.56/MWh-flex for 100 and 200 MWh BESS after ancillary revenue. Monte Carlo results favour TES in about 70% of cases, under adopted assumptions.**

**Keywords:** NGCC+CCS; thermal energy storage; battery energy storage; marginal cost of flexibility; Great Britain

## Introduction

Great Britain's low-carbon transition requires dispatchable generation that can operate flexibly with variable renewables. NGCC plants with post-combustion CCS can provide firm capacity, but solvent regeneration requires reboiler heat, normally supplied by steam extraction. This heat-power coupling reduces peak net output, worsens part-load performance and weakens market responsiveness [1], [4].

Plant-integrated TES can shift regeneration heat while preserving capture operation [2]. Grid-scale BESS provides fast electricity-side flexibility but does not directly address the CCS thermal bottleneck [3]. This work compares the two pathways using net marginal cost of flexibility, £/MWh-flex, defined as annualised net storage cost divided by annual market-facing flexible electrical output. It provides a UK-focused service-denominator comparison and uncertainty-based ranking for CCS-retrofitted NGCC flexibility.

## Experimental/Simulation

A representative 630 MW-class NGCC+CCS plant with post-combustion MEA capture is evaluated using 2024 GB hourly wholesale electricity prices. The no-storage baseline as Case 0 and BESS configurations as case 2A and Case 2B (100 and 200 MWh with a fixed 50 MW converter) are simulated in HOMER Pro, with dispatch, NPC and operating cost taken as HOMER outputs [5]. The TES pathway (Case 1) is represented externally as a reduced-order reboiler heat-shift proxy derived from the Case 0 hourly export; this methodological asymmetry is retained as a limitation.

In the central TES proxy, charging is assigned to hours at or below the 30th percentile of the 2024 GB price series, while discharge is assigned to hours at or above the 70$^{th}$ percentile. The proxy shifts 20% of eligible CCS reboiler thermal load and applies a 90% thermal round-trip efficiency. Useful thermal discharge is converted to electrical-equivalent flexibility using an adopted steam-extraction penalty factor of 0.28 kWe/kWth. Costs are annualised over a 25-year project life using an 8% discount rate. The BESS pathway is costed using the HOMER delta-NPC method, with external O&M and a blended ancillary-services credit included in post-processing. Uncertainty testing varies TES capex, BESS NPC, discount rate, steam-extraction penalty and BESS ancillary credit.

## Results and discussion

The TES and BESS pathways provide flexibility through fundamentally distinct physical mechanisms and at different system boundaries. TES shifts regeneration heat away from high-price periods, increasing net electrical output without relaxing capture rate. BESS adds grid-interface flexibility through charge-discharge operation but

https://doi.org/10.18745/pb.27307

leaves the plant's internal heat-power coupling unchanged. HOMER generator output is unchanged across Case 0 and both BESS cases, while the TES proxy acts on the reboiler heat demand that defines the CCS duty cycle.

Fig. 1 and Table 1 summarise the central case. TES delivers 25,863 MWh-flex/yr at £43.56/MWh-flex. After the blended ancillary-services credit, the 100 MWh BESS delivers 172,926 MWh-flex/yr at £57.03/MWh-flex, while the 200 MWh BESS delivers 174,020 MWh-flex/yr at £60.56/MWh-flex. BESS therefore provides more annual flexible output, but at a higher net unit cost under the adopted assumptions.

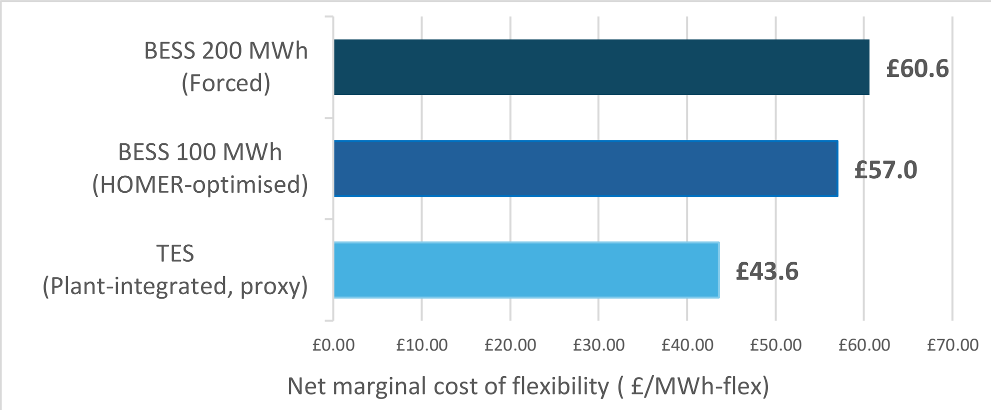



*Fig. 1. Net marginal cost of flexibility across the three compared configurations under central-case assumptions. BESS values include the blended ancillary-services revenue offset (£60,000/MW/yr × 50 MW = £3.00 million/yr) applied to the converter rating.*

*Table 1. Central-case comparison. NPC and generator output are HOMER-derived; annualised cost, ancillary credit and £/MWh-flex are post-processed values.*

| **Metric** | **NGCC+CCS** (Baseline) | **TES** (Plant-Integrated) | **BESS 100 MWh** (Grid-level) | **BESS 200 MWh** (Grid-level) |
|---|---|---|---|---|
| NPC (£M) | 709.7 | — | 844.3 | 851.6 |
| ΔNPC vs Case 0 (£M) | — | — | 134.6 | 141.9 |
| Annualised net cost (£M/yr) | — | 1.13 | 9.86 | 10.54 |
| BESS ancillary credit applied (£M/yr) | — | — | 3.00 | 3.00 |
| Annual MWh-flex (electrical-equivalent) | — | 25 863 | 172 926 | 174 020 |
| **Marginal cost of flexibility (£/MWh-flex)** | **—** | **43.56** | **57.03** | **60.56** |
| Generator output (GWh/yr, context) | 5 208.96 | 5 208.96 | 5 208.96 | 5 208.96 |

Sensitivity and Monte Carlo testing show that the ranking is conditional. TES capex, steam-extraction penalty, BESS NPC and BESS ancillary credit are the main drivers. TES is lower-cost than the optimised 100 MWh BESS in 69.61% of 10,000 samples and lower-cost than the forced 200 MWh BESS in 77.79%.

## Conclusion

Plant-integrated TES gives the lowest central net marginal cost of flexibility for the modelled CCS-retrofitted NGCC plant, but the result is not universal. Its stronger contribution is mechanistic: TES targets the reboiler/steam-side coupling at the plant, whereas BESS compensates externally at the grid interface. Flexibility

https://doi.org/10.18745/pb.27307

investments around CCS-equipped gas plants should therefore be assessed against the physical origin of the operational constraint, not only through generic storage-cost or arbitrage metrics.

## References

[1] J. Rúa, M. Bui, L. O. Nord, and N. Mac Dowell, “Does CCS reduce power generation flexibility? A dynamic study of combined cycles with post-combustion $CO_2$ capture,” International Journal of Greenhouse Gas Control, vol. 95, p. 102984, 2020.

[2] B. J. Limb et al., “Economic viability of using thermal energy storage for flexible carbon capture on natural gas power plants,” Journal of Energy Storage, vol. 55, p. 105836, 2022.

[3] A. Vannoni and A. Sorce, “Electrical storage or gas power plant flexibility by thermal storage? A market analysis,” Journal of Energy Storage, vol. 126, Article 116823, 2025, doi: 10.1016/j.est.2025.116823.

[4] T. J. P. Hersbach, M. D. Mastrandrea, and M. W. Wara, “Flexible operation and fugitive methane emissions limit the potential of power plant carbon capture and storage,” The Electricity Journal, vol. 38, p. 107494, 2025.

[5] S. Hughes et al., “Methodology for estimating performance and cost of natural gas combined cycle plants with carbon capture,” National Energy Technology Laboratory, DOE/NETL-2023/3874, 2023.

https://doi.org/10.18745/pb.27307

# Human Activity Recognition: Classifying Physical Actions Using Smartphone Inertial Sensors

**Bala Vijay Sai Ummaleti[1], and Okechi Onuoha[1*]**

[1]*Computer Science Department, University of Hertfordshire*

*Corresponding author: o.onuoha@herts.ac.uk

**This paper presents a systematic multi-model Human Activity Recognition (HAR) framework using the UCI HAR Dataset. Three classifiers - Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbours (KNN) are trained and optimised via GridSearchCV (3-fold cross-validation) for the study. The tuned SVM achieves the highest test accuracy (98.74%, F1=0.987) whilst Random Forest offers the best inference-latency trade-off (0.12 s). PCA is applied for 2D visualisation, confirming strong activity separability. Ten-fold cross-validation yields: SVM 95.49% (±2.46%), RF 94.17% (±2.56%), KNN 90.95% (±1.91%). The results provide a reproducible baseline for accuracy-efficiency-interpretability optimisation in mobile HAR.**

Keywords: Human Activity Recognition; SVM; Random Forest; SHAP; Smartphone Sensors.

## Introduction

Smartphone-based Human Activity Recognition (HAR) is a rapidly growing field enabling passive, continuous monitoring of physical states from inertial sensor data. Applications span fall detection, rehabilitation monitoring, healthcare analytics, and fitness tracking [1], [2]. Modern smartphones embed tri-axial accelerometers and gyroscopes sampled at 50Hz, generating rich sensor streams from which physical activities can be automatically inferred.

Despite strong benchmark results, mobile HAR faces three core challenges: (i) high-dimensional feature spaces (e.g., 561 features) create bottlenecks for distance and kernel-based classifiers; (ii) accuracy-efficiency trade-offs are rarely studied alongside interpretability; and (iii) black-box model behaviour limits clinical trust [3]. This paper addresses all three within a single reproducible framework.

**Research Hypotheses**: H1: PCA-based dimensionality reduction reduces latency whilst retaining accuracy ≥95%. H2: SHAP identifies Y-axis vertical acceleration as dominant for stair-related activity discrimination. H3: Sequential models (LSTM/GRU) improve F1 by ≥10% over static classifiers (planned phase).

## Methodology

The framework follows a six-phase design: (1) dataset acquisition and audit, (2) preprocessing and Exploratory Data Analysis (EDA), (3) baseline model training, (4) hyperparameter optimisation via GridSearchCV, (5) cross-validation and latency analysis, and (6) SHAP explainability.

**Dataset**. The UCI HAR Dataset [1] is used. It contains 10,299 samples from 30 participants (ages 19–48) performing six activities at 50Hz. The 561-feature vector is derived from fixed-width sliding windows applied to raw accelerometer and gyroscope signals. **Preprocessing**. The features.txt file contained 29 duplicate column names (multiple statistical transforms on the same sensor axis). A custom renaming function appended numerical suffixes, producing 561 unique column names. An 80/20 stratified train-test split yielded 8,239 training and 2,060 test samples. StandardScaler (fitted on train only) standardised all features to zero mean and unit variance, essential for SVM kernel distances and PCA. **PCA Usage**. PCA was used solely for 2D visualisation of activity separability. Dimensionality reduction for latency optimisation was not applied to the model training pipeline in this study, all classifiers were trained on the full 561-feature space. **Hyperparameter Optimisation**. GridSearchCV with 3-fold cross-validation searched the following spaces — SVM: C ∈ {0.1,1,10}, kernel ∈ {rbf, linear}, γ ∈ {scale, auto} (12 candidates, 36 fits); RF: n_estimators ∈ {100,200}, max_depth ∈ {10,20,None}, min_samples_split ∈ {2,5} (12 candidates, 36 fits); KNN: k ∈ {3,5,7,9}, weights ∈ {uniform, distance}, metric ∈ {euclidean, manhattan} (16 candidates, 48 fits).

https://doi.org/10.18745/pb.27307

## Results and Discussion

**Baseline Performance**. Table 1 shows baseline test-set results. RF achieves the highest accuracy (98.11%) with fastest inference (0.096 s). SVM is competitive at 97.86% but exhibits critical latency: 149.57s training and 7.40s inference on 561 features, a deployment barrier for real-time mobile systems. KNN reaches 96.26% but scales poorly with test-set size (0.31s inference).

**Tuned Model Performance**. The results of the tuning tests show their performance on the 20% test set. The SVM model which received tuning now shows the highest accuracy of 98.74% with F1 score 0.9874 while RF and KNN models show lower results with 98.20% and 97.57% accuracy and F1 scores of 0.9820 and 0.9757 respectively. SVM shows higher inference latency at 3.81 seconds compared to RF which takes 0.12 seconds which represents a major difference essential for mobile real-time applications. The accuracy and F1 scores are presented in Figure 1 as a visual comparison.

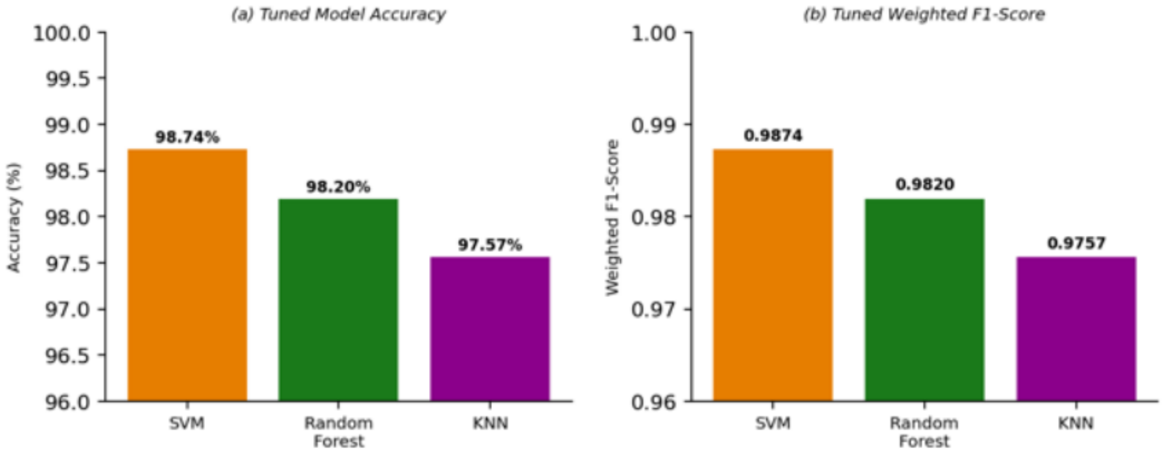



*Fig. 1: Tuned model performance: SVM leads in accuracy (98.74), RF is optimal for latency.*

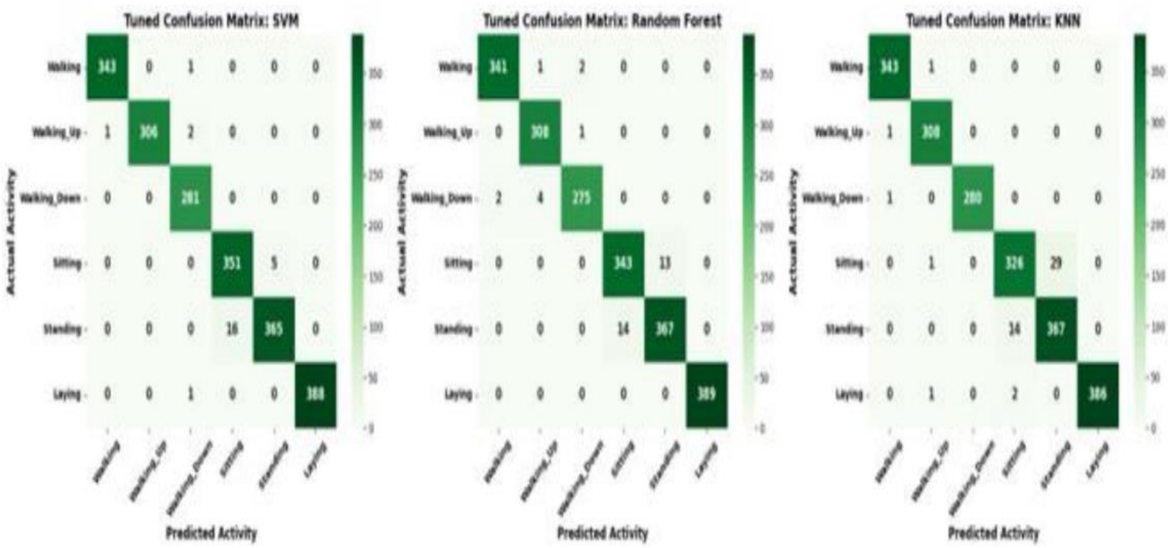



*Fig. 2: Tuned confusion matrices (SVM, RF, KNN). Primary confusion: Sitting↔Standing across all models.*

*Table. 1: Baseline Model Performance (Test Set, 561 Features)*

| Model | Acc. (%) | F1 | Train (s) | 0Infer (s) |
|---|---|---|---|---|
| Random Forest | 98.11 | 0.981 | 4.12 | 0.096 |
| SVM | 97.86 | 0.979 | 149.57 | 7.40 |
| KNN | 96.26 | 0.963 | - | 0.31 |

**Tuned Confusion Matrices**. Figure 2 shows the tuned confusion matrices for all three models. The three models show their confusion matrices which have been adjusted for their correct operation. The first key observation shows that Laying achieves perfect classification through all tested models because its horizontal gravity profile creates a distinct identification pattern. The Sitting/Standing confusion exists as the main mistake because SVM misclassifies 16 Standing→Sitting and 5 Sitting→Standing samples which RF misclassifies as 14 Standing→Sitting and 13 Sitting→Standing samples while KNN shows the highest confusion with Sitting/Standing through 29 Sitting→Standing misclassifications. The Walking Downstairs motion shows slight identification problems with Walking Upstairs because RF tests show four samples which SHAP analysis confirmed as matching jerk energy profiles.

## Conclusion

This paper presented a fully reproducible multi-model HAR framework on the UCI HAR Dataset. Core findings: (1) Tuned SVM achieves the highest test-set accuracy (98.74%, F1=0.9874) and 10-fold CV generalisation (95.49%), confirming RBF kernel's strong margin-based discriminability. (2) Tuned RF (98.20%, 0.12 s inference) provides the best accuracy-latency trade-off — the recommended choice for real-time mobile deployment. (3) KNN's tuned latency (24.71 s) disqualifies it for real-time use despite competitive accuracy. (4) PCA confirms strong activity separability in 2D projection; dimensionality reduction for latency improvement is an identified direction for future work. (5) SHAP confirms H2: tGravityAcc-iqr()-Y and Y-axis gravity features are globally dominant, with particular importance for stair-related activities. Sitting/Standing confusion remains the principal unresolved challenge across all models.

Future work: LSTM/GRU sequential model evaluation (H3); PCA-integrated SVM pipeline for latency reduction (H1 full validation); SHAP waterfall plots for individual prediction auditing; and evaluation on demographically diverse datasets beyond the 30-participant UCI benchmark

## References

[1] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in Proc. 21st European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, 2013.

[2] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, "Deep learning for sensor-based human activity recognition: overview, challenges and opportunities," ACM Computing Surveys, vol. 54, no. 4, Art. 77, 2021.

[3] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in Advances in

Neural Information Processing Systems (NeurIPS), vol. 30, pp. 4766–4775, 2017.

https://doi.org/10.18745/pb.27307

# Visual Classification of Mechanical Linkage Trajectories Using Vector Field Overlays

**Raimund Kirner *1, Olga Tveretina 1**
*1 University of Hertfordshire, Computer Science, Hatfield, United Kingdom*
*Corresponding author: r.kirner@herts.ac.uk

This extended abstract presents a study on visual classification of planar four bar linkage trajectories using vector fields plotted along the path. Instead of judging candidate mechanisms only from shape similarity to waypoints, the method overlays coupler orientation, velocity, total acceleration, and tangential acceleration along the end effector trajectory. These fields expose motion signatures that distinguish smooth, robust, and highly dynamic trajectories. We show how vector field patterns support rapid visual comparison between linkage designs and reveal hidden trade offs between kinematic accuracy, smoothness, and compactness.

Keywords: mechanical linkage; four bar linkage; trajectory analysis; vector fields; visual classification

## Introduction

Planar linkage synthesis is usually judged from kinematic path generation: a mechanism is considered successful when its end effector passes close to target waypoints or reproduces a desired curve [1], [2]. That view is incomplete. Two mechanisms can generate trajectories that look similarly acceptable in the workspace while having very different motion behaviour along the same path. Large changes in orientation, speed, or acceleration can imply higher dynamic loads, stronger vibrations, and lower mechanical robustness, even before a full dynamic simulation is performed [3]-[6].

This work takes a physics-informed view of trajectory interpretation. Instead of treating a trajectory as a static geometric curve, it is interpreted as a path carrying local directional and dynamic information. When this information is drawn directly on the path as vector fields, different classes of linkage behaviour become visually separable. The approach is lightweight, visually intuitive, and suitable for early stage design screening, where many candidate mechanisms must be compared quickly.

The objective of this extended abstract is to show how vector field overlays can classify linkage trajectories into practically meaningful categories such as low rotation versus high rotation, nearly uniform speed versus peaked speed, and smooth transport versus highly dynamic motion. The method is intended to complement optimisation and synthesis workflows by adding an interpretable visual layer between geometric path quality and mechanically relevant motion behaviour.

## Experimental/Simulation

The study considers a cyclic four bar crank rocker mechanism with an end effector attached to the coupler or output link. For each candidate design, the input crank angle is swept through a full rotation and the end effector positions are sampled to form a trajectory. Candidate mechanisms may be generated either from kinematic synthesis alone or from multi objective optimisation that also penalises compactness and higher order trajectory variation, following common trends in modern linkage design [1]-[6]. Here the emphasis is not on the solver itself, but on the trajectory representation used for visual classification.

https://doi.org/10.18745/pb.27307

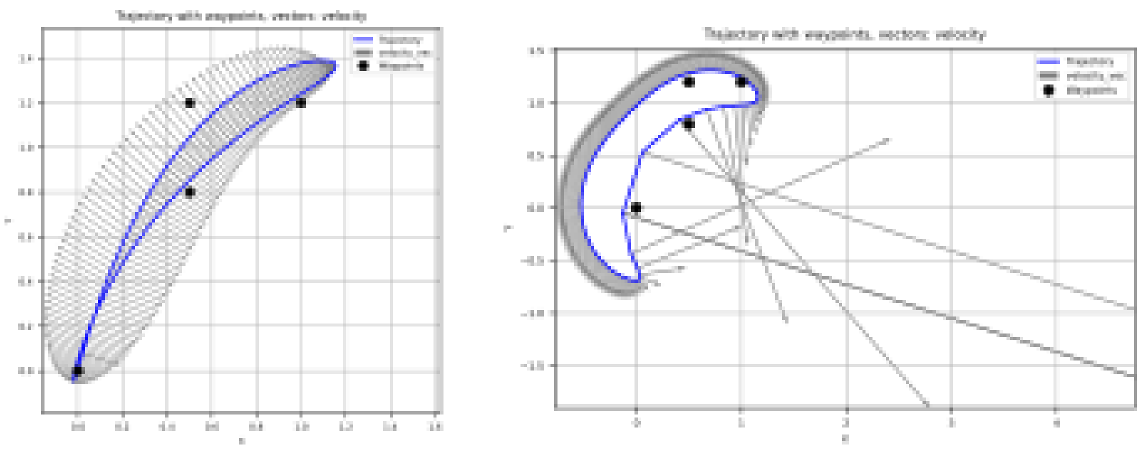
Four vector fields are used for classification. First, a coupler normal field represents local end effector orientation along the path. Second, a velocity field uses vector length to encode speed magnitude; for clarity, the vectors are drawn normal to the trajectory rather than in the true velocity direction. Third, a total acceleration field highlights regions where inertial effects are likely to be larger. Fourth, a tangential acceleration field separates local acceleration from deceleration and reveals how strongly the mechanism changes speed along the path. Together these overlays create a compact visual descriptor of trajectory behaviour.

Classification is qualitative but systematic. A trajectory is labelled orientation stable when the normal field rotates only modestly along the path. It is labelled speed regular when the velocity vectors have comparable lengths over most of the curve. It is labelled dynamically mild when total and tangential acceleration remain small except near limited local features. Conversely, strong angular sweep, large isolated vector peaks, or abrupt sign changes in tangential acceleration indicate more aggressive motion classes. These classes are intended for engineering interpretation, not for statistical learning.

**Results and discussion**

Figure 1 shows a representative comparison between two linkage trajectories that satisfy the same waypoint set, yet display markedly different motion behaviour of the end effector. Plot 1 has a narrow path with pointed ends. Plot 2 appears visually smoother because its ends are more rounded. If the decision were based only on geometric shape, Plot 2 could easily be judged as the gentler mechanism.

The vector field classification shows the opposite for several criteria. In the coupler normal field, Plot 1 exhibits relatively constant velocity, while Plot 2 sweeps through a much larger velocity range. This immediately separates a velocity stable class from a velocity varying class. For tasks that require smoothness in motion, that distinction matters even when the traced path itself looks acceptable.



a) Plot 1    b) Plot 2

Fig. 1. Vector field overlay comparing two linkage trajectories. The blue curve shows the end effector path and the grey vectors show the velocity along the trajectory. Despite similar waypoint satisfaction, the two trajectories belong to different motion classes because their local vector patterns differ strongly.

Other interesting overlays are acceleration and tool orientation overlays. Plot 1 maintains a comparatively limited range in tool alignment and acceleration, whereas Plot 2 provides a higher variation in tool alignment and acceleration. The visual lesson is important: rounded geometry does not guarantee smooth motion. Because linkage kinematics map crank rotation to workspace motion

https://doi.org/10.18745/pb.27307

non uniformly, visually softer corners can still correspond to highly non uniform traversal speed. The tangential acceleration field further identifies where the mechanism is alternately speeding up and slowing down, providing a direct visual indicator of motion aggressiveness.

This classification perspective also clarifies the trade off between geometric tracking and mechanical smoothness. A trajectory selected only for low waypoint error can still be longer, less compact, and dynamically less uniform. In contrast, solutions that also account for velocity or acceleration related criteria tend to move toward a more compact and mechanically robust class, even if some geometric matching is sacrificed [3]-[6]. The main advantage of the method is speed of insight. Compared with scalar objective values alone, the overlays expose where along the curve problematic behaviour occurs. Compared with full dynamic simulation, they are far cheaper to compute, although they remain proxy descriptors rather than replacements for complete force or joint analysis [4]-[6].

## Conclusion

Vector field overlays turn a linkage trajectory into an interpretable motion signature. The method supports visual classification of candidate four bar mechanisms according to orientation change, speed regularity, and dynamic severity. The key result is that trajectories that appear similar in shape can belong to very different motion classes once local vector information is shown. This makes the approach valuable for early design screening, optimisation result interpretation, and communication of trade offs between tracking accuracy and mechanical smoothness. A natural next step is to convert these visual classes into quantitative descriptors for automated ranking or learning based comparison of linkage designs.

## References

[1] S. Sleesongsom and S. Bureerat, “Four bar linkage path generation through self adaptive population size teaching learning based optimization,” Knowledge Based Systems, vol. 135, pp. 180-191, 2017.
[2] Y. H. Kang, J. W. Lin, and W. C. You, “Comparative study on the synthesis of path generating four bar linkages using metaheuristic optimization algorithms,” Applied Sciences, vol. 12, no. 15, 2022.
[3] R. M. Rayner, M. N. Sahinkaya, and B. J. Hicks, “Improving the design of high speed mechanisms through multi level kinematic synthesis, dynamic optimization and velocity profiling,” Mechanism and Machine Theory, vol. 118, pp. 100-114, 2017.
[4] C. Llopis Albert, F. Rubio, C. Devece, and S. Zeng, “Multiobjective optimization framework for designing a steering system considering structural features and full vehicle dynamics,” Scientific Reports, vol. 13, 2023.
[5] J. Huang, C. Li, H. Long, L. Yin, B. Wen, and J. Wang, “Multi objective optimization design of a heavy duty folding mechanism and self discharging equipment development,” Mechanics and Industry, vol. 24, 2023.
[6] P. Flores, J. Ambrosio, J. P. Claro, and H. M. Lankarani, Kinematics and Dynamics of Multibody Systems with Imperfect Joints: Models and Case Studies. Berlin, Germany: Springer, 2008.

https://doi.org/10.18745/pb.27307

# Trade-Offs Between Realism and Interactivity in Extended Reality Telepresence for Remote Skill Acquisition

**Alex Stanley[1*], Yasir Iqbal[1], Giuseppe Morana[1], Alessandro Pugliesi[2], Alessandro Zocco[2], Salvatore Livatino[1]**

*[1]Dept. Engineering, University of Hertfordshire, UK*

*[2]ELT Group, Italy*

*Corresponding author: as22agz@herts.ac.uk

**Extended reality (XR) telepresence systems for remote skill acquisition must balance seeing a task environment as realistically as possible with allowing meaningful user action inside it. This paper presents a focused state-of-the-art investigation based on a structured review of research published between 2015 and 2025. The literature shows that high-fidelity telepresence strengthens shared presence and manipulation, whereas 360° media preserves scene authenticity but remains largely observational. Between these poles, collaborative and hybrid XR systems emerge as the most promising direction for guidance, rehearsal, and collaborative understanding.**

Keywords: extended reality; telepresence; realism; interactivity; remote skill acquisition

## Introduction

Extended reality-mediated training is increasingly investigated for domains in which direct, co-located instruction is costly, hazardous, or logistically constrained. In industrial maintenance, healthcare, construction, and defence-related operations, trainees often need access to procedures, tools, and expert guidance without being physically present. A core difficulty in these settings is establishing common ground: remote participants must not only observe the same scene, but also understand actions, intentions, and object relations in a shared spatial frame [1]. Comparative XR collaboration studies show that richer representations can improve task performance and usability, yet the value of these systems depends less on technical novelty alone than on how convincingly they support both situated perception and purposeful action [2].

This paper presents a focused state-of-the-art investigation of trade-offs between realism and interactivity in XR telepresence for remote skill acquisition. The review asks how far realism must be pursued before it contributes to training value, and what forms of interactivity matter most for guidance, rehearsal, and collaborative understanding. The aim is not to rank systems from best to worst, but to clarify how current approaches distribute visual authenticity, shared presence, and user agency across remote training scenarios.

## Experimental/Simulation

The investigation is based on the structured literature review described in the source report. Searches were carried out across Scopus, IEEE Xplore, and the ACM Digital Library using Boolean combinations of terms related to training, telepresence, VR, AR, XR, HMDs, and skill acquisition. Publications from 2015 to 2025 were screened for relevance to remote collaboration, immersive media, and training-oriented XR applications.

For synthesis, studies were interpreted through a case-based framework describing alternative configurations of trainer, trainee, and equipment across physical and virtual settings. The main analytical lens for this shorter paper is the relationship between realism and interactivity. Realism refers here to the authenticity of spatial perception, visual context, and co-presence; interactivity refers to the extent to which users can manipulate, annotate, guide, or coordinate actions within the shared environment.

## Results and discussion

The reviewed literature shows that collaborative AR and shared-workspace XR establish an important baseline for the realism-interactivity trade-off. Rather than reconstructing entire scenes at maximum fidelity, these systems use spatial cues, viewpoint sharing, and annotations to help remote users build common ground [1]. User studies comparing 2D video, 360° views, and richer 3D interfaces indicate that improved spatial

https://doi.org/10.18745/pb.27307

representation can reduce effort and improve procedural performance [2]. In this strand of work, realism is valuable because it clarifies action: the most useful systems are not simply more lifelike, but better at connecting what users see to what they can do together.

High-fidelity telepresence systems sit at the realism-interactivity extreme. Holoportation, omnistereoscopic telepresence, holographic classroom systems, and more recent room-scale approaches all attempt to strengthen co-presence through depth-rich capture, free-viewpoint perception, or embodied remote-user representation [3]-[6]. Their central contribution is to bring visual realism closer to shared agency: users do not merely observe remote actors, but perceive them as spatially situated collaborators. Yet the literature also suggests that realism is not automatically equivalent to effective training. Some systems excel at social presence, while others better support procedural manipulation, indicating that different forms of realism serve different pedagogical purposes.

*Table 1. Comparative summary of how representative XR approaches position realism and interactivity in remote skill acquisition*

| Approach | Representative studies | Realism profile | Interactivity profile | Main implication |
|---|---|---|---|---|
| Collaborative AR and shared-workspace XR | [1], [2] | Shared spatial cues and task context; richer than flat video | Moderate to high via annotation and guidance | Best for guided remote assistance |
| High-fidelity volumetric telepresence | [3], [4], [6] | Strong depth, co-presence, and free-viewpoint perception | High shared agency and richer manipulation | Best for co-action and spatial demonstration |
| Holographic classroom/educational telepresence | [5] | Strong social presence but lighter physical depth | Moderate participation; weak tool interaction | Good for communication-led learning |
| 360°/180° immersive video | [7], [8], [9], [10] | High visual authenticity from captured scenes | Low; mainly gaze control with limited manipulation | Good for exposure and reflection |

At the opposite end, 360° and 180° immersive video prioritise scene authenticity over direct manipulation. These media are valuable where exposure to a real environment matters more than co-action, such as reflective practice, contextual familiarisation, or access to hazardous and difficult-to-reach sites [7]-[9]. However, their interactivity remains limited: users can look around, but they generally cannot change viewpoint beyond the capture position, manipulate objects, or engage with robust occlusion and depth relations [10]. The contrast with high-fidelity systems shows that visual realism alone is insufficient for telepresence; training value depends on whether realism supports actionable understanding rather than passive observation.

Taken together, the state of the art is best understood as a design continuum. Maximum realism is most useful when it directly supports shared action, guidance, or object understanding. Where this is not necessary, lighter approaches may offer a better balance between immersion and usability. The most promising research direction is therefore not absolute realism, but selective realism paired with selective interaction: hybrid XR environments that preserve authentic visual context while adding task cues, annotations, guided attention, or limited object-level interaction.

## Conclusion

This focused review shows that the central scientific issue in XR telepresence for remote skill acquisition is not realism alone, but the relationship between realism and interactivity. High-fidelity systems demonstrate the

https://doi.org/10.18745/pb.27307

benefits of shared spatial presence, while 360° media demonstrates the value of authentic captured context. Between them lies the most productive direction for future work: hybrid systems that transform visual realism into meaningful, task-relevant action.

## References

[1] B. Marques, S. Silva, J. Alves, T. Araujo, P. Dias, and B. S. Santos, "A Conceptual Model and Taxonomy for Collaborative Augmented Reality," IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 12, pp. 5113-5133, 2022.

[2] E. Chang, Y. Lee, and B. Yoo, "A User Study on the Comparison of View Interfaces for VR-AR Communication in XR Remote Collaboration," International Journal of Human-Computer Interaction, pp. 1-16, 2023.

[3] S. Orts-Escolano et al., "Holoportation: Virtual 3D Teleportation in Real-time," in Proceedings of the 29th Annual Symposium on User Interface Software and Technology, pp. 741-754, 2016.

[4] F. Kaynar, M. Hofbauer, A. MacWilliams, J. Newman, A. Hutter, and E. Steinbach, "AR in VR: Omnistereoscopic Telepresence with Holograms for Remote Maintenance and Collaboration," IADIS International Journal on WWW/Internet, vol. 20, no. 2, pp. 99-116, 2022.

[5] B. Dominguez-Dager, F. Gomez-Donoso, R. Roig-Vila, F. Escalona, and M. Cazorla, "Holograms for Seamless Integration of Remote Students in the Classroom," Virtual Reality, vol. 28, article 24, 2024.

[6] X. Huang, D. Frehlich, Z. Xia, P. Gholami, and R. Xiao, "GaussianNexus: Room-Scale Real-Time AR/VR Telepresence with Gaussian Splatting," in Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology, pp. 1-18, 2025.

[7] A. Shojaei, S. Rokooei, A. Mahdavian, L. Carson, and G. Ford, "Using Immersive Video Technology for Construction Management Content Delivery: A Pilot Study," Journal of Information Technology in Construction, vol. 26, pp. 886-901, 2021.

[8] N. Walshe and P. Driver, "Developing Reflective Trainee Teacher Practice with 360-Degree Video," Teaching and Teacher Education, vol. 78, pp. 97-105, 2019.

[9] J. Pirker and A. Dengel, "The Potential of 360° Virtual Reality Videos and Real VR for Education-A Literature Review," IEEE Computer Graphics and Applications, vol. 41, no. 4, pp. 76-89, 2021.

[10] J. Reyna, "The Potential of 360-Degree Videos for Teaching, Learning and Research," in INTED2018 Proceedings, pp. 1448-1454, 2018.

https://doi.org/10.18745/pb.27307

# Task-Dependent Benefits of Immersive VR in Pilot Training: Search, Procedure, and Teleoperation

**Giuseppe Morana[1*], Phil Gainley[2], Yasir Iqbal[1], Alessandro Zocco[3], Salvatore Livatino[1]**

[1]*Dept. Engineering, University of Hertfordshire, UK*

[2]*Dept. Computer Science, University of Hertfordshire, UK*

[3]*ELT Group, Italy*

Corresponding author: g.morana3@herts.ac.uk

**This paper examines whether the value of immersive virtual reality in pilot training depends on task type rather than display technology alone. A comparative study between desktop monitor and head-mounted display conditions is used to analyse three training activities: instrument Search & Find, procedural Check List execution, and constrained drone tele-guiding. The results indicate a differentiated pattern. Immersive viewing benefits spatially demanding search and remote navigation tasks, whereas desktop display remains advantageous for simple procedural recall. The paper argues for task-led deployment of XR in aviation training instead of universal replacement of conventional simulators.**

**Keywords:** virtual reality; pilot training; immersion; visual search; teleoperation

## Introduction

Pilot training requires rapid interpretation of instruments, accurate search, procedural execution, and timely decision-making. For this reason, simulation remains central to aviation education. Head-mounted displays are attractive because they can increase the sense of presence, widen the usable field of regard, and support more natural scene engagement than conventional desktop systems [1], [2]. Yet immersive learning theory also suggests that technological richness does not guarantee better outcomes for every task; benefit depends on how immersion interacts with cognition, attention, and task structure [3].

A more useful academic question is therefore not whether virtual reality is globally better than desktop simulation, but where immersive presentation materially improves training. This issue is especially important in pilot education, where search efficiency, procedural memory, and spatial control are all relevant but do not rely on the same cognitive processes [4], [5]. The present paper addresses that gap by comparing three task classes within the same training context: spatial search, checklist procedure, and constrained teleoperation.

## Experimental

A within-subject design was used in which twelve participants completed the same aviation tasks on a desktop monitor and on a head-mounted display. The virtual cockpit reproduced a Cessna 152 dashboard, while the navigation scenario required remote control of a drone moving through a narrow tunnel. The sample included users with none to good prior experience of virtual reality and computer games, and presentation order was counterbalanced to reduce learning and fatigue effects.

Three tasks were analysed. In Search & Find, participants identified cockpit instruments and indicated their locations. In Check List, participants reproduced an ordered sequence of cockpit actions involving switches, yoke movement, and throttle operation. In Tele-Guide, participants piloted a drone through a constrained environment while avoiding collisions. Performance was assessed through completion time and precision; subjective factors included presence, comfort, and ease of use.

https://doi.org/10.18745/pb.27307

## Results and discussion

The results show that immersive benefit was strongly task-dependent. In Search & Find, the head-mounted display reduced mean completion time from 29.069 s on the desktop monitor to 19.322 s, a reduction of approximately 34%. Precision also improved from 2.083 to 2.583, with both time and precision differences reaching significance ($p = 0.047$ and $p = 0.046$, respectively). This pattern indicates that immersion supports tasks that depend on distributed visual attention, spatial localisation, and active cockpit scanning.

A different pattern emerged in the Check List task. Mean completion time was almost unchanged between desktop and immersive viewing (12.864 s versus 13.080 s; $p = 0.447$), but procedural precision was significantly higher on the desktop monitor (2.667 versus 2.083; $p = 0.022$). For simple ordered action sequences, the stable and allocentric presentation of a desktop display appears to remain more effective than immersive observation. The educational implication is that procedural rehearsal does not necessarily benefit from stronger presence or wider viewing.

The Tele-Guide task produced a third pattern. Mean navigation time differed little between conditions (27.578 s on desktop and 28.467 s in the head-mounted display; $p = 0.383$), but mean precision favoured immersive viewing, with a 55% improvement in the disengagement-after-collision metric (4.919 versus 2.265; $p = 0.282$). Although this latter difference was not significant, it points to a practical control advantage in dynamic spatial guidance. The subjective findings are consistent with these results: immersive viewing scored higher on spatial presence and involvement, suggesting that head-mounted displays are most valuable when task success depends on scene comprehension, search, and spatial control rather than on simple procedural repetition [1], [3], [5].

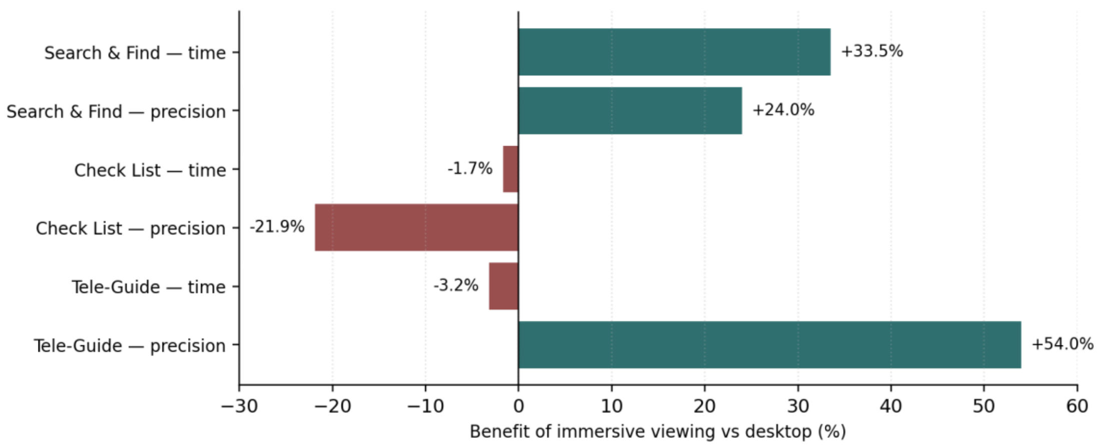



*Fig. 1. Direction-normalised summary of immersive-viewing benefit across the three pilot-training tasks.*

## Conclusion

Immersive virtual reality should be treated as a task-specific training resource rather than as a universal replacement for desktop simulation. Its clearest benefits appeared in spatial search and dynamic guidance, whereas the desktop monitor remained superior for straightforward procedural recall. A balanced aviation curriculum should therefore allocate desktop displays to basic checklist rehearsal and reserve immersive systems for tasks in which scene engagement, spatial encoding, and control precision are central to learning.

https://doi.org/10.18745/pb.27307

## References

[1] M. Slater and S. Wilbur, “A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments,” Presence: Teleoperators and Virtual Environments, vol. 6, no. 6, pp. 603-616, 1997.

[2] N. Nilsson, R. Nordahl, and S. Serafin, “Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence,” Human Technology, vol. 12, pp. 108-134, 2016.

[3] G. Makransky and G. B. Petersen, “The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality,” Educational Psychology Review, vol. 33, no. 3, pp. 937-958, 2021.

[4] W. T. Walters and J. Walton, “Efficacy of virtual reality training for pilots: A review of links between user presence, search task performance, and collaboration within virtual reality,” Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 65, no. 1, pp. 919-922, 2021.

[5] B. Olk, A. Dinu, D. J. Zielinski, and R. Kopper, “Measuring visual search and distraction in immersive virtual reality,” Royal Society Open Science, vol. 5, no. 5, p. 172331, 2018.

https://doi.org/10.18745/pb.27307

# Beyond Trust: Governance Requirements for Safe Real-World Deployment of Assistive AI-Driven Robotic Dermatology

**Fereshteh Amooei[1*], Ali Fallahi[2], and Abolfazl Zaraki[2]**

*1School of Health, Medicine and Life Sciences, University of Hertfordshire, Hatfield, United Kingdom*

*2Robotics Research Group, University of Hertfordshire, Hatfield, United Kingdom*

*Corresponding author: f.amooei2@herts.ac.uk

**Assistive AI-driven robots are increasingly deployed in direct patient contact settings, yet governance requirements for safe real-world use remain underdeveloped, particularly in intimate clinical applications such as dermatology. This paper synthesises a trust-centred interaction framework with a governance-centred ethical framework to propose implementable controls for ongoing consent, dignity protection, data governance, fairness assurance, accountability, and human oversight in AI-driven robotic dermatology. Key deployment concerns and practical implications are provided to support responsible clinical deployment.**

Keywords: human-robot interaction; assistive robotics; robotic dermatology; ethical governance; medical AI

## Introduction

Assistive robots that interact directly with patients are being explored across a range of clinical settings, from rehabilitation and mobility support to cognitive assistance and personal care [1,2]. A common challenge is that real-world deployment quickly surfaces requirements that laboratory evaluations miss: patients behave unpredictably, clinical workflows impose constraints, and the ethical dimensions of close human–robot contact become salient in ways purely technical assessments cannot anticipate [3,4]. Dermatology presents one of the most demanding instances of this challenge. An assistive dermatology robot must perform visual assessment, navigate physical proximity, and in some configurations initiate or guide touch, all within a clinical encounter that may be experienced as intimate, exposing, or vulnerable [5]. Figure 1 (a) presents a conceptual illustration of an assistive AI-driven dermatology robot operating in a clinical examination room.

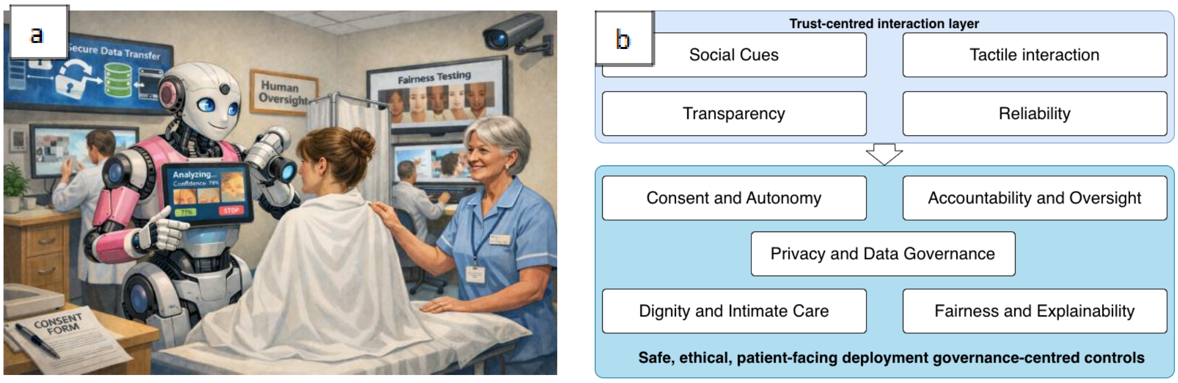



*Figure 1. (a) Conceptual illustration of assistive AI-driven dermatology robot in clinical examination room. Image generated using ChatGPT (OpenAI). (b) Integrated trust and governance architecture for robotic dermatology. Authors' own conceptual illustration, adapted from [6].*

Our earlier work proposed a trust-centred interaction framework for dermatology robots [6], identifying four primary trust dimensions: social cues (greetings, eye contact, expressive gesture); tactile interaction (warm, gentle, controllable touch); transparency (real-time explanations and confidence indication); and reliability (diagnostic accuracy, human check pathways). This work complements and extends that foundation. Real-world deployment of patient-facing assistive robots raises a further layer of requirements around consent management, data governance, dignity protection, fairness, and accountable oversight, requirements that must

https://doi.org/10.18745/pb.27307

be addressed before a system enters clinical use [7,8]. This paper synthesises both frameworks into an integrated contribution: a governance-centred ethical framework for intimate AI-driven robotic dermatology, grounded in WHO guidance on AI for health [7,8], professional standards for intimate examinations [9] and dermatology image handling [10], and the emerging literature on fairness in dermatology AI [11,12]. Figure 1 (b) summarises how the identified trust-centred interaction layer can be combined with governance-centred controls to support safe clinical deployment.

## Framework Development

The framework was developed by synthesising AI ethics guidance, clinical standards, and dermatology-specific literature, and extends the trust dimensions in [6] into five governance dimensions that translate ethical principles into practical design controls. **1. Consent and Autonomy:** Autonomy: Treat consent as ongoing, with stop controls and renewed permission for intimate steps [7,9,10]; **2. Dignity and Intimate Care:** Support modesty, emotional comfort, and patient control throughout the encounter [9,10]; **3. Privacy and Data Governance:** Secure image capture, storage, access, deletion, and secondary-use governance [7,8,10]; **4. Fairness and Explainability:** Validate across groups, display uncertainty, and trigger human review when needed [7,8,11,12]; **5. Accountability and Oversight:** Allocate responsibility clearly and support audit, escalation, and redress [7,8].

## Results and Discussion

The integrated framework is summarised in Table 1, which maps each governance dimension to its key deployment concern and practical design implication. Together with the trust dimensions from [6], this creates a complete design framework for patient-facing dermatology robots that are socially effective, technically reliable, ethically sound, and manageable throughout their lifecycle.

*Table 1. Integrated trust and governance framework for AI-driven assistive robotic dermatology. Rows marked * are reproduced from [6] with attribution.*

| Dimension | Key deployment concern | Practical implication |
|---|---|---|
| **Social cues [6] *** | Robotic silence or lack of eye contact erodes comfort | Greetings, friendly voice, expressive gestures, responsive speech |
| **Tactile interaction [6] *** | Cold or excessive pressure causes distrust | Warm, gentle, adjustable touch; immediate correction on discomfort |
| **Transparency [6] *** | Patients do not understand what the robot is doing | Real-time scan feedback, confidence indication, lay-language explanations |
| **Reliability [6] *** | Diagnostic errors or lack of human check pathways undermine trust | Validated diagnostic accuracy, clear human check pathways, and escalation protocols |
| **Consent & autonomy [7,9,10]** | Patient unable to pause or stop the examination | Ongoing consent process, visible stop controls, renewed permission at intimate stages |
| **Dignity & intimate care [9,10]** | Exposure or proximity causes distress or loss of control | Chaperone option, privacy-preserving positioning, patient-paced workflow |
| **Privacy & data governance [7,10]** | Sensitive images captured, stored, or reused insecurely | Encrypted transfer, access control, defined deletion policies, secondary-use governance |
| **Fairness & explainability [7,11,12]** | Unequal performance across skin tones or rare conditions | Subgroup validation, uncertainty display, automatic escalation to human review |
| **Accountability & oversight [7,8]** | Responsibility unclear when harm occurs | Audit logs, incident escalation, redress mechanisms, clear role allocation |

The framework addresses three gap areas that existing guidance does not cover for embodied AI systems. **First**, dermatology AI may perform unevenly across darker skin tones and rare conditions, creating unequal and potentially harmful patient outcomes [11,12]. **Second**, intimate robotic care raises dignity and distress concerns beyond standard privacy issues, making trust-repair measures part of governance as well as interaction design

[6]. **Third**, shared responsibility across multiple actors makes clear role allocation, audit trails, and redress essential for accountable deployment [8].

## Conclusion

This paper presented an integrated trust and governance framework for assistive AI-driven robotic dermatology, synthesising a social and tactile trust framework [6] with a governance-centred ethical framework grounded in WHO guidance [7,8] and professional standards for intimate examinations and dermatology image handling [9,10]. The framework maps nine dimensions onto key deployment concerns and practical implications. As a limitation, the framework is conceptual in nature and has not yet been evaluated in real-world clinical deployment; future work should test the governance controls in practice over time.

## Acknowledgments

The authors acknowledge the use of ChatGPT (OpenAI) for AI-assisted writing support in drafting and refining portions of this manuscript, and for generating Figure 1(a).

## References

[1] S. Ozturkcan and E. Merdin-Uygur, “Humanoid service robots: The future of healthcare?,” Journal of Information Technology Teaching Cases, vol. 12, no. 2, pp. 163–169, 2022.

[2] J. Holland et al., “Service robots in the healthcare sector,” Robotics, vol. 10, no. 1, p. 47, 2021.

[3] D. Townsend and A. MajidiRad, “Trust in human–robot interaction within healthcare services: A review study,” in International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 86281, American Society of Mechanical Engineers, 2022.

[4] J. Torrent-Sellens, A. I. Jiménez-Zarco, and F. Saigí-Rubió, “Do people trust in robot-assisted surgery? Evidence from Europe,” International Journal of Environmental Research and Public Health, vol. 18, no. 23, p. 12519, 2021.

[5] M. Muddassir et al., “Robotics meets cosmetic dermatology: Development of a novel vision-guided system for skin photo-rejuvenation,” IEEE/ASME Transactions on Mechatronics, vol. 27, no. 2, pp. 666–677, 2021.

[6] F. Amooei, A. Fallahi, and A. Zaraki, "Designing Trustworthy Dermatology Robots: A Framework for Social, Tactile, and Transparent Interaction," in TRUST 2025: SCRITA and RTSS @ RO-MAN 2025, 2025.

[7] World Health Organization, Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva: WHO, 2021. https://www.who.int/publications/i/item/9789240029200

[8] World Health Organization, Ethics and Governance of Artificial Intelligence for Health: Large Multi-Modal Models. Geneva: WHO, 2024. https://www.who.int/publications/i/item/9789240084759

[9] General Medical Council, "Intimate Examinations and Chaperones," 2024. https://www.gmc-uk.org/

[10] British Association of Dermatologists, "UK Guidance on the Use of Mobile Photographic Devices in Dermatology," 2017. https://www.bad.org.uk

[11] R. Daneshjou et al., “Disparities in dermatology AI performance on a diverse, curated clinical image set,” Science Advances, vol. 8, no. 31, 2022, Art. no. eabq6147.

[12] R. Daneshjou et al., “Lack of transparency and potential bias in artificial intelligence datasets and algorithms: a scoping review,” JAMA Dermatology, vol. 157, no. 11, pp. 1362–1369, 2021.

https://doi.org/10.18745/pb.27307

# Presence as a Predictor of Search Efficiency and Procedural Recall in Immersive Cockpit Training

**Phil Gainley[1*], Alessandro Zocco[2], Giuseppe Morana[3], Yasir Iqbal[3], Salvatore Livatino[3]**

[1]*Dept. Computer Science, University of Hertfordshire, UK*
[2]*ELT Group, Italy*
[3]*Dept. Engineering, University of Hertfordshire, UK*
*Corresponding author: p.gainley@herts.ac.uk

**This work investigates whether presence in immersive cockpit training is merely a positive user-experience outcome or a factor linked to performance. Using desktop monitor and head-mounted display conditions, the study connects presence ratings with instrument search and checklist-based procedural recall in a virtual Cessna cockpit. Immersive viewing increased spatial presence and substantially improved visual search, but did not improve performance on the ordered checklist task. The findings suggest that presence is most beneficial when training depends on spatial attention and target localization, rather than on short procedural sequences that can be rehearsed efficiently on a stable desktop display.**

**Keywords:** virtual reality; presence; pilot training; memorization; visual search

## Introduction

Presence is frequently described as the subjective sensation of "being there" in a mediated environment [1], while broader reviews of immersion emphasise that presence emerges from the interaction between system properties, perception, and user action [2]. In XR-based education, this matters because immersive learning outcomes depend not only on display fidelity but on whether the medium supports attention, encoding, and task-relevant processing [3]. In pilot training, this issue is particularly important, since cockpit activity combines visual search, rapid interpretation, and remembered procedures [4].

Previous research suggests that immersive systems can improve visual search in complex scenes [5], and that immersive or embodied learning can support recall under appropriate conditions [6], [7]. However, these benefits are unlikely to be uniform across all cockpit tasks. The present paper therefore tests a narrower claim: if immersive presence matters, it should strengthen performance most clearly in search-intensive cockpit actions, whereas simple ordered procedural recall may remain equally good or better on a stable desktop display.

## Experimental

Twelve participants completed the same virtual cockpit tasks in two display conditions: a 55-inch desktop monitor and a head-mounted display. The cockpit reproduced a Cessna 152 dashboard. The design was within-subject and counterbalanced, with practice before testing and questionnaire completion after task execution. This structure allowed direct comparison of presence ratings and behavioral performance across immersive and non-immersive viewing.

Two tasks were used for the present analysis. In Search & Find, participants searched for specified cockpit instruments and indicated their locations. In Check List, participants reproduced a learned sequence of cockpit operations, making the task a practical measure of short procedural recall. Objective measures included completion time and precision, while subjective measures included spatial presence, involvement, comfort, and ease of use.

Figure 1 shows the virtual Cessna 152 dashboard used for the cockpit search and checklist tasks.

https://doi.org/10.18745/pb.27307

## Results and discussion

The immersive condition produced higher presence ratings overall, with the clearest gain appearing in spatial presence. The strongest reported difference occurred on the spatial-presence item assessing perceived surrounding of the virtual space (Q5, $p = 0.006$). This confirms that the head-mounted display increased the feeling of being located within the cockpit environment rather than merely observing it on a screen.

That increase in presence was accompanied by a large improvement in visual search. In Search & Find, mean completion time fell from 29.069 s on the desktop monitor to 19.322 s in the head-mounted display, while mean precision rose from 2.083 to 2.583. Both differences were significant ($p = 0.047$ for time and $p = 0.046$ for precision). These results are consistent with the interpretation that presence contributes to better target localisation and more efficient cockpit scanning when the task requires spatially distributed attention [1], [3], [5].

The pattern did not generalise to procedural recall. In the Check List task, mean completion time remained essentially unchanged between desktop and immersive viewing (12.864 s versus 13.080 s; $p = 0.447$), but precision was significantly better on the desktop monitor (2.667 versus 2.083; $p = 0.022$). This suggests that higher presence is not automatically advantageous for simple ordered procedural sequences. Instead, the immersive benefit appears selective: presence supports search efficiency and scene engagement, whereas short recall-based action chains may still be better served by the visual stability and lower attentional load of a desktop display [3], [6], [7].

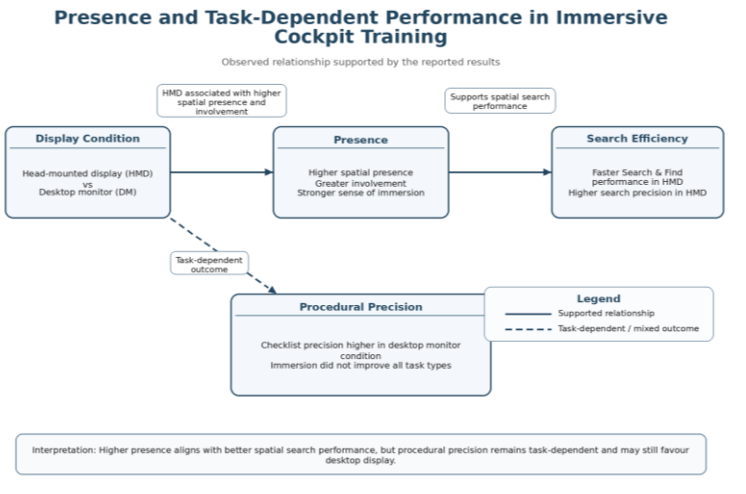



Figure 1 Left: shows the virtual Cessna 152 dashboard used for the cockpit search and checklist tasks. Right: Observed relationship between display condition, presence, search efficiency, and task-dependent procedural precision in immersive cockpit training.

## Conclusion

Presence should not be treated as a decorative user-experience metric in immersive pilot training. The data indicates that it is meaningfully associated with better visual search, yet not with improved performance on a simple procedural-recall task. The educational implication is that immersive cockpit systems should be targeted at activities requiring spatial attention and localization, while conventional displays remain valuable for rapid rehearsal of ordered procedures.

https://doi.org/10.18745/pb.27307

## References

[1] M. Slater and S. Wilbur, “A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments,” Presence: Teleoperators and Virtual Environments, vol. 6, no. 6, pp. 603-616, 1997.

[2] N. Nilsson, R. Nordahl, and S. Serafin, “Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence,” Human Technology, vol. 12, pp. 108-134, 2016.

[3] G. Makransky and G. B. Petersen, “The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality,” Educational Psychology Review, vol. 33, no. 3, pp. 937-958, 2021.

[4] W. T. Walters and J. Walton, “Efficacy of virtual reality training for pilots: A review of links between user presence, search task performance, and collaboration within virtual reality,” Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 65, no. 1, pp. 919-922, 2021.

[5] B. Olk, A. Dinu, D. J. Zielinski, and R. Kopper, “Measuring visual search and distraction in immersive virtual reality,” Royal Society Open Science, vol. 5, no. 5, p. 172331, 2018.

[6] E. Krokos, C. Plaisant, and A. Varshney, “Virtual memory palaces: immersion aids recall,” Virtual Reality, vol. 23, no. 1, pp. 1-15, 2019.

[7] M. C. Johnson-Glenberg, H. Bartolomea, and E. Kalina, “Platform is not destiny: Embodied learning effects comparing 2D desktop to 3D virtual reality STEM experiences,” Journal of Computer Assisted Learning, vol. 37, no. 5, pp. 1263-1284, 2021.

https://doi.org/10.18745/pb.27307

# Tiny-mBERT: A Lightweight Multi-Lingual Model for Intent Classification

**Xuanming Liu[1], Li Meng[1*], and Na Helian[1]**
[1]*University of Hertfordshire*
*Corresponding author: l.1.meng@herts.ac.uk

**Pre-trained multi-lingual models provide a shared representation space that can effectively support downstream tasks across multiple languages. However, deploying such models on resource-constrained devices requires a lightweight configuration, which is particularly challenging when training from scratch with limited task-specific data. This paper proposes a knowledge distillation framework designed to adapt a compressed lightweight model for intent classification across multiple target languages. Experimental results demonstrate that at a compression ratio of 4.35, the lightweight model retains an accuracy of 86.73% on English data while simultaneously achieving an average accuracy of 75.52%±3.39 across eight other target languages.**

Keywords: knowledge distillation (KD); multi-lingual adaptation; intent classification; lightweight model; cross-lingual alignment

## Introduction

The rapid development of dialogue systems and virtual assistants has boosted the demand for efficient natural language understanding, where intent classification plays a crucial role in enabling conversational agents to accurately interpret user demands. As these systems are designed to support multilingual and cross-platform users, the intent classification component must achieve both high accuracy and efficient deployment.

Large pre-trained language models such as mBERT [1] demonstrate strong performance across a wide range of NLP tasks in multilingual scenarios. However, their massive parameter sizes cause high memory usage, computational costs and latency, making them unsuitable for resource-limited devices such as mobile and edge devices. Meanwhile, transformers are proven to have structural redundancy, making model compression feasible. Among the common approaches for model compression, including pruning, quantisation and knowledge distillation (KD), KD is most effective. Nevertheless, most existing BERT compression studies focus only on English monolingual tasks. Although multilingual adaptation methods improve cross-lingual semantic alignment, they often overlook model compression, leaving a key challenge: how to compress multilingual models while retaining semantic representation and cross-lingual alignment.

To resolve this challenge, this paper proposes a three-stage knowledge distillation pipeline for multilingual intent classification, comprising model compression, task adaptation and multilingual alignment through dual-distillation and Multiple Negatives Ranking Loss (MNRL) [2]. The result lightweight model is named as Tiny-mBERT, which unifies model compression with multilingual alignment. The key contributions include the development of this unified framework and a comprehensive set of multilingual experiments demonstrating its effectiveness.

## Experimental Setup/Simulation

We propose a three-stage knowledge distillation framework as illustrated in Fig. 1. The framework is designed to progressively transfer multi-lingual intent classification knowledge from various teacher models to a student model – the Tiny-mBERT. We use the MASSIVE dataset to provide downstream task data. MASSIVE (Multilingual Amazon SLU Resource Package) is a large-scale multilingual corpus released by Amazon for slot-filling, intent classification, and virtual assistant evaluation [3]. The dataset was first collected in English, and then each utterance in the dataset was translated in parallel into 51 other languages. The dataset contains 16521 utterances per language, spanning across 60 intent classes. The distribution of utterances across intent classes is imbalanced, with the largest class (e.g., *calendar_set*) containing 1,150 samples, while the smallest (e.g., *cooking_query*) includes only 6. Let the English data in the MASSIVE dataset be denoted as $\mathbf{E}$, and the augmented English dataset is $\mathbf{E_a}$. Each English utterance $E_i \in \mathbf{E}$ has a corresponding translation ($T_i$) in each target language, forming aligned multilingual pairs for training.

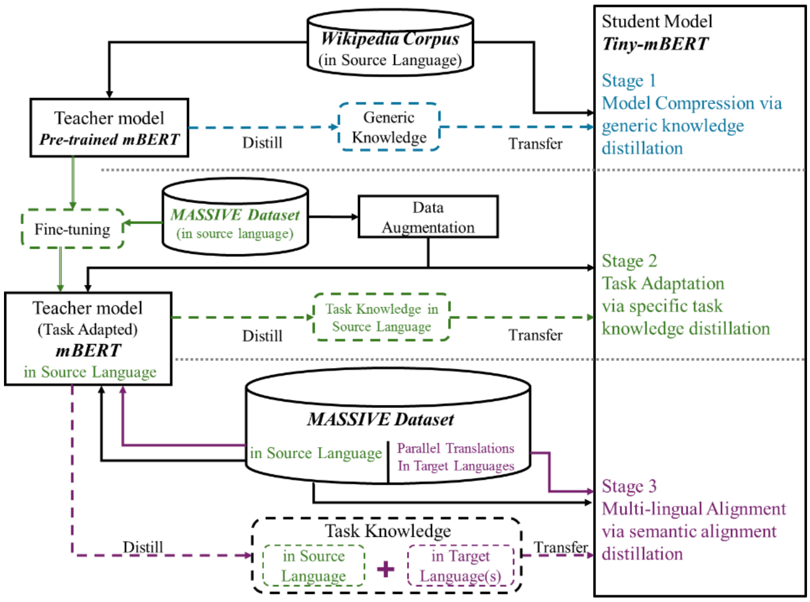



*Fig. 1. The proposed framework consists of three distinct knowledge distillation stages.*

Stage 1 of the proposed framework focuses on model compression through KD, aiming to retain general linguistic knowledge while significantly reducing the number of parameters. A pre-trained mBERT model is utilised as the teacher in the KD process, whereas the lightweight student model is randomly initialised. The Wikipedia corpus is used to distil general English language knowledge from the teacher to the student. This stage provides a robust and computationally efficient initialisation, forming a strong foundation for the subsequent distillation stages.

Stage 2 performs task adaptation for the source language (i.e., English for the MASSIVE dataset and hence in our work). The pre-trained mBERT model is first fine-tuned on the dataset $\mathbf{E}$, which is then augmented to create $\mathbf{E}_a$ by expanding each utterance fivefold. The student model is further trained on $\mathbf{E}_a$to learn the new downstream task of intent classification. This stage enables the student model to acquire task-specific knowledge while leveraging the compressed representations obtained during Stage 1.

Stage 3 addresses multilingual semantic alignment. In this stage, a dual-distillation mechanism integrated with Multiple Negatives Ranking Loss (MNRL) is introduced to enforce semantic consistency across languages. The teacher model from stage 2 is retained, and the student model continues its training to acquire multi-lingual knowledge. During each training iteration, input pairs $(E_i, E_i)$ and $(E_i, T_i)$ are fed into the model, and their respective losses are aggregated. This approach facilitates effective cross-language knowledge transfer while preserving a compact model configuration ideal for efficient deployment.

## Results and discussion

By leveraging both general linguistic knowledge and task-specific knowledge, the student model achieves approximately 86% accuracy on English data by the end of Stage 2. In our experiments, the number of attention heads was fixed at 12, and the intermediate size remains 4 times the hidden size in the student model, consistent with the original transformer configuration.

An uncompressed configuration with 12 transformer layers and a hidden size of 768 (denoted as 12L-768H) is used as the baseline for the student model. Various compressed configurations of the student model are then evaluated for comparison, as summarised in Table 1. The results indicate that performance on English intent classification remains robust even under substantial model compression. In particular, the student models achieve performance comparable to the baseline even when reduced to just 2 transformer layers with a hidden size of 204 (denoted as 2L-204H), demonstrating the effectiveness of the proposed distillation framework in preserving task performance while significantly reducing model complexity.

*Table 1. Comparison of accuracy across different student model configurations on English data after model compression and task adaptation. The compression ratio is calculated as the number of parameters in the 12L-768H baseline model divided by that in each compressed configuration.*

https://doi.org/10.18745/pb.27307

| Student model Configuration | 12L-768H | 4L-768H | 4L-312H | 3L-312H | 2L-312H | 2L-204H |
|---|---|---|---|---|---|---|
| # of layers | 12 | 4 | 4 | 3 | 2 | 2 |
| Hidden size | 768 | 768 | 312 | 312 | 312 | 204 |
| # of parameters | 178M | 122M | 42.1M | 40.9M | 39.7M | 25.7M |
| Compression ratio | 1.0(uncompress) | 1.46 | 4.23 | 4.35 | 4.48 | 6.93 |
| Accuracies on English | 87.65 | 87.15 | 86.53 | 86.73 | 86.35 | 86.17 |

Stage 3 experiments were conducted using all compressed student model configurations listed in Table 1. For each configuration, the student model was trained on the parallel translations across all target languages. The trained student model was then evaluated on each target language individually before an average performance is computed across all target languages.

As illustrated in Fig. 2, the average accuracy across different model configurations exhibits a stair-step decreasing trend, with models sharing the same hidden size achieving very similar performance. This observation suggests that, within the explored compression settings, hidden size has a more significant impact on multilingual alignment performance than model depth.

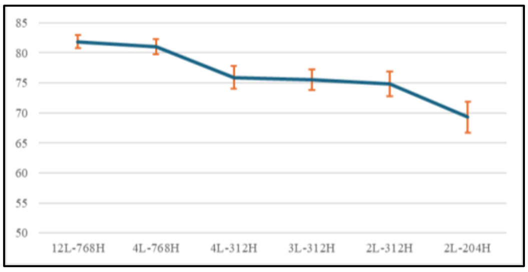



*Fig. 2. Comparison of the average performance obtained by different student model configurations.*

## Conclusion

This study demonstrates the effectiveness of the proposed three-stage distillation framework for multi-lingual intent classification. Experimental results show that the compressed student model maintains performance comparable to the uncompressed baseline on English intent classification while achieving effective knowledge transfer across multiple target languages. Although reducing model depth results in only marginal performance degradation, a reduction in hidden size leads to a more distinct drop in accuracy.

## References

[1] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding', in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), J. Burstein, C. Doran, and T. Solorio, Eds., Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[2] Piperno, Ruben, Luca Bacco, Felice Dell'Orletta, Mario Merone, and Leandro Pecchia. 'Cross-Lingual Distillation for Domain Knowledge Transfer with Sentence Transformers'. Knowledge-Based Systems 311 (February 2025): 113079. https://doi.org/10.1016/j.knosys.2025.113079.

[3] J. FitzGerald et al., 'MASSIVE: A 1M-Example Multilingual Natural Language Understanding Dataset with 51 Typologically-Diverse Languages', in Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), A. Rogers, J. Boyd-Graber, and N. Okazaki, Eds., Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 4277–4302. doi: 10.18653/v1/2023.acl-long.235.

https://doi.org/10.18745/pb.27307

# Bioaerosol Sensing Capabilities of Hydrogels

**Garrad, T.[1*], Dimov, N.[1], Johnson, I.[1], Coudron, L.[1], Cook, M.[2]**

[1]*University of Hertfordshire (UH)*

[2]*University College London (UCL)*

*Corresponding author: t.garrad@herts.ac.uk

**Hydrogel-based biosensors utilise the biocompatibility and properties of three-dimensional crosslinked polymer networks, enabling rapid, real-time, quantifiable assays that produce signals detectable and usable for continuous aerosol monitoring systems for environmental and military personnel. Experimentation will include stringent testing of the immunoassay implementation in hydrogel matrices and hydrogel modification to optimise porosity and, therefore, the diffusion of pathogens through the material. Preliminary data show effective baselines, which will be vital for future development. Hydrogel adhesion or well plate design will need to be modified for future runs; however, the assays themselves are functioning qualitatively and provide a basis for future quantification.**

Keywords: Hydrogel; polymer; immunoassay; biodetection; bacteriophage.

## Introduction

Immunoassays are among the most common methods for detecting and quantifying pathogens and other biomarkers, using antibodies to bind to specific proteins or epitopes on the antigens being analysed. Antigens can include viruses, bacteria or selective biomarkers within a sample [1], [2]. Immunoassays are commonly used to diagnose diseases in both qualitative point-of-care (POC) and quantitative laboratory settings. The most common immunoassays are enzyme-linked immunosorbent assays (ELISAs) or lateral flow assays (LFAs). An example application is cancer screening, which can measure tissue and serum biomarkers to allow medical practitioners to make important decisions about treatment plans [3]. ELISAs typically use labelled antibodies, which can be read in a plate reader and quantified by comparing the output to a standard concentration curve. This is also a limitation, as the concentration curve must be run and observed in the same sitting due to day-to-day differences in the readouts [4].

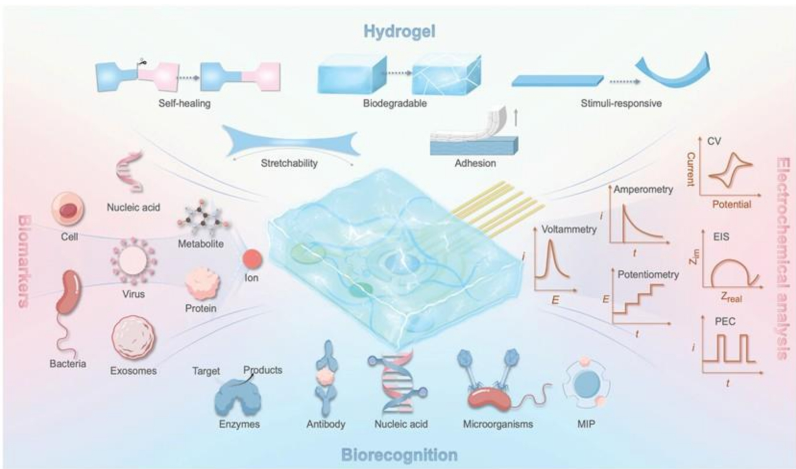



Figure 1: Variations of hydrogel-based electrochemical biosensors. Image taken from Huang et al., 2025 [5].

The integration of hydrogels aims to enhance immunoassays by improving assay sensitivity and reliability, even beyond the standard assay. Figure 1 shows some of the electrochemical sensors that can already be integrated into hydrogels [5]. Hydrogels are three-dimensional crosslinked polymer networks utilised for their ability to absorb solvent volumes many times their own dry weight [6]. Hydrogels can be synthetically derived from petroleum-based sources or from natural sources such as plant cellulose, which is a key consideration for using greener chemistries to extend the longevity of finite resources worldwide [7]. Among the polymer types, composite hydrogels have been used as a base to provide the required chemical functional groups that confer

https://doi.org/10.18745/pb.27307

both crosslinking properties and binding sites for the capture and detection of substances [8]. Typical hydrogel-based immunoassays occur on the surface of the gel, utilising the complex chemistries to improve the binding affinity of capture and detection substances, such as antibodies [9].

Altering the strength of chemical crosslinking can enhance the assay's capabilities by shifting the temperature at which the sol-gel transition occurs [10]. Targeting specific temperatures for the hydrogel's sol-gel point can enable the pathogens and detection elements of the assay to fully integrate into the overall system, changing state with temperature cycling in polymerase chain reaction (PCR) or during incubation periods in ELISAs [11], [12]. Chemically modified hydrogels with specific temperature-sol-gel points could be designed for use as integrated isothermal amplification assays, such as loop-mediated isothermal amplification (LAMP). Porosity is a major consideration in hydrogel-based assays; mobility and diffusion are key factors for pathogen detection, and having pores sufficiently large to allow the passage of pathogens whilst preventing contamination from larger sources is very important [13].

This experimentation aims to improve upon previously reported hydrogel-based immunoassays, optimising the detectors to reduce detection time to a level suitable for real-time pathogen detection for use as personal sampling devices. Once optimised, these detection devices will be readable by any unspecialised personnel, ensuring the results are freely available to anyone with the sensor.

## Proposed experimental

96-well plates, 3,3',5,5'-tetramethylbenzidine solution (TMB), horseradish peroxidase (HRP) antibody labelling kits and stop solution required for ELISAs were purchased from ABCAM (Cambridge, UK). Emesvirus zinderi (MS2) bacteriophage and anti-MS2 antibodies were obtained from DSTL (Porton Down, UK). Phosphate-buffered saline (PBS) tablets, bovine serum albumin (BSA) and general lab-ware were purchased from Sigma-Aldrich (Merck, UK). DI water was obtained from a Millipore Milli-Q system. Carbopol, cellulose, gelatine, and agarose, used as hydrogel precursors, were obtained from University College London (London, UK).

Initial experimentation includes testing the bacteriophage's binding site viability between samples of considerable storage age and freshly propagated samples. Evaluating the differences in signal strength and levels of antibody binding using a standard ELISA immunoassay. This, therefore, creates a baseline against which to compare the hydrogel assay research. This would be followed by binding antibodies directly to different hydrogels and evaluating whether there is a measurable difference in signal output between hydrogel-based and the standard assays, completed using optically labelled antibodies.

Other factors to consider are whether transparent hydrogels are necessary for the assay to work effectively, whether the difference in absorbance between the blanks and samples is sufficient, and whether the hydrogel's porosity affects the mobility of the pathogens within the crosslinked matrix. The transparency of the blank gels will be analysed using a plate reader at 450 nm, the wavelength of the labelled detection antibodies used in the standard ELISA, and subtracted from the hydrogel-based assay reading to obtain a final absorbance result. The hydrogel porosity is directly linked to crosslinking density, so the pathogen diffusion will be analysed by varying these properties and evaluating whether there is a measurable difference in absorbance.

## Preliminary results and discussions

Whilst the preliminary findings are completely insignificant, a few key points can be extracted from the data. Including that MS2 viability after long-term storage at 4 °C remains strong, and whilst reproduction is not possible, immunoassay testing with these samples is still possible. This result is supported by Olson *et al.* (2004), who studied medium- to long-term storage conditions and found no degradation of the active binding sites [14]. Initial ELISAs show good detection across concentration curves and random samples, with high reliability and repeatability. Whilst this data is qualitative only, it provides valuable insights into assay limits and where integration of the hydrogel will be most effective.

Hydrogel adhesion to the well plate was an issue during the washing steps, as the material kept falling out of the wells, invalidating the experiments. Changes in the solvent volume within the hydrogel led to more adhesive gels, but this was insufficient to prevent them from falling from the well during the more strenuous impacts required by the ELISA protocol. Adhesion could be improved by chemically bonding the hydrogel to the wells or by using modified 3D-printed wells with a hook-like implement that holds the material in place while stress is applied. Both options should be explored, and the best results should be used for further analysis.

## Conclusion

Preliminary findings are promising for future successes, and more work will be completed using different hydrogels and modified wells. Either a fresh or a stored bacteriophage sample has been shown to have functioning binding sites, providing sufficient signal strength for future assay testing and validation. Hydrogels have some ability to maintain the viability of pathogens, even if they are not bound to the capture entities within the sensor. Hydrogel mobility and adhesion will be key for assay development, and chemical modifications will be required for further study, either to the hydrogel or to the well plate surface.

## References

[1] R. M. Lequin, ‘Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA)’, *Clin. Chem.*, vol. 51, no. 12, pp. 2415–2418, Dec. 2005, doi: 10.1373/CLINCHEM.2005.051532.

[2] D. Wild, *The Immunoassay Handbook: Theory and applications of ligand binding, ELISA and related techniques*. 2013. Accessed: Jan. 12, 2026. [Online]. Available: https://books.google.com/books/about/The_Immunoassay_Handbook.html?id=xuYf6tcVdqYC

[3] S. Jeong, M. J. Park, W. Song, and H. S. Kim, ‘Current immunoassay methods and their applications to clinically used biomarkers of breast cancer’, *Clin. Biochem.*, vol. 78, no. 1, pp. 43–57, Apr. 2020, doi: 10.1016/j.clinbiochem.2020.01.009.

[4] S. Hosseini, P. Vázquez-Villegas, M. Rito-Palomares, and S. O. Martinez-Chapa, ‘Advantages, Disadvantages and Modifications of Conventional ELISA’, *SpringerBriefs in Applied Sciences and Technology*, no. 9789811067655, pp. 67–115, 2018, doi: 10.1007/978-981-10-6766-2_5.

[5] H. Deng *et al.*, ‘Chitosan hydrogel membrane-integrated ELISA platform for high-sensitive detection of alpha-fetoprotein in human serum’, *Talanta*, vol. 298, p. 128883, Feb. 2026, doi: 10.1016/J.TALANTA.2025.128883.

[6] P. Mehta, M. Sharma, and M. Devi, ‘Hydrogels: An overview of its classifications, properties, and applications’, *J. Mech. Behav. Biomed. Mater.*, vol. 147, p. 106145, Nov. 2023, doi: 10.1016/J.JMBBM.2023.106145.

[7] R. Kundu, P. Mahada, B. Chhirang, and B. Das, ‘Cellulose hydrogels: Green and sustainable soft biomaterials’, *Current Research in Green and Sustainable Chemistry*, vol. 5, p. 100252, Jan. 2022, doi: 10.1016/J.CRGSC.2021.100252.

[8] G. Bergamaschi *et al.*, ‘Composite Peptide–Agarose Hydrogels for Robust and High-Sensitivity 3D Immunoassays’, *ACS Appl. Mater. Interfaces*, vol. 14, no. 4, pp. 4811–4822, Feb. 2022, doi: 10.1021/acsami.1c18466.

[9] M. Z. Quazi, J. Hwang, Y. Song, and N. Park, ‘Hydrogel-Based Biosensors for Effective Therapeutics’, *Gels 2023, Vol. 9, Page 545*, vol. 9, no. 7, p. 545, Jul. 2023, doi: 10.3390/GELS9070545.

[10] E. Manso Castillo, W. Brett, J. Tarrio Saavedra, S. Naya Fernández, J. Lopez Beceiro, and A. Ponton, ‘Modelling sol-gel transition and mechanical properties of hydrogels using statistical learning’, *Rheologica Acta 2025 65:1*, vol. 65, no. 1, pp. 31–50, Nov. 2025, doi: 10.1007/S00397-025-01528-5.

[11] Y. Ye, W. Wang, X. Liu, Y. Chen, S. Tian, and P. Fu, ‘A Sol–Gel Transition and Self-Healing Hydrogel Triggered via Photodimerization of Coumarin’, *Gels*, vol. 10, no. 1, p. 21, Jan. 2024, doi: 10.3390/GELS10010021/S1.

[12] J. Li, A. Harada, and M. Kamachi, ‘Sol–Gel Transition during Inclusion Complex Formation between α-Cyclodextrin and High Molecular Weight Poly(ethylene glycol)s in Aqueous Solution’, *Polym. J.*, vol. 26, no. 9, pp. 1019–1026, 1994, doi: 10.1295/POLYMJ.26.1019;KWRD.

[13] R. Foudazi, R. Zowada, I. Manas-Zloczower, and D. L. Feke, ‘Porous Hydrogels: Present Challenges and Future Opportunities’, *Langmuir*, vol. 39, no. 6, pp. 2092–2111, Feb. 2023, doi: 10.1021/ACS.LANGMUIR.2C02253.

[14] M. R. Olson, R. P. Axler, and R. E. Hicks, ‘Effects of freezing and storage temperature on MS2 viability’, *J. Virol. Methods*, vol. 122, no. 2, pp. 147–152, Dec. 2004, doi: 10.1016/J.JVIROMET.2004.08.010.

https://doi.org/10.18745/pb.27307

# Electrostatic aerosol trapping device

**Etelka Chung[1*], Ian Munro[1], Ian Johnston[1], Loic Coudron[1]**
[1]*Wolfson Centre for Biodetection Instrumentation Research (WCBIR), School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK*
*Corresponding author: e.chung@herts.ac.uk

**Airborne bioaerosols pose a risk to health and agriculture, requiring effective and biologically safe sampling methods. This study investigates ozone-free electrostatic aerosol collection by eliminating corona discharge, as used in existing electrostatic precipitators. Using ITO-coated slides under varying voltages (±10 kV), aerosolised fluorescent 1 µm particles were sampled in a controlled chamber. Positively and negatively biased slides significantly enhanced collection, achieving up to 17.0x and 8.5x increases compared to uncharged control biases. A fan-assisted collection device obtained 59.1% collection efficiency. Results demonstrate that ozone-free electrostatic actuation enables efficient aerosol capture with the potential to preserve sample integrity for downstream analysis.**

Keywords: Aerosol; electrostatic; microbeads; air sampling; microbes.

## Introduction

Airborne transmission of pathogenic bioaerosols represents a considerable threat to both public and animal health, with significant implications for agriculture and the economy [1]. Effective monitoring and capture of these aerosols are therefore essential to limit the spread of infection and support early detection systems. Conventional electrostatic precipitation (ESP) has proven efficient for aerosol collection [2], but relies on corona discharge, which generates ozone as a byproduct. Ozone exposure can damage biological materials, reduce microorganism viability [3] and thereby compromise downstream analysis such as cultivation. To overcome this limitation while maintaining the advantages of electrostatic collection, this study explores ozone-free electrostatic actuation, removing corona discharge, as a viable alternative for bioaerosol sampling.

## Experimental

Indium tin oxide (ITO)-coated microscope slides (Diamond Coatings) served as the collection substrates. Each slide was subjected to one of four electrical conditions: positive bias, negative bias, grounded, or no electrical connection control (floating). The experiments were conducted in an 8 m$^3$ aerosol test chamber (Fig. 1a) with voltages ranging from -10 kV to +10 kV. Fluorescent polystyrene microbeads (1 µm diameter, PSM) were used as model aerosol particles. The particle concentration in the aerosol chamber was continuously monitored using four OPC-N3 optical particle counters (AlphaSense) positioned near the slides. Aerosols were generated via nebulisation for 15 minutes, followed by a 10-minute sampling period during which voltages were applied. After each test, the chamber air was cleaned using a HEPA-filtered extraction system, and the collected samples were imaged using an EVOS M7000 fluorescence microscope. Particle counts were quantified using Celeste 6 analysis software and normalised against the OPC-N3 count. Additionally, a fan-assisted collection unit (Fig. 1b) was constructed by laser cutting acrylic sheets (2mm thick) to direct airflow towards the slide. Both fan speed and distance from the collection substrate were varied to evaluate their effect on collection efficiency.

https://doi.org/10.18745/pb.27307

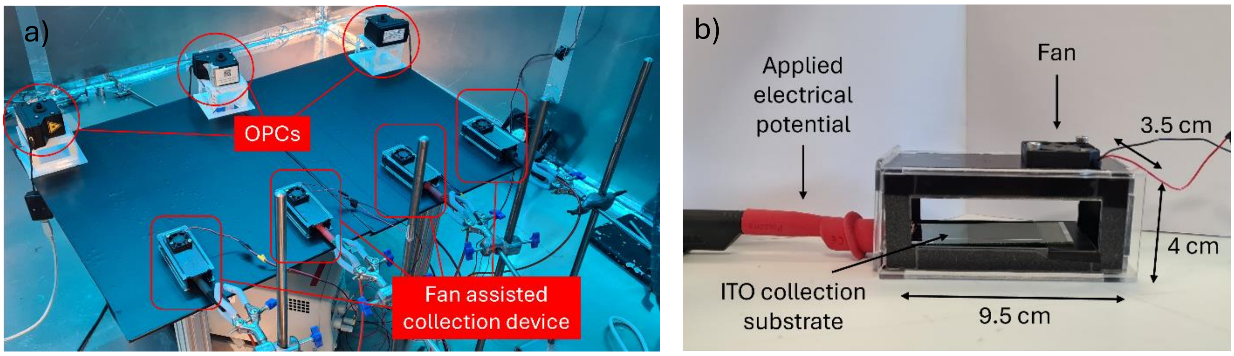



*Fig. 1. a) Aerosol chamber with experimental set-up, b) fan-assisted collection device*

## Results and discussion

Electrostatic capture of PSM aerosols was successfully achieved without the need for corona discharge. Positively and negatively biased slides collected up to 17.0 and 8.5 times more particles, respectively, compared with control slides, grounded and floating. The collection of PSM increased with the magnitude of the applied voltage, indicating that stronger electric fields enhanced particle attraction and adhesion. The highest collection performance was observed at +10 kV, representing a 1917.3% enhancement compared with the mean PSM count recorded under floating and grounded control conditions. Attraction of PSM is suggested to be governed by Coulomb forces.

The fan-assisted system initially produced a collection efficiency of 31.5%. Investigation on the airflow parameters revealed that lower fan speeds and shorter distances between the fan and slide enhanced performance, yielding a maximum efficiency of 59.1% at a flow rate of 10.9 L/min. Lower fan speed would result in lower particle velocities, thus a longer residence time in the electric field, and consequently a greater opportunity to capture particles compared to those with higher velocity. A shorter distance between the fan and the collection slide had enhanced collection efficiency due to a stronger electric field near the collection slide.

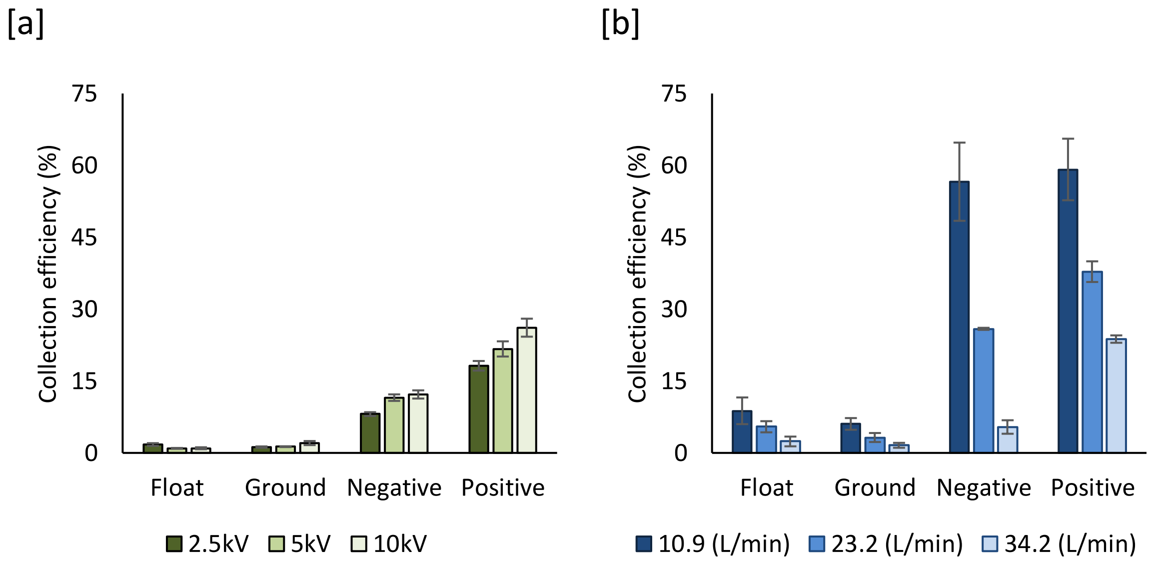



*Fig. 2. a) Collection efficiency from experiment using 4 v fan speed with fan at 3 cm distance from collection substrate with alternating applied potential (in kV), b) collection efficiency of the 4 tested electrical biases at different fan speeds (measured in airflow, L/min).*

## Conclusion

These findings demonstrate that ozone-free electrostatic actuation can provide effective aerosol collection with potential of viability and particle integrity preservation. The approach shows promise for future applications in environmental bioaerosol monitoring, public health surveillance, and agricultural biosecurity. Ongoing work will focus on assessing bioaerosol viability under these collection conditions to confirm suitability for sampling that requires bioaerosol viability.

This research was supported by the Biodetection Technologies Hub, which is funded by Research England's Expanding Excellence in England (E3) fund and the Engineering and Physical Sciences Research Council [grant number EP/X017591/1].

## References

[1] Z. Xu, Y. Wu, F. Shen, Q. Chen, M. Tan, and M. Yao, "Bioaerosol science, technology, and engineering: Past, present, and future," *Aerosol Science and Technology*, vol. 45, no. 11, pp. 1337–1349, Nov. 2011, doi: 10.1080/02786826.2011.593591.

[2] T. G. Foat *et al.*, "A prototype personal aerosol sampler based on electrostatic precipitation and electrowetting-on-dielectric actuation of droplets," *J. Aerosol Sci.*, vol. 95, pp. 43–53, May 2016, doi: 10.1016/J.JAEROSCI.2016.01.007.

[3] H. Ouyang *et al.*, "Control technologies to prevent aerosol-based disease transmission in animal agriculture production settings: a review of established and emerging approaches," *Front. Vet. Sci.*, vol. 10, p. 1291312, 2023, doi: 10.3389/FVETS.2023.1291312.

https://doi.org/10.18745/pb.27307

# Dynamic Load Modelling and Rule-Based Supervisory Controller for a 10 kW Hybrid Solar PV–Wind Off-Grid Microgrid

**Derek Iyamu*, Sikiru O. Ismail, Oluyomi Simpson, and Richard Addo-Tenkorang**
*Department of Engineering, School of Physics, Engineering and Computer Science,*
*University of Hertfordshire, Hatfield, AL10 9AB, England, United Kingdom*
*Corresponding author: d.iyamu@herts.ac.uk

**This study presents a 10 kW hybrid solar photovoltaic (PV)-wind off-grid microgrid developed in MATLAB/Simulink with a rule-based supervisory controller, as its main contribution. Unlike most comparable studies that rely on fixed resistive loads, the proposed controller is validated using a three-phase dynamic load block that represents realistic 24-hour residential demand. The results obtained show that the system successfully switches between PV and wind sources, as the load changes. The hybrid model achieves stable DC-link voltages of around 600 V for both the solar PV and the wind subsystems, a balanced 230 V root mean square (RMS), 400 V line-to-line AC output, and a total harmonic distortion (THD) of 2.88%, confirming compliance with IEEE 519 under dynamic operating conditions.**

Keywords: **Hybrid solar photovoltaic; Wind model; Dynamic load; Perturb and observe; Maximum power point tracking.**

## Introduction

Hybrid renewable energy systems, especially the combination of solar and wind sources, have recently attracted several studies. They provide an alternative power supply in areas without grid access. However, the research gaps in existing studies is that supervisory controllers are rarely tested under dynamic load conditions. Studies, such as Okokpujie et al. [1] and Halmous et al. [2], relied on fixed resistive loads and constant weather inputs, meaning their controllers are never exposed to the demand variation in a real off-grid household experiences throughout the day, and as such, it remains unclear whether these systems would switch correctly between sources when the load condition changes.

Therefore, this study addresses this gap by developing a 10 kW hybrid solar photovoltaic (PV)-wind microgrid model in Simulink, which implements separate DC-AC inverters for the solar and wind models and a three-phase dynamic block driven by an hourly household demand profile over a full 24-hour period. The objectives are to: (i) design and optimise a rule-based supervisory power dispatch between the PV and wind subsystems, (ii) validate stable DC-link voltage regulation under Perturb and Observe (P&O) maximum power point tracking (MPPT), and (iii) demonstrate a reliable three-phase AC output under droop control.

## Simulation

The hybrid solar PV-wind system was modelled in MATLAB/Simulink and designed as two 5 kW subsystems. The solar PV subsystem is based on the Trina Solar Vertex S (415 W per module), configured as a 6 series × 2 parallel PV array to output a total power of 5 kW at an open-circuit voltage of 300 V [3]. A DC–DC boost converter, which was regulated by a P&O MPPT controller, steps the output up to a 600 V DC bus. The wind subsystem integrates a 5 kW Aeolos-H horizontal-axis turbine with a permanent magnet synchronous generator (PMSG) [4]. A three-phase bridge rectifier converts the PMSG output to DC, and a second boost converter takes the voltage to the same 600 V DC bus.

Each DC bus feeds into the DC-AC inverter controlled by droop control, which maintains the frequency and three-phase voltage without disturbances between the inverters. A rule-based supervisory controller implements a solar-priority dispatch strategy, which means the solar PV inverter supplies the load whenever its output is sufficient, while the wind inverter compensates only when the load demand exceeds the 5 kW rated capacity of the solar PV system. The load demand was modelled, using a 1-D lookup table from the Simulink library in

MATLAB. The purpose of this block is to allow the implementation of hourly residential consumption values and observe in real time as it switches in the simulation.

## Results and discussion

**DC-link voltage stability**: The solar PV and wind turbine DC-DC boost converters maintained stable DC-bus voltages throughout the simulation. The solar PV DC link settled in the range of 600–613 V, while the wind DC link maintained a steady voltage of 585–600 V with minimal ripple, shown in Figs 1(a) and (b).

**Dynamic load analysis and supervisory control:** To test the system against realistic household demand, the load was set to vary from a 0.7 kW morning baseline up to a 6.4 kW midday peak, as depicted in Fig. 1(c). The result shows that the supervisory controller actively manages these changes. A 30% power constraint on the solar PV output forced the wind inverter to compensate early at time (t) = 5 hours. This wind generation scaled up further between 10 -14 hours, as the midday demand exceeded the restricted solar capacity.

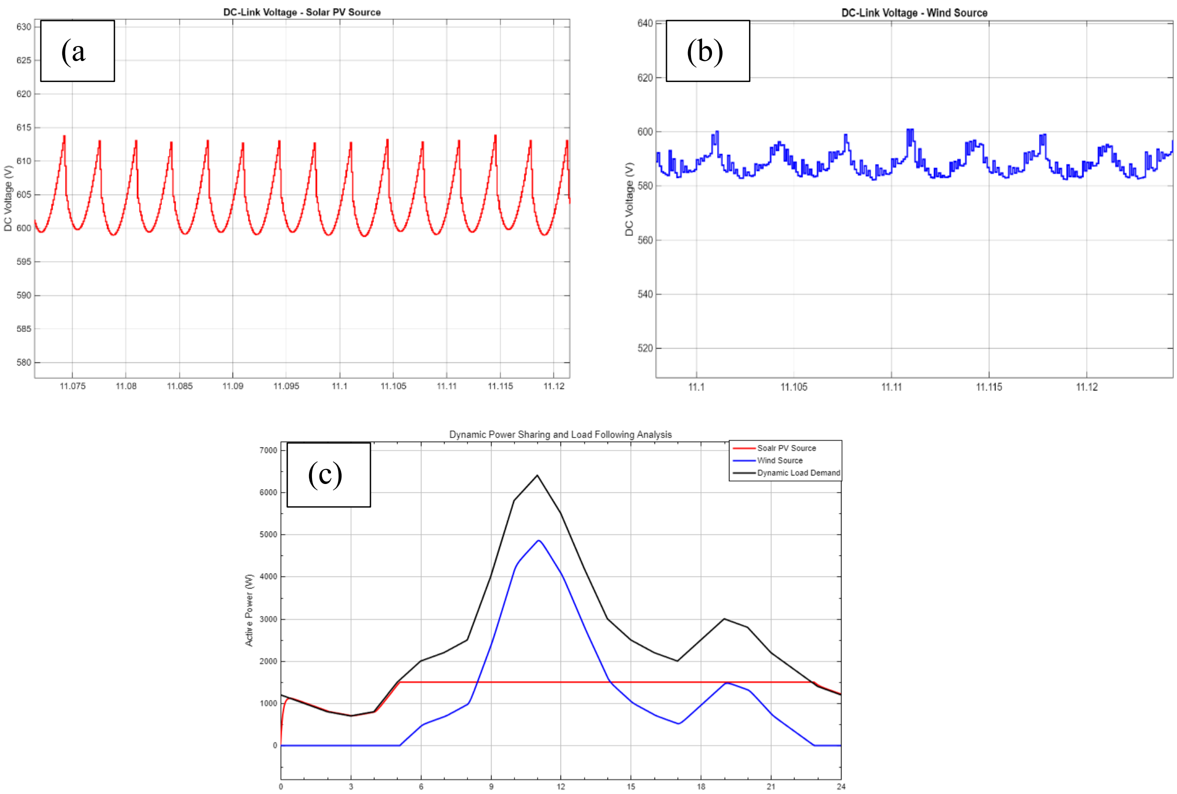



*Fig. 1. Simulation results for the hybrid solar PV-wind microgrid, showing the DC-link voltage of (a) solar PV and (b) wind sources as well as (c) dynamic power sharing and load analysis over a full 24-hour period, demonstrating wind compensation when solar PV output is constrained to a 1.5 kW limit (30%).*

## Conclusion

This study designed and validated a MATLAB/Simulink model of a 10 kW hybrid solar PV-wind off-grid microgrid, implementing a three-phase dynamic load block to capture realistic full 24-hour residential demand. The P&O MPPT technique was able to maintain a stable DC-link voltage in both the solar PV and the wind inverters. The rule-based supervisory controller successfully maintained solar power priority, ensuring that when the dynamic load exceeded the rated capacity of 5 kW, the wind inverter compensated for the additional demand during the simulation. Future work will incorporate a battery energy storage system along with variable irradiance and wind speed profiles to improve the hybrid model analysis under real-world conditions.

## References

[1] I. P. Okokpujie *et al.*, “Modelling and Simulation of Hybrid-Power Generation via Solar and Wind Turbine Using MATLAB/Simulink for Sustainable Energy Development,” *NJSTR*, vol. 7, no. 3, pp. 304–318, Aug. 2025, doi: 10.37933/nipes/7.3.2025.1622.

[2] A. Halmous, Y. Oubbati, M. Lahdeb, P. K. Balachandran, and S. Kannan, “Optimizing control and management of hybrid power system, consisting of PV-wind and battery-super capacitor, using COOT algorithm,” *Sci Rep*, vol. 15, no. 1, p. 33342, Sep. 2025, doi: 10.1038/s41598-025-12585-4.

[3] Trina Solar Limited, “Datasheet: Vertex S TSM-DE09R.05,” Trina Solar Limited, www.trinasolar.com, TSM_EN_2022_PA1, 2022. Accessed: Oct. 30, 2025. [Online]. Available: https://static.trinasolar.com/sites/default/files/Datasheet_Vertex_S_DE09R.05_EN_2022_PA1.pdf

[4] Aeolos Wind Turbine Company, “5kW Wind Turbines – 5000w Wind Generator – Home Wind Energy Turbine.” Accessed: Apr. 11, 2025. [Online]. Available: https://www.windturbinestar.com/5kwh-aeolos-wind-turbine.html

https://doi.org/10.18745/pb.27307

# HYDRA – Hydrogels for Bioaerosol Capture

**James Summers[1*], Ian Johnston[1], Michael Cook[2], and Loic Coudron[1]**

[1]*School of Physics, Engineering and Computer Science, University of Hertfordshire, U.K.*

[2]*UCL School of Pharmacy, University College London, London, WC1N 1AX, U.K.*

*Corresponding author: js23acc@herts.ac.uk

**Abstract**

Three hydrogels are known to be capable of capturing a wide range of infectious bioaerosols with significant efficiencies, and then sustaining the viability of these bioaerosols' pathogenic content to a useful degree. However, these hydrogels are all limited in this regard by, for instance, poor thermal stability, poor reproducibility, and/or dehydration issues. Moreover, no clear insight into the effects of any material properties on either of these performance parameters exists for all these materials. We propose that unexplored hydrogels can be developed as bioaerosol capture substrates without these limitations, and as model materials to gain further insight for this field.

Keywords: hydrogel; bioaerosol; capture; viability; materials

## Introduction

A polymer hydrogel is a three-dimensional polymer network that is capable of both absorbing and retaining a substantial amount of water while resisting dissolution. [1] These materials generally exhibit various properties, including high biocompatibility, mechanical softness, flexibility, and viscoelasticity. [2] Moreover, they often exhibit other properties, such as electroconductivity, [3] and/or transparency. [4] In addition, precise approaches, such as the modulation of the degree of cross-linking, for the engineering of one or more of their properties, such as degree of swelling, are commonly accessible. [5] Therefore, polymer hydrogels are key materials in various fields, including biomedical, environmental, and electronic devices.

Infections aerosols are aerosols that are capable of transmitting pathogens to host(s). Their properties generally include particle sizes between 1 µm to 10 µm, high-water content, and the significant presence of other biomatter. These aerosols often exhibit other properties, including a high-risk level, and/or anthropogenicity. [6]

Three hydrogels (mucus, [7] agar, [8] and gelatin [9]) are capable of capturing a wide range of these infectious aerosols with significant efficiencies, and then sustaining the viability of their pathogenic content to a useful degree. However, they are all limited in this regard by, for instance, poor thermal stability, poor reproducibility, and/or dehydration issues. Moreover, no clear insight into the effects of any material properties on either of these performance parameters exists for all these materials. The scope of this project is:

1. To explore whether hydrogels without all of the limitations specifically just stated have potential as high-collection-efficiency, viability-sustaining bioaerosol-capture substrates.
2. To provide some clear insight into the effect of both softness and hydration on both capture efficiency and viability factors for hydrogel materials.

This is targeted at being achieved through the following five objectives:

1. To identify a leading class of hydrogels for both of these aims through a thorough literature review.
2. To generate four candidates from this leading class with systematically varying softness.
3. To scope these four candidates' capture efficacity and viability factors for infectious aerosols.
4. To determine a leading candidate through analysis.
5. To explore whether one or both of bioaerosol-capture performance parameters can potentially be enhanced for the leading candidate through systematic variation in hydration.

https://doi.org/10.18745/pb.27307

## Experimental

Reaction mixtures were produced in 1.5 ml plastic Eppendorf tubes using an Opentrons OT2 liquid-handling robot according to scripts detailed at the GitHub repository, JamesSummers240. After generation, each reaction mixture was mixed for 10 s using a Vortex mixer. Reaction mixtures were then irradiated with a Kessil A160 Tuna Blue lamp (40 W) for their predetermined time within a bespoke safety enclosure. The generation of reaction mixtures and subsequent syntheses were all conducted under ambient atmospheres.

## Results and discussion

Poly(2-hydroxyethyl methacrylate) hydrogels (p(HEMA)) were chosen as the leading candidate class through a thorough and systematic literature review, over, for instance, poly(acrylamide) and poly(ethylene glycol) hydrogels. Subsequently, we determined a facile, general synthetic procedure to make these hydrogels using visible-light initiation and 1.5 ml plastic eppendorfs as reaction vessels, under ambient atmospheres.

A semi-automated workflow was subsequently developed to accurately map this specific polymerization process through five rounds of Diverse Mini-Batch Active Learning (contracted results shown in Figures 1 and 2). This system evaluated whether a reaction system will gel or not and, if it will gel, what the chemical yield, degree of swelling and optical transparency will be, with each dependent variable being mapped across five identical independent reaction variables.

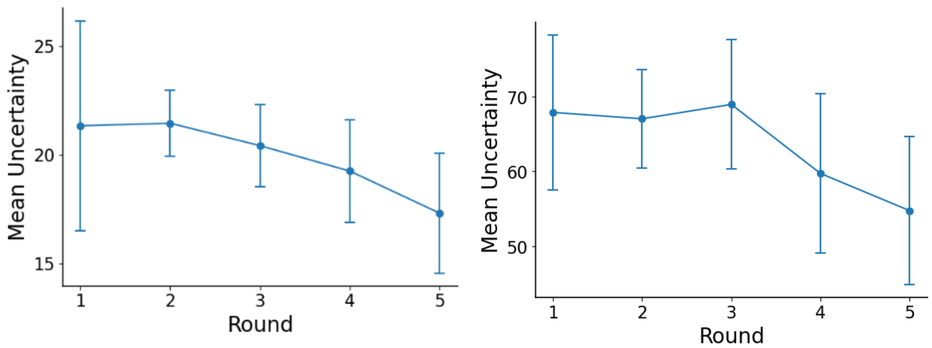



*Fig. 1. The means and corresponding standard deviations for the mean uncertainty for our yield and swelling models (left and right columns, respectively), for our specific test programme and for 70/30 test/train splits across 15 split seeds.*

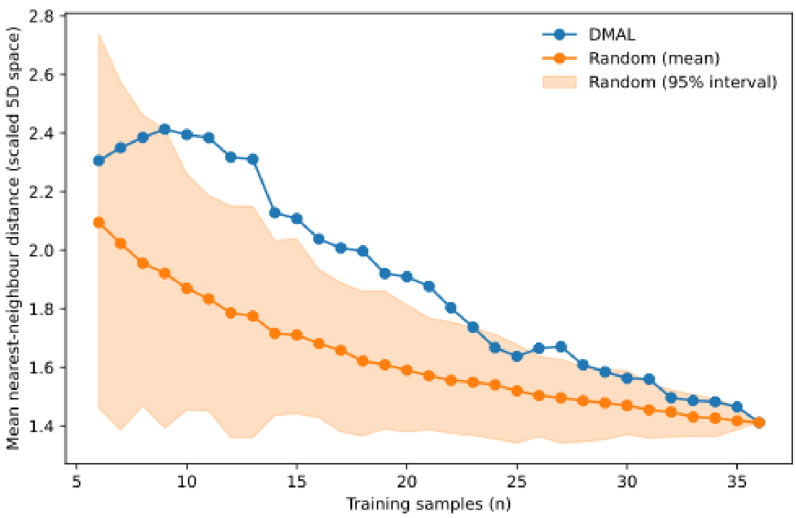



*Fig 2: Post-hoc analysis comparing DMAL order of sample selection to random selection (n = 500).*

https://doi.org/10.18745/pb.27307

The models generated were then applied to yield 10 initial pre-candidates for objective 2. This was done through:

1. Choosing suitable predictive design criteria for each of the predictive models.
2. Together, applying a soft scoring method to rank the entire polymer synthesis landscape by these design criteria.
3. Synthesising the initial top 10 pre-candidates and then collecting all of their associated initial experimental data.

The pre-candidates demonstrated the most of required criteria for evaluation, with, for instance, good transparency and yields being exhibited by all of the products. However, evaluation of softness is still outstanding prior to progression of any of the initial pre-candidates.

Simultaneously, effective moulds to produce hydrogel products with the required form-factor for integration with our Bespoke University of Hertfordshire Hydrogel Impactor (namely, a cylinder with diameter = 14 mm and height = 6 mm) were developed. Specifically, two 3D-printed silicone designs were found to have the required flexibility and inertness for the desired shaping during synthesis, and facile subsequent removal. In contrast, three PLA mould designs were found to lack the required flexibility and structural integrity for proper application.

## Conclusion

10 hydrogel pre-candidates have been experimentally generated with our required properties for evaluation of capture efficiency and viability factors of infectious bioaerosols, or of relevant simulants. Moreover, these pre-candidates exhibit the required form-factor for integration with our Bespoke University of Hertfordshire Hydrogel Impactor. However, evaluation of their mechanical softness is yet to be conducted, which will be required for any further project progression, although the feasibility of various indentation-based methods for this end is currently being investigated.

## References

The reference list must be provided using the IEEE citation format in Calibri 10 pt, single spacing.

[0] A. Author1, A. Author2, A. Author3. “Title of paper” Journal Name, vol. X, no. X, pp. range, year.
[1] T. C. Ho, C. C. Chang, H. P. Chan, T. W. Chung, C. W. Shu, K. P. Chuang, T. H. Duh, M. H. Yang, Y. C. Tyan. “Hydrogels: Properties and Applications in Biomedicine” Molecules, vol. 27, pp. 2902, 2022.
[2] N. A. Peppas, J. Z. Hilt, A. Khademhosseini, R. Langer. “Hydrogels in Biology and Medicine: From Molecular Principles to Bionanotechnology” Advanced Materials, vol. 18, no. 11, pp. 1345–1360, 2006.
[3] J. Li, D. J. Mooney. “Designing Hydrogels for Controlled Drug Delivery” Nature Reviews Materials, vol. 1, pp. 16071, 2016.
[4] E. Caló, V. V. Khutoryanskiy. “Biomedical Applications of Hydrogels: A Review of Patents and Commercial Products” European Polymer Journal, vol. 65, pp. 252–267, 2015.
[5] E. M. Ahmed. “Hydrogel: Preparation, Characterization, and Applications” Journal of Advanced Research, vol. 6, no. 2, pp. 105–121, 2015.
[6] R. Tellier, Y. Li, B. J. Cowling, J. W. Tang. “Recognition of Aerosol Transmission of Infectious Agents: A Commentary” BMC Infectious Diseases, vol. 19, pp. 101, 2019.
[7] B. Button, L.-H. Cai, C. Ehre, M. Kesimer, D. B. Hill, J. K. Sheehan, et al. “A Periciliary Brush Promotes Lung Health by Separating the Mucus Layer from Airway Epithelia” Science, vol. 337, no. 6097, pp. 937–941, 2012.
[8] A. A. Andersen. “New Sampler for the Collection, Sizing, and Enumeration of Viable Airborne Particles” Journal of Bacteriology, vol. 76, no. 5, pp. 471–484, 1958.
[9] SKC Inc. “Gelatin Filters (Cat. Nos. 225-9551 and 225-9552) Operating Instructions / Product Information” SKC Inc., product manual.

University of Hertfordshire UH
https://doi.org/10.18745/pb.27307

# Simulation-Based Assessment in the Era of Generative AI: Enhancing Decision-Making and Academic Integrity

**Alix Vargas[1*], Angela Traill[1], Aland Escudero[1] and Neil Vaz[1]**
[1] School of Physics Engineering and Computer Science, University of Hertfordshire
*Corresponding author: a.vargas@herts.ac.uk

**This study investigates the role of simulation-based assessment as a response to the challenges posed by generative AI in higher education. While AI tools raise concerns regarding academic integrity, cognitive offloading, and assessment validity, they also offer opportunities for innovation. Using a mixed-method approach combining survey data, simulation performance metrics, and learning analytics, the study evaluates how simulation environments support authentic assessment and decision-making skills. Findings suggest that simulation-based assessment reduces AI misuse risks while enhancing engagement, critical thinking, and professional competencies, offering a viable framework for assessment redesign in AI-driven educational contexts.**

Keywords: Generative AI; Simulation-based learning; Assessment design; Academic integrity; Decision-making

## Introduction

Generative AI tools such as ChatGPT are rapidly reshaping higher education, challenging traditional assessment paradigms that rely heavily on static, text-based outputs. While these tools offer opportunities for enhanced learning and personalised support, they also introduce significant risks related to academic integrity, assessment validity, and cognitive dependency. The increasing reliance on AI-generated outputs has been associated with cognitive offloading, potentially reducing higher-order thinking engagement (Gerlich, 2025).

Empirical evidence suggests that generative AI can improve performance but may also reduce mental effort if not carefully integrated (Deng et al., 2025), highlighting the need to rethink assessment strategies.

This study addresses this gap by proposing simulation-based assessment as a robust alternative. Unlike traditional approaches, simulation environments require students to engage in iterative decision-making, trade-off analysis, and real-time problem-solving. The paper investigates whether such process-oriented assessment can mitigate AI-related risks while enhancing engagement, critical thinking, and professional competencies.

## Methodology

A mixed-method research design was employed to evaluate simulation-based assessment within an AI-influenced learning environment. The study draws on three complementary data sources: (1) perception data from student and academic surveys examining attitudes towards generative AI and simulation-based learning; (2) performance data derived from simulation outcomes, including KPI-based measures such as Weighted ROI (Return on Investment) across multiple decision-making rounds; and (3) engagement data captured through attendance, participation, and output quality using learning analytics.

The study was conducted with MSc students (*n ≈ 30*) and academics (*n = 4*) involved in a simulation-based module using *The Fresh Connection* (InChainge, 2026; Kok and van Zyl-Cillié, 2024)(InChainge, 2026; Kok and van Zyl-Cillié, 2024). In this simulation of a hypothetical fresh fruit beverage manufacturing company, students were assigned specific operational KPIs and engaged in six iterative rounds of cross-functional decision-making, involving decisions relating to Sales, Purchasing, Operations, and Supply Chain. This required them to evaluate trade-offs and adapt strategies dynamically in response to evolving simulation conditions.

Quantitative and comparative analyses were conducted to examine relationships between engagement and performance. This approach enables an integrated evaluation of student behaviour and perceived learning impact.

https://doi.org/10.18745/pb.27307

## Results and discussion

The findings demonstrate that simulation-based assessment provides a robust approach to addressing challenges introduced by generative AI, by shifting focus from static outputs to performance-based evaluation. Three key dimensions are examined. First, the progression of student performance across simulation rounds is analysed to capture learning development over time. Second, differences in student and academic perceptions of assessment robustness are explored. Third, the relationship between engagement and performance outcomes is investigated to understand how sustained participation influences learning effectiveness.

1. Engagement as a driver of performance

Simulation-based assessment enables the observation of how student performance evolves through iterative decision-making processes. To examine this, simulation outcomes were tracked across multiple rounds using average ROI as a business-related key performance indicator that is aligned to the simulation's subject matter. The results indicate a clear pattern of progressive improvement. Students begin with negative performance outcomes, reflecting initial uncertainty in the subject matter and limited familiarity with the simulation environment. However, performance steadily improves across subsequent rounds, approaching break-even and, in some cases, positive results. This trend suggests that students refine their strategies, through iterative engagement with the simulation environment, over time learning from previous decisions and adapting to system feedback. The evolution of simulation performance across rounds is presented in Fig. 1.

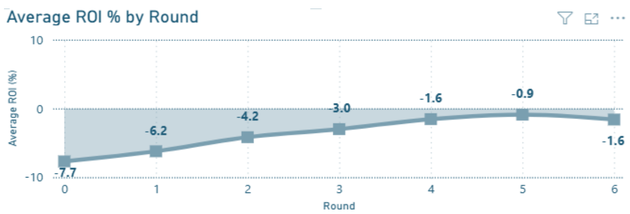



*Fig. 1. Evolution of Average Simulation Performance (ROI) Across Iterative Rounds*

2. Perceptions of Assessment Robustness in the Context of Generative AI

Survey data from students (n ≈ 30) and academics (n = 4) were analysed to examine perceptions of the new assessment robustness with simulation at its core, in the context of generative AI. The findings reveal a divergence between the two groups. Academics predominantly perceive newly introduced assessment approaches as robust, whereas student responses are more distributed across categories, indicating greater uncertainty regarding the reliability of the new assessment methods. Though requiring further investigation, this variation suggests that students may be more aware of the potential vulnerabilities of traditional assessment formats to AI-supported task completion; indeed latter observations indicated AI was used to clarify subject-specific terms and identify decision-making strategies. Academics, however, retain greater confidence in the new assessment approach. A comparison of these perceptions across both groups is presented in Fig. 2.

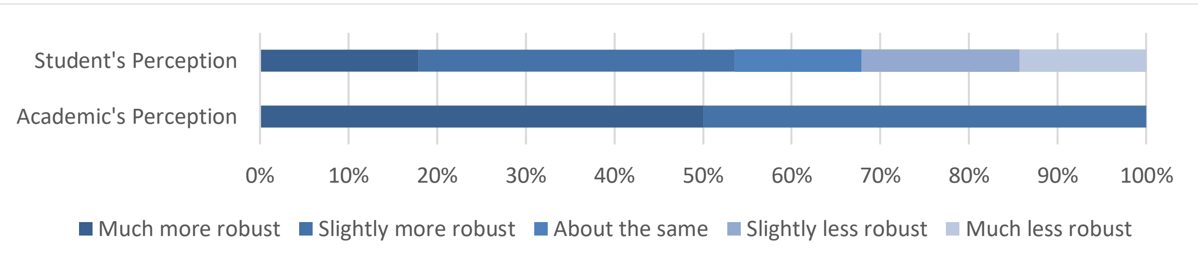



*Fig. 2. Comparison of Student and Academic Perceptions of Assessment Robustness in the Context of Generative AI*

3. Engagement as a driver of performance outcomes

The findings highlight the importance of student engagement in shaping performance within simulation-based assessment. While overall performance improves over time (Fig. 1), variations across students indicate that engagement plays a key role in determining outcomes. Performance was analysed across different engagement levels, measured through attendance and participation. Students with lower engagement consistently demonstrate weaker results, including negative ROI, whereas highly engaged students achieve stronger, often positive, performance outcomes. This pattern suggests that performance is not solely driven by exposure to the simulation, but by active participation in iterative decision-making processes. Higher engagement enables students to better understand system dynamics and refine strategies over time. A comparison of performance across engagement levels is presented in Fig. 3.

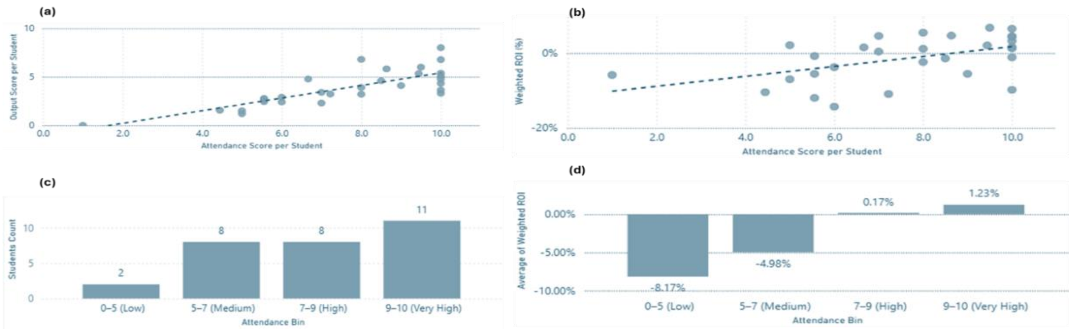


*Fig. 3. Impact of student engagement on performance outcomes: (a) attendance and academic output quality, (b) attendance and simulation ROI, (c) distribution of engagement levels, and (d) average ROI by engagement level.*

## Conclusion

This study demonstrates that simulation-based assessment provides an effective response to the challenges introduced by generative AI in higher education. By focusing on decision-making processes rather than static outputs, it enhances authenticity and supports more valid evaluation of student learning.

The findings support a shift towards human-in-the-loop, process-oriented assessment models, where AI is used as a support tool rather than a substitute for learning. This approach positions simulation-based assessment as a resilient framework for maintaining academic integrity while fostering higher-order skills in AI-driven learning environments. However, the findings should be interpreted within the context of a small-scale pilot study conducted within a specific MSc module, which may limit broader generalisation. In addition, KPI-based measures such as ROI may not fully capture the quality of reasoning, teamwork, or reflective judgement. Future research should therefore explore scalability across larger and more diverse cohorts, long-term learning impact, and the integration of AI-enhanced feedback systems while preserving authentic assessment practices.

## References

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers and Education*, *227*. https://doi.org/10.1016/j.compedu.2024.105224

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies*, *15*(1). https://doi.org/10.3390/soc15010006

InChainge. (2026). *The Fresh Connection*. https://game.thefreshconnection.eu/

Kok, T. J., & van Zyl-Cillié, M. (2024). The Fresh Connection As An Experiential Learning Tool In Industrial Engineering Education: A Case Study. *South African Journal of Industrial Engineering*, *35*(3), 135–147. https://doi.org/10.7166/35-3-3087

https://doi.org/10.18745/pb.27307

# A Synthetic Data Pipeline for Instruction-Following Compliance in Clinical Conversational AI

**Akhila Anil[1*], Eamonn Swanton[2], Farshid Amirabdollahian[1] and Yi Sun[1]**
*[1]School of Physics, Engineering and Computer Science, University of Hertfordshire*
*[2]Heales Medical*
*Corresponding author: a.anil@herts.ac.uk

**This work presents a synthetic data pipeline designed to fine-tune a Structured Clinical Conversational Assistant (CA) that can follow structured instructions while communicating naturally with patients, without requiring a separate governance layer to validate outputs at inference time. The pipeline covers 27 consultation phases spanning the full clinical workflow, encoding domain-specific constraints directly into the training data to ensure the model learns compliant behavior from the ground up rather than relying on inference-time correction. The generated dataset was evaluated across four complementary metrics, n-gram leakage, response length compliance, BERTScore F1, and LLM-as-judge pass rate which covers instruction independence, response conciseness, semantic alignment, and clinical boundary compliance, with results indicating that the dataset meets the quality threshold suitable for fine-tuning.**

Keywords: fine-tuning; synthetic data generation; clinical NLP; instruction following; LLM-as-judge

## Introduction

Clinical Conversational Agents (CAs) developed to assist medical professionals must satisfy two requirements simultaneously: follow a strict conversation protocol and communicate naturally with patients. Getting both right at the same time is harder than it looks. A common solution is a governance layer at inference time that checks and regenerates non-compliant responses, but this adds latency and complexity that make real-time deployment difficult.

This work started from a practical problem in a CA that conducts structured patient evaluations based on instructions provided by the medical professional. A key constraint of the system is that no recommendation or diagnosis is permitted, as a human clinician is always expected to remain in the loop and act on the CA's conclusions.

The first model trialled, Llama 4 [6], followed the clinical protocol reasonably well but was highly repetitive. It defaulted to the same opening phrases across turns, and in most cases the response had to be regenerated due to the frequency with which it was breaking the governance rules. This highlighted that a fine-tuned model specific to the CA use case would be far more beneficial than a general-purpose large language model. The second model, Mistral 8B [7] fine-tuned on medical conversational dialogue data, was more fluent but drifted from the clinical protocol. It would sometimes ignore the structured instruction it was given and respond as a general chatbot. In some cases it was even willing to provide a diagnosis to the patient, which was completely against the guidelines of the system.

Both problems share the same root cause: neither model was trained on data that reflected the specific constraints of a structured clinical consultation. The hypothesis explored in this work is that if the training data is designed to cover those constraints, through phase-specific instructions, diverse patient inputs, and examples of correct boundary behaviour, the model will learn to comply with the requirements of the system without needing a governance layer at inference time, making the system faster and more practical to deploy.

## Methodology

The pipeline was developed following an initial evaluation of two baseline models, Llama 4 and a conversational fine-tune of Mistral 8B , both deployed in the live CA system. The failure modes identified during this evaluation, namely high phrase repetition and instruction non-compliance, led to the design of the synthetic data generation pipeline.

https://doi.org/10.18745/pb.27307

A synthetic data generation pipeline [1] was built using MedGemma 27B [5], selected for its domain-specific medical pretraining to produce clinically grounded patient-assessor exchanges. MedGemma 27B was prompted with phase-specific system instructions and few-shot examples to generate realistic patient-assessor dialogue pairs across 27 interaction phases spanning the full consultation workflow, including authentication, GP visit enquiry, medication collection, symptom reporting, severity rating, follow-up questioning, adversarial boundary enforcement, and consultation closure.

Triple variance augmentation was applied across the system instruction, patient input, and assistant response of each generated patient-agent exchange. System instructions were paraphrased across records to prevent the model from memorising a specific instruction string and instead learn the underlying intent. Patient inputs were rewritten from clean laboratory responses such as "yes" or "01/01/1985" into natural conversational equivalents, making the model robust to real patient language. Assistant responses were paraphrased to prevent the model from defaulting to the same opening phrase across turns, which was one of the key failure modes observed in the baseline model in the initial evaluation.

Dataset quality was assessed using four complementary metrics. *N-gram leakage* measures the degree to which response text directly overlaps with the system instruction [3], with scores ranging from 0 (no overlap) to 1 (complete overlap), a lower score indicates greater instruction independence, meaning the model has learned to follow intent rather than copy phrasing. *Response length compliance* evaluates whether responses respect the word limits specified per phase, expressed as a ratio from 0 to 1 where 1 indicates the response exactly meets the word limit and values below 1 indicate more concise responses. *BERTScore F1 [2]* captures semantic alignment between each response and its instruction without penalising natural paraphrasing, with scores ranging from 0 to 1 where a score in the range of 0.85 to 0.92 indicates the response is semantically relevant to the instruction. *LLM-as-judge [4]* scoring provides a clinically grounded pass rate ranging from 0% to 100%, each generated record was evaluated by a separate Llama 4 judge model against 28 failure rules defined to systematically cover the major identifiable failure modes across the 27 consultation phases, spanning five categories: clinical boundary violations, instruction following failures, conversation flow failures, tone and style failures, and completeness failures. Only high-confidence failures were removed to maintain dataset volume.

## Results and Discussion

The final training set of 626 records was obtained from an initial generation of 222 records. The 222 records were generated across 27 consultation phases, each producing multiple exchanges from varied patient inputs paired with randomly sampled instruction paraphrases, with low-quality responses filtered out at generation time,this is expanded to 888 through triple variance augmentation, and reduced to 626 following LLM-as-judge filtering.

Evaluation across 626 records indicates strong instruction independence, response variation, and constraint adherence across all consultation phases, as summarized in Table 1. N-gram leakage scores confirm that responses are largely independent from their system instructions, with the majority showing no direct word-for-word overlap. Response length compliance confirms that responses consistently respect the word limits specified per phase. BERTScore F1 falls within the target range, confirming semantic alignment between responses and instructions without direct reproduction of instruction text. LLM-as-judge evaluation reflects strong overall clinical compliance, with the remaining failures concentrated in structurally complex phases.

LLM-as-judge evaluation using Llama 4 against 28 clinical failure rules produced an overall pass rate of 94.2%, with the remaining 5.8% of failures spread across five categories. The most common failure was over-questioning, where the model asked multiple questions in a single turn despite being instructed to ask only one. This was followed by instruction leakage, where response text closely mirrored the wording of the system instruction rather than expressing the same intent naturally. A smaller proportion of failures involved the model omitting information it was required to ask for, or continuing to collect symptoms after the patient had indicated they were done. In a few cases it also crossed clinical boundaries by implying a diagnosis.

| Metrics | Value | Indicates |
|---|---|---|
| N-gram leakage score | 0.0167 | Responses are largely independent from instruction wording |
| Length compliance rate | 99.5% | Responses consistently respect the word limit specified per phase |
| BERTScore F1 | 0.880 | Responses are semantically aligned with instructions without verbatim copying |
| LLM judge pass rate | 94.2% | 94 in 100 responses meet all the clinical compliance rules |

*Table 1. Quality evaluation of 626 synthetic training records across automatic and LLM-as-judge metrics.*

## Conclusion

This work presents a constraint-aware synthetic data pipeline for fine-tuning a CA, motivated by two observed failure modes in existing approaches: high phrase repetition in the Llama 4 baseline and instruction non-compliance in the conversational fine-tune, and demonstrates that these behaviors can be systematically targeted through training data design. The resulting dataset of 626 training records achieving 94.2% clinical compliance, strong semantic alignment (BERTScore F1: 0.880), and near-perfect length compliance (99.5%). While the quantitative metrics presented here provide an objective measure of dataset quality, human evaluation remains the gold standard for assessing the naturalness and clinical appropriateness of conversational AI outputs. As next steps, the fine-tuned model will be evaluated in a live CA system against a Llama 4 baseline, measuring governance fallback rate, repetition, and clinical boundary compliance.

### Acknowledgment

This work is part of a Knowledge Transfer Partnership between Business Partner and the University of Hertfordshire (KTP reference 13794 ).

## References

[1] H. Murtaza, M. Ahmed, N. F. Khan, G. Murtaza, S. Zafar and A. Bano, "Synthetic Data Generation: State of the Art in Health Care Domain," *Computer Science Review*, vol. 48, p. 100546, 2023. doi: 10.1016/j.cosrev.2023.100546.

[2] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, “BERTScore: Evaluating Text Generation with BERT,” *arXiv:1904.09675 [cs]*, Feb. 2020, Available: https://arxiv.org/abs/1904.09675

[3] A. Purpura, L. Wang, S. Badyal, E. Beaufrand and A. Faulkner, "Deconstructing Instruction-Following: A New Benchmark for Granular Evaluation of Large Language Model Instruction Compliance Abilities," *arXiv*, 2024. Available: https://arxiv.org/abs/2601.18554.

[4] L. Zheng, W. L. Chiang, Y. Sheng, S. Zhuang, Z. Wu, Y. Zhuang, Z. Lin, Z. Li, D. Li, E. Xing, H. Zhang, J. E. Gonzalez and I. Stoica, "Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena," in *Proc. NeurIPS*, 2023.

[5] D. Golden and R. Pilgrim, "MedGemma: Our Most Capable Open Models for Health AI Development," Google Research Blog, Jul. 2025. Available: https://research.google/blog/medgemma-our-most-capable-open-models-for-health-ai-development/

[6] Meta AI, "Llama 4," Meta AI Blog, 2025. Available: https://ai.meta.com/blog/llama-4-multimodal-intelligence

[7] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. de las Casas et al., "Mistral 7B," arXiv:2310.06825, 2023. Available: https://arxiv.org/abs/2310.06825

https://doi.org/10.18745/pb.27307

# Experimental Evaluation of LoRa Communication Performance in Enclosed Environments for HUMS Applications

**Chris Aldhous-Grayce[12*], Dr Arooj Siddiqui[1]**
[1]*School of Physics, Engineering and Computer Science, University of Hertfordshire, UK*
[2]*MBDA UK Ltd, Stevenage, UK*
*Corresponding author: ca22acc@herts.ac.uk

**This study evaluates the impact of enclosure materials and antenna configuration on LoRa communication performance for Health and Usage Monitoring Systems (HUMS). A controlled indoor setup was developed, with a LoRa sensor node placed within various test enclosures and a LoRa gateway positioned externally under line-of-sight conditions. RSSI and SNR were measured at separation distances of 1 m, 5 m and 10 m. Results highlight that enclosure material significantly affects signal attenuation with glass-reinforced plastic (GRP) exhibiting the highest losses, while antenna configuration influences performance. Despite reductions in RSSI, SNR remained relatively stable demonstrating the robustness of LoRa communication.**

Keywords: LoRa; RF propagation; enclosure effects; antenna configuration; HUMS

## Introduction

LoRa (Long Range) communication is widely used for low-power, long-range wireless sensor networks [1] including monitoring applications such as Health and Usage Monitoring Systems (HUMS) due to its ability to provide reliable communication with low power consumption [2].

In practical deployments sensor nodes are often enclosed within protective housings, which can affect radio frequency (RF) propagation through attenuation, reflection and absorption. These effects depend on enclosure material properties and antenna configuration [3] and are particularly relevant in safety-critical applications where nodes are deployed within sealed environments.

This study experimentally evaluates the influence of enclosure materials, antenna configuration and distance on LoRa communication performance under controlled indoor line-of-sight (LoS) conditions. Performance is assessed using Received Signal Strength Indicator (RSSI) and Signal-to-Noise Ratio (SNR) at separation distances of 1 m, 5 m and 10 m.

The objectives of this study are to: (i) quantify the impact of enclosure materials on signal attenuation; (ii) evaluate the influence of antenna configuration and (iii) assess the effect of distance on RSSI and SNR.

This study provides a comparative analysis of key factors affecting LoRa performance in enclosed environments, offering practical insights for reliable HUMS deployment.

## Experimental Setup

A controlled indoor experimental setup was developed to evaluate the impact of enclosure materials and antenna configurations on LoRa communication performance. The system consisted of a LoRa module acting as the transmitter and a LoRa gateway as the receiver.

Three enclosure materials were investigated: aluminium, glass-reinforced plastic (GRP) and medium-density polyethylene (MDPE), representing a range of conductive and non-conductive properties. A baseline configuration without an enclosure was also considered.

The overall experimental setup is illustrated in Fig. 1, showing the relative positioning of the sensor node, enclosure and gateway under line-of-sight (LoS) conditions.

https://doi.org/10.18745/pb.27307

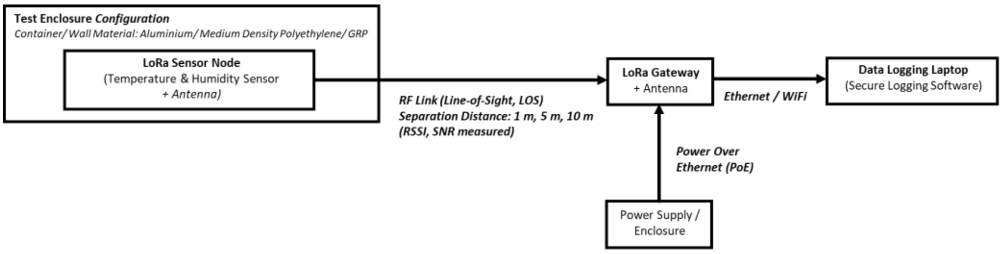



*Fig. 1. Schematic representation of the experimental setup showing the LoRa sensor node within the enclosure and the externally positioned gateway under line-of-sight conditions.*

Multiple antenna configurations were evaluated to assess the effect of antenna type and placement. Module antennas included a wire monopole, a flexible printed circuit (FPC) dipole, an embedded FPC antenna and a fourth-generation Long-Term Evolution (LTE) puck antenna. The monopole, FPC dipole and embedded antennas were integrated within the enclosure, while the LTE puck antenna was positioned externally to the module but remained within the enclosure. Two gateway antennas were used: an omnidirectional fibreglass antenna and a directional Yagi antenna.

Measurements were conducted at separation distances of 1 m, 5 m and 10 m. For each configuration RSSI and SNR were recorded over fixed-duration intervals using the gateway interface, enabling consistent logging and data extraction. All tests were performed in a controlled indoor environment to minimise external interference.

A total of 96 test cases were carried out covering all combinations of enclosure material, antenna configuration and separation distance.

## Results and discussion

Communication results highlight the influence of enclosure material, antenna configuration and distance on LoRa communication performance. Figures 2 and 3 present the measured RSSI and SNR across a sample of the tested configurations.

As shown in Fig. 2, RSSI decreases with increasing distance for all configurations consistent with expected propagation losses. Attenuation varies with enclosure material with GRP exhibiting the highest attenuation, producing consistently lower RSSI values compared to aluminium and MDPE. This behaviour is attributed to increased absorption and scattering within the material. In contrast, the baseline (no enclosure) configuration maintains higher RSSI values across all distances.

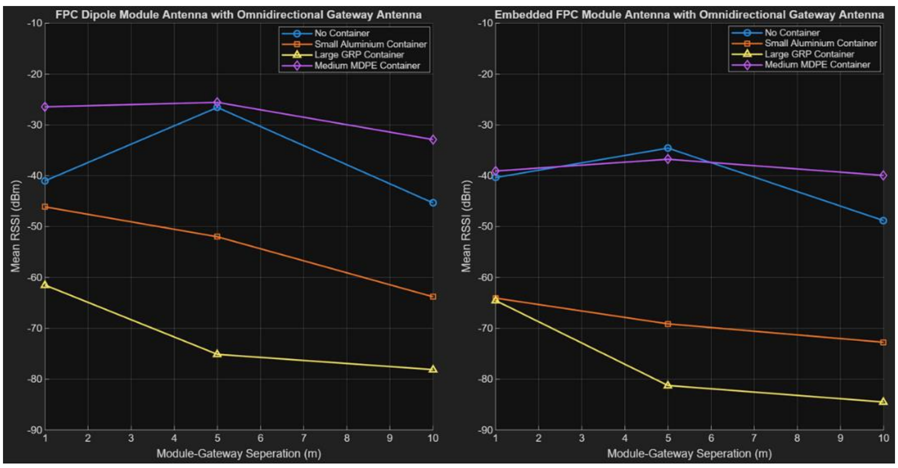



*Fig. 2. RSSI variation with distance for different enclosure materials and antenna configurations.*

Antenna configuration also influenced RSSI performance. The results presented focus on the FPC dipole and embedded FPC antennas, which demonstrated noticeable differences in signal behaviour across enclosure materials and distances. The FPC dipole antenna generally exhibited more stable RSSI performance likely due

https://doi.org/10.18745/pb.27307

to its radiation characteristics, while the embedded FPC antenna showed greater sensitivity to enclosure effects. This highlights the importance of antenna design and integration within constrained environments, where the surrounding materials can significantly influence signal propagation.

The SNR results shown in Fig. 3, demonstrate a different trend. While some variation was observed across configurations and distances, SNR remained relatively stable compared to RSSI. This indicates that despite reduced signal strength with distance and enclosure effects, the overall link quality remains sufficiently robust for reliable communication. This behaviour reflects the resilience of LoRa modulation, which is designed to operate effectively under low signal conditions [4].

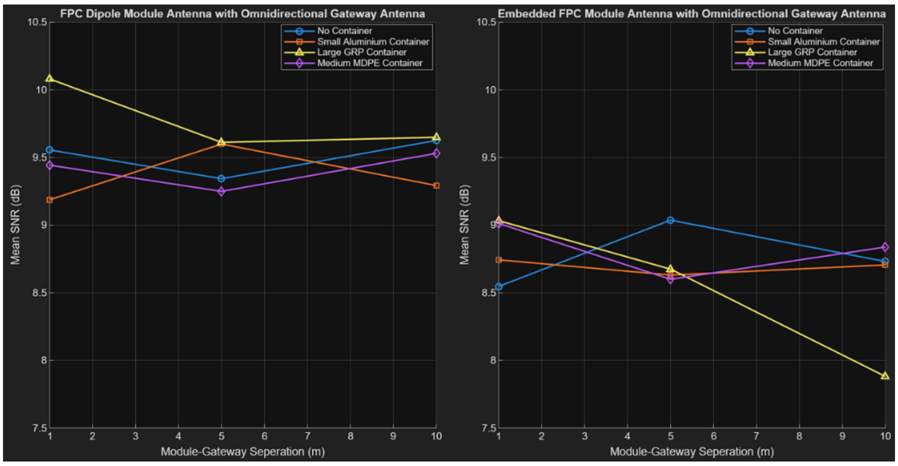



*Fig. 3. SNR variation with distance for different enclosure materials and antenna configurations.*

Overall enclosure material is the dominant factor affecting signal attenuation, with antenna configuration providing an additional influence. Distance contributes to attenuation as expected, but its impact is less significant within the tested range. These findings highlight the importance of material selection and antenna integration when deploying LoRa-based HUMS systems in enclosed environments.

A limitation of this study is the controlled indoor environment, which may not fully represent more complex real-world deployment conditions.

## Conclusion

This study demonstrates that enclosure material and antenna configuration significantly influence LoRa communication performance in enclosed environments. While RSSI decreases with distance, enclosure material is the dominant factor with GRP exhibiting the highest attenuation. Antenna configuration also affects signal behaviour, particularly within constrained environments. Despite reductions in signal strength, SNR remains relatively stable highlighting the robustness of LoRa communication.

These findings emphasise the importance of material selection and antenna integration in the design of reliable LoRa-based HUMS deployments. Future work should consider more complex environments, including larger separation distances, non-line-of-sight and outdoor conditions.

## References

[1] M. Islam, H. Jamil, S. Pranto, R. Das, A. Amin, and A. Khan, “Future Industrial Applications: Exploring LPWAN-Driven IoT Protocols,” *Sensors*, vol. 24, no. 8, Apr. 2024.
[2] M. Ballerini, D. Brunelli, M. Magno, and L. Benini, “NB-IoT vs. LoRaWAN: An experimental evaluation for industrial applications,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7802–7811, 2020.
[3] A. I. Griva et al., “LoRa-based IoT network assessment in rural and urban environments considering antenna gain and network parameters,” *Sensors*, vol. 23, no. 3, 2023.
[4] Semtech Corporation, *“LoRa Modulation Basics,”* Application Note AN1200.22, Rev. 2, 2015.

https://doi.org/10.18745/pb.27307

# Evaluating Body Pose Grouping via Clustering Algorithms

**Mahendra Nath Reddy E, Priyanka Babu Rajasree[1] and Paul Moggridge[1*]**

[1]*University of Hertfordshire*

*Corresponding author: p.moggridge@herts.ac.uk

**MediaPipe and OpenPose have demonstrated the easy exaction of body pose data from video streams. While classification models are highly effective at classifying body poses [1], they require labelled data. Clustering offers a potential solution to accelerate the labelling process by grouping the poses in the continuous stream of coordinates into discrete clusters. Thus this leads to our research question. How effective are clustering algorithms (K-Means, DBSCAN, and Agglomerative clustering) for grouping body pose data? Our initial findings suggest that clustering can find discrete poses with minimal data preprocessing.**

Keywords: clustering; body pose; k-means; DBSCAN; agglomerative;

## Introduction

Prior studies have examined clustering algorithms for grouping body pose data through feature extraction and unsupervised learning. K-means clustering has been used to group similar skeletal poses by leveraging feature vectors from joints and angles, reducing redundancy and identifying representative postures [2]. Clustering applied to motion arc features extracted from 3D skeletal data demonstrates the ability to capture structure in human movement and distinguish behavioral patterns across subjects [3]. Recent work combines latent representation with clustering, showing meaningful action groups can be identified using unsupervised methods, with performance dependent on feature representation and algorithm choice, notably hierarchical clustering is shown to outperform K-means and other methods [4]. However, existing studies rely on engineered or latent features without directly evaluating clustering effectiveness on raw pose data. Research on skeleton-based action recognition highlights that pose data contains discriminative spatio-temporal patterns, suggesting that clustering performance can be improved through advanced feature representations like frequency-aware embeddings [5]. Clustering has been used to model intra-class variability in motion data, capturing stylistic differences beyond action categories in motion-derived features [8]. This study systematically evaluates the effectiveness of clustering algorithms for grouping body pose data, proposing an approach for understanding their performance on Indian classical dance.

## Experiment Setup

The Indian Classical Dance (Bharatanatyam) was selected as the primary data source because of its predefined, discrete movement profile and repeatable geometric structures [6], which provide natural "ground-truth" clusters within a continuous coordinate stream [8]. While MediaPipe is computationally efficient, it exhibits a significant limitation known as "Pose Hallucination" [9] which leads to significant noise into the dataset, as the "estimated" points do not reflect the dancer's actual pose in the video.

To mitigate the hallucination errors, frames were discarded if ankle indices' (27 and 28) visibility scores were below 0.65 after manual inspection of the extracted pose sequences and if the normalised vertical distance between the mid-hip to ankles was less than 0.18 of the image height (collapsed skeletons, hallucinations). This reduced the raw dataset from 13,474 to a high-fidelity subset of 4,789 frames.

K-Means was implemented as the primary clustering method due to its scalability and performance as a baseline for unsupervised pose grouping [10]. Before clustering, the coordinates were normalised using a StandardScaler to ensure zero mean and unit variance across the 99-feature vector (33 landmarks × 3 dimensions). The Elbow Method (using the Kneedle algorithm) identifies a distinct inflection point at (k=6) and determines the number of discrete pose clusters as k=6.

https://doi.org/10.18745/pb.27307

Following the K-means baseline, DBSCAN with OPTICS was applied to the StandardScaled coordinates. A reachability distance threshold of 2.4 was selected to produce five distinct clusters, plus one noise cluster, based on the prominent valleys observed in the plot. This distance was specifically chosen to ensure comparability with the K-Means results.

An advantage of DBSCAN is its ability to handle noise and outliers. In this context, the algorithm effectively ignores intermediary poses and transient inaccuracies in the MediaPipe coordinates. Qualitative evaluation of the 3D scatter plots in Fig 2 confirms that DBSCAN successfully identifies unique postures. Notably, clusters C1 and C4 represent similar poses performed at different spatial coordinates within the video frame.

Finally, Agglomerative (hierarchical) clustering was performed on the scaled dataset. This algorithm was trialled as it has been shown to outperform other methods [4]. To maintain consistency with the K-Means elbow point, a distance threshold of 175 was applied. To validate the resulting clusters, Fig 1 shows a t-SNE embedding of the high-dimensional pose vectors in a 2D space where the clusters can be visualised and evaluated subjectively, demonstrating that Agglomerative clustering produces a grouping similar to that of K-Means.

## Results and discussion

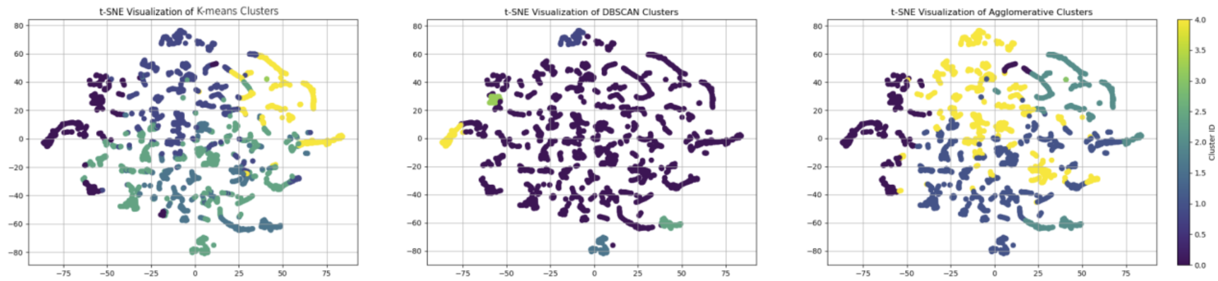



*Fig. 1. Clustering results shown in a 2d t-Distributed Stochastic Neighbour Embedding (t-SNE) with plexity=30. Each data point is a frame in the video and the colour represents the cluster assignment.*

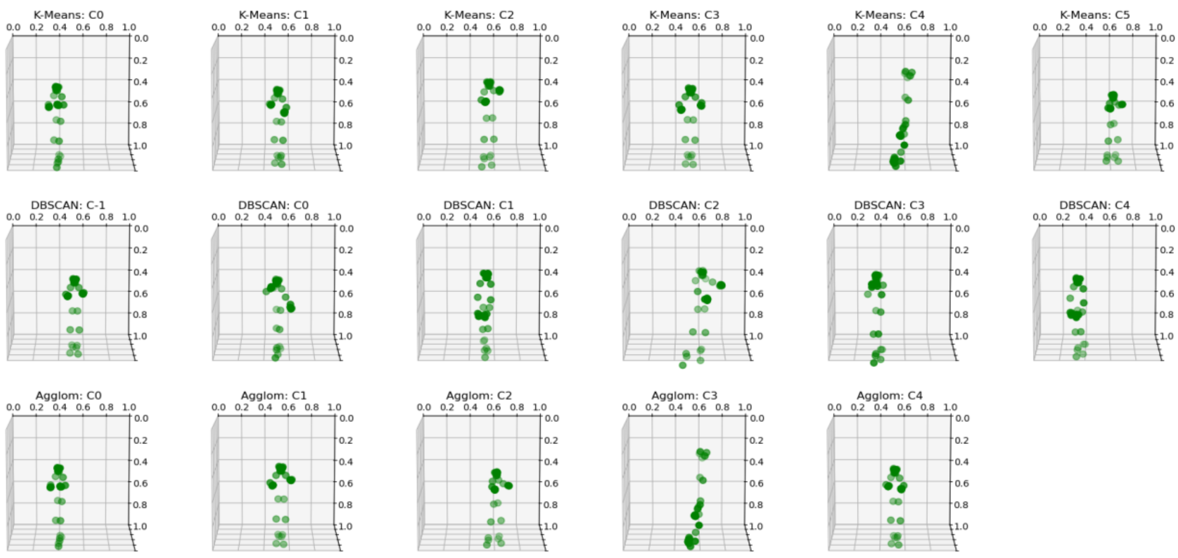



*Fig. 2. The clusters (3d poses) found. (Top) K-Means' clusters 0 to 5. (Middle) DBSCAN clusters' 0 to 4. Noise is denoted as cluster -1. (Bottom) Agglomerative (hierarchical) clusters' 0 to 4.*

https://doi.org/10.18745/pb.27307

## Conclusion

This work demonstrates that K-Means, DBSCAN, and Agglomerative clustering can identify poses in MediaPipe body landmark data from video. DBSCAN was effective at isolating temporary inaccuracies and transitional movements as noise. Our methodology didn't perform spatial normalisation thus similar dance poses in different areas of the frame are considered distinct. Challenges remain in clothing obscuring body landmark detection. Future work includes using motion features to enhance action recognition [3][7]. The work could accelerate the annotation process for training pose classification models for monitoring and fall-detection for the elderly.

## References

[1] J. Jangade and K. S. Babulal, "Study on Deep Learning Models for Human Pose Estimation and its Real Time Application," *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, Mathura, India, 2023, pp. 1-6, doi: 10.1109/ISCON57294.2023.10112004.

[2] S. R. Krishna, M. Pavithra, A. P. A, J. G. Thatipudi, G. D and K. Rajkumar, "MNCL: Empirical Development of Human Action Recognition Based on Postures Using Modified Neural Classification Logic," 2025 International Conference on Visual Analytics and Data Visualization (ICVADV), Tirunelveli, India, 2025, pp. 1299-1306, doi: 10.1109/ICVADV63329.2025.10961249.

[3] C. Qu, R. Chen and X. Che, "Motion Arc Analysis in Virtual Reality Environment," 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Prague, Czech Republic, 2022, pp. 3384-3389, doi: 10.1109/SMC53654.2022.9945281.

[4] W. Vexo, C. Jeenanunta, S. Chanyachatchawan, A. Tunpan and N. Sirimarnkit, "Unsupervised Human Action Recognition for Quality Control in Industrial Environments," 2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 2024, pp. 541-546, doi: 10.1109/IICAIET62352.2024.10730750.

[5] X. Wu, P. Zhang, J. Wu, M. Wang, W. Qin and T. Liu, "Spectral Channel Attention for Skeleton Action Recognition in Lubricant Filling," 2025 6th International Conference on Control, Robotics and Intelligent System (CCRIS), Guangzhou, China, 2025, pp. 1-5, doi: 10.1109/CCRIS65764.2025.11193097.

[6] “Neeli Meghamula lo dance | Abhirami | Bharathanatyam,” YouTube, uploaded by Abhirami, [Online]. Available: https://www.youtube.com/watch?v=n7lwFNWm1a0

[7] R. Rastgoo, K. Kiani and S. Escalera, "Diffusion-Based Continuous Sign Language Generation with Cluster-Specific Fine-Tuning and Motion-Adapted Transformer," 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Nashville, TN, USA, 2025, pp. 4088-4097, doi: 10.1109/CVPRW67362.2025.00393.

[8] T. Mallick, P. P. Das, and A. K. Majumdar, "Posture and sequence recognition for Bharatanatyam dance performances using machine learning approaches," Journal of Visual Communication and Image Representation, vol. 87, p. 103548, 2022.

[9] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint arXiv:1906.08172, 2019.

[10] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297, 1967.

https://doi.org/10.18745/pb.27307

# SaP-AI Framework supporting GenAI E&T Accelerated Courses

**Dr. Muhammad Jamro***

*Dept. of Electrical & Electronics Engineering*

*Corresponding author: M.Jamro@herts.ac.uk

## Abstract:

**This paper proposes a flexible 'Students as Partners' and AI Engagement (SaP-AI) framework for engineering and technology (E&T) education in response to Generative AI. Drawing on degree apprenticeship "day-release" and traditional "block delivery" principles, the model proposes reduced on-campus delivery with AI-supported, cloud-enabled, and Industrialised Internet of Things (IIoT)-driven project-based learning to deliver regular or accelerated E&T degrees. It supports hands-on training and skills development by integrating controlled in-person assessments with structured remote experiential learning. The framework enhances technical competence, employability, and scalable modernisation in post-92, business-facing universities.**

**Keywords:** GenAI, SaP, CDIO, Scrum, Project-based learning, Accelerated courses

## Introduction

The rapid evolution of Generative AI, alongside the anticipated emergence of Artificial General Intelligence (AGI) by next decade is reshaping learning, teaching, and assessment in higher education [1]. While pandemic-driven online delivery improved accessibility, it exposed vulnerabilities in academic integrity due to AI-enabled misconduct. During COVID-19, the UH-LTIC (now CEdSS) contributed to the EU-funded OP4RE project, implementing ProctorU remote invigilation for home-based examinations. In parallel, many UK universities distributed engineering kits to support remote practical learning [2]. Such mechanisms suit disruptive, AI-augmented delivery; however, the delivery of future AI-supported courses must be SMART (specific, measurable, achievable, reasonable, and time-bound). Thus, an accelerated course style is strongly advised for future popular courses to improve learners' engagement and productivity. As AI systems increasingly replicate cognitive outputs, traditional assessment models lose validity, shifting emphasis toward synthesis and human-centric competencies [3]. AI grading tools (e.g., AI-Grader, SmartGradr, SmartEducator, EduSageAI, Gradescope (by Turnitin), Pensieve Grader) can refine submissions through iterative feedback aligned with rubrics, but overreliance risks performance-driven outputs that do not reflect learning outcomes. In response, universities have reintroduced controlled on-campus assessments such as exams and oral viva, though these are resource-intensive increasing staff-workload (i.e., submitting scanned, recorded files as evidence, and completing various MS Forms). The administrative and management system must adopt professional, AI-based, and smart automated solutions that meet AQA compliance while reducing the time demands on academic and technical staff.

The hybrid models using project-based learning (PjBL), remote laboratories, and cloud platforms show potential, although some practical activities still require on-campus "block" or "day-release" delivery [4]. Students are positioned as co-creators and partners rather than customers through the Students as Partners (SaP) approach, promoting collaborative engagement in learning and curriculum design. For example, aligned with this, many post-92 UK's universities, including UH, offer optional in-module retrieval (IMR), allowing failed students to retake assessments before the exam board; it is capped to a pass and does not improve existing grades but supports vulnerable students, plan ahead, accordingly. The IMR is generally acceptable to accreditation bodies (e.g., UK Engineering Council, IET, IMechE, and ABET etc), provided standards are maintained. Ironically, post-92 universities emerged from vocational, work-based colleges, and now the industry is moving towards skills-based degrees and professional certificate courses; hence, E&T education providers must develop their courses by revisiting their past successes.

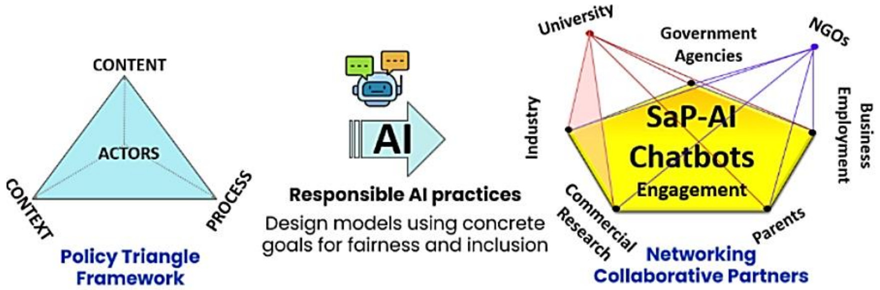
## Students as Partner and AI-Engagement (SaP-AI) Framework for E&T Education

The SaP-AI framework addresses existing challenges through flexible, technology-enabled pedagogies and integrity-driven assessment. It supports hybrid delivery models aligned with industry needs and digital transformation priorities. Central to the framework is the constructive integration of GenAI, where students-as-partners co-create curricula and interdisciplinary programmes within project-based learning environments [4, 5]. This approach can reduce programme delivery time, as active learners engage with relevant free online courses gaining lifelong skills and nationally recognised, accredited qualifications, gaining lifelong skills alongside

https://doi.org/10.18745/pb.27307

nationally recognised, accredited qualifications. However, effective implementation requires strong collaboration among academia, industry, policymakers, professional bodies, and other key stakeholders.

The proposed framework adopts a scalable networking model, enabling multiple stakeholders to establish strategic partnerships, as illustrated in Figure 1. These interactions can be structured through a policy framework encompassing actors, context, content, and process, thereby supporting continuous programme development and revision




| | |
|---|---|
| ACTORS | Policymakers, individuals, groups, whose interests? Whose idea? other stakeholders. |
| CONTENT | Ideas for Policy, Type, Scale of policy, objectives, planning, gaps, etc |
| PROCESS | Strategies, agendas, policy formulation, implementation, evaluation, etc |
| CONTEXT | Situational, structural, cultural, environmental, globalisation, sustainability, business conditions, limitations, vulnerabilities, etc. |

**Figure 1** Strategic Networking Collaborative Partnerships for SaP and GenAI Engagement

The proposed framework can be applied to develop accelerated courses as given in the Figure 2. It presents a modular nine-month accelerated degree structure, where each module is delivered in intensive eight-week blocks combining online, face-to-face, and project-based learning. The first week is for induction as block delivery, followed by six weeks of learning and teaching, one week for assessment, and one week for the project. Greater staff input is required for developing continuous professional development (CPD) based projects or training workshop, enabled by GenAI to enhance learning efficiency and support continuous progression towards an AI-augmented future. This transformative style accelerated course a way forward, and aligns with the growing relevance of executive education, bootcamps, and industry-aligned E&T certification programmes (e.g., Cisco, Wireshark, AWS, MS Learn, Google, Micro-credential).

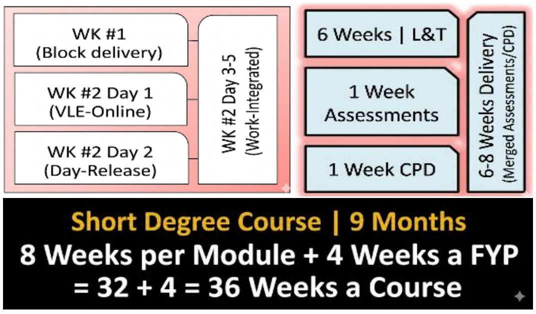



**Figure 2** Short (accelerated) Degree Course: Duration 9 months (4 Core modules + Major Project).

## Proposed Design and Delivery of SaP-AI Collaborative Accelerated Courses

Degree apprenticeships in E&T use only 80% and 20% of their time for employment and education provision, respectively. Mostly, they attend university for just one day a week. The UH-academic quality assurance (AQA) following the best delivery options, make this model a success for active learners. For example, the “day-release” demonstrates the value of integrating study with applied experience, while “block delivery” and “flipped classroom” approaches optimise use of resources. The SaP-AI framework builds on these by shifting from attendance-based to outcomes-based delivery that supports flexible engagement and employability [4, 6-8]. The learners will develop their viable ‘Courses on Demand’ for sustainable employability. The future-proof courses must be short, accelerated (duration: 9-12 months), focussing on sector-wise competence and employers’ demands. The proposed model adopts a hybrid structure in which students attend campus one to two days per week for laboratories, workshops, and controlled assessments, after completing intense induction week. At the end of the proposed short degree course, an AI-supported investigation and commercially viable individual project-work ensures the qualitative contribution to research. Curriculum design separates knowledge acquisition from application, where the theoretical content is delivered digitally, while in-person sessions focus

https://doi.org/10.18745/pb.27307

on experiential learning. A hybrid Scrum-CDIO project-based learning (PjBL) approach forms a core component, supported by cloud platforms and IIoT technologies for remote experimentation and collaboration building teamwork, which is essential requirement of an industry business [2, 7, 8]. Assessment emphasises authenticity through controlled in-person evaluations and structured remote tasks, including project design and reflective portfolios. Process-based evaluation, iterative submissions, and oral defence mechanisms support academic integrity in AI-rich environments. Framework advocates effective staff–student resources with balanced workload, including AI-automated rubric-based marking, feedback, reducing double-blind marking and manual evidence uploads [9, 10]. The streamlined quality processes, reduces various board meetings, Templates, replacing with AI-supported dashboards and assessment and moderation.

**Work-Integrated and Applied Learning:** The UK Government and education providers have invested heavily in Degree Apprenticeships, and should offer equivalent support to traditional learners, ensuring inclusive education in line with EDI principles and the Equality Act 2010. The proposed framework enables students to combine study with employment, internships, or entrepreneurship. Where aligned with programme outcomes, these are assessed through structured reflection, enhancing employability and industry alignment. Inclusivity is supported through digital tools and academic guidance. Effective implementation requires resource developing and stakeholders' engagement, synchronising with hierarchical managers, learning technologists, academic, technical staff, and Students' Union.

**Conclusion:** The SaP-AI Engagement model provides a strategic response to AGI-driven transformation in higher education. By integrating apprenticeship-inspired delivery with AI-enabled and project-based learning, it offers a flexible, resilient, and industry-aligned framework, thereby enabling accelerated degrees in E&T disciplines. The model identifies resource challenges and the need for network development with stakeholders, while preparing graduates for an evolving professional landscape.

## References

[1] M. Jamro and F. Jamro, *AI Matters: Chatbot Integration in Holistic STEM Education*. 2024, ISBN: 978-1-4461-7585-9.
[2] M. Jamro, S. Xiao, A. Hutton-Mckenzie, and A. Yip, "Mini-Projects: Day-release Practical Lab Approach for STEM related disciplines for the 'new normal' Education System," *UH J. BliP*, Autumn 2021
[3] Q. Xia, X. Weng, F. Ouyang, T. J. Lin, and T. K. F. Chiu, "A scoping review on how generative artificial intelligence transforms assessment in higher education," *Int. J. Educ. Technol. High. Educ.*, vol. 21, no. 1, p. 40, 2024.
[4] J. Weldon and L. Konjarski, *Block Teaching Essentials: A Practical Guide*. Springer, 2024, ISBN: 9819774705.
[5] A. Tlili, B. Shehata, M. A. Adarkwah, A. Bozkurt, D. T. Hickey, R. Huang, and B. Agyemang, "What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education," *Smart Learning Environments*, vol. 10, no. 1, p. 15, 2023.
[6] W. F. Yu, "AI as a co-creator and a design material: Transforming the design process," *Design Studies*, vol. 97, no. 2, Art. no. 101303, 2025.
[7] M. Jamro, "µ-Projects based on Hybrid CDIO-Scrum Methodology for STEM business," in *Proc. UH-SPECS Conf.*, 2023.
[8] M. Jamro, "IoT Security with QoS: Game changer for Industry and STEM Education," in *Proc. IEEE ICCST*, 2021, pp. 1–4.
[9] M. Wang, M. Wang, X. Xu, L. Yang, D. Cai, and M. Yin, "Unleashing ChatGPT's Power: A Case Study on Optimizing Information Retrieval in Flipped Classrooms via Prompt Engineering," *IEEE Trans. Learn. Technol.*, vol. PP, no. 99, pp. 1–13, 2023.
[10] B. B. Tomić, A. D. Kijevčanin, Z. V. Ševarac, and J. M. Jovanović, "An AI-based approach for grading students' collaboration," *IEEE Trans. Learning Technol.*, vol. 16, no. 3, pp. 292–305, 2022.

https://doi.org/10.18745/pb.27307

# On The Present And Future Changes In Clear Air Turbulence In A Changing Climate

**Anukrishna Pullanhiyodan[1*,2], P.R. Tiwari[1,2], Boyan Tatarov[1,2] and Detlef Mueller[1,2]**
[1]*Centre for Atmospheric and Climate Physics Research, Department of Physics, Astronomy and Mathematics, SPECS, University of Hertfordshire, Hatfield, United Kingdom AL10 9AB*
[2]*Centre for Climate Change Research, Department of Physics, Astronomy and Mathematics, SPECS, University of Hertfordshire, Hatfield, United Kingdom AL10 9AB*
*Corresponding author: a.pullanhiyodan@herts.ac.uk

**Clear-air turbulence (CAT) is a persistent aviation hazard, yet long-term changes in its dynamical drivers remain poorly understood. We quantify wind shear changes over 85-year period (1940–2025) using ERA5 reanalysis, relating these to multidecadal jet-stream shifts. We evaluated CMIP6 models against ERA5 for robustness, then examined future wind shear changes under SSP2-4.5 and SSP5-8.5 scenarios, emphasizing turbulence-prone international flight corridors. Our results reveal a significant increase in the Ellrod index magnitude along North Atlantic flight corridors with T1 index values increasing by up to 30 × $10^{-9}$ $s^{-2}$ under SSP5-8.5 relative to pre-industrial conditions. This analysis provides a physically grounded basis for characterizing emerging turbulence risk profiles as the climate continues to change.**

Keywords: Aviation turbulence, Jet stream, Wind shear, Climate change

## Introduction

Atmospheric turbulence at aircraft cruising altitudes is an important hazard in aviation industry worldwide [1]. There are different kinds of turbulence that impact aviation, like convective turbulence which occurs near clouds but can easily detected by pilots using RADAR [2]. In contrast, CAT is difficult to be detected by remote sensing techniques. This is primarily because CAT is caused by vertical wind shear associated with the jet stream in cloud free regions [3],[4],[5]. Recent studies indicate climate change have a larger impact on North Atlantic jet stream, where increasing vertical shear is consistent with intensification of shear driven CAT [1]. To quantify these risks, this research utilizes the Ellrod T1 index [7], a diagnostic tool for CAT that combines vertical wind shear and deformation to identify regions of potential atmospheric instability.

The scope of this research involves a multi-decadal analysis of these shear instabilities using high-resolution data to quantify historical shifts and project future risks. By examining the relationship between jet stream strength and shear instability, this study aims to provide a clearer picture of how flight corridors are evolving under a warming climate. The primary objectives of this work are to evaluate long-term trends in vertical wind shear and CAT frequency using the ERA5 reanalysis dataset from 1940 to 2025, while simultaneously assessing the robustness of CMIP6 (Coupled Model Intercomparison Project Phase 6) [8] models in capturing these dynamical changes compared to historical observations. By applying the Ellrod T1 index to both reanalysis and model data, this study seeks to characterize the emerging risk profile under different climate scenarios, such as SSP2-4.5 and SSP5-8.5 [8], to provide a physically grounded basis for future aviation safety and flight planning.

## Experimental Framework

The methodology of this study uses ERA5 and CMIP6 models to evaluate atmospheric turbulence. Historical atmospheric conditions were characterized using the ERA5 reanalysis dataset, spanning 85-year period from 1940 to 2025. The analysis focused on key meteorological variables, specifically air temperature (ta) and wind components zonal wind (ua) and meridional wind (va), to calculate the Ellrod T1 index and vertical wind shear (VWS) across the historical record and future emission scenarios. To assess future trends and model reliability, a suite of state-of-the-art CMIP6 models (CESM2-WACCM, CNRM-CM6-1, CanESM5, EC-Earth3, FGOALS-F3-L, IITM-ESM, IPSL-CM6A-LR, KACE-1-0-G, MIROC6, MPI-ESM1-2-HR, MRI-ESM2-0, GFDL-CM4, NorESM2-MM, TAIESM1, and UKESM1-0-LL) was employed. These models were evaluated against the ERA5 baseline to ensure

https://doi.org/10.18745/pb.27307

their robustness in simulating shear instability. Future projections were examined under two distinct Shared Socioeconomic Pathways: SSP2-4.5 (intermediate scenario) and SSP5-8.5 (high-emission).

**Zonal Wind Shear in the zonal wind:**

We define the vertical shear magnitude of the zonal wind (u) between the 200 hPa and 300 hPa as [1]:

$$\left|\frac{du}{dp}\right| = \frac{u(200\ hPa) - u(300\ hPa)}{100\ hPa} \qquad \text{(i)}$$

where the result is expressed in m s$^{-1}$ hPa$^{-1}$. The 200–300 hPa pressure layer was selected as it corresponds to typical aircraft cruising altitudes (~10–12 km) and encompasses the core of the upper-tropospheric jet stream, where shear-driven CAT is most prevalent.

**Ellrod T1 Index:**

It is defined as the product of total vertical shear magnitude and horizontal kinematic properties of the flow [7]:

$$T1 = VWS * DEF \qquad \text{(ii)}$$

where T1 has units of s$^{-2}$, $DEF$ is the total deformation and $VWS$ is the total vertical shear magnitude converted to height coordinates using the Hypsometric equation

## Results and discussion

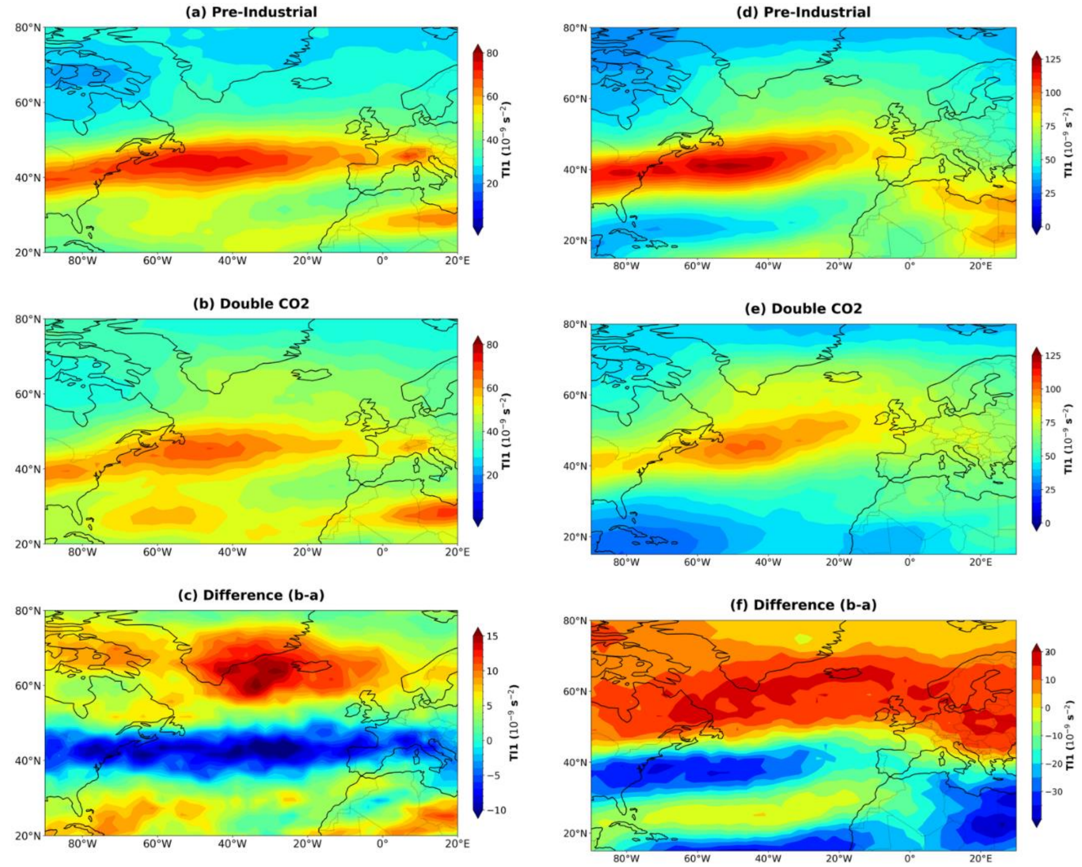



Figure. 1 (a) to (c) is Ellrod T1 index of GFDL-CM2.1 CMIP3 model at 200 hPa which serves as baseline of previous studies [6]. It shows spatial patterns of North Atlantic flight-level winter clear air turbulence in a changing climate. The higher values of index indicate more clear air turbulence. Figure. 1 (d) to (f) is GFDL-CM4 CMIP6

*Fig. 1 shows the median of Ellrod T1 Index from GFDL CM2.1 (CMIP3) on left and GFDL-CM4 (CMIP6) on right. The top two panels (a), (d) and (b), (c) are from pre-industrial and doubled $CO_2$ integrations, and the bottom panel (c), (f) is the difference.*

strengthening of turbulence-prone regions, particularly between 50°N-70°N. The observed northward shift and strengthening of turbulence-prone regions are statistically significant at the 95% confidence level, assessed using a Welch's t-test applied to the ensemble of 14 CMIP6 models. This is consistent with previous findings [1] suggesting that a warming climate enhances the vertical wind shear in the upper-level jet stream. While the full ensemble of 14 CMIP6 models was evaluated for consistency, the GFDL-CM4 model is presented here as a representative high-resolution example to illustrate the characteristic spatial shifts and intensification of the T1 index.

Both model generations identify the North Atlantic flight corridor as a high-risk zone. But CMIP6 data suggests a more significant increase in the T1 magnitude (reaching differences up to $30 * 10^{-9}s^{-2}$) compared to the CMIP3 estimates. The observed increase in T1 index is linked to the strengthening of meridional temperature gradient in the upper troposphere and lower stratosphere [1]. This in turn accelerates the jet stream and increases the shear induced CAT.

## Conclusion

Our study shows how the clear air turbulence evolved from pre-industrial to double $CO_2$ integrations. The results indicate that the frequency and intensity of CAT increases significantly on North Atlantic flight corridors. Our ongoing work uses multi-model CMIP6 ensemble to evaluate the consistency of projected turbulence trends across various climate scenarios.

## References

[1] Lee, S. H., Williams, P. D., & Nicoll, K. A. (2019). Increased shear in the North Atlantic upper-level jet stream over the past four decades. Nature, 572(7771), 639-642.

[2] Storer, L. N., Williams, P. D., & Gill, P. G. (2019). Aviation Turbulence: Dynamics, Forecasting, and Response to Climate Change. Oxford Research Encyclopaedia of Climate Science

[3] Storer, L. N., Williams, P. D., & Joshi, M. M. (2017). Global response of clear-air turbulence to climate change. Geophysical Research Letters, 44(19), 9976-9984.

[4] Prosser, M. C., Williams, P. D., Marlton, G. J., & Harrison, R. G. (2023). Evidence for large increases in clear-air turbulence over the past four decades. Geophysical Research Letters, 50(11), e2023GL103814.

[5] Foudad, M., Sanchez-Gomez, E., Jaravel, T., Rochoux, M. C., & Terray, L. (2024). Past and future trends in clear-air turbulence over the northern hemisphere. Journal of Geophysical Research: Atmospheres, 129, e2023JD040261.

[6] Williams, P. D., & Joshi, M. M. (2013). Intensification of winter transatlantic aviation turbulence in a warming climate. Nature Climate Change, 3(7), 644-648.

[7] Ellrod, G. P. & Knapp, D. L. (1992). An objective clear-air turbulence forecasting technique: Verification and operational use. Weather Forecast. 7, 150–165.

[8] Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958.

https://doi.org/10.18745/pb.27307

# Precedent-Driven Liability in Autonomous Mobility: From Self-Driving Cars to Advanced Air Mobility with Unmanned Air Systems

**Dr Ivan Sikora [1], Dr Ori Igwe [2*]**
[1]*University of Hertfordshire*
[2]*University of West London*
*Corresponding author: Ori.Igwe@uwl.ac.uk

**Abstract**
**This research examines how legal precedents from autonomous ground vehicle litigation, including Tesla Autopilot crash cases[4] and the 2018 Uber Tempe[5] fatal collision, can inform liability and regulatory frameworks for Advanced Air Mobility (AAM) using autonomous Unmanned Air Systems (UAS). Through doctrinal analysis of tort and product liability principles, the study identifies transferable standards for fault allocation, operator oversight, and data-logging obligations. It argues that aviation regulators must adapt these precedents to address three-dimensional airspace risks, remote-piloting accountability, and cyber-command vulnerabilities, thereby supporting harmonised safety certification, coherent liability attribution, and robust incident-response frameworks for the emerging autonomous aerial mobility ecosystem.**
**Keywords**: liability; regulatory framework; Advanced Air Mobility (AAM); autonomous Unmanned Air Systems (UAS); tort

## Introduction

### 1. Context of Work

Autonomous mobility faces legal issues with frameworks struggling to assign fault [1]. Complex socio-technical systems challenge control and responsibility [2]; [3]. Scholars call for new liability models based on system accountability. Liability in ground vehicles is complicated by programming and multiple responsible entities, as seen in Tesla Autopilot and a 2018 Uber crashes. In 2023, a Los Angeles court found Tesla not at fault, citing manufacturer warnings and user duty. However, in 2025, Tesla was found to be one-third liable, with $243 million in damages over safety claims, highlighting the overstatement of safety. The legal frameworks for AAM are underdeveloped and require reform to support responsible growth. The March 2024 Government Accountability Office report[6] states that current airspace rules, designed for piloted aircraft, may be insufficient for autonomous aircraft, causing gaps in negligence standards, fault attribution, and operator accountability. Scholarship shows a gap between fast tech advances and slow legal adaptation. Liability laws struggle to address distributed responsibility in autonomous flights; fragmented international standards and a lack of governance models exacerbate the issues. Without reforms, AAM risks legal fragmentation and safety problems.

---

[4]*Hsu v Tesla, Inc* (Superior Court of California, April 2023) reported in:
'Tesla wins Autopilot crash case in California' (TechCrunch, 21 April 2023) https://techcrunch.com/2023/04/21/tesla-wins-autopilot-crash-case-in-california/ accessed 1 May 2026.

Canellas, M. & Haga, R. (2020). *"Unsafe At Any Level: NHTSA's Levels of Automation Are a Liability for Autonomous Vehicle Design and Regulation."* Communications of the ACM, Vol. 63, No. 3, pp. 31–34. DOI: 10.1145/3342102
Yacoub, A. R. (2022). *"Liability and Regulatory Oversight of Semi-Autonomous and Autonomous Vehicles."* Boston University Law Review.
'Tesla found partly to blame for fatal Autopilot crash' (BBC 2023)
https://www.bbc.co.uk/news/articles/c93dqpkwx4xo

[5] David Shepardson, 'Backup driver in 2018 Uber self-driving crash pleads guilty' (Reuters, 28 July 2023)
https://www.reuters.com/business/autos-transportation/backup-driver-2018-uber-self-driving-crash-pleads-guilty-2023-07-28/ accessed 1 May 2026.
'Backup driver in fatal Uber autonomous car crash pleads guilty to endangerment' (CBS News, 28 July 2023)
https://www.cbsnews.com/sanfrancisco/news/uber-autonomous-self-driving-fatal-car-crash-backup-driver-pleads-guilty-endangerment-tempe-arizona/ accessed 1 May 2026.

[6] US Government Accountability Office (2024). *Advanced Air Mobility: Legal Authorities and Issues to Consider for Operations* (GAO-24-106451).

https://doi.org/10.18745/pb.27307

In human-machine collaboration, the AI "black box' and decision complexity make accident probabilities unpredictable and prevention costs hard to internalise, weakening incentives through extended decision chains and multiple liable parties [4]. Opacity increases in AAM, where autonomous flight may occur without human supervision during Beyond Visual Line of Sight (BVLOS) missions, creating accountability gaps that current tort frameworks can not address without change. These risks [5] emphasise the need for strong governance with ethical safeguards, cybersecurity resilience, and strict regulation to ensure safe AI use in aviation. Cybercriminals can exploit vulnerabilities in airport systems, air traffic control, data links, Wi-Fi, satellite navigation, and aircraft [6].

Integrating AI into transportation raises accountability, ethical concerns about safety, security, and governance [7]. Autonomous AI systems may reduce human oversight, thereby increasing the risk of malfunctions, accidents, and fatalities. Cyber threats target aviation infrastructure, including aircraft, air traffic control, airports, and entertainment systems, exploiting unpatched software, legacy vulnerabilities, and data manipulation [8]. Legally, AI decision-making can lead to accidental deaths, with questions of accountability because blame may fall on manufacturers, programmers, operators, or retailers. This research examines these legal issues, suggesting that lessons from autonomous ground-vehicle law can inform the development of liability and certification frameworks for autonomous aerial mobility, offering a clearer, more accountable regulatory approach than isolated doctrinal development.

## 2. Scope

This research focuses on autonomous mobility systems in mixed-traffic environments involving ground-based autonomous vehicles and airborne UAS, where fault attribution, liability, and operator accountability are either judicially established or regulated. It primarily examines Anglo-European legal systems, particularly case law and regulations from the US, the UK, and the EU. Key legislative instruments include the UK Automated Vehicles Act 2024, which shields the "user-in-charge" from criminal liability when in self-driving mode, holding manufacturers or developers liable instead, along with EASA's 2024 Machine Learning Concept Papers, EU U-space Regulation 2021/664, UK CAA's framework, FAA's Part 107, and the FAA AAM Plan 2023. The judicial review covers 2018-2026, with a focus on 2023-2026, the most active period for autonomous vehicle liability cases.

## 3. Research Objectives

The rapid deployment of autonomous vehicles, from SAE Level 4 self-driving cars to advanced UAS-enabled urban air mobility, has outpaced the development of consistent legal doctrine. Regulators are addressing ground-vehicle autonomy through liability, safety certification, and data-recording rules, thereby creating case law and guidance that are valuable for AAM-UAS applications. Meanwhile, AAM moves toward commercial operations such as air taxis, cargo drones, and BVLOS UAS fleets, raising questions about how traditional legal concepts apply in complex 3D airspace. This research focuses on autonomous mobility in mixed-traffic environments, where laws governing fault, liability, and responsibility are established or under negotiation. It emphasises Anglo-European civil and common law jurisdictions, recent case law, regulatory instruments on autonomous driving, and emerging policies from aviation safety authorities, including EASA, the UK CAA, and the US FAA.

The objectives are threefold:

**Objective 1:** To identify, systematise, and critically evaluate the key legal precedents and doctrinal developments arising from autonomous ground-vehicle litigation, and to assess their applicability to the construction of liability frameworks for Advanced Air Mobility (AAM) and autonomous Unmanned Aerial Systems (UAS).

**Objective 2:** To examine how existing legal doctrines and precedents can inform the development of aviation-specific regulatory frameworks governing AAM-UAS operations.

**Objective 3 :** To develop a theoretically grounded and practically applicable framework for liability and regulatory governance in AAM-UAS systems, integrating legal doctrine, aviation safety regulation, and cybersecurity resilience.

https://doi.org/10.18745/pb.27307

## Experimental/Simulation

This study evaluates the transfer of liability principles from autonomous ground vehicles to AAM systems using UAS. It combines doctrinal analysis [9] with legal simulation, analysing laws, case law, and safety guidance to find convergence in fault, responsibility, and safety mechanisms. A simulation framework models hypothetical UAS incidents by drawing on autonomous vehicle precedents, replicating variations in autonomy, operator involvement, and system failures. Outputs are evaluated against legal liability doctrines to assess consistency, gaps, and adaptation options for aviation law.

## Results and discussion

The analysis demonstrates significant doctrinal overlap between autonomous ground vehicle jurisprudence and emerging AAM regulatory challenges, particularly regarding product liability, negligence, and system-level accountability. Case-based simulation results indicate that liability allocation becomes increasingly manufacturer-centric as autonomy levels increase, reinforcing the need for system-based responsibility models in aviation contexts.

Table 1 summarises key doctrinal principles derived from autonomous vehicle case law and their applicability to AAM governance, including failure-to-warn obligations, design defect reasoning, and post-deployment software update duties. The findings suggest that existing aviation regulatory structures are insufficiently equipped to accommodate distributed autonomy in airspace operations. In particular, the absence of a clear equivalent to automotive “user-in-charge” frameworks in UAS regulation creates uncertainty about incident attribution and enforcement. However, the convergence of legal reasoning across jurisdictions indicates a strong foundation for developing harmonised, cross-domain liability standards that integrate both terrestrial and aerial autonomous systems.

### Table 1: Principles from Autonomous Vehicle Case Law and Their Applicability to AAM Governance

| Doctrinal Principle | Autonomous Vehicle Application | AAM/UAS Application | Regulatory Implication for Aviation |
|---|---|---|---|
| Failure to Warn | Manufacturers must communicate system limitations | Operators must understand UAS autonomy limits | Mandatory disclosure standards in certification |
| Design Defect | Unsafe design under foreseeable use leads to liability | Faulty detect-and-avoid or navigation systems | Loss Of Certification, Operational Grounding, or Re-Regulation Of Permissible Autonomy Levels |
| Post-Deployment Software Updates | Duty to fix defects via updates | Continuous UAS software and cybersecurity updates | Binding post-certification update obligations |
| Comparative Fault | Shared liability between driver and manufacturer | Shared liability between pilot, operator, and developer | Clear fault allocation frameworks are needed |
| System-Level Responsibility | Shift toward manufacturer/software liability | Greater responsibility for UAS autonomy providers | Supports authorised system operator models |
| Data Logging and Evidence | Use of event data recorders | Flight data and command logs essential | Mandatory aviation-grade data recording |
| Foreseeability of Misuse | Manufacturers anticipate user overreliance | Design must account for operator misuse | Human factors integration in certification |

https://doi.org/10.18745/pb.27307

## Conclusion

Autonomous ground and aerial vehicles together with AI in transport presents opportunities and risks, necessitating strong regulation to address ethical, legal, and operational issues. Cybersecurity threats are critical, as malicious actors could exploit autonomous systems to cause harm. Judicial and legislative actions are vital to deter misconduct and define liability for operators, manufacturers, and developers. Judicial rulings can clarify liability and support regulation. Effective regulation must balance benefits like efficiency and safety with responsible, ethical use. It should evolve with AI developments, especially in light of societal concerns about superintelligent systems. However, regulation should prioritise current, proven risks such as cybersecurity, system reliability, and safety in autonomous transport.

## References

[1] R. Abbott, "The reasonable computer: Disrupting the paradigm of tort liability," *Georgetown Law Journal*, vol. 108, no. 1, pp. 1–44, 2020. **https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2877380**
[2] R. Calo, "Robotics and the lessons of cyberlaw," *California Law Review*, vol. 103, no. 3, pp. 513–563, 2015. https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=1022&context=faculty-articles
[3] G. Wagner, "Liability for artifcial intelligence and autonomous systems," *Journal of European Tort Law*, vol. 11, no. 2, pp. 156–176, 2020. https://share.google/gt8CTAZZAnt1bbsLY
[4] X. Tian, "How should autonomous vehicles allocate accident liability? Rethinking the application of the Hand formula," *Asian Journal of Law and Economics*, vol. 16, no. 3, 2025. DOI: https://doi.org/10.1515/ajle-2025-0037
[5] Elbasyouny and O. Dababneh, "Feasibility and challenges of pilotless passenger aircraft: Technological, regulatory, and societal perspectives," *Future Transportation*, vol. 6, no. 1, p. 3, 2026, doi: 10.3390/futuretransp6010003.
[6] I. Sikora and O. Igwe, "AI and Aviation (Crewed and Uncrewed): An Analysis of Cyber Security Vulnerabilities", *ASLT 2024*. Rzsezsow: University of Rzsezsow, 2024.
[7] O. Igwe, "Artificial intelligence: A twenty first century international regulatory challenge," *Athens Journal of Law*, vol. 10, no. 4, pp. 737–764, Oct. 2024, doi: 10.30958/ajl.10-4-19.
[8] S. Ishtiaq and N. A. A. Rahman, "Cyber security vulnerabilities and defence techniques in the aviation industry," in *Atlantis Insights in Computer Sciences: Proc. 3rd Int. Conf. on Integrated Computing, Communication & Security*, 2021. [Online]. Available: https://www.researchgate.net/publication/3749851
[9] T. Hutchinson, "Doctrinal research: Researching the jury," in *Research Methods for Law*, M. McConville and W. H. Chui, Eds., 2nd ed. Edinburgh, UK: Edinburgh University Press, 2017

https://doi.org/10.18745/pb.27307

# Modelling High Intensity Focussed Ultrasound For Breast Tumour Ablation

**Oziegbe.E.Onokhua[1*], Yi Sun[1], Nada Yousif[1]**
[1]*University of Hertfordshire, School of Physics, Engineering and Computer Science*
*Corresponding author: oziegbeenduranceonokhua@gmail.com

**Breast cancer remains a leading cause of death among women. Conventional invasive and minimally invasive treatments may result in pain, scarring, and complications. High-Intensity Focused Ultrasound (HIFU) offers a non-invasive alternative which may precisely targets tumours while sparing healthy tissue. However, maintaining thermal energy stability and compensating for breast boundary irregularities due to patient movement remain key challenges. This study develops a conservative linear wave solver integrated with a k-space pseudospectra method to solve coupled PDEs equivalent to the Westervelt equation. Thermal effects are modelled using Pennes and dual-phase lag equations, with tumour destruction quantified via the Arrhenius integral.**

Keywords: Breast cancer, High intensity focussed ultrasound (HIFU), Westervelt equation, Pennes bio-heat transfer equation (BHTE), Dual-phase lag (DPL) model, Anatomical realistic breast phantom (ARBP)

## Introduction

Over the years, breast cancer has been the leading cause of death among women. Without a complete understanding regarding the reasons behind the overall causes of breast cancer, Vaidya and Joseph [1] put forward various risks associated with the development of breast cancer. However, the descriptions given in [1] can only provide an insight towards breast cancer risks. In 2010 in the UK, a total number of 49,564 cases of breast cancer among women and 397 among men were reported in [1], with the number of deaths resulting from these cases among the diagnosed women and men given as 11,556 and 77 respectively. One study [2] predicts there could be 69,900 projected cases of breast cancer every year by 2038-2040.

The current or previous treatment methods for breast cancer involve biopsy which is an invasive surgical procedure and may be traumatic to patients. A minimally invasive technique for predicting the overall pathological response of breast cancer via core needle biopsy (CNB), under the guidance of ultrasound has been reported in [3]. The diagnostic prediction of breast cancer involving high-grade carcinoma from core biopsy has been studied in [4]. These methods are described as invasive and minimally invasive and therefore highlight the need for a non-invasive method.

High intensity focussed ultrasound (HIFU) uses focused energy to treat disease without the need for surgical excision or tissue opening to reach the targeted region (or region of interest). HIFU had its first potential clinical application in the 1950s [5, 6]. It successfully achieved its first therapeutic application when [7] designed some special features of the HIFU system for the treatment of psychiatric disorders. It was clinically applied for the treatment of breast cancer and thyroid cancer in 1964 [8]. One of the major challenges facing the clinical application of HIFU is the heating control. Dubinsky et al [9], emphasised that the heating control mechanism of HIFU is one of the drawbacks of its clinical application. Dubinsky et al [9] further stressed that uniform heating and temperature maintenance in the tumour volume was not adequate at temperatures greater than $43^0$C.

In recent research, Rapheima et al [10] modelled HIFU for pulse wave propagation through an anatomically realistic breast phantom for the ablation of breast cancer. However, their methods suffered several setbacks. The system modelled could not account for 1) even temperature distribution through tissue, 2) coagulative tissue necrosis (thermal dose adoption), 3) heat and temperature control (dual phase lag) and the accountability of the heterogeneity of tissue structure. Hence, in this study we set out to examine how a 3D HIFU system can be optimised to deliver uniform acoustic energy onto region of interest (tumour). We present a preliminary study where we have developed a 3D HIFU model capable of achieving uniform acoustic energy and controlled thermal distribution within an anatomically realistic breast phantom representing tumour structures. This model

https://doi.org/10.18745/pb.27307

can quantify tissue necrosis using thermal dose metrics and Arrhenius-based damage modelling for accurate prediction of ablation efficacy.

## Experimental/Simulation

Ultrasound simulation of biological tissues requires tissue segmentation and careful modelling of the transducer interface. More so, it heavily relies on the mathematical models and the computational domain used. In this study we have used the anatomically realistic breast phantom freely available for breast cancer research in the UWCEM repository [11]. We have considered three layers such as fibroconnective, glandular, and fatty layers of grid points (Nx = 292, Ny = 264, and Nz = 180), provided and developed in MATLAB by [12] for our preliminary HIFU study. The transducer geometry used in our simulation was modelled and designed to be consistent with those used in a previous study [10]. This was necessary to achieve a tight focus which is expected to produce focal intensity in a clinically reliable zone. The computation and communication hurdle expected when simulating a complex system such as the breast was resolved by using k-space pseudospectra method [13] via the k-wave toolbox [14] which can solve a set of three coupled PDEs equivalent to the generalised form of the Westervelt equation.

The first part of our simulation uses a conservative linear homogenous form [15] for pressure wave propagation and generation as well as an embedded power law absorption term [16]. The sound intensity produced by the acoustic vibration due to the pressure field, density and sound speed and of which heat $Q_s$is generated during absorption can be determined according to [18]. Finally, the heat generated by the acoustic energy is distributed in the acoustic field by using the Pennes bio-heat transfer equation [19].

The effect of the temperature rise during thermal ablation of tumour and the associated damage to tumour can be accounted for by implementing the Arrhenius integral [20]. To characterise the degree of damage incurred by tumour tissue, Arrhenius value equal to 1 indicates that 63% of tumour cells have been destroyed while Arrhenius value equal to 4.6 denotes 99% of the tumour cells have been necrosed.

## Results and discussion

The waveform along the 3D grid points of the breast phantom was simulated. The sound intensity and volumetric heating source obtained was $7.990 \text{ x } 10^4 \text{ W}/m^2$ and $1.047 \text{ x } 10^8 \text{ W}/m^3$ respectively. As a result of this volumetric heating source, we obtained a maximum temperature rise of approximately $56.27℃$ within the tumour region. The maximum temperature is restricted and focused onto the tumour centre. With this temperature, we achieved Arrhenius value of 4.7 which indicates that 99% of the tumour cells have been destroyed. The temperature rise over time is consistent with increasing value of Arrhenius integral. As the temperature falls to the baseline temperature value, the Arrhenius value is seen to be consistent throughout the duration of the sonication time. This therapeutic behaviour demonstrates that our system is quantifying the biological effects in tissue to a considerable level. Comparing our value to the Arrhenius baseline value of 4.6, we estimated an error propagation in measurement of $3.7\%$ which shows that our measured value is still within the baseline value of the Arrhenius damage integral.

## Conclusion

In this preliminary simulation, we have been able to achieve a foundational numerical simulation through an anatomically realistic breast phantom. A more robust approach will be to implement a non-linear wave equation which is equivalent to the Westervelt wave equation. Also, in terms of the heat distribution, an extended approach will be to implement the dual phase lag model which is an upgrade to the Pennes bio heat transfer equation. With these implementations, we will be able to build a sophisticated HIFU system which is capable of simulating complex and heterogenous breast tissues. Also, our future numerical methods will solve the limitations seen in previous research work [10] which is the key objective of this research.

## References

[1] J. S. Vaidya and D. Joseph, *Fast Facts: Breast Cancer*. Basel, SWITZERLAND: S. Karger AG, 2014. [Online]. Available: http://ebookcentral.proquest.com/lib/herts/detail.action?docID=1727117
[2] *Cancer Research UK*. Accessed: Jan. 18, 2026. [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer
[3] Z. Shi, P.-F. Qiu, Y. Wang, and H. Liu, 'Minimally invasive biopsy technique predicting breast pathological complete response after neoadjuvant therapy for breast cancer', *Gland Surg.*, vol. 14, pp. 1263–1271, Jul. 2025, doi: 10.21037/gs-2025-103.
[4] S. Salzano *et al.*, 'p63 immunohistochemical expression in tumor cells of high-grade invasive breast carcinomas on core biopsy: a potential diagnostic pitfall', *Ann. Diagn. Pathol.*, vol. 80, p. 152554, Aug. 2025, doi: 10.1016/j.anndiagpath.2025.152554.
[5] W. Fry and F. Fry, 'Fundamental neurological research and human neurosurgery using intense ultrasound', *IRE Trans. Med. Electron.*, no. 3, pp. 166–181, 1960.
[6] W. J. Fry, J. W. Barnard, F. J. Fry, and J. F. Brennan, 'Ultrasonically produced localized selective lesions in the central nervous system', *Am. J. Phys. Med. Rehabil.*, vol. 34, no. 3, pp. 413–423, 1955.
[7] W. G. Bradley Jr, 'MR-Guided Focused Ultrasound: A Potentially Disruptive Technology', *J. Am. Coll. Radiol.*, vol. 6, no. 7, pp. 510–513, Jul. 2009, doi: 10.1016/j.jacr.2009.01.004.
[8] M. OKA, 'Clinical use of ultrasonics and related biological research in Japan.', *Am. J. Phys. Med.*, vol. 37, no. 4, pp. 210–218, Aug. 1958.
[9] T. J. Dubinsky, C. Cuevas, M. K. Dighe, O. Kolokythas, and J. H. Hwang, 'High-Intensity Focused Ultrasound: Current Potential and Oncologic Applications', *American Journal of Roentgenology*, vol. 190, no. 1. pp. 191–199, 2008.
[10] R. Rahpeima and C.-A. Lin, 'A comprehensive numerical procedure for high-intensity focused ultrasound ablation of breast tumour on an anatomically realistic breast phantom', *PLOS ONE*, vol. 19, no. 10, p. e0310899, Oct. 2024, doi: 10.1371/journal.pone.0310899.
[11] M. J. Burfeindt *et al*., "MRI-Derived 3-D-Printed Breast Phantom for Microwave Breast Imaging Validation," in *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 1610-1613, 2012, doi: 10.1109/LAWP.2012.2236293.
[12] R. O. Mays, "Microwaves in Heterogeneous Tissue," University of Wisconsin-Madison, 2016.
[13] B. E. Treeby, J. Jaros, A. P. Rendell, and B. T. Cox, "Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method," *J. Acoust. Soc. Am.,* vol. 131, no. 6, pp. 4324-4336, 2012.
[14] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave-fields," *J. Biomed. Opt.,* vol. 15, no. 2, p. 021314, 2010.
[15] Hamilton, M.F., Morfey, C.L. (2024). Model Equations. In: Hamilton, M.F., Blackstock, D.T. (eds) Nonlinear Acoustics. Springer, Cham. https://doi.org/10.1007/978-3-031-58963-8_3
[16] ter Haar, >Gail, & Coussios, C. (2007). High intensity focused ultrasound: Physical principles and devices. International Journal of Hyperthermia, 23(2), 89–104. https://doi.org/10.1080/02656730601186138
[17] Kim J. Architectural Acoustics. Sejin Co. 2014.
[18] Huang J, Holt RG, Cleveland RO, Roy RA. Experimental validation of a tractable numerical model for focused ultrasound heating in flow-through tissue phantoms. The Journal of the Acoustical Society of America. 2004;116(4):2451–8. pmid:15532675
[19] Pennes HH. Analysis of tissue and arterial blood temperatures in the resting human forearm. Journal of applied physiology. 1948;1(2):93–122. pmid:18887578
[20] M. Paruch, 'Mathematical Modeling of Breast Tumour Destruction Using Fast Heating during Radiofrequency Ablation', *Materials*, vol. 13, no. 1, p. 136, Dec. 2019, doi: 10.3390/ma13010136.

https://doi.org/10.18745/pb.27307

# Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy

**Tanojit Paul[1*], Shan Goh[2], Javier Espinosa-Montiel[1], Shamil Rafeeq[1], Gozde Isik[1], Chris Stopford[1], Loic Coudron[1], Nikolay Dimov[1], Ian Johnston[1]**

[1]*School of Physics, Engineering & Computer Science, University of Hertfordshire, College Ln, Hatfield AL10 9AB*
[2]*Health, Medicine and life sciences, University of Hertfordshire, College Ln, Hatfield AL10 9AB*

*Corresponding author: t.paul@herts.ac.uk

**Indoor air contains many microbes that can move as bioaerosols and affect human health. However, the low amount of microbial DNA in air makes metagenomic study difficult. This study evaluated six air samplers and multiple DNA extraction kits across café and gym environments. The CUB Sampler consistently recovered the highest viable counts (up to 375 CFU/m$^3$) and the greatest microbial diversity, capturing over 1,500 species. Results emphasize that careful selection of samplers and extraction protocols is essential for characterizing indoor air microbiomes.**

Keywords: Bioaerosols; Air sampling; Metagenomics; DNA extraction; Indoor air quality

## Introduction

People spend almost 90% of their time indoors [1], where air often contains diverse microorganisms. These bioaerosols can cause significant health issues, including pneumonia, allergies, and infections [2,3,4]. While indoor spaces can contain over 100,000 microbial cells per cubic meter [5,6], the total biological material is several orders of magnitude lower than in soil or the human body. This low biomass necessitates high-volume sampling and optimized DNA extraction protocols for metagenomic sequencing. This study evaluates various sampling technologies to establish a robust workflow for characterizing indoor air microbial communities.

## Experimental/Simulation

Sampling was conducted at the Spectra Café and Sports Village Gym at the University of Hertfordshire (UH) using six devices: BioSampler, Coriolis µ, Coriolis Compact, CUB Sampler, Compact Personal Air Sampler (CPAS), and Electrostatic Precipitators (ESP). Four extraction kits were tested, with Qiagen PowerWater selected for its consistency and purity (0.85–41.8 ng/µL). Following extraction, Illumina amplicon sequencing was performed to generate taxonomic profiles. Parallel culture-based analysis was conducted using 72-hour incubation at 30°C to assess microbe viability.

## Results and discussion

Evaluation of DNA extraction kits identified the Qiagen PowerWater kit as superior for low-biomass air samples, yielding 32.2 ng/µL for a mock community compared to 5.3 ng/µL for PowerSoil Pro. While direct DNA yields from air were consistently limited (up to 0.029 ng/µL), absolute abundance analysis of 16S rRNA gene sequences revealed significant differences between samplers. The CUB sampler achieved the highest recovery, reaching over 7,500 16S rRNA gene copies/µL in gym environments.
Illumina amplicon sequencing provided a high-resolution view of the microbial community, as visualized in the microbial composition plot (Fig. 1(a)). Alpha diversity analysis showed that the CUB captured a complex environmental mix, identifying approximately 1,500 observed species in the café and 1,200 in the gym. Fig. 1(a) Microbial composition across different sampling devices and environments, highlighting the diverse taxa captured by the different samplers.
Taxonomic profiling revealed distinct "biological fingerprints" for each environment. The gym was dominated by Staphylococcus and Corynebacterium (associated with human skin and sweat), alongside Micrococcus (dust),



a

https://doi.org/10.18745/pb.27307

while the café was characterized by background flora such as Micrococcus and Enhydrobacter. Parallel culture analysis confirmed these findings; the CUB recovered the highest counts, reaching 375 CFU/m³ in the café. (Fig. 1(b)).

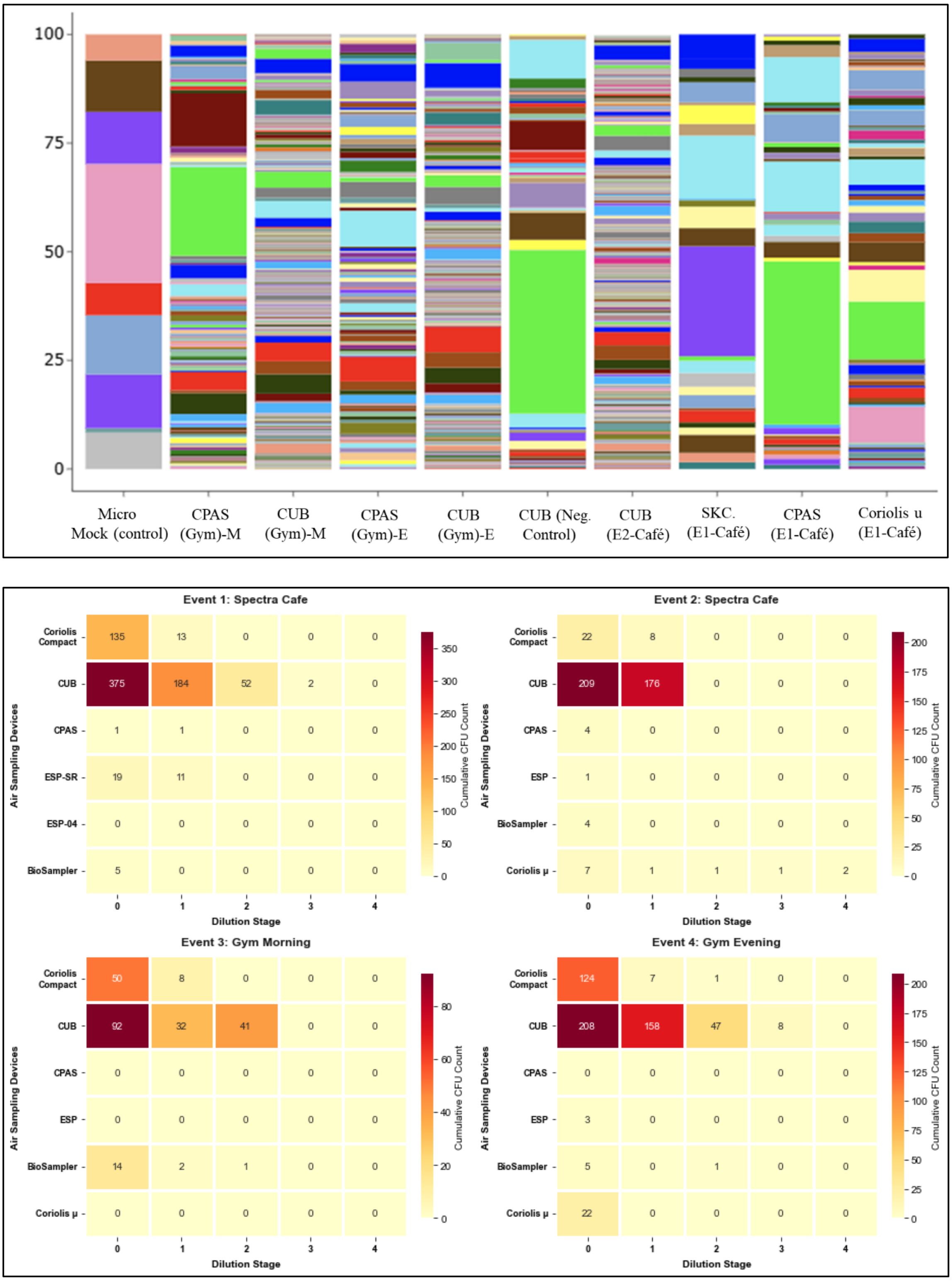



*Fig. 1(a) Microbial composition across various sampling devices, (b)Heatmaps showing cumulative CFU counts for various samplers across four sampling events at different dilution levels.*

https://doi.org/10.18745/pb.27307

## Conclusion

This study establishes an optimized metagenomic workflow for indoor air by integrating diverse sampling technologies and optimized extraction protocols. The CUB Sampler provided high microbial diversity and viable counts. The characterization of skin and dust associate microbes in gym and café highlights the potential of this multi-sampler approach for comprehensive indoor air quality assessment and public health monitoring.

## Acknowledgement

This work is supported by the Biodetection Technologies Hub funded by Research England and EPSRC Centre for Doctoral Training in Aerosol Science.

## References

[1] P. Hoppe and I. Martinac, “Indoor climate and air quality,” Int J Biometeorol, vol. 42, no. 1, pp. 1-7, 1998.
[2] K. H. Kim, K. E. Kabir, and S. A. Jahan. “Airborne bioaerosols and their impact on human health” J. Environ. Sci., vol. 67, pp. 23–35, 2018.
[3] M. Alighadri, M. Alipour, H.R. Ghaffari, A. Zarei, N. Gharari, B. Alizadeh, A. Sadigh, A. Asghari, M. Fazlzadeh, R. Rostami. “Indoor air pollution due to bacterial bioaerosols in beauty salons of Ardabil, Iran: characterization, influencing factors and health risk assessment” *Journal of Environmental Health Science and Engineering*, vol. 23, no. 1, p. 3, 2024.
[4] P. Srikanth, S. Sudharsanam, and R. Steinberg. “Bioaerosols in indoor environment: composition, health effects and analysis” Ind. J. Med. Biol., vol. 26, pp. 302–312, 2008.
[5] M. Toivola, A. Nevalainen, and S. Alm. “Personal exposures to particles and microbes in relation to microenvironmental concentrations” Indoor Air, vol. 14, no. 5, pp. 351–359, 2004.
[6] X. Jiang, C. Wang, J. Guo, J. Hou, X. Guo, H. Zhang, J. Tan, M. Li, X. Li, H. Zhu. “Global Meta-Analysis of Airborne Bacterial Communities and Associations with Anthropogenic Activities” *Environmental Science & Technology*, vol. 56, no. 14, pp. 9891–9902, 2022.

https://doi.org/10.18745/pb.27307

# CFD Optimisation of Airfoil Fin Heat Sinks for High Performance Electronics Cooling

**Karan Kumar[1]***

*[1]Department of Engineering, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, United Kingdom*

*Corresponding author: kk23aac@herts.ac.uk

**This study investigates airfoil shaped pin fin heat sinks for high power density compact electronic systems using Computational Fluid Dynamics. A straight fin baseline was first developed for mesh independence testing, validation and comparison. Airfoil designs were then assessed by varying chord length, orientation angle and airfoil type. Response Surface Methodology, ANOVA and desirability based optimisation were used to balance Nusselt number and pressure drop. Compared with the straight fin baseline, the final modified airfoil geometry increased heat transfer by 4.17% and reduced pressure drop by 7.18%.**

Keywords: Computational Fluid Dynamics; airfoil fin heat sink; electronics cooling; Nusselt number; pressure drop

## Introduction

Modern electronic systems operate at increasingly high power densities, making thermal management a major design challenge. Early microchannel cooling research showed that compact liquid cooled heat sinks can remove high heat fluxes effectively, but also highlighted the importance of pressure drop and pumping demand in small passages [1]. Later reviews of heat sink design have also emphasised that thermal enhancement must be assessed together with hydraulic cost, because a geometry that improves heat transfer may still be inefficient if it creates excessive flow resistance [2].

Airfoil shaped pin fins offer a possible solution because their streamlined geometry can reduce wake formation and support smoother flow through compact cooling passages. Recent studies on airfoil pin fin heat sinks have shown that airfoil arrangement, orientation and profile shape can strongly affect the balance between heat transfer and pressure loss [3], [4]. This study therefore investigates whether airfoil fin geometries can improve the thermal hydraulic performance of a compact heat sink compared with a straight fin baseline, with the aim of supporting more efficient cooling for high power density compact electronic systems. The work focuses on three design variables: chord length, orientation angle and airfoil type.

## Experimental/Simulation

A three dimensional steady state CFD model was developed in ANSYS Fluent to simulate single phase forced convection with conjugate heat transfer. The heat sink was modelled as aluminium and the coolant as water. A velocity inlet, pressure outlet and volumetric heat generation in the heater region were applied, allowing heat to conduct through the base and fins before being removed by the coolant.

The study began with a straight fin baseline, used for mesh independence testing, solver stabilisation and validation of post processing definitions. Mesh independence was assessed using the volume averaged heater temperature, and the final mesh was selected once this value stabilised. The baseline produced a Nusselt number of 7.354 and a pressure drop of 144.241 Pa, and was used as the reference for all airfoil cases. Nusselt number represented heat transfer performance, while pressure drop represented the hydraulic penalty.

The airfoil design space was studied using a structured Design of Experiments. The factors were chord length, orientation angle and airfoil type. Chord lengths of 8 mm, 10 mm and 12 mm were considered with NACA 0012 and NACA 0020 profiles. The CFD dataset was analysed using Response Surface Methodology and ANOVA to identify factor effects and interactions. Desirability based optimisation was then applied with equal importance given to increasing Nusselt number and reducing pressure drop.

https://doi.org/10.18745/pb.27307

## Results and discussion

The CFD results showed a clear trade off between heat transfer and pressure drop. Run 10 produced the highest Nusselt number, 8.609, which represented a 17.06% increase in heat transfer compared with the straight fin baseline. However, its pressure drop increased to 218.638 Pa, which was 51.58% higher than the baseline. This case showed that stronger flow redirection and mixing can improve heat transfer but may also create a significant hydraulic penalty.

Run 2 produced the lowest pressure drop, 117.228 Pa, which represented an 18.73% reduction compared with the baseline. However, its Nusselt number fell to 6.433, which was 12.52% lower than the baseline. This confirmed that reducing pressure loss alone does not give the best overall cooling design. The most suitable design therefore had to be selected as a compromise between thermal improvement and hydraulic resistance.

The equal importance desirability analysis selected a balanced solution corresponding most closely to Run 3. This case gave a Nusselt number of 7.117 and a pressure drop of 137.893 Pa. Although it was not the strongest case for either response individually, it provided the best compromise and was therefore refined further. The ANOVA results supported this interpretation. The Nusselt number response was influenced mainly by orientation related terms, suggesting that thermal performance depended strongly on flow organisation and mixing. In contrast, pressure drop was affected more strongly by chord length and airfoil type, showing that hydraulic loss was mainly controlled by fin size, profile thickness and blockage effects.

The modified Run 3 geometry improved both responses at the same time. Its Nusselt number increased from 7.117 to 7.661, while pressure drop reduced from 137.893 Pa to 133.882 Pa. Using the straight fin baseline as the comparison point, the final modified design achieved a 4.17% increase in heat transfer and a 7.18% reduction in pressure drop. The 3 Mod contour results show a more effective use of the available flow region, with improved flow distribution through the fin array and fewer weak circulation regions. This explains how the modified design improved heat transfer without increasing hydraulic resistance. The final geometry therefore achieved a genuine thermal hydraulic improvement rather than simply shifting the trade off from one response to the other.

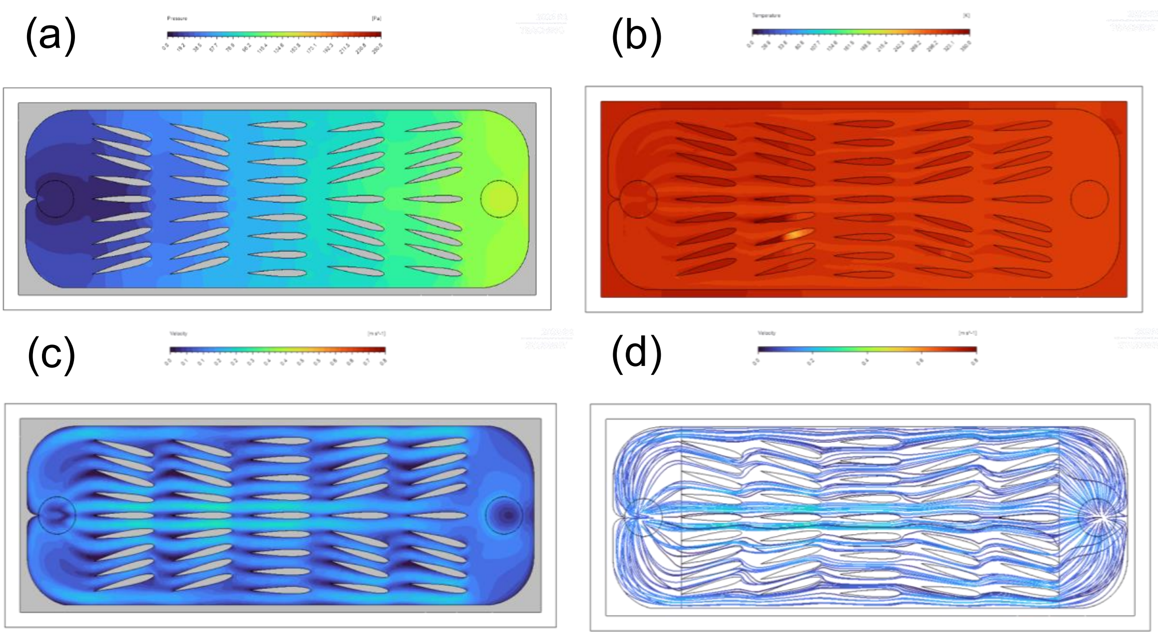



*Fig. 1. Modified Run 3 results, showing (a) pressure, (b) temperature, (c) velocity and (d) streamlines. The final geometry improved flow distribution while increasing Nusselt number and reducing pressure drop compared with the straight fin baseline.*

*Table. 1. Key CFD cases for thermal hydraulic comparison.*

| Case | Description | Chord length, mm | Airfoil shape | Nu | ΔP, Pa |
|---|---|---|---|---|---|
| Run 0 | Straight fin baseline | N/A | Straight fin | 7.354 | 144.241 |
| Run 2 | Lowest pressure drop | 8 | NACA 0012 | 6.433 | 117.228 |
| Run 3 | Balanced selected case | 8 | NACA 0012 | 7.117 | 137.893 |
| Run 10 | Highest heat transfer | 10 | NACA 0020 | 8.609 | 218.638 |
| 3 Mod | Final modified design | 8 | NACA 0012 | 7.661 | 133.882 |

The comparison in Table 1 shows why the final design was not selected only from the maximum or minimum response values. The highest Nusselt number case and the lowest pressure drop case represented opposite ends of the design space, while the modified Run 3 geometry gave the most balanced improvement. This makes the final result more useful for practical cooling design, where heat transfer gain must be achieved without a large increase in pumping requirement.

## Conclusion

This study shows that airfoil shaped pin fin heat sinks can improve compact electronics cooling when the design is selected using a balanced thermal hydraulic approach. The airfoil cases showed that the highest heat transfer and lowest pressure drop occurred in different geometries, confirming that neither response should be optimised alone. This was important because Run 10 improved heat transfer substantially but introduced a large pressure penalty, while Run 2 reduced pressure drop but weakened heat transfer. The final design therefore had to be selected by considering both responses together.

The final modified airfoil design improved both responses compared with the straight fin baseline, increasing Nusselt number by 4.17% and reducing pressure drop by 7.18%. This supports the use of CFD, Design of Experiments, ANOVA and desirability based optimisation as a practical method for developing more efficient compact heat sinks. The results also suggest that airfoil orientation and profile selection can improve the use of internal flow passages by reducing ineffective flow regions while maintaining useful fluid interaction with the fin surfaces.

The main limitation of this work is that the results are based on steady state single phase CFD. Future work should include experimental validation, additional Reynolds number ranges, and further airfoil profiles to confirm whether the same thermal hydraulic improvement is maintained under wider operating conditions. Overall, the study provides a clear basis for further development of airfoil based heat sink designs for compact electronics cooling.

## References

[1] D. B. Tuckerman and R. F. W. Pease, "High-performance heat sinking for VLSI," IEEE Electron Device Letters, vol. 2, no. 5, pp. 126–129, 1981.
[2] H. E. Ahmed, B. H. Salman, A. S. Kherbeet and M. I. Ahmed, "Optimization of thermal design of heat sinks: A review," International Journal of Heat and Mass Transfer, vol. 118, pp. 129–153, 2018.
[3] H. Babar, H. Wu, H. M. Ali and W. Zhang, "Hydrothermal performance of inline and staggered arrangements of airfoil shaped pin-fin heat sinks: A comparative study," Thermal Science and Engineering Progress, vol. 37, 101616, 2023.
[4] K. D. Wu and H. C. Weng, "Thermal performance of open microchannel heat sink with NACA airfoil shaped pin fins," Journal of Mechanics, vol. 39, pp. 229–244, 2023.

https://doi.org/10.18745/pb.27307

# Experimental Validation of a Dual-Band mmWave Antenna Using a Scaled Proof-of-Concept Approach

**Silan Panthi**
*Corresponding author: panthi.silan@gmail.com

**This paper presents the design and validation of a compact dual-band millimetre-wave antenna for 5G mobile applications operating at 28 GHz and 39 GHz. A stacked microstrip patch antenna with coaxial feeding was optimised in Ansys HFSS, achieving return losses higher than 16 dB and gains of approximately 6 dBi within a compact footprint suitable for integration on small electronic devices. A scaled proof-of-concept prototype operating at 11.2 GHz and 15 GHz was fabricated on an FR4 substrate to validate the design. Measured results confirm dual-band operation while highlighting trade-offs in substrate losses, isolation, and fabrication tolerances, demonstrating the feasibility of a practical dual-band mmWave antenna implementation.**

Keywords: Dual-band antenna; Millimetre-wave (mmWave); 5G communications; Microstrip patch antenna; Proof-of-concept

## Introduction

The design of 5G mobile communication antennas is advancing quickly for mobile terminals and base stations, aiming for higher transmission rates, reduced latency, and enhanced access rates, which offer significant advantages over the existing 4G system [1]. By targeting two key frequency bands within the mmWave spectrum, the antenna is optimised for flexible, high-performance operation under varying conditions [2]. With increasing user density and data traffic in urban environments, ensuring efficient, reliable communication poses a significant challenge [3]. This paper addresses the challenges of compact and efficient antenna design for 5G mmWave communication by developing a dual-band stacked microstrip antenna operating at 28 GHz and 39 GHz. The design aims to enhance network capacity, coverage, and performance within the size constraints of smartphone integration. A combined approach of simulation-driven optimisation using Ansys HFSS and practical validation is adopted.

To verify the design, a scaled proof-of-concept prototype operating at 11.2 GHz and 15 GHz is developed and fabricated using an FR4 substrate. The prototype demonstrates dual-band operation while meeting key performance metrics, including return loss, gain, and radiation characteristics. Experimental testing enables evaluation under real-world conditions, highlighting practical trade-offs between performance, material limitations, and fabrication constraints.

## Design and Simulation

The proposed antenna consists of a compact dual-band stacked microstrip patch structure designed to operate at 28 GHz and 39 GHz for 5G mmWave applications. The design employs a two-layer substrate configuration with radiating patches on each layer and a common ground plane, enabling dual-band, dual-polarised operation. A coaxial probe feeding technique is used to achieve efficient impedance matching at 50 Ω while maintaining a compact footprint of 4 mm × 4 mm × 0.96 mm, suitable for smartphone integration, as shown in Fig. 1.

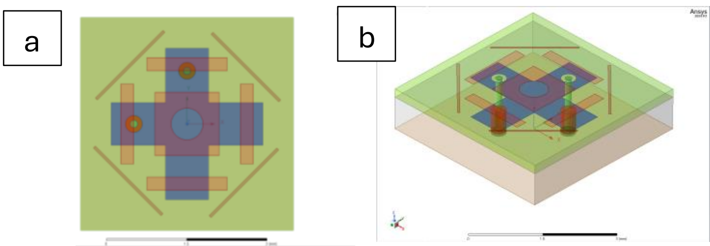



*Fig. 1. Proposed Design; Top View (a) and Isometric View (b)*

https://doi.org/10.18745/pb.27307

The antenna was modelled and optimised using Ansys HFSS through parametric analysis of key design variables, including patch dimensions, substrate thickness, and feed position. The optimisation process aimed to achieve resonance at the target frequencies with return losses higher than 10 dB and improved radiation characteristics. The final design achieved simulated return losses of 19.88 dB at 28.275 GHz and 16.46 dB at 39.385 GHz, as shown in Fig.2 (a), confirming effective impedance matching and dual-band performance. The antenna exhibited directional radiation patterns with simulated gains of 6.7 dBi at 28.3 GHz and 6 dBi at 39.4 GHz, as shown in Fig.2 (b) and (c).

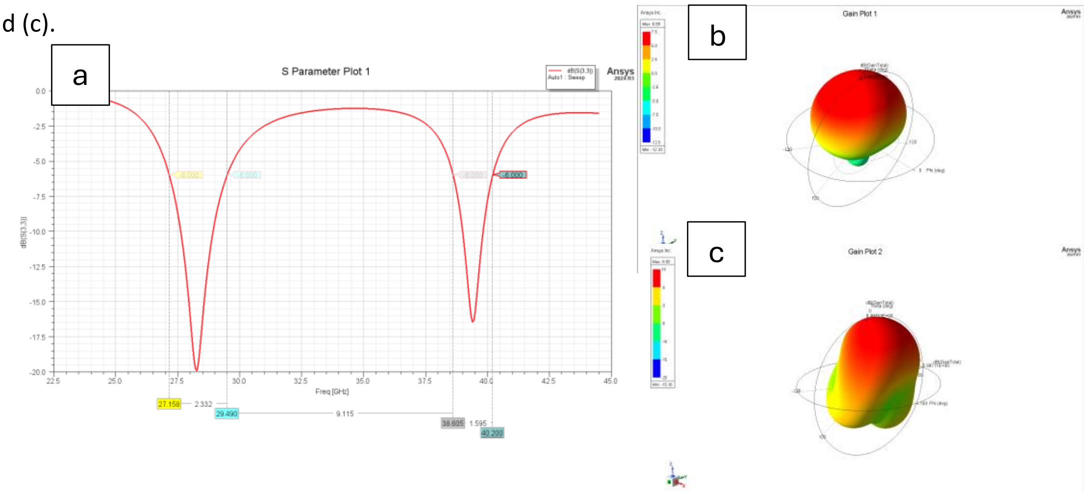



*Fig. 2. Dual Band Design S11 Parameter (a); Total Gain Plot at 28.275 GHz(b) and Total Gain Plot at 39.385 GHz (c)*

To validate the design methodology under practical constraints, a scaled proof-of-concept antenna was developed at 11.2 GHz and 15 GHz. The POC preserved a similar frequency ratio to the original 28/39 GHz design while accommodating fabrication limitations. It used an FR4 substrate with overall dimensions of 16 mm × 16 mm × 1.6 mm. The POC was also simulated in Ansys HFSS, achieving return losses of 22.80 dB at 11.2 GHz and 19.54 dB at 15 GHz, with simulated gains of 4.20 dBi and 4.54 dBi, respectively, as shown in Fig. 3 (b) and (c). These results confirmed that the scaled prototype maintained effective dual-band operation before fabrication and experimental testing.

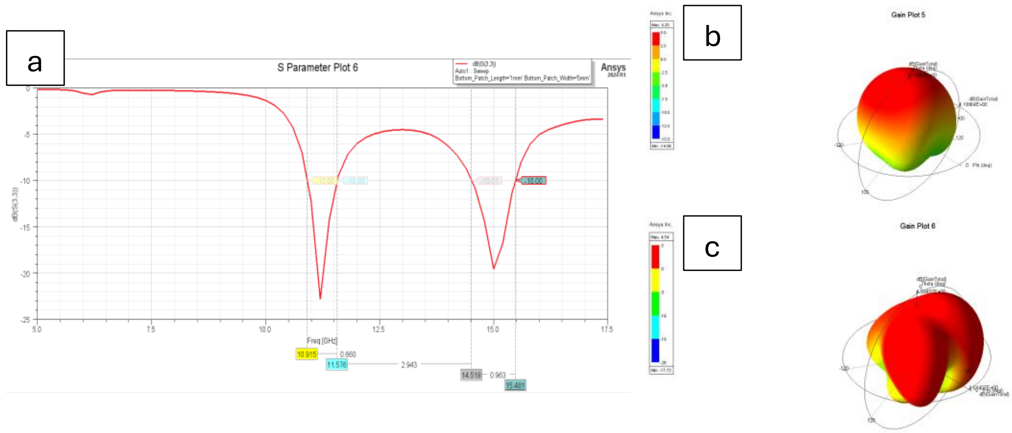



*Fig. 3. Proof-of-Concept Design S11 Parameter (a); Total Gain Plot at 11.2 GHz(b) and Total Gain Plot at 15 GHz (c)*

## Fabrication and Testing

The proof-of-concept (POC) antenna was fabricated using an FR4 substrate via PCB milling on an LPKF ProtoMat S63 system. Due to software limitations, the antenna geometry was exported from Ansys HFSS as DXF files and converted into Gerber and drill files for fabrication. The stacked antenna structure consisted of separate top and bottom substrates, each with radiating patches, a ground plane, and alignment holes, as shown in Fig. 4(a).

https://doi.org/10.18745/pb.27307

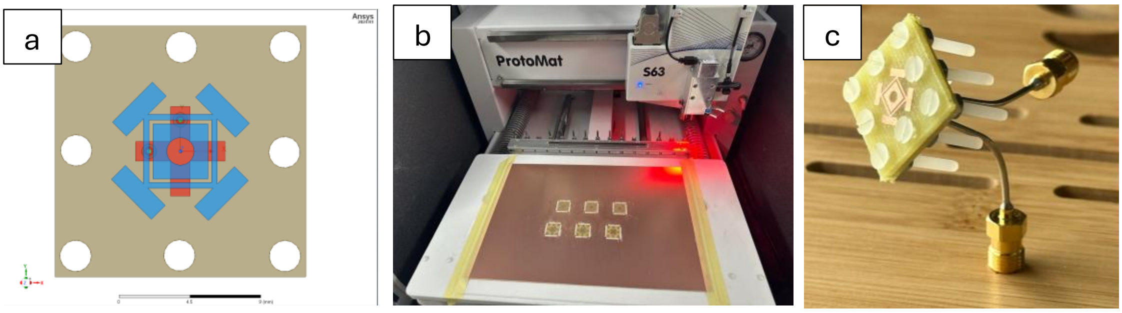



*Fig. 4. Proof-of-Concept Design (a); Fabrication Process using LPKF ProtoMat S63 (b) and Fabricated POC Antenna (c)*

The fabrication was carried out in two phases, including an initial trial to validate milling accuracy and a final fabrication stage with optimised parameters. The process required precise manual soldering of the coaxial feed to the bottom patch, as well as careful alignment of the stacked layers using M2 pins to minimise air gaps, as shown in Fig. 4(a) and (c). Maintaining accurate layer alignment and minimising air gaps were critical to preserving impedance matching and overall antenna performance.

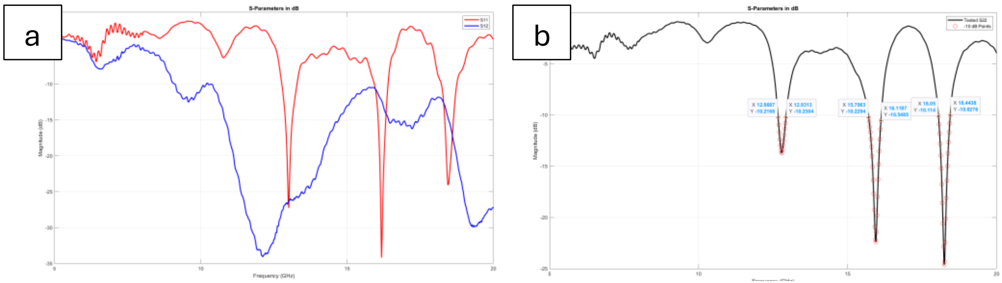



*Fig. 5. S Parameter plot obtained using MATLAB and VNA (a) and S11 Parameter plot obtained using MATLAB and VNA (b)*

The fabricated antenna was characterised using a Vector Network Analyser (VNA) to measure its S-parameters. The measured results confirmed dual-band operation, with resonant frequencies observed at approximately 12.8 GHz (−13.64 dB) and 15.9 GHz (−22.4 dB), along with an additional resonance near 18.4 GHz (≈−24 dB), as shown in Fig. 5. The S-parameter data were exported and processed in MATLAB to obtain results in decibels for comparison with the simulation. Minor deviations between simulated and measured responses were observed, primarily due to fabrication tolerances and material inconsistencies.

## Conclusion

This work presented the design, simulation, and experimental validation of a compact dual-band mmWave antenna for 5G mobile applications. The proposed stacked microstrip configuration achieved dual-band operation at 28 GHz and 39 GHz, with good impedance matching and gain, within a compact form factor suitable for smartphone integration. A scaled proof-of-concept prototype operating at 11.2 GHz and 15 GHz was fabricated and tested, confirming the design approach's feasibility under practical constraints. While discrepancies between simulation and measurement highlighted the impact of substrate losses and fabrication tolerances, the results demonstrate the viability of dual-band mmWave antenna implementation and provide a foundation for future improvements using low-loss materials and advanced integration techniques.

## References

[1] H. Wang, R. Zhang, Y. Luo and G. Yang, "Compact Eight-Element Antenna Array for Triple-Band MIMO Operation in 5G Mobile Terminals," *IEEE Access,* vol. 8, pp. 19433-19449, 2020.

[2] A. Azari, A. Skrivervik, H. Aliakbarian and R. Sadeghzadeh, "A Super Wideband Dual-Polarized Vivaldi Antenna for 5G mmWave Applications," *IEEE Access,* vol. 11, pp. 80761-80768, 31 07 2023.

[3] Z. Pi and F. Khan, "An Introduction to millimeter wave mobile broadband systems," *IEEE Communications Magazine,* vol. 49, no. 6, pp. 101-107, 11 06 2011.

https://doi.org/10.18745/pb.27307

# Mechanical and Durability Performance of Low-carbon Concrete Utilizing Recycled Concrete Coarse Aggregates and Fly Ash for Road Pavement Construction

**Goda K. Baral and Mike Otieno**

*Department of Engineering, School of Physics, Engineering and Computer Science, Centre for Engineering research, University of Hertfordshire*

*Corresponding author: gb25aao@herts.ac.uk

**Abstract**

**Concrete production using construction and demolition waste (CDW) together with low-carbon binders such as fly ash (FA) can significantly reduce the negative environmental impact associated with the construction, repair and maintenance of durable and sustainable road pavements. This study focused on the mechanical and durability performance of concretes made using recycled concrete coarse aggregates (RCCA) and FA. Specimens made using 70/30 CEM II/FA and 50/50 RCCA/NA (natural aggregates) at a 0.50 w/b ratio were tested for mechanical (compressive and flexural strength, and elastic modulus) and durability (chloride penetration and freeze-thaw damage) performance. Preliminary results indicate a potential to use RCCA and FA in pavement construction.**

Keywords: Construction and demolition waste; Recycled concrete coarse aggregates; Natural aggregates; Fly ash

## Introduction

Concrete is one of the most widely used construction material, but its production has significant environmental impacts – consuming large quantities natural resources [1]. In addition, the construction sector generates vast quantities of CDW estimated at over 450 million tons annually in the EU, with nearly 75% still ending up in the landfills despite its strong recycling potential [2]. These issues highlight the need for sustainable construction practices. This study aims to assess the feasibility of using RCCA and FA in rigid road pavement concretes and the outcomes are expected to promote sustainable construction practices.

## Method

Table 1 presents a summary of the concrete mix proportions used. Compressive strength (BS EN 12390-3) [3], flexural strength (BS EN 12390-5) [4] and modulus of elasticity (BS EN 12390-13:2021) [5] were determined to assess the respective mechanical properties of concrete. Durability performance was assessed by carrying out freeze-thaw resistance test (CEN/TS 12390-9:2016) [6], carbonation resistance test (BS EN 12390-12) [7] and rapid chloride penetration test (RCPT, ASTM C1202).

*Table.1. Concrete mix proportions (kg/m$^3$) [8]*

| Mix label | Binder | | Coarse aggregate | | Fine aggregate | Water |
|---|---|---|---|---|---|---|
| | CEM II/A-L 32.5 R (70%) | Fly ash (30%) | Natural (50%) | Recycled (50%) | | |
| Reference concrete | 475 | – | 1000 | – | 710 | 237.5 |
| RCCA-concrete | 332.5 | 142.5 | 500 | 500 | 710 | 237.5 |

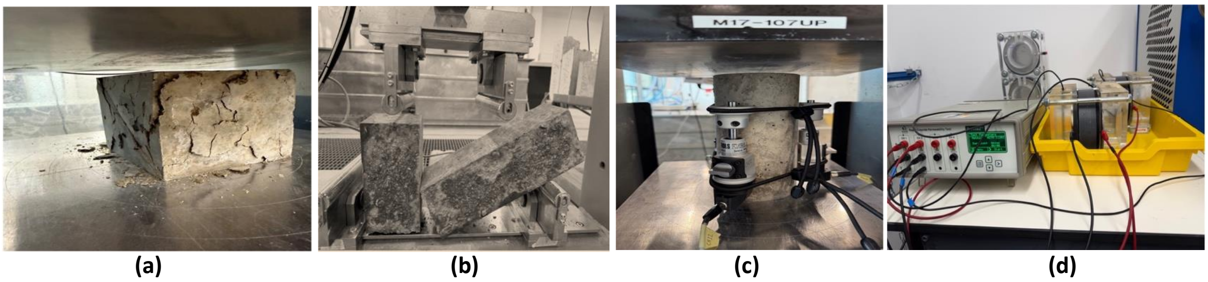



*Fig. 18. Tests performed – (a) compressive strength, (b) flexural strength, (c) modulus of elasticity, (d) chloride penetration*

## Preliminary Results

The compressive strength and flexural strength tests were evaluated to determine the load bearing capacity and ability of resistance to cracking under flexural stress of the pavement. Modulus of Elasticity was measured to understand the stiffness of the concrete pavement under wheel load. Durability of the concrete was assessed by carrying out accelerated 28-day and 56-day chloride penetration test (rapid chloride penetration test, RCPT) and by exposing the concrete to 56 freeze-thaw cycles (–15 to 20 °C). Preliminary results show that RCCA concrete improved mechanical properties. As expected, a steady increase in compressive strength was observed between 7 and 56 days. A similar trend was observed in the flexural strength and elastic modulus (E-mod) results – see Fig. 2 (a)-(c). Negligible (0.02%) mass loss was recorded after 56 freeze-thaw cycles relative to the reference concrete.

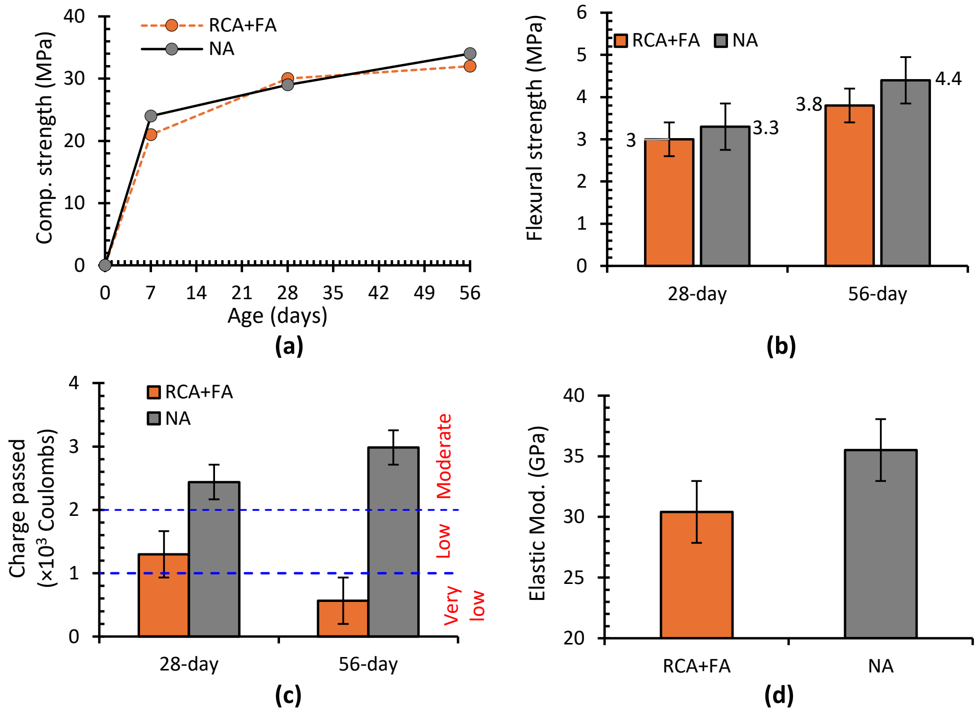



*Fig. 2. (a) Compressive strength, (b) flexural strength, (c) Chloride penetration (RCPT) and (d) Elastic modulus results*

## Conclusion and way forward

Based on the preliminary results, the following conclusions can be made: (i)The RCCA blended concrete (50% RCCA and 30% FA replacement) demonstrates a consistent strength development in terms of both compressive and flexural strength, with time meeting adequate structural requirement for concrete pavements; (ii)The mix exhibits good durability performance, specially high resistance to chloride and negligible mass loss in 56 freeze-thaw cycles, indicating superior surface integrity under environmental conditions; (iii) The replacement of natural aggregates with RCCA combined with FA provides a viable sustainable alternatives for road pavement construction without compromising the essential mechanical and durability properties. In general, the results show satisfactory mechanical and durability performance of FA-blended RCCA concrete and provide a basis for further tests to ascertain its suitability for use in concrete road pavements. Ultimately, a critical assessment will be made on its suitability to contribute to reducing the negative environmental impact associated with the construction, repair and maintenance of durable and sustainable road pavements. The study is ongoing.

## References

[1] M. Alexander, S. Mindness and M. Otieno, (2025), Aggregates in Concrete (2$^{nd}$ Edition), Taylor and Francis, ISBN: 9781032221304, doi.org/10.1201/9781003271253.

[2] A. Kanellopoulos, D. Nicolaides and M. Petrou, (2014), Mechanical and durability properties of concretes containing recycled lime powder and recycled aggregates, Construction and Building Materials, Vol. 53, pp. 253-259, 2014.

[3] BS EN 12390-3:2019, (2019), Testing hardened concrete – Part 3: Compressive strength of test specimens, British Standards Institution.

[4] BS EN 12390-5:2019, (2019), Testing hardened concrete – Part 5: Flexural strength of test specimens, British Standards Institution.

[5] BS EN 12390-13:2021, (2021), Testing hardened concrete – Part 13: Determination of secant modulus of elasticity in compression, British Standards Institution.

[6] PD CEN/TS 12390-9:2016, (2016), Testing hardened concrete – Part 9: Freeze-thaw resistance with de-icing salts- Scaling British Standards Institution, British Standards Institution.

[7] BS EN 12390-12:2020, (2020), Testing hardened concrete – Part 12: Determination of carbonation resistance of concrete – Accelerated carbonation method British Standards Institution, British Standards Institution.

[8] G. Owens, (2013), Fundamentals of Concrete, Midrand: The Concrete Institute, South Africa.

https://doi.org/10.18745/pb.27307

# Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber

**G.Isik[1], L.Coudron[1], I. Johnston[1], R. Kaye[1], C. Stopford[1], W. Stanley[1], L. Weeasiri[1], S. Goh[2], N. Dimov[1]**

[1]*Wolfson Centre for Biodetection Instrumentation Research (WCBIR), School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK*

[2]*Wolfson Centre for Biodetection Instrumentation Research (WCBIR), Department of Clinical, Pharmaceutical and Biological Science, University of Hertfordshire, Hatfield, AL10 9AB, UK*

*Corresponding author: gi24aag@herts.ac.uk

Keywords: Air sampling devices; DNA recovery; Aerosol chamber; eDNA detection

## Introduction

Airborne DNA (eDNA) sampling is a promising tool for monitoring microorganisms, pathogens, and biodiversity in the environment [1, 2]. Despite its potential, reliable detection remains challenging due to low DNA loads in the air and variability in sampling efficiency [3,4].

The performance of aerosol collection devices plays a critical role in DNA recovery and downstream detection sensitivity. Various bioaerosol samplers which are filter-based, electrostatic precipitator, cyclones, and liquid-based systems are available. However, their comparative efficiency under controlled conditions is not well established [5,6].

Controlled aerosol chamber experiments using nebulized cells provide a reproducible approach to evaluate sampler performance. This study compares the following four aerosol samplers in terms of cell recovery, reproducibility, and suitability for low-biomass airborne eDNA applications: Innovaprep CUB sampler, The Electrostatic Precipitator (ESP), the SKC Biosampler, and the Compact Personal Aerosol Sampler (CPAS).

## Experimental/Simulation

Air sampling experiments were conducted in a controlled 8 $m^3$ aerosol chamber using four different aerosol samplers: CUB, ESP, SKC, and CPAS. All the devices were positioned on a table in the middle of chamber at a height of 37 cm above the chamber floor to ensure consistent sampling conditions.

Prior to each experiment, the chamber air was cleaned for 20 minutes with all devices operating. Subsequently, a 3 mL cell solution ($10^8$ cells) containing a mixture of prokaryotic and eukaryotic cells was nebulised while all samplers remained active for another 40 minutes bringing the total effective operating time for each device to 60 minutes. Data were normalised based on the total sampling duration.

Following sample collection, DNA was extracted using the Qiagen DNeasy PowerWater Kit according to the manufacturer's protocol. The DNA concentration was quantified using the Qubit dsDNA High Sensitivity assay. The DNA concentration was further evaluated using quantitative PCR (qPCR) with universal primers for detecting Prokaryotes (Pro) and Eukaryotes (Euk).

All experiments were performed in trieplicates, including appropriate blanks and negative controls to assess contamination.

https://doi.org/10.18745/pb.27307

## Results and discussion

DNA recovery varied significantly across the tested aerosol samplers, demonstrating clear differences in collection efficiency under controlled aerosol chamber conditions (Figure 1d). Qubit measurements showed that the CUB sampler achieved the highest DNA yield (0.542 ng/µL), followed by the ESP (0.059 ng/µL) and the SKC (0.034 ng/µL). In contrast, the CPAS with glass filter, glass filter and coated slide as passive sampler yielded a DNA concentrations below the detection limit with Qubit (Table 1a,b).

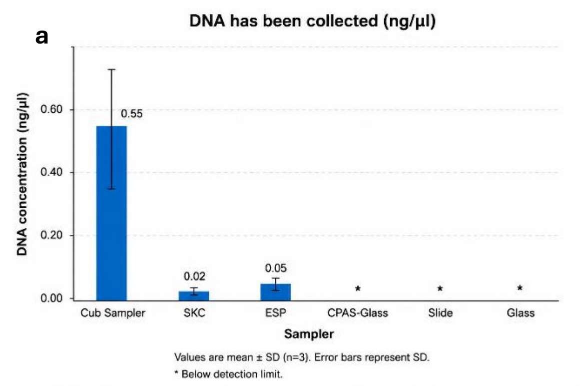
When normalised to the total sampled air volume, the DNA concentration (ng/m$^3$) followed a similar trend, with the CUB sampler significantly outperforming other devices (Fig. 1c). The ESP and the SKC showed moderate but consistent recovery, whereas the results obtained from CPAS devices were below the limit of detection in this experiment.

Although the Innovaprep Cub sampler showed the highest DNA recovery, this result should not be interpreted only because of higher flow rate. The total amount of air processed is an important parameter, but a sampler's performance also depends on particle-size-dependent collection efficiency, filter material, elution efficiency, elution buffer and compatibility with downstream extraction [6,7]. Therefore, the higher DNA yield observed for the CUB sampler reflects a combination of all of these.

It is also important to distinguish between total DNA measured by Qubit and DNA measured by qPCR. Qubit measures the total amount of double-stranded DNA in the extract, including bacterial, fungal, eukaryotic and extracellular or fragmented DNA, whereas prokaryotic and eukaryotic qPCR assays measure only targeted species-specific sequences [8]. Therefore, qPCR-derived eDNA values are not equivalent to the total DNA concentration measured by Qubit.

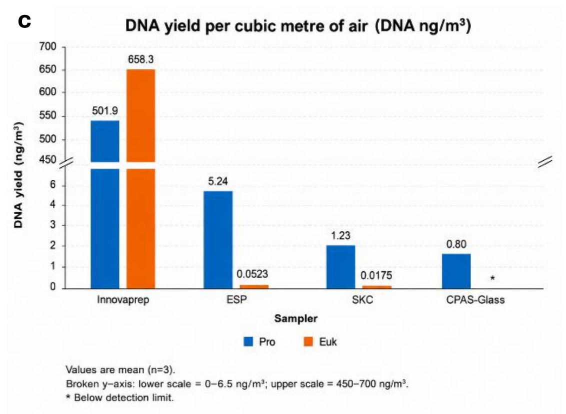
The recovered DNA, derived from collected cells, represents only a fraction of the aerosolised cells introduced into the aerosol chamber. Losses may occur during aerosolisation, deposition onto chamber walls and fans, collection of samples, recovery of samples and DNA extraction losses. Therefore, differences between aerosolised DNA and recovered DNA are expected and should be considered when interpreting samplers' efficiency (Figure 1d).

These findings demonstrate that device selection has a substantial impact on DNA recovery from devices and downstream detection sensitivity. The results emphasise the need for standardisation and optimisation of bioaerosol sampling strategies, particularly for applications requiring reliable detection of low-concentration DNA.




b

| Devices | DNA has been collected (ng/µl) |
|---|---|
| Cub Sampler | 0.542 |
| SKC | 0.034 |
| ESP | 0.059 |
| CPAS-Glass Filter | < 0.50 ng/ml |
| Slide | < 0.50 ng/ml |
| Glass Filter | < 0.50 ng/ml |



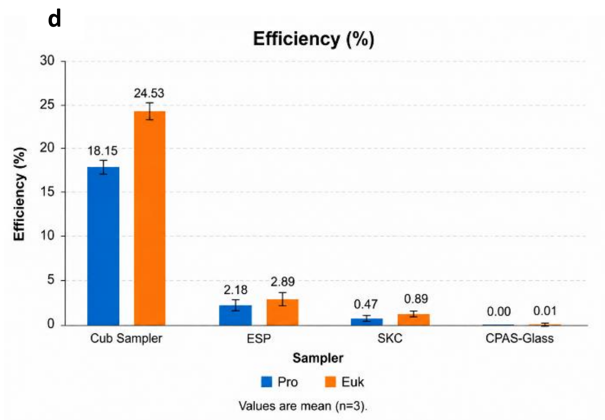




***Figure 1: DNA recovery across aerosol samplers,****(a,b) Total DNA concentration measured by Qubit (ng/mL), (c) DNA concentration normalised to sampled air volume (ng/m³) quantified by qPCR for prokaryotic (Pro) and eukaryotic (Euk) targets, (d) Sampling efficiency estimated based on qPCR-derived DNA concentration relative to the initial aerosolised input generated from a 3 mL suspension containing $10^8$ CFU.*

## Conclusion

Aerosol sampler selection significantly influenced airborne eDNA recovery under controlled conditions. The Innovaprep Cub sampler demonstrated the highest DNA yield and detection sensitivity for both prokaryotic and eukaryotic targets, although greater variability between replicates indicated lower reproducibility. In contrast, ESP and SKC showed moderate but more consistent performance, suggesting potential advantages for applications requiring reproducibility and sample preservation. CPAS-based approaches exhibited limited recovery efficiency under the tested conditions and may require further optimisation. Overall, these findings highlight the importance of optimising and standardising aerosol sampling strategies for low-biomass airborne eDNA applications. Future work should focus on evaluating sampler performance under real-world environmental condition. While the CUB sampler demonstrated superior DNA recovery, it also showed higher variability across replicates, as indicated by larger standard deviation values compared to other samplers, and may be less suitable for certain field applications due to its reliance. In contrast, ESP and SKC exhibited more consistent performance, which may be advantageous depending on the application.

## References

[1] Johnson CM, et al. Airborne environmental DNA for terrestrial vertebrate community monitoring. *Current Biology*. 2021;31(20):1–8.
[2] Clare E, et al. Measuring biodiversity from DNA in the air. *Current Biology*. 2022;32(3):1–6.

[3] Deiner K, et al. Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular Ecology*. 2017;26:5872–5895.
[4] Johnson MS, Cox RD, Barnes MA, et al. Airborne environmental DNA metabarcoding detects more diversity, with less sampling effort, than a traditional plant community survey. *BMC Ecology and Evolution*. 2021;21:218.
[5] Šantl-Temkiv A, et al. Bioaerosol sampling: sampling mechanisms, bioefficiency and field studies. *Aerosol Science and Technology*. 2020;54:1–29.
[6] Mainelis G. Bioaerosol sampling: classical approaches, advances, and perspectives. *Aerosol Science and Technology*. 2020.
[7] InnovaPrep. *AirPrep™ Cub Sampler User Guide*. 2021.
[8] Smith CJ, Osborn AM. Advantages and limitations of quantitative PCR (qPCR)-based approaches in microbial ecology. *FEMS Microbiology Ecology*. 2009;67(1):6–20.

This work is supported by the Biodetection Technologies Hub funded by Research England.

UKRI Research England Biodetection Technologies Hub

https://doi.org/10.18745/pb.27307

# Practical Investigation on the Use of Synthetic Dataset for Academic research framework and simulation of realistic scenarios while using machine learning.

**Cheima Ali Bensaad[*1], Sumama Islam [2], *Zehra S. Qureshi* [3] ,Ye Lin Soe [4]**
[1,2,3,4]*University of Hertfordshire, Computer science department*
*Corresponding author: c.ali-bensaad@herts.ac.uk

## Abstract:

This paper investigates the use of synthetic datasets in academic research, with a focus on machine learning applications and the simulation of realistic scenarios. It examines the characteristics, limitations, and generation methods of synthetic data, alongside challenges related to reliability, viability, realism, performance, and data quality. Building on findings reviewed and synthesised from several practical projects, the study proposes a practical framework consolidating the key findings and recurring methodological patterns, to guide researchers for the effective use of synthetic data. The framework covers key stages including data quality assessment, balancing techniques, machine learning implementation, evaluation, optimisation, regularisation, and distribution alignment. Ultimately, this work aims to support informed research practices and academic teaching, particularly for computer science and data science students undertaking final year projects.

**Keywords:** Synthetic data; balancing techniques, Framework, data quality and reliability., Machine learning models.

## Introduction:

The proposed pipeline is derived from a synthesis of common outcomes observed across approximately ten projects conducted by the author, primarily focused on the application of synthetic data within the financial domain, particularly in fraud detection. By consolidating recurring methodological patterns and key findings, the pipeline reflects a generalised and practical approach to the effective use of synthetic datasets. The central research problem addressed is**: *"How can the use of synthetic datasets provide a realistic and effective environment for evaluating machine learning models within an academic research framework?"*** The paper also evaluates synthetic data as a viable alternative to real world datasets, particularly where access to sensitive or restricted data is limited. This work is intended to support researchers and final year students in understanding how to use synthetic data to simulate real-life scenarios, avoid common implementation and methodological mistakes to help produce better quality research.

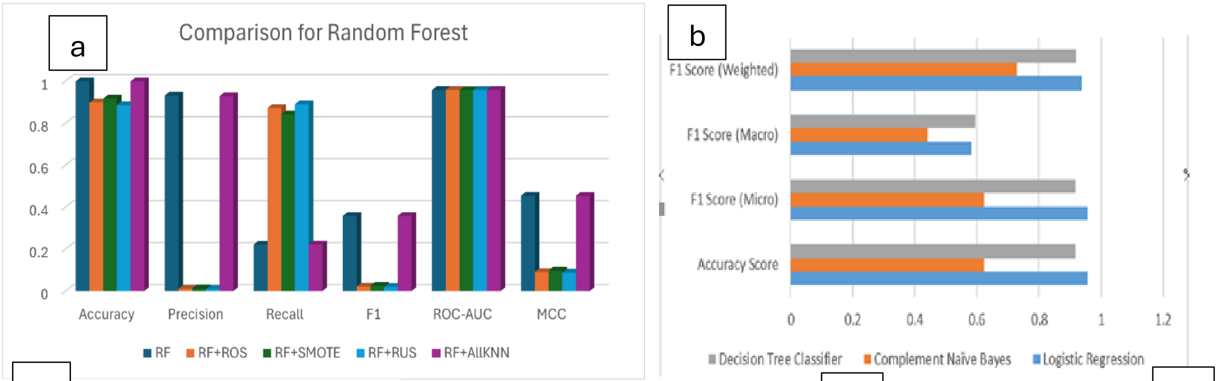
## Experimental/Simulation:

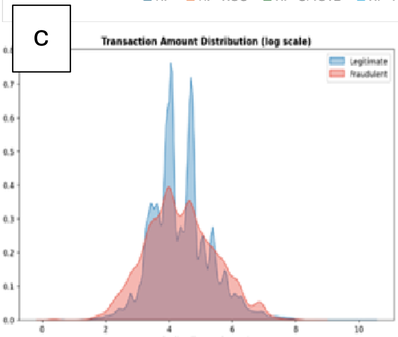
These projects applied machine learning models and balancing techniques to several synthetic financial fraud datasets (all the ones available on Kaggle), drawing from both MSc dissertations and others previously published work in which synthetic data was not the primary focus, though recurring observations across both revealed a significant methodological gap. The ML models used, individually or in combination, include Logistic Regression, Decision Tree, Random Forest, KNN, SVM, and Naïve Bayes. Balancing techniques include SMOTE, TVAE, RUS, ROS, and SMOTEENN. Evaluation relied on ROC, AUC, learning curves, regression metrics, and classification reports, while optimisation employed hyperparameter tuning, grid search, and cross-validation. Although some projects share datasets, most draw from different sources; the target variable, fraud detection, remains consistent across all studies (C.Ali Bensaad and all, 2024-2026) . This work does not cover the full preprocessing pipeline; rather, it emphasises the key stages that inform the development of a practical methodological framework and checklist for researchers working with synthetic and imbalanced datasets.

## Results:

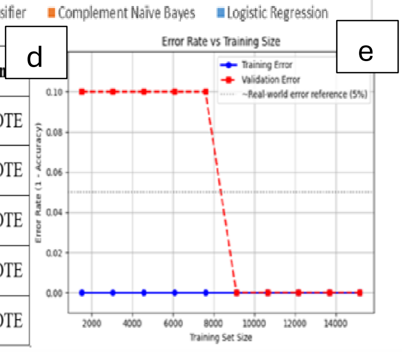
The scenarios assume that the goal is not to compare model performance, but rather to identify anomalies, similarities, and key characteristics across methods and datasets like the Learning curve can show that most errors rely in the testing set. Overall observations across projects align with (C. Ali Bensaad, Yasir Iqbal, 2025), showing that integrating data balancing with machine learning significantly improves fraud detection performance (Fig.A). The results are summarized through the following figures:

https://doi.org/10.18745/pb.27307




| Metric for synthetic data quality | SMOTE | TVAE | Winn |
|---|---|---|---|
| Mean Absolute Difference (MAD) | 12.94 | 7,007.18 | SMOTE |
| Std Divergence (STD-MAD) | 33.40 | 1,450.42 | SMOTE |
| Frobenius Distance (Correlation) | 0.2350 | 2.0843 | SMOTE |
| KS-Test Features Won | 7/12 | 5/12 | SMOTE |
| Mean KS Statistic | 0.1496 | 0.2170 | SMOTE |



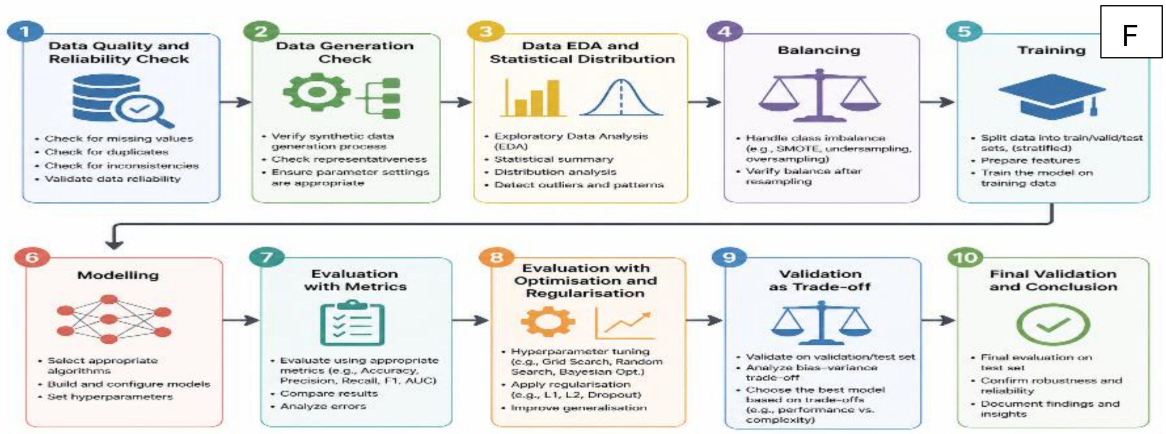




***Fig. 1****. Results of the use of synthetic data and recommended framework:((Fig A) Comparison of models with balancing (Fig.B) Ml models without balancing. (Fig.c) Data distribution and risk analysis.(FigD) Data quality assessment. FigE learning curve. (Fig.F) proposed framework. (Source: Authors results of several output projects).*

Random Forest consistently achieved the strongest results across most datasets use. Learning curves further reveal that most generalisation errors concentrate in the test set, indicating that model complexity, quality of the data and balancing choices directly affect out-of sample performance (Fig. E). A notable paradox emerges when models are trained on unbalanced data: accuracy appears high (Fig. B), yet this is misleading. Models default to favouring the majority class, failing to detect the minority, precisely the class of interest in fraud detection. This underscores the necessity of appropriate balancing prior to training. Data quality has been assessed when SMOTE produced more statistically faithful synthetic fraud samples than TVAE across all quality metrics, achieving lower mean absolute difference, standard deviation divergence, and correlation distance, while outperforming in KS-tests on certain features. Even using deep generative methods such as TVAE , where datasets were sourced externally, data quality was verified through EDA (Fig.C) via distribution and pattern analysis. More interpretation is presented in the next point.

https://doi.org/10.18745/pb.27307

## Discussion and Conclusion:

The following conclusions reflect on the interpretation of the methodology and proposed framework (Fig. F) of main 10 steps derived from our experimental studies that researchers need to consider:

Overall, Models trained on synthetic data demonstrate performance comparable to real data, and in some cases surpass it, that renders them a viable alternative to real dataset for simulation real case scenarios, , where it is valid alternative in situations of dealing with ethical consideration or when it is hard to have access to dataset(Preeti patel,2024) .

Synthetic data generation should aim to capture the key statistical properties of the original/real dataset, with particular attention to whether minority and anomalous classes are adequately preserved. This can be explored using tools such as SDV, Synthpop, DataSynthesizer, and TVAE (Fig. D), as well as DBSCAN or KNN-based approaches. Where researchers collect rather than generate synthetic data, the creator should be consulted regarding generation methodology and reliability. Further verification should be conducted through exploratory data analysis, including class distribution, risk score analysis (Fig. C), and assessment against the ONS Synthetic Data Spectrum for fidelity and disclosure risk (Fig. D). The EthicalOS framework should also be considered at this stage along with practical investigation of the data bias in term of feature engineering, feature importance.

The choice of balancing technique depends on dataset size and the generation method employed. For large datasets, under-sampling is generally preferable, as oversampling can amplify data size and noise. Notably, generation methods such as KNN inherently resemble oversampling by producing new samples from existing data points (Fig. D). It is crucial to address data balancing as an initial step, prior to data splitting and model training, to prevent bias and results errors.

Supervised models are the most commonly appropriate for anomalies, fraud and minority class detection (Fig. A, B). It is essential that researchers understand how these models operate, particularly algorithms that tends to fail predicting/detecting minority class. Generally, Random Forest and Decision Trees consistently demonstrate strong performance, while SVM, despite being common in binary classification, can underperform on minority classes. Researchers must therefore define the target class carefully, assigning the minority class as the positive label, within their implementation (Ali Bensaad & Iqbal, 2025).

Beyond model performance, robust and additional evaluation is essential (Fig. E). Accuracy alone is insufficient; assessment should incorporate F1-score, recall, and the confusion matrix to examine false positive and negative rates. Model selection should reflect a balanced trade-off across these metrics, complemented by optimisation and regularisation analysis including learning curves, hyperparameter tuning, and cross-validation across training, validation, and test sets, to find out how the models are doing in terms of the errors.

In conclusion, synthetic datasets exhibit patterns and characteristics sufficiently close to real data to serve as a reliable research alternative. However, the characteristics, limitations, and strengths of each generation techniques, balancing must be carefully investigated relative to the specific dataset and research context in prior of implementing the models and methods.

## References :

[1] Cheima. Ali Bensaad., Yasir Iqbal., Kashish. "Comparative analysis of the impact of hybrid balancing techniques on various machine learning models for fraud data classification". 13th International Conference on Frontiers of Intelligent Computing: Theory and Applications, FICTA 2025, London, UK Springer proceeding Smart innovations and technologies.

[2] Patel, P."Synthetic data". *Business Information Review*, *41*(2), 48-52. (2024).

[3] Cheima.Ali Bensaad, Sumama Islam , Zehra S. Qureshi ,Ye LinSoe. "Fraud detection using ML", 2024-2026. Msc project, University of Hertdforshire.

[4] Llugiqi, M., Mayer, R. "An Empirical Analysis of Synthetic-Data-Based Anomaly Detection"International Cross-Domain Conference for Machine Learning and Knowledge Extraction, 2022.

https://doi.org/10.18745/pb.27307

# Modelling Treatment For Essential Tremor: A Biophysical Model Of Ultrasound Induced Heating In The Thalamus

**Samar Abbass[1], Ayesha Jameel[2,3], Joely Smith[4,5,6], Peter Bain[7,8], Dipankar Nandi[2,7], Brynmor Jones[3], Rebecca Quest[3,6], Wladyslaw Gedroyc[2,3], Roman Borisyuk[9], Nada Yousif[10*]**

[1] *School of Engineering and Computer Science, Global Academic Foundation, Egypt;*

[2] *Department of Surgery and Cancer, Imperial College London, UK;*

[3] *Department of Imaging, Imperial College Healthcare NHS Trust, UK;*

[4] *University College London Hospitals NHS Foundation Trust, UK;*

[5] *Institute of Neurology, University College London, UK;*

[6] *Department of Bioengineering, Imperial College London, UK*

[7] *Department of Brain Sciences, Imperial College London, UK;*

[8] *Department of Neurosciences, Imperial College Healthcare NHS Trust, UK;*

[9] *Department of Mathematics and Statistics, University of Exeter, UK;*

[10] *School of Physics, Engineering and Computer Science, University of Hertfordshire, UK;*

*Corresponding author: n.yousif@herts.ac.uk

**MR guided focused ultrasound is used as a clinical therapy to alleviate the symptoms of essential tremor by using high energy ultrasound to ablate a nucleus in the thalamus. We used a single compartment, biophysical model of a thalamic neuron to examine the impact of tissue heating on neuronal firing. We found that the neuron was gradually inhibited with increasing temperature, whether in isolation or in a small network. By inspecting the individual ionic currents in the model, we found that the amplitude of the T-type calcium current decreased with increasing temperature due to an acceleration of the inactivation rate.**

Keywords: Essential tremor; Magnetic resonance guided focused ultrasound; Thalamus; Biophysical model; Network model.

## Introduction

Essential tremor is the most common movement disorder worldwide, affecting approximately one million people in the UK [1-3]. ET is characterised by an uncontrollable shaking of both arms and can extend to the rest of the body [4]. ET affects everyday tasks and many patients rely on carers for activities of daily living [5,6]. Magnetic resonance guided focused ultrasound (MRgFUS) is being increasingly adopted as a primary treatment for ET [7]. MRgFUS uses low frequency and high energy ultrasound, known as sonications, to non-invasively heat and ablate the thalamus, a deep grey matter structure in the brain, to disrupt the tremor pathways and cause tremor suppression. During this clinical treatment, trial sonications are used to heat the tissue to sub-ablative temperatures, to allow for better targeting and testing for side effects.

Low energy ultrasound which can be used for neuromodulation rather than ablation [8] has been shown in previous modelling work to impact on the excitation of neurons via the mechanical modulation of the cell membrane [9,10]. However, modelling the impact of high energy ultrasound on neuronal activity is limited, and yet better understanding of the impact of heating during MRgFUS for ablation, would allow better monitoring of side effects during the clinical treatment and therefore lead to better outcomes [11]. In this study we used an existing model of a thalamocortical neuron (TC) [12] to examine the change in neuronal activity induced by heating during MRgFUS.

## Experimental/Simulation

The magnetic resonance thermography data collected during a MRgFUS treatment was used to define the temperature range reached at the targeted tissue during both ablative and sub-ablative trial sonications. We used an existing biophysical model of a thalamocortical neuron consisting of a single compartment [12], via the

https://doi.org/10.18745/pb.27307

NEURON programming environment [13]. The model includes a number of ionic currents which lead to the oscillatory firing patterns expected in a thalamocortical neuron. Specifically: a low-threshold (T-type) calcium current ($I_T$), a hyperpolarisation-activated cation current ($I_h$), a potassium leak current, and sodium and potassium currents. We simulated the effect of increasing tissue temperature via the 'celsius' parameter in NEURON [13] and analysed the impact on neuronal firing. In addition, we ran the same simulations, with the thalamocortical neuron connected in a small network which consisted of two TC neurons and two inhibitory thalamic neurons from the reticular nucleus (RE) [12].

## Results and discussion

The single TC cell model was previously shown to replicate three firing patterns at the default 36°C: resting state (no firing), waxing and waning bursts of action potentials, and slow oscillations (regular spiking in the delta frequency range). We found that increasing the temperature above 36°C did not change the firing in the resting state, inhibited the firing of the bursts in the waxing and waning state and reduced the frequency of the action potentials in the slow oscillation delta state (Figure 1a). For all three of these firing patterns, we found that no action potentials were observed when the temperature exceeded 40°C.

We looked more closely at the impact of the ionic channels on this behaviour, and focussed on the impact on the regular slow oscillation state. We found that the amplitude of the T-type calcium current decreased as the temperature increased, which we did not observe for the $I_h$ current. We examined the kinetics of $I_T$ and found that this was driven by an acceleration of the inactivation rate (figure 1b). We changed the conductance of $I_T$, while simultaneously changing the temperature, and this demonstrated that while the conductance of $I_T$ has an impact on the firing frequency, the cell remains silent above 40°C (Figure 1c). Interestingly, we observed a similar pattern when the TC neuron was connected into a small network of four neurons. However, we found that by manipulating the conductance of the $I_T$ current, the neuron was able to continue to fire action potentials at higher temperatures.

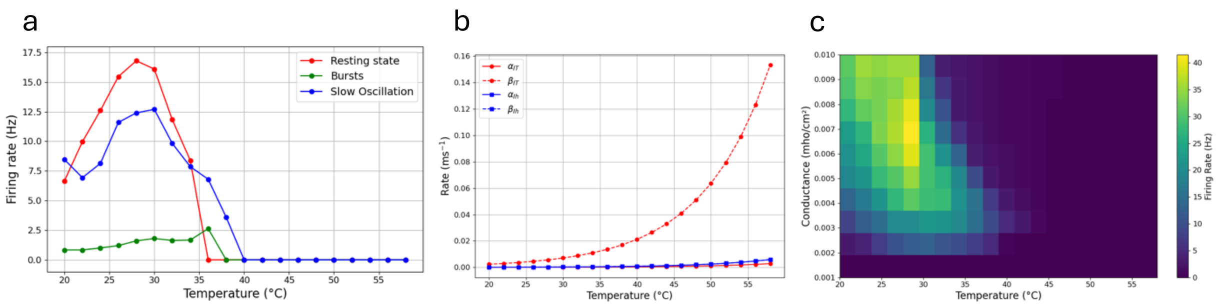



*Fig 1. (a) Spiking rate of a single TC cell vs temperature in three regimes: Resting state, Bursts and Slow oscillations in the delta frequency range. (b) Impact of temperature variation on the activation ($a_{IT}$ and $a_{Ih}$) and inactivation ($b_{IT}$ and $b_{Ih}$) gating of the $I_T$ channel (red) and the $I_h$ channel (blue) included in the TC cell model. (c) Spiking rate variation with $I_T$ conductance and temperature for a single TC cell in the slow oscillation regime.*

## Conclusion

We show by using a single compartment neuron model of a thalamocortical neuron, the cell's firing patterns are inhibited when temperature increases. Using the model, we assessed the influence of the different currents on this behaviour and found that the T-type calcium current was impacted by increasing temperature. These results indicate that clinical observations made during trial sonications may be explained by neuronal inhibition due to small increases in temperature. Here we used a small network containing only four neurons, future work will extend the network approach and connect these findings to a population-level model of essential tremor [14-16] to make testable predictions about the impact of heating during MRgFUS on tremor amplitude.

## References

[1] E. D. Louis and J. J. Ferreira, “How common is the most common adult movement disorder? Update on the worldwide prevalence of essential tremor,” *Movement Disorders*, vol. 25, no. 5, pp. 534–541, Feb. 2010, doi: https://doi.org/10.1002/mds.22838.
[2] B. Thanvi, N. Lo, and T. Robinson, “Essential tremor—the most common movement disorder in older people,” *Age and Ageing*, vol. 35, no. 4, pp. 344–349, Apr. 2006, doi: https://doi.org/10.1093/ageing/afj072.
[3] “Essential Tremor.” Available: https://tremor.org.uk/images/downloads/information/Essential_Tremor_Pamphlet.pdf
[4] M. F. Brin and W. Koller, “Epidemiology and Genetics of Essential Tremor,” *Movement Disorders*, vol. 13, no. S3, pp. 55–63, Oct. 2008, doi: https://doi.org/10.1002/mds.870131310.
[5] A. I. Tröster, R. Pahwa, J. A. Fields, C. M. Tanner, and K. E. Lyons, “Quality of life in Essential Tremor Questionnaire (QUEST): Development and initial validation,” *Parkinsonism & Related Disorders*, vol. 11, no. 6, pp. 367–373, Sep. 2005, doi: https://doi.org/10.1016/j.parkreldis.2005.05.009.
[6] D. Lorenz, D. Schwieger, H. Moises, and G. Deuschl, “Quality of life and personality in essential tremor patients,” *Movement Disorders*, vol. 21, no. 8, pp. 1114–1118, 2006, doi: https://doi.org/10.1002/mds.20884.
[7] A. Jameel, P. Bain, D. Nandi, B. Jones, and W. Gedroyc, “Device profile of exAblate Neuro 4000, the leading system for brain magnetic resonance guided focused ultrasound technology: an overview of its safety and efficacy in the treatment of medically refractory essential tremor,” *Expert Review of Medical Devices*, pp. 1–9, May 2021, doi: https://doi.org/10.1080/17434440.2021.1921572.
[8] T. Deffieux, Y. Younan, N. Wattiez, M. Tanter, P. Pouget, and J.-F. Aubry, “Low-Intensity Focused Ultrasound Modulates Monkey Visuomotor Behavior,” *Current Biology*, vol. 23, no. 23, pp. 2430–2433, Dec. 2013, doi: https://doi.org/10.1016/j.cub.2013.10.029.
[9] M. Plaksin, E. Kimmel, and S. Shoham, “Cell-Type-Selective Effects of Intramembrane Cavitation as a Unifying Theoretical Framework for Ultrasonic Neuromodulation.,” vol. 3, no. 3, May 2016, doi: https://doi.org/10.1523/eneuro.0136-15.2016.
[10] T. Lemaire, E. Neufeld, N. Kuster, and S. Micera, “Understanding ultrasound neuromodulation using a computationally efficient and interpretable model of intramembrane cavitation,” *Journal of Neural Engineering*, vol. 16, no. 4, p. 046007, Jul. 2019, doi: https://doi.org/10.1088/1741-2552/ab1685.
[11] M. Ciocca *et al.*, “Illusions of Self-Motion during Magnetic Resonance-Guided Focused Ultrasound Thalamotomy for Tremor,” *Annals of neurology*, May 2024, doi: https://doi.org/10.1002/ana.26945.
[12] A. Destexhe, T. Bal, D. A. McCormick, and T. J. Sejnowski, “Ionic mechanisms underlying synchronized oscillations and propagating waves in a model of ferret thalamic slices,” *Journal of Neurophysiology*, vol. 76, no. 3, pp. 2049–2070, Sep. 1996, doi: https://doi.org/10.1152/jn.1996.76.3.2049.
[13] M. Hines, “NEURON and Python,” *Frontiers in Neuroinformatics*, vol. 3, 2009, doi: https://doi.org/10.3389/neuro.11.001.2009.
[14] N. Yousif, M. Mace, N. Pavese, R. Borisyuk, D. Nandi, and P. Bain, “A Network Model of Local Field Potential Activity in Essential Tremor and the Impact of Deep Brain Stimulation,” *PLoS Computational Biology*, vol. 13, no. 1, pp. e1005326–e1005326, Jan. 2017, doi: https://doi.org/10.1371/journal.pcbi.1005326.
[15] N. Yousif, P. G. Bain, D. Nandi, and R. Borisyuk, “A Population Model of Deep Brain Stimulation in Movement Disorders From Circuits to Cells,” *Frontiers in Human Neuroscience*, vol. 14, Mar. 2020, doi: https://doi.org/10.3389/fnhum.2020.00055.
[16] N. Yousif, P. G. Bain, D. Nandi, and R. Borisyuk, “Non-conventional deep brain stimulation in a network model of movement disorders,” *Biomedical Physics & Engineering Express*, vol. 11, no. 1, p. 015042, Dec. 2024, doi: https://doi.org/10.1088/2057-1976/ad9c7d.

https://doi.org/10.18745/pb.27307

# Safety Implications of Educational Fuel Cell Systems in Academic Settings

**Mahmoud Chizari**[1]
[1] *School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK*
*Corresponding author: m.chizari@herts.ac.uk

**The adoption of low-power proton exchange membrane (PEM) fuel cells in education is increasing, supported by compact hydrogen storage and desktop electrolysers. Despite their low power (<60 W), these systems introduce safety risks associated with hydrogen handling, storage, and user interaction. This study synthesises risk assessments, COSHH evaluations, and operational procedures for classroom-scale fuel cell systems using refillable hydrogen cartridges and benchtop electrolysers. Key hazards are identified and evaluated, and a practical safety framework is proposed to guide safe implementation in schools and universities, supporting the responsible integration of hydrogen technologies into engineering education.**

**Keywords:** PEM fuel cells; hydrogen safety; engineering education; risk assessment; laboratory safety

## Introduction

The increasing emphasis on sustainable energy technologies has led to the integration of hydrogen and fuel cell systems into engineering education. Educational proton exchange membrane (PEM) fuel cells, typically operating below 60 W, are widely used to demonstrate electrochemical energy conversion and hydrogen-based energy storage. These systems are often configured with small-scale hydrogen storage cartridges and desktop electrolysers, enabling a complete instructional cycle from hydrogen generation to electricity production [1], [2]. A typical educational setup is illustrated in Figure 1, showing the fuel cell stack, hydrogen storage cartridge, and desktop refilling station. Such configurations enable hands-on learning but also introduce safety considerations associated with hydrogen handling and system operation. Although these systems are designed for instructional use, their operation introduces hazards associated with hydrogen, a highly flammable gas, as well as electrical and thermal effects. In educational environments, where users are often inexperienced and operational controls may vary, these risks require structured evaluation. The perceived safety of low-power systems can obscure underlying hazards, particularly those linked to hydrogen handling and user interaction. Established safety frameworks for hydrogen systems emphasise the need for systematic hazard identification and risk control even at small scales [3], [4].


a b c

Fig. 1. (a) low-power PEM fuel cell stack (<60 W) used for electrochemical conversion of hydrogen into electrical energy; (b) refillable metal hydride hydrogen storage cartridge (HYDROSTIK type) for low-pressure hydrogen storage and supply; (c) desktop electrolyser and hydrogen refilling station (HYDROFILL type) used for on-site hydrogen generation.

## Methods

A structured qualitative risk assessment methodology was applied to evaluate the safety of educational PEM fuel cell systems operating below 60 W. The analysis integrates manufacturer guidance, institutional risk assessments (RA), COSHH evaluations, and standard operating procedures (SOP), together with established hydrogen safety principles [1,3,4]. The system under consideration consists of a PEM fuel cell stack, refillable metal hydride hydrogen storage cartridges, and a desktop electrolyser used for hydrogen generation, consistent with typical instructional setups [1,2]. System boundaries include hydrogen production, storage, transfer, and utilisation, as

https://doi.org/10.18745/pb.27307

well as user interaction during setup, operation, and shutdown. Hazard identification was conducted through systematic review of the documentation, focusing on chemical, mechanical, electrical, and operational risks. Hydrogen-related hazards were prioritised due to its wide flammability range and potential inhalation risks in confined environments [4]. Additional hazards include electrical faults, thermal effects, and procedural deviations during refilling and operation. Risk evaluation was performed qualitatively using likelihood–severity considerations aligned with institutional risk matrices and safety guidance [4]. Control measures such as ventilation, use of fume cupboards, and adherence to operating procedures were assessed for effectiveness in mitigating identified risks.

## Results and Discussion

The analysis indicates that hydrogen handling is the dominant contributor to system risk. Hydrogen leakage, although typically limited in volume, may result in flammability hazards or oxygen displacement in poorly ventilated environments [4]. Potential leakage points include tubing connections, cartridge interfaces, and seals, particularly under repeated use or improper handling. Hydrogen generation using desktop electrolysers introduces additional hazards related to gas accumulation and operational conditions. The use of fume cupboards and adequate ventilation is identified as a primary control measure to mitigate these risks [4]. Manufacturer guidelines further emphasise correct filling procedures, system monitoring, and adherence to operational limits during hydrogen refilling [1]. Temperature control is also critical, as hydrogen cartridges must not be exposed to temperatures exceeding 50°C to prevent ignition risks [4], consistent with manufacturer recommendations for safe storage and operation [1].

Electrical hazards are generally limited due to low-voltage operation; however, improper use of equipment or poor cable management may lead to accidents, including electric shock or trip hazards [4]. Water handling during electrolysis also introduces risks of spills and slips, requiring appropriate laboratory practices.

User interaction is a key factor influencing safety outcomes. Risk assessments highlight that students, staff, and other laboratory users may be exposed if procedures are not followed correctly [2]. Standard operating procedures reinforce the importance of supervision, correct handling, and adherence to defined workflows during hydrogen refilling and system operation [5]. Safe implementation therefore requires a combination of engineering controls, procedural discipline, and user training. Systems should be operated in controlled environments with appropriate ventilation, regular inspection, and adherence to manufacturer instructions [1].

## Safety Documentation Requirements (RA, COSHH, and SOP)

The development, implementation, and enforcement of risk assessments (RA), COSHH evaluations, and standard operating procedures (SOP) should be undertaken at the institutional level, with clearly defined responsibilities. Primary accountability typically lies with laboratory supervisors, course instructors, or designated technical staff responsible for practical teaching activities. These individuals must ensure that all safety documentation is prepared, validated, and aligned with institutional and regulatory requirements. In addition, users of the equipment, including students, are responsible for adhering strictly to the defined procedures and safety controls. Effective implementation therefore requires a combination of institutional oversight, competent supervision, and informed user participation [5].

## Risk Assessment (RA)

- Perform a formal risk assessment covering hydrogen generation, storage, and fuel cell operation.
- Identify hazards such as hydrogen leakage, fire risk, and electrical hazards affecting users [2].
- Evaluate risks using structured likelihood–consequence frameworks.
- Implement control measures including ventilation, equipment checks, and operational restrictions.
- Review assessments periodically and after any changes or incidents [2].

## COSHH Assessment (Hydrogen Gas)

- Classify hydrogen as a flammable gas with inhalation risks in confined spaces [4].
- Assess exposure routes, including inhalation and fire hazards.
- Ensure engineering controls such as fume cupboards for ventilation.
- Specify PPE requirements including safety glasses and lab coats [4].

- Define safe storage conditions, including temperature limits (<50°C) and appropriate storage locations [4].

## Standard Operating Procedure (SOP)

- Provide step-by-step procedures for hydrogen refilling and system operation in alignment with manufacturer instructions [1], [5].
- Require PPE use (goggles, lab coat, appropriate clothing) during operation [5].
- Define correct setup and operation of the electrolyser and hydrogen cartridge.
- Ensure monitoring during refilling and proper shutdown procedures.
- Specify safe handling, storage, and transport of equipment [5].
- Require cleaning of workspace and proper disposal of materials after use [5]

## Educational Organisation Safety Policy and Regulatory Considerations

Educational institutions intending to implement PEM fuel cell systems should establish formal safety policies aligned with institutional laboratory regulations and national health and safety requirements. The introduction of hydrogen-based technologies into teaching laboratories should be supported by documented procedures covering risk assessment, COSHH compliance, emergency response, equipment maintenance, and user training. Responsibility for implementation should be assigned to competent personnel, including laboratory supervisors, technical staff, and course leaders. Institutions should ensure compliance with applicable safety frameworks such as ISO/TR 15916 for hydrogen systems and relevant national regulations governing hazardous substances and laboratory operation [3], [6]. Access to equipment should be restricted to trained users, and all activities should be conducted under appropriate supervision. Safety documentation should be reviewed periodically and updated following incidents, procedural changes, or equipment modifications. In addition, organisations should maintain records of inspections, training activities, and maintenance schedules to support traceability, accountability, and continuous safety improvement within educational environments.

## Conclusion

Educational PEM fuel cell systems offer a valuable and accessible platform for introducing hydrogen-based energy technologies within engineering and science education. Despite their relatively low power output, these systems present important safety considerations associated with hydrogen generation, storage, and handling. The findings of this study demonstrate that operational procedures, user competence, and environmental controls have a greater influence on overall safety than system power rating alone. By integrating manufacturer guidance, institutional risk assessments, COSHH evaluations, and standard operating procedures, a practical safety framework has been developed for educational implementation. The study highlights the importance of structured supervision, adequate ventilation, regular equipment inspection, and formal safety documentation in reducing operational risks. The proposed approach supports the safe and reliable integration of PEM fuel cell technologies into academic laboratories while aligning with established hydrogen safety principles and regulatory expectations [3], [6]. This framework can assist educational organisations in adopting sustainable energy technologies responsibly and in developing safe learning environments for future hydrogen and fuel cell education.

## References

[1] Horizon Educational, “HYDROFILL Pro User Manual and Safety Guidelines,” Horizon Fuel Cell Technologies, Singapore. [Online]. Available: https://www.cienytec.com/PDF/horizon-hydrofill-pro-user-manual.pdf. [Accessed: May 3, 2026].

[2] M. Chizari, “Risk Assessment – Water Electrolysing Use of HYDROFILL PRO,” Internal Laboratory Risk Assessment Report, SPECS, University of Hertfordshire, Hatfield, U.K., 2025.

[3] International Organization for Standardization, ISO/TR 15916:2015, Basic Considerations for the Safety of Hydrogen Systems. Geneva, Switzerland: ISO, 2015. doi: 10.3403/30330317.

[4] M. Chizari, “COSHH Risk Assessment – Hydrogen Gas Use with HYDROFILL PRO,” Internal COSHH Assessment Report, SPECS, University of Hertfordshire, Hatfield, UK., 2025.

[5] M. Chizari, “Standard Operating Procedure for Refilling HYDROSTIK PRO Using HYDROFILL PRO,” Internal Standard Operating Procedure (SOP), SPECS, University of Hertfordshire, Hatfield, UK., 2025.

[6] National Fire Protection Association, NFPA 2: Hydrogen Technologies Code. Quincy, MA, USA: NFPA, 2023.

https://doi.org/10.18745/pb.27307

# Modelling Delamination using J-Integral in Smoothed Particle Hydrodynamics

**Fiona Jepleting[1*], Tom De Vuyst[1], Yiding Liu[1], Bhishm Dewangan[1] and Faustine Ondore[2]**

*[1]University of Hertfordshire*

*[2]Technical University of Kenya*

*Corresponding author: f.m.jepleting@herts.ac.uk

**Delamination and interfacial debonding can severely degrade stiffness and load-bearing capacity of bonded structures. This work implements a J-integral-based delamination formulation within a Total-Lagrangian Smoothed Particle Hydrodynamics (TLSPH) framework providing an energy-driven description of crack initiation and growth. A Mode I Double Cantilever Beam (DCB) benchmark is simulated and compared against a finite element tied contact. The SPH formulation reproduces the finite element J-integral values within approximately 8.4% and captures the dominant crack-tip stress fields. These results demonstrate that the proposed approach offers a robust basis for mesh-free modelling of delamination in future composite applications.**

Keywords: Delamination; J-integral; Smoothed Particle Hydrodynamics (SPH); Finite Element Method (FEM), fracture mechanics.

## Introduction

Delamination and interfacial debonding are critical failure mechanisms in many layered and bonded structures as they often initiate at weak interfaces and propagate under loading, leading to progressive loss of stiffness and ultimately structural failure. Because they are governed by the initiation and growth of cracks along interfaces, delamination and debonding are therefore naturally analysed within the framework of fracture mechanics. Classical fracture mechanics provides parameters such as the stress intensity factor $K$, the strain-energy release rate $G$, and the $J$-integral to quantify the conditions required for crack growth and structural failure [1], [2].

The $J$-integral, introduced by Rice[2], offers a path-independent measure of the crack-tip driving force. For two-dimensional, quasi-static problems it can be written in contour form as

$$J = \int_{\Gamma} \left[ U n_1 - \sigma_{ij} n_i \frac{\partial u_j}{\partial x_1} \right] \mathrm{d}s,$$

where $U$ is the strain-energy density, $\sigma_{ij}$ is the Cauchy stress tensor, $u_j$ is the displacement vector, $n_i$ is the outward unit normal to the contour $\Gamma$, and $x_1$ denotes the direction of crack growth. For conservative material systems in the absence of crack-face tractions, body forces, and thermal strains, $J$ is path independent and may be evaluated on any contour enclosing the crack tip [2], [3]. For isotropic linear elastic materials, the equality $G = J$ links this numerical quantity directly to the strain-energy release rate, providing an energetic measure for Mode I delamination and interfacial debonding [1], [2].

Numerical methods enable the evaluation of the stress and deformation fields required for fracture assessment. Within FEA, delamination is modelled using Cohesive Zone Models (CZM), which employs traction-separation law, Virtual Crack Closer Technique (VCCT), which computes energy release rates from nodal forces at the crack tip, or Extended Finite Element Method (XFEM) which modifies shape functions to represent discontinuities [3], [4]. However, mesh-based methods may encounter difficulties when modelling evolving discontinuities, large deformations, or complex crack paths, which are characteristic of delamination and debonding problems.

Smoothed Particle Hydrodynamics (SPH) is a mesh-free method that offers inherent advantages for modelling fracture and failure, as displacement discontinuities and crack growth can be represented without remeshing [5], [6]. In particular, the Total-Lagrangian SPH formulation, in which all kernel operations are evaluated in the reference configuration, has been shown to improve numerical stability in tension-dominated problems and to provide a robust framework for solid mechanics and fracture applications [7]. Within this context, delamination

https://doi.org/10.18745/pb.27307

behaviour is analysed using an energy-based fracture mechanics perspective. The present work employs the $J$-integral as a fracture parameter, allowing the energetic driving force for crack growth to be quantified directly from the computed stress and deformation fields.

## Numerical Simulation

In this work, the contour formulation of the $J$-integral is transformed into an equivalent domain(area) integral through an introduction of a smooth weighting function $q_1$, the $J$-integral can be written as

$$J = \int_A \left[\sigma_{ij} \frac{\partial u_i}{\partial x_1} - U\, \delta_{1j}\right] \frac{\partial q_1}{\partial x_j}\, \mathrm{d}A,$$

where $A$ is a finite region surrounding the crack tip and $\delta_{1j}$ is the Kronecker delta [3], [4]. This domain formulation avoids the need for explicitly defined integration contours and is particularly well suited to mesh-free discretisations.

Within a Total-Lagrangian SPH framework, the domain integral is evaluated as a discrete summation over particles contained within the $J$-integral region [7], [8].

$$J = \sum_{a \in A} V_a \left[\sigma_{ij} \frac{\partial u_i}{\partial x_1} - U\, \delta_{1j}\right]_a \left\langle \frac{\partial q_1}{\partial x_j} \right\rangle_a ,$$

where $V_a$ denotes the particle volume and $\left\langle \frac{\partial q_1}{\partial x_j} \right\rangle_a$ is the gradient of the weighting function evaluated at particle $a$. This formulation enables the $J$-integral to be computed directly from SPH-derived stress and displacement fields without requiring precise crack-tip localisation.

To assess the performance of the proposed delamination algorithm with the $J$-integral formulation, a Mode I Double Cantilever Beam (DCB) configuration shown in Fig 1a and b is analysed and compared against a finite element (FEM) model. The specimen has a total length of 125 mm and a uniform thickness of 2.5 mm and is modelled as a two-dimensional isotropic material with Young's modulus $E = 210$ GPa, Poisson's ratio $\nu = 0.3$, and density $\rho = 7800$ kg/m³. A pre-crack of length 62.5 mm is introduced at the mid-plane of the beam. In the FEM model, delamination is represented using a tied contact between the interfaces.

The $J$-integral is evaluated using a domain-based formulation with a Wendland C2 weighting function for $q_1$. The SPH discretisation uses 19 particles through the thickness of the beam, while the reference FEM model uses 20 finite elements through the thickness, ensuring comparable spatial resolution in the vicinity of the crack tip. Both models employ the same pre-crack length to allow a consistent comparison of the crack-tip fields.

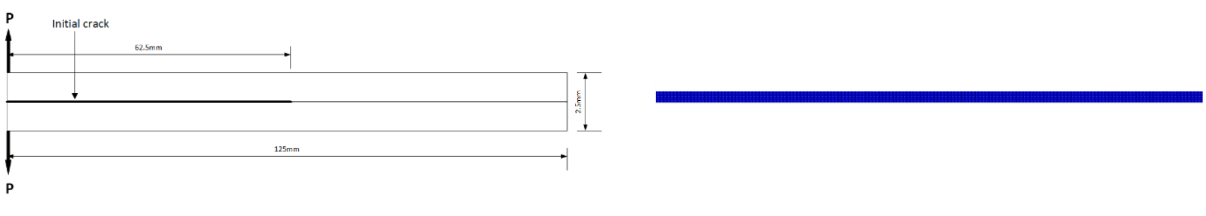



*Fig. 1. a) DCB configuration b) SPH DCB representation*

## Results and discussion

Qualitative comparison of the $\sigma_{xx}$ stress contours from the FEM and SPH simulations (Fig. 2a and b) shows very similar overall stress distributions and localisation of the tensile stress field ahead of the crack tip. The SPH formulation captures the dominant crack-tip stress features observed in the FEM reference solution.

https://doi.org/10.18745/pb.27307

Quantitative comparison of the $J$-integral values show the FEM model predicts 2230 J/m², while the SPH simulation yields 2042 J/m²—a percentage difference of 8.4% using 220 particles in the support domain. Given SPH's mesh-free nature and coarse through-thickness discretisation (19 particles), this agreement is good. The combined agreement in both the crack-tip stress distribution and the $J$-integral values demonstrates that the proposed SPH-based formulation provides a robust and physically consistent estimate of the strain-energy release rate for Mode I delamination.

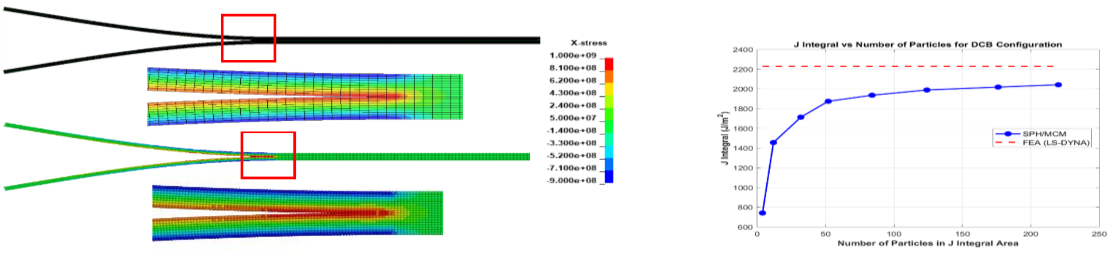



*Fig. 2 $\sigma_{xx}$ stress contours: (a) FEM reference and SPH simulation, (b) Comparison of J-integral plots.*

## Conclusion

A $J$-integral based delamination formulation has been implemented within a Total-Lagrangian Smoothed Particle Hydrodynamics framework and validated against a Mode I Double Cantilever Beam benchmark. The method enables the strain-energy release rate to be computed directly from SPH-derived stress and deformation fields, providing an energy-based measure of the crack-tip driving force without the need for remeshing or explicitly tracked contours. The results show that the SPH model reproduces finite element $J$-integral values within approximately 8.4% and accurately captures the crack-tip stress fields. These findings demonstrate that the proposed SPH $J$-integral formulation offers a robust basis for energy-driven delamination modelling and provides a natural pathway toward progressive delamination growth, mixed-mode fracture, and delamination in composite laminates.

## References


[1] T. L. Anderson, *Fracture Mechanics: Fundamentals and Applications*, 4th ed. Florida, CRC Press, 2017.

[2] J. R. Rice, "A path independent integral and the approximate analysis of strain concentration by notches and cracks," *J. Appl. Mech.*, vol. 35, pp. 379–386, 1968.

[3] C. F. Shih, B. Moran, and T. Nakamura, "Energy release rate along a three-dimensional crack front in a thermally stressed body," *Int. J. Fract.*, vol. 30, pp. 79–102, 1986.

[4] F. Z. Li, C. F. Shih, and A. Needleman, "A comparison of methods for calculating energy release rates," *Eng. Fract. Mech.*, vol. 21, pp. 405–421, 1985.

[5] L. B. Lucy, "A numerical approach to the testing of the fission hypothesis," *Astron. J.*, vol. 82, pp. 1013–1024, 1977.

[6] J. J. Monaghan, "Smoothed particle hydrodynamics," *Reports on Progress in Physics*, vol. 68, pp. 1703–1759, 2005.

[7] R. Vignjevic, J. R. Reveles, and J. Campbell, "SPH in a total Lagrangian formalism," *Computational Methods in Engineering Science*, vol. 14, no. 3, pp. 181–198, 2006.

[8] T. De Vuyst, R. Vignjevic, N. Djordjevic, M. Gintalas, and K. Hughes, "Fracture Mechanics in Smoothed Particle Hydrodynamics: An algorithm to calculate the J-Integral," *Eng. Fract. Mech.*, vol. 333, p. 111833, 2026.

https://doi.org/10.18745/pb.27307

# Early Comparative Evaluation of Transformer Models for Multilingual Software Vulnerability Detection

**Fiza Naseer[1*], Javad Khan[1,] Muhammad Yaqoob[1], and Alexios Mylonas[1]**

*School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK*

Corresponding author: f.naseer@herts.ac.uk, j.a.khan@herts.ac.uk, m.yaqoob3@herts.ac.uk, a.mylonas@herts.ac.uk.

## Abstract

**Software vulnerability detection is increasingly important as modern applications combine multiple programming languages. This paper presents an early comparative evaluation of BERT, RoBERTa, and CodeBERT for binary vulnerability detection across HTML, Python, JavaScript, and PHP using the CVEFixes dataset and language-wise three-fold stratified cross-validation. The results show clear performance differences across languages, indicating that multilingual vulnerability detection requires more language-aware and robust transformer-based modelling strategies.**

**Keywords:** Software vulnerability detection; Transformer models; CodeBERT; Multilingual code; Binary classification

## Introduction

Software vulnerabilities are a significant source of security incidents, financial damage, and increased software maintenance effort. Their impact has become more severe as software is increasingly deployed in critical sectors such as healthcare [2] and finance [8], where vulnerabilities can lead to privacy violations, service disruption, unauthorised access, and substantial economic loss [10]. The Common Weakness Enumeration (CWE) catalogue contains 943 weaknesses in version 4.17 [1], demonstrating the breadth of security issues in modern software systems.

To address these risks, researchers have proposed a range of vulnerability detection techniques, including static, dynamic, and hybrid analysis, as well as machine learning and deep learning approaches [9]. Traditional program-analysis techniques are often affected by high false-positive rates, limited code coverage, or poor practical efficiency [9]. Learning-based methods can improve automation and reduce manual effort; however, traditional machine learning approaches rely heavily on handcrafted features, while many deep learning methods still perform vulnerability prediction at a coarse-grained file or function level [11].

More recently, transformer-based models have shown promising results in the field of software vulnerability detection because of their ability to capture contextual information and long-range dependencies in source code [ 6, 7, 12, 14]. Many researchers have reported encouraging results using transformer architectures for vulnerability-related tasks [3, 14]. Nevertheless, most existing work focuses on one or only a few programming languages, for example C/C++, which limits applicability in multilingual software environments [5]. Since modern software systems increasingly combine multiple programming languages, effective vulnerability detection methods must generalise across different syntax, semantics, and coding patterns. Although several studies have explored multilingual settings, they remain limited in terms of language coverage and achieved performance, which calls for better generalisation [4]. Inspired by this gap, we present an early language-wise comparison of BERT, CodeBERT, and RoBERTa for vulnerability detection in HTML, Python, JavaScript, and PHP.

The contributions of this study are as follows:

- An early multilingual transformer-based pipeline for software vulnerability detection.
- A language-wise comparison of BERT, CodeBERT, and RoBERTa.
- Empirical evidence that transformer-model performance varies substantially across programming languages.

https://doi.org/10.18745/pb.27307

## Experimental/Simulation

The study views software vulnerability detection as a binary classification problem, as shown in Figure 1, where each source-code sample is labelled as vulnerable or safe. The CVEFixes dataset [13] contains source-code samples from multiple programming languages. This study focuses on HTML, Python, JavaScript, and PHP.

The selected subset of the CVEFixes dataset contains source-code samples from HTML, Python, JavaScript, and PHP. For this study, three fields were used: code, language, and safety. The safety field was treated as the binary class label, with vulnerable samples encoded as 1 and safe samples encoded as 0. After preprocessing, the final subset contained 285 HTML samples, including 143 vulnerable and 142 non-vulnerable samples (50.18% and 49.82%); 1,564 Python samples, including 782 vulnerable and 782 non-vulnerable samples (50.00% each); 1,562 JavaScript samples, including 781 vulnerable and 781 non-vulnerable samples (50.00% each); and 5,589 PHP samples, including 2,795 vulnerable and 2,794 non-vulnerable samples (50.01% and 49.99%). This near-balanced distribution reflects the selected experimental subset used in this study and may not represent the complete class distribution of the full CVEFixes dataset.

HTML, Python, JavaScript, and PHP were selected because they are commonly used in web-based and server-side software systems. JavaScript and PHP are frequently used in dynamic web applications. Python is widely used for backend services, automation, and data-driven applications, while HTML appears in web-facing and template-related code. The selected languages also differ in syntax, structure, and typical vulnerability patterns, which makes them suitable for an early language-wise comparison. However, they do not represent the full range of programming languages. Therefore, the findings should be interpreted as an early multilingual evaluation rather than a fully generalisable multilingual vulnerability detection benchmark.

During preprocessing, incomplete records and missing or empty code samples were removed. Only the code, language, and safety fields were retained for the final experiments. Source-code samples were tokenised using a maximum sequence length of 512 tokens. BERT, RoBERTa, and CodeBERT were evaluated with their corresponding tokenisers. Training was performed using language-wise three-fold stratified cross-validation to preserve class distribution across folds. The main hyperparameters included learning rate, batch size, number of epochs, and weight decay. Fine-tuning was implemented using the Hugging Face Trainer framework, and performance was measured using accuracy, precision, recall, and F1-score.

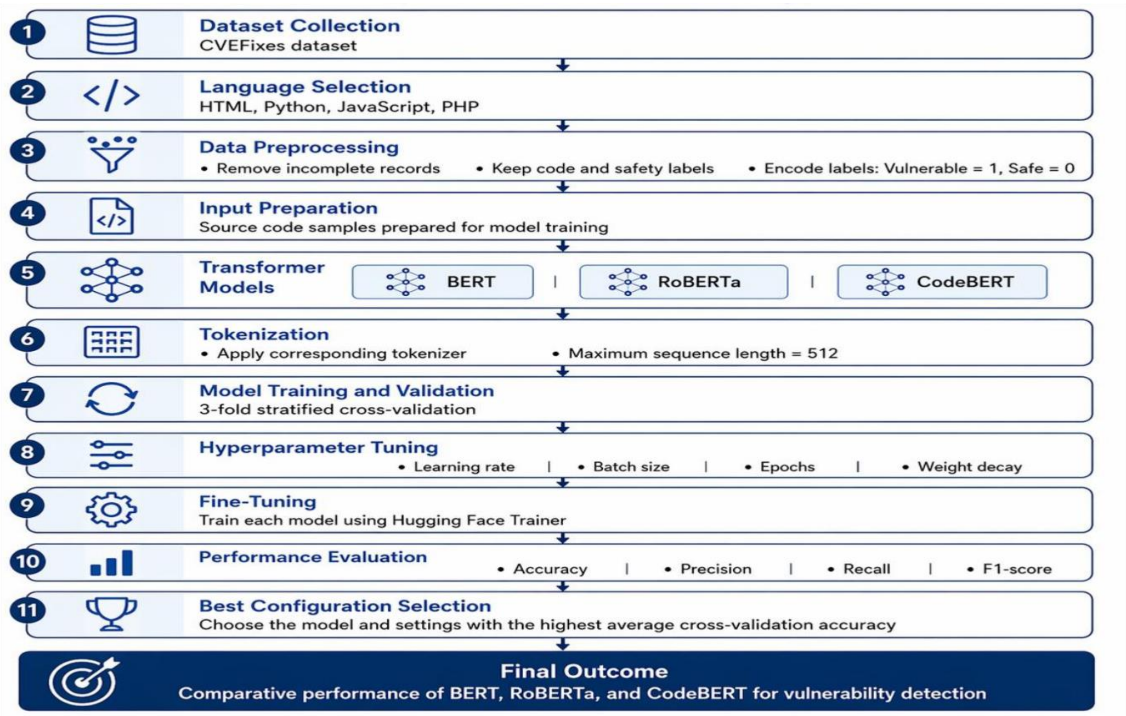



*Figure 1. Methodology Diagram*

https://doi.org/10.18745/pb.27307

## Results and Discussion

Table 1 summarises the early language-wise results for BERT, RoBERTa, and CodeBERT. The results show that transformer performance varies substantially across programming languages. The strongest overall performance was obtained for HTML, where CodeBERT achieved an F1-score of 0.7171. BERT and RoBERTa also performed relatively well on HTML, with F1-scores above 0.70. In contrast, Python, JavaScript, and PHP produced lower F1-scores, all below 0.45, suggesting that these languages remain more challenging for the evaluated models in the current experimental setting.

*Table 1. Early language-wise results for BERT, RoBERTa, and CodeBERT.*

| Language | Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|
| HTML | BERT | 0.7208 | 0.7673 | 0.7208 | 0.7064 |
| HTML | RoBERTa | 0.7246 | 0.7854 | 0.7246 | 0.7078 |
| HTML | CodeBERT | 0.7359 | 0.8208 | 0.7359 | 0.7171 |
| Python | BERT | 0.5473 | 0.7563 | 0.5473 | 0.4314 |
| Python | RoBERTa | 0.5473 | 0.7454 | 0.5473 | 0.4330 |
| Python | CodeBERT | 0.5480 | 0.7293 | 0.5480 | 0.4367 |
| JavaScript | BERT | 0.5333 | 0.7354 | 0.5333 | 0.4060 |
| JavaScript | RoBERTa | 0.5397 | 0.7042 | 0.5397 | 0.4289 |
| JavaScript | CodeBERT | 0.5352 | 0.7044 | 0.5352 | 0.4177 |
| PHP | BERT | 0.5212 | 0.6379 | 0.5212 | 0.4036 |
| PHP | RoBERTa | 0.5182 | 0.7501 | 0.5182 | 0.3728 |
| PHP | CodeBERT | 0.5196 | 0.7498 | 0.5196 | 0.3758 |

The results indicate that no single model consistently outperforms the others across all languages. CodeBERT achieved the best F1-score for HTML and Python, RoBERTa achieved the highest F1-score for JavaScript, and BERT achieved the strongest F1-score for PHP. This variation suggests that vulnerability patterns are language dependent and that multilingual vulnerability detection may require language-specific tuning, improved class-imbalance handling, and richer source-code representations. These findings also support the need for future studies using broader datasets and more robust evaluation measures, including vulnerable-class recall, macro F1-score, Matthews correlation coefficient, and PR-AUC.

## Conclusion

This paper presented an early comparative evaluation of BERT, RoBERTa, and CodeBERT for multilingual software vulnerability detection across HTML, Python, JavaScript, and PHP. The findings show that transformer-based models can support binary vulnerability detection, but their performance varies considerably across programming languages. The best results were achieved on HTML, while Python, JavaScript, and PHP remained more challenging in the current experimental setting. These differences may be influenced by language syntax, coding patterns, dataset quality, and class distribution. Overall, the results suggest that transformer models may not generalise equally well across all programming languages. Future work will expand the language coverage, investigate mixed-language and language-specific training strategies, apply stronger imbalance-handling techniques, and evaluate additional transformer-based models using more robust vulnerability-focused metrics.

## References

[1] CWE, “CWE vulnerabilities,” Accessed: Nov. 18, 2025. [Online]. Available: https://cwe.mitre.org/.
[2] J. E. Ameh, A. Otebolaku, A. Shenfield, and A. Ikpehai, “C3-vulmap: A dataset for privacy-aware vulnerability detection in healthcare systems,” Electronics, vol. 14, no. 13, Art. no. 2703, 2025.
[3] A. Bahaa, A. E.-R. Kamal, H. Fahmy, and A. S. Ghoneim, “DB-CBIL: A DistilBERT-based transformer hybrid model using CNN and BiLSTM for software vulnerability detection,” IEEE Access, 2024.
[4] Y. Cao and Y. Dong, “Multi-source cross-domain vulnerability detection based on code pre-trained model,” Information and Software Technology, Art. no. 107764, 2025.

[5] Y. Chen and Z. Liu, “HLT: A hierarchical vulnerability detection model based on transformer,” in Proc. 4th Int. Conf. Data Intelligence and Security (ICDIS), 2022, pp. 50–54.
[6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proc. Conf. North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019, pp. 4171–4186.
[7] Z. Feng et al., “CodeBERT: A pre-trained model for programming and natural languages,” arXiv preprint arXiv:2002.08155, 2020.
[8] C. Gao, W. Yang, J. Ye, Y. Xue, and J. Sun, “sguard+: Machine learning guided rule-based automated vulnerability repair on smart contracts,” ACM Transactions on Software Engineering and Methodology, vol. 33, no. 5, pp. 1–55, 2024.
[9] S. M. Ghaffarian and H. R. Shahriari, “Software vulnerability analysis and discovery using machine-learning and data-mining techniques: A survey,” ACM Computing Surveys, vol. 50, no. 4, pp. 1–36, 2017.
[10] N. S. Harzevili, A. B. Belle, J. Wang, S. Wang, Z. M. Jiang, and N. Nagappan, “A systematic literature review on automated software vulnerability detection using machine learning,” ACM Computing Surveys, vol. 57, no. 3, 2025.
[11] B. Wu and F. Zou, “Code vulnerability detection based on deep sequence and graph models: A survey,” Security and Communication Networks, vol. 2022, Art. no. 1176898, 2022.
[12] X. Zhang, F. Zhang, B. Zhao, B. Zhou, and B. Xiao, “Vuldtransformer: Source code vulnerability detection via transformer,” in Proc. 14th Asia-Pacific Symposium on Internetware, 2023, pp. 185–193.
[13] G. P. Bhandari, A. Naseer, and L. Moonen, “CVEfixes: Automated collection of vulnerabilities and their fixes from open-source software,” in Proc. 17th Int. Conf. Predictive Models and Data Analytics in Software Engineering (PROMISE ’21), 2021, pp. 1–10.
[14] F. Naseer, J. A. Khan, M. Yaqoob, A. Mylonas, and I. Gambo, “A Systematic Literature Review for Transformer-Based Software Vulnerability Detection,” arXiv preprint arXiv:2604.24822, 2026. https://doi.org/10.48550/arXiv.2604.24822

https://doi.org/10.18745/pb.27307

# Preliminary Comparison between Granger causality, delay Symbolic Transfer Entropy and Cross Permutation-entropy in assessing directionality of interaction between physiological signals

**Deepak Panday[1*], Harikala Kandel[2], Tony Steffert[3], and David Mayor[4]**

[1] *Lecturer, School of Physics, Engineering and Computer Science, UH*

[2] *School of Computing and Mathematical Sciences, Birkbeck, University of London*

[3] *Chief Science Officer, SONA, Tunbridge Wells*

[4] *Visiting Research Fellow (Physiotherapy), School of Health and Social Work, UH*

*Corresponding author: d.panday@herts.ac.uk

**Abstract**

**Granger causality (GC), delay Symbolic Transfer Entropy (dSTE) and Cross Permutation-entropy (XPE) were used to explore directionality of causal interaction between cardiac activity (ECG) or respiration (RESP) and the involuntary 3-dimensional head movements observed in a study at the University of Hertfordshire investigating the effects of transcutaneous electroacupuncture stimulation (TEAS) frequency on human physiology. GC was the most informative method used, but dSTE may provide more detail on interactions with the different dimensions (axes) of 3D head movement. No method indicated unequivocally which dimensions of HM were more affected by breathing (RESP) or cardiac activity (ECG). Further research is planned.**

Keywords: Causal interactions; Granger causality; delay Symbolic Transfer Entropy; Cross Permutation-entropy; Head movement

## Introduction

Granger causality (GC) [1,2], the related cross-frequency causal interaction measure 'delay Symbolic Transfer Entropy' (dSTE) based on vector quantisation [3,4], and Cross Permutation-entropy (XPE) [5] are all methods which have been used to assess causal interaction (directionality of information flow) between physiological signals.

Here, they were calculated for datasets collected during a study at UH investigating the effects of different frequencies of transcutaneous electroacupuncture stimulation (TEAS) on the brain (EEG), heart (ECG) and other aspects of human physiology [6-9]. Our objectives for this conference presentation were (1) to investigate causal interactions between the ECG and involuntary head movements (3D HM) observed during data collection and (2) those between Respiration (RESP) and 3D HM. A third objective was to assess, if possible, which of the three methods used provided the most informative results.

Future objectives are to assess causal interactions (1) within the EEG, and (2) those – if any – between the EEG and participants' involuntary head movements. As a comparison, CG, dSTE and XPE will also be used to re-examine the known interactions between ECG and RESP. In addition, Welch's periodogram method will be used to explore whether predominant ECG, Respiration and EEG frequencies occur in the 3D HM data (or vice versa) – a much-simplified (and very possibly simplistic) alternative to the cross-spectral methods employed by Granger and Dimitriadis.

## Materials and Methods

64 participants originally enrolled in the study, with analysable data from 48. Demographics, data collection and preprocessing methods are described elsewhere [6-9]. Head movement data were gathered using two head-mounted sensors – initially a 1-dimensional Pico Micro Movement Sensor (Unimed Electrodes, Farnham, Surrey, UK), and then a 3-dimensional sensor developed in the UH School of Physics, Engineering & Computer Science:

https://doi.org/10.18745/pb.27307

3D HM data were available from 33 participants. Because of technical issues (mostly practical problems with data recording), numbers of recordings were further reduced. Eventually, for both ECG/3D HM and RESP/3D HM, after removing inadvertently duplicated results, 3328 rows of data were analysed for GC, dSTE and XPE (832 for each of the 3 head movement dimensions, plus 832 for the resultant composite Euclidean displacement). Data were resampled so that sample rates were consistent at 250 Hz across all data types (ECG, RESP, HM). Data were not normalised, but were resampled so that sample rates were consistent at 250 Hz across the three data types. Stationarity was confirmed using the Augmented Dickey-Fuller (ADF) test for 98.9% ECG, 7.9% RESP and 87.6-96.4% HM data.

For GC computation, lags were used corresponding approximately to timescales suggested for reflex postural coupling (≤0.35 s) [10], voluntary motor coordination (≤2 s) [11] and postural adaptation (≤7.5 s) [12]. For a test sample of 300 files of ECG and 3D head movement data, Shapiro-Wilk tests indicated that both GC and the computation times required to calculate GC were not normally distributed. Median run times for the three lags were of the order of 0.12, 4.53 and 130.73 s per file (IQR 0.07-0.15, 3.05-5.14 and 115.09-145.26 s, respectively). To lessen computing time, 'postural adaptation' lags range were not examined further.

MATLAB® 2025b [Mathworks, Natick, MA] and Excel® LTSC MSO (V 2408) [MicroSoft , Redmond, Washington] were used for all computations. For GC, the MATLAB function Granger_Cause_1 was used [13], code for dSTE was taken from GitHub [14], and that for XPE from EntropyHub [15].

Following the GC, dSTE and XPE calculations, results for the spline-interpolated data were compared in IBM SPSS Statistics v. 31.0.0.0 (117) [International Business Machines, Armonk, NY], using Pearson's *R* and point-biserial correlation methods.

## Results and discussion

GC is considered to 'predict' y from x based on values of the F-statistic being greater than a certain critical value (here, 0.05). Counts were made of such occurrences, for both xày and yàx. In our initial ECG/3D HM test for GC with 300 sample results, greatest differentiation between GC (ECàHM) and GC (HMàECG) was found in the midrange timescale. Subsequent analysis of the whole dataset confirmed this. For the RESP/3D HM interaction, in contrast, greatest differentiation between GC (RESPàHM) and GC (HMàRESP) was found for the shorter timescale. Disappointingly, GC did not really indicate which HM axes (directions) were more affected by breathing (RESP) or cardiac activity (the ECG).

For dSTE, values of vector 'quantisation' between x and y based on a neural-gas network [16] could be positive or negative, the former indicating direction of information flow xày, and the latter yàx. Here, such positive and negative values were counted, and their medians calculated. These appear to indicate that displacements along the Z (up/down) HM axis were particularly affected by the ECG, whereas displacements along X, Y and Z axes were all similarly affected by respiration [Figure 1]. Further statistical analysis is required to confirm this.

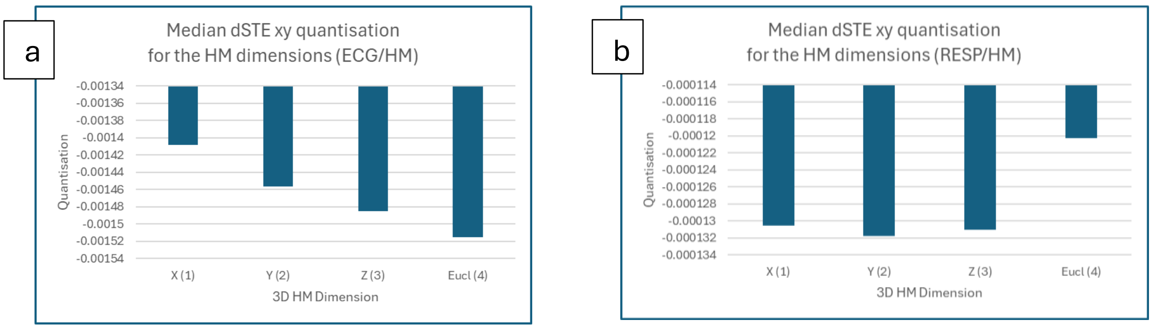



*Figure 1. Median dSTE xày quantisation appears to differ for the HM axes more for ECG/HM [a] than for RESP/HM [b]*

XPE results can be interpreted to indicate (as for dSTE) that HM may affect ECG for Y, Z and Euclidean distance, but that ECG affects HM displacements in the X (back and forth) direction. Results for RESP/HM are less clear.

## Conclusion

For ECG/HM, some strong (negative) correlations were found between dSTE and GC ($p < 10^{-60}$ and $< 10^{-80}$), and for RESP/HM, between GC and XPE ($p < 10^{-12}$) – all still significant after Bonferroni correction. Different methods of assessing causal interaction between data streams may thus be appropriate for different types of data. Granger causality is currently more generally used than the other two approaches described here, but dSTE may provide more detail on interactions with the different directions (axes) of 3D head movement. Data 'differencing' and normalisation are recommended when comparing results using GC, dSTE and XPE.

## References

[1] C.W.J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424-438, August, 1969.
[2] A. Shojaie, E.B. Fox, "Granger causality: a review and recent advances," *Annual Review of Statistics and Its Application*, vol. 9, pp. 289-319. https://doi.org/10.1146/annurev-statistics-040120-010930.
[3] S. Dimitriadis, Y. Sun, N. Laskaris, N. Thakor, A. Bezerianos, "Revealing cross-frequency causal interactions during a mental arithmetic task through symbolic transfer entropy: a novel vector-quantization approach," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 10, pp. 1017-1028, Jan, 2016. https://doi.org/10.1109/TNSRE.2016.2516107.
[4] S.I. Dimitriadis, "Universal lifespan trajectories of source-space information flow extracted from resting-state MEG data," *Brain Sciences*, vol. 12, no. 10, p. 1404, Oct, 2022. https://doi.org/10.3390/brainsci12101404.
[5] W.B. Shi, P.J. Shang, A.J. Lin, "The coupling analysis of stock market indices based on cross-permutation entropy," *Nonlinear Dynamics*, vol. 79, no. 4, pp. 2439-2447, March, 2015. https://doi.org/10.1007/s11071-014-1823-1.
[6] R. Lopes Alves, M. Zortea, D. Mayor, T. Watson, T. Steffert, "Effect of different frequencies of transcutaneous electrical acupoint stimulation (TEAS) on EEG source localization in healthy volunteers: A semi-randomized, placebo-controlled, crossover study," *Brain Sciences*, vol. 15, no. 3, p. 270. March, 2025. https://doi.org/ 10.3390/brainsci15030270.
[7] D. Mayor, T. Steffert, P. Steinfath, T. Watson, N. Spence, D. Banks, "Effects of Transcutaneous Electroacupuncture Stimulation (TEAS) on eyeblink, EEG, and Heart Rate Variability (HRV): A non-parametric statistical study investigating the potential of TEAS to Modulate physiological markers," *Sensors*, vol. 25, no. 14, p. 4468. July, 2025. https://doi.org/10.3390/s25144468.
[8] Ç. Uyulan, D. Mayor, T. Steffert, T. Watson, D. Banks, "Classification of the central effects of transcutaneous electroacupuncture stimulation (TEAS) at different frequencies: A deep learning approach using wavelet packet decomposition with an entropy estimator," *Applied Sciences*, vol. 13, no. 4, p. 2703. Feb, 2023. https://doi.org/10.3390/app13042703.
[9] T. Vasei, H. Gediya, M. Ravan, "Investigating brain responses to transcutaneous electroacupuncture stimulation: A Deep Learning approach," *Algorithms*, vol. 17, no. 11, p. 477, Oct, 2024. https://doi.org/10.3390/a17110477.
[10] J.V. Jacobs, F. Horak, "Cortical control of postural responses," *Journal of Neural Transmission*, vol. 114, no. 10, pp. 1339-1348. Oct, 2007. https://doi.org/ 10.1007/s00702-007-0657-0.
[11] M.G. Tucker, J.J. Kavanagh, R.S. Barrett, S. Morrison, "Age-related differences in postural reaction time and coordination during voluntary sway movements," *Human Movement Science*, vol. 27, no. 5, pp. 728-737, Oct, 2008. https://doi.org/10.1016/j.humov.2008.03.002.
[12] Y. Suzuki, H. Geyer, "A neuro-musculo-skeletal model of human standing combining muscle-reflex control and virtual model control," *Annual International Conference of the IEEE Engineering in Biology and Medicine Society*, vol.2018, pp. 5590-5593, July, 2018. https://doi.org/10.1109/EMBC.2018.8513543.
[13] Robert, "Granger_Cause_1," Mathworks. [Online]. Available: https://uk.mathworks.com/matlabcentral/fileexchange/59390-granger_cause_1. [Accessed: 30 April 2026].
[14] S. Dimitriadis, "Symbolic_Transfer_Entropy," GitHub. [Online]. Available: https://github.com/stdimitr/Symbolic_Transfer_Entropy. [Accessed 30 April 2026].
[15] M.W. Flood, B. Grimm, "EntropyHub: An open-source toolkit for entropic time series analysis," *PloS ONE*, vol. 16, no. 11, p. e0259448, Nov, 2021. https://doi.org/10.1371/journal.pone.0259448.
[16] T.M. Martinetz, S.G. Berkovich, K.J. Schulten, "'Neural-gas' network for vector quantization and its application to time-series prediction," *IEEE Transactions on Neural Networks*, vol. 4, no. 4, pp. 558-569, 31 July, 1993.

University of Hertfordshire UH
https://doi.org/10.18745/pb.27307

# Cross-Domain Tuberculosis Detection in Chest X-rays

**Rimsha Javed*[1], Muhammad Yaqoob[1], and Volker Steuber[1]**
[1]*School of Physics, Engineering and Computer Science, University of Hertfordshire, UK*
*Corresponding author: r.javed3@herts.ac.uk

**A key challenge in cross-domain generalization for automated tuberculosis (TB) detection in chest X-rays is that models often learn dataset-specific features rather than disease relevant characteristics. In this paper, the cross-domain performance of an ImageNet-pretrained DeiT-tiny vision transformer for TB classification is evaluated. The model trained on the Shenzhen dataset performs well on in-domain data but demonstrates reduced performance on external datasets, including Montgomery and TBX11K due to domain shift. Although the model exhibits moderate transferability across datasets, the results highlight limitations in robustness and the control of false-positive predictions. These findings suggest that improved domain adaptation, feature alignment, and calibration strategies are required to achieve clinically reliable and generalizable TB detection according to different imaging conditions.**

Keywords: Tuberculosis; X-ray Imaging; domain adaptation; deep learning; computer vision

## Introduction

Chest X-ray interpretation by expert radiologists is expected to generalize across clinical settings. A radiologist can assess tuberculosis-related abnormalities regardless of whether the image was acquired in a different hospital, scanner system, demographic group, age category, or patient population. In contrast, many machine learning models perform well only when training and testing images come from the same dataset, however, their performance decreases on external datasets [1]. This suggests that such models may learn not only TB-relevant radiographic patterns, i.e., upper-lobe opacities, cavitation, infiltrates, fibrosis, or pleural changes, but also dataset-specific and clinically irrelevant signals, including image contrast, acquisition protocol, scanner characteristics, cropping, markers, resolution, annotation style, or population-specific biases. Therefore, there is a need to develop models that focus more specifically on disease-relevant features and remain robust across domains [2]. This work employs an ImageNet-pretrained DeiT-tiny vision transformer to assess the feasibility of cross-domain tuberculosis classification using chest radiographs

## Experimental Setup

As preliminary setup, we employed an ImageNet-pretrained DeiT-tiny [3] transformer backbone for binary pulmonary TB classification. DeiT-a lightweight vision transformer architecture that can model global image relationships for small imaging datasets. Chest X-ray images are resized to 224 × 224 pixels, normalized using ImageNet statistics, and augmented during training with random crop, horizontal flip, rotation, and color jitter. The DeiT-tiny backbone extracts image-level feature embeddings, which are passed to a custom fully connected classification head containing linear layers, batch normalization, ReLU activation, dropout, and a final single-logit output for TB versus non-TB prediction.

Training is performed on the Shenzhen TB dataset [4] using a stratified train/validation split. Specifically, a StratifiedShuffleSplit with a 0.2 validation ratio is applied, resulting in 528 training images and 133 validation images while preserving the original class distribution. The DeiT-tiny backbone is initially frozen so that only the classification head is trained; after six epochs, the backbone is unfrozen and fine-tuned using a smaller learning rate than the classification head. The model is trained with AdamW using differential learning rates, cosine annealing scheduling, focal loss with label smoothing, mixup augmentation class-balanced weighted sampling, gradient clipping, mixed-precision training, and gradient accumulation. Early stopping is monitored using validation AUC, and the best checkpoint is saved according to the highest validation AUC.

To address class imbalance during training, we employ a weighted random sampling strategy that ensures balanced mini-batches between TB and non-TB classes. Moreover, focal loss is used to emphasize harder and minority samples, further mitigating class imbalance effects during optimization.

For cross-domain evaluation, the best Shenzhen-trained checkpoint is tested on external datasets without any additional fine-tuning. The Montgomery subset [4] contains 138 images (80 normal and 58 TB). For evaluation, a balanced subset of 116 images (58 normal and 58 TB) is constructed via random under-sampling to ensure equal class representation. The TBX11K subset [5] consists of 192 images (88 normal and 104 TB) and is evaluated in its original distributed without re-sampling to reflect real world class imbalance conditions.

Decision thresholds are derived exclusively from held-out Shenzhen validation predictions and then applied unchanged to all external datasets. Performance is reported using AUC-ROC curves, precision-recall curves, and confusion matrices. We evaluate the transferability of a Shenzhen-trained DeiT-tiny TB classifier to independent chest X-ray datasets without tuning on external domains.

## Results and discussion

The DeiT-tiny model achieved the best performance at epoch 13, with a validation AUC of 0.9301 and validation loss of 0.5384 on the Shenzhen TB dataset. However, the training AUC was lower (0.7295) than the validation AUC, which may reflect the use of mix-up, data augmentation and weighted sampling during training. Therefore, the training AUC is not directly comparable to the validation AUC, which is computed on unaugmented images.

To ensure clinically meaningful decision-making, threshold calibration was performed using ROC analysis on the Shenzhen validation set. The optimal operating threshold was selected using Youden's index to maximize the trade-off between sensitivity and false positive rate. In addition, a sensitivity-constrained threshold was also derived to prioritize high TB detection rates. These calibrated thresholds were fixed and applied unchanged during external validation to ensure consistent cross-domain evaluation.

In cross-domain evaluation, performance decreased compared with Shenzhen validation, which is expected because Montgomery and TBX11K differ in acquisition source, image appearance, population, and label distribution. On the balanced Montgomery subset, the model achieved an AUC of 0.7754 and F1-score of 0.7188. The confusion matrix showed 46 out of 58 TB cases correctly detected and 34 out of 58 non-TB cases correctly classified with sensitivity of 0.79 and specificity of 0.59. This indicates reasonable TB detection ability, but a relatively high false-positive rate, with 24 non-TB cases classified as TB.

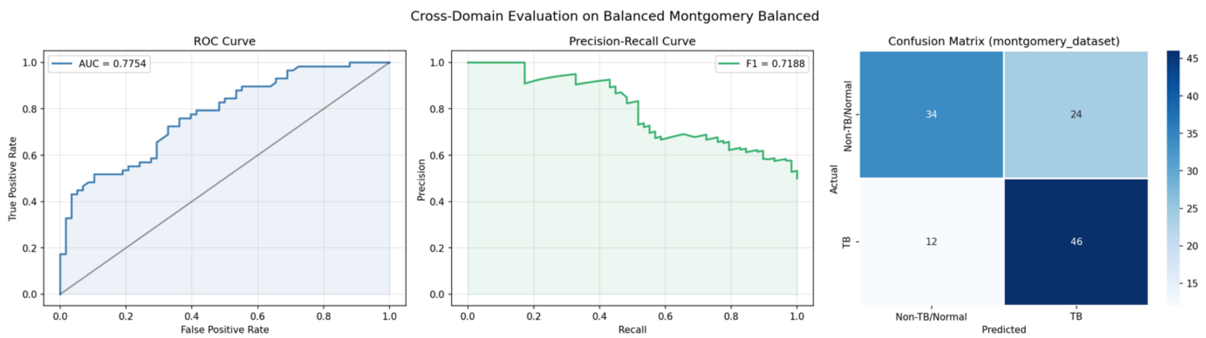



*Fig. 1. Cross-domain evaluation of the Shenzhen-trained DeiT-tiny model on the balanced Montgomery TB dataset. The figure shows the ROC curve, precision-recall curve, and confusion matrix*

Performance was better on the TBX11K subset, with an AUC of 0.8489 and F1-score of 0.7948. The model correctly identified 91 out of 104 TB cases and 54 out 88 non-TB cases with a sensitivity of 0.88 and specificity of 0.61. These preliminary results suggest that DeiT-tiny transfers better to TBX11K than to Montgomery, although specificity remains limited across both datasets. The preliminary results show that the model appears more sensitive than specific i.e., it tends to detect TB cases but also misclassifies a notable proportion of non-TB images as TB.

https://doi.org/10.18745/pb.27307

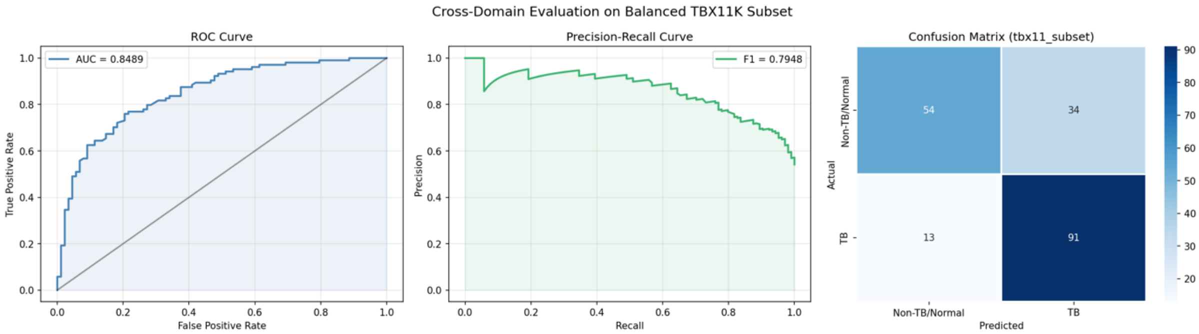



*Fig. 2. Cross-domain evaluation of the Shenzhen-trained DeiT-tiny model on the TBX11K subset. The figure shows the ROC curve, precision-recall curve, and confusion matrix*

In general, the model maintains a consistent sensitivity–specificity trade-off, throughout the domains. Sensitivity in the screening test is better than having a large number of false negatives (missed TB cases) since it would lead to a lower proportion of true positives (specificity) and a higher proportion of false positives. This is particularly the case for external data sets where various imaging protocols and data distributions lead to even more false positive predictions, resulting from domain shift.

## Conclusion

The preliminary results show that an ImageNet-pretrained DeiT-tiny model can learn useful TB-discriminative features from the Shenzhen dataset while maintaining moderate cross-domain performance on independent datasets. The high validation AUC on the Shenzhen dataset demonstrate effective source-domain learning, however, the reduced performance on Montgomery and TBX11K highlights the persistent challenge of domain shift in chest X-ray-based TB classification. Future work should focus on enhancing cross-domain robustness and reducing false positive predictions. A possible future direction could be kernel-based domain adaptation e.g., maximum mean discrepancy (MMD) or kernel alignment losses, to reduce feature-distribution mismatch between datasets. Furthermore, future work could explore feature-level alignment, calibration methods, larger external validation testing, uncertainty estimation, and explainability methods to confirm the model focus on clinically meaningful lung regions.

## References

[1]. Ravin, N., Saha, S., Schweitzer, A., Elahi, A., Dako, F., Mollura, D. and Chapman, D., 2022. Mitigating domain shift in AI-based TB screening with unsupervised domain adaptation. IEEE Access, 10, pp.45997-46013.

[2]. Çallı, E., Sogancioglu, E., Van Ginneken, B., Van Leeuwen, K.G. and Murphy, K., 2021. Deep learning for chest X-ray analysis: A survey. Medical image analysis, 72, p.102125.

[3]. Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A. and Jégou, H., 2021, July. Training data-efficient image transformers & distillation through attention. In International conference on machine learning (pp. 10347-10357). PMLR.

[4]. Jaeger, S., Candemir, S., Antani, S., Wáng, Y.X.J., Lu, P.X. and Thoma, G., 2014. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery, 4(6), p.475.

[5]. Liu, Y., Wu, Y.H., Zhang, S.C., Liu, L., Wu, M. and Cheng, M.M., 2023. Revisiting computer-aided tuberculosis diagnosis. IEEE transactions on pattern analysis and machine intelligence, 46(4), pp.2316-2332.

https://doi.org/10.18745/pb.27307

# Multimodal Brain Tumour Classification Using Feature Fusion

**Wajih ul Islam[1*], Muhammad Yaqoob[1], Javed Ali Khan[1] and Volker Steuber[1]**

[1]*School of Engineering and Computer Science, University of Hertfordshire, UK*

*Corresponding author: w.u.islam@herts.ac.uk

**Clinicians diagnose brain tumors by synthesizing patient symptoms, medical history, and quantitative imaging data from modalities such as MRI and CT scans into a unified clinical judgement. However, most deep learning models rely on MRI/CT images alone, failing to replicate the clinicians' multimodal reasoning. We explore a two-branch multimodal network combining raw MRI scans with 91 extracted radiomic features (intensity, texture, shape, and boundary descriptors) to classify brain tumors into glioma, meningioma, pituitary, and no-tumor. A pre-trained CNN backbone encodes the image stream whereas a dedicated MLP encodes the radiomic stream. Both streams are fused via concatenation, gated, or bidirectional cross-modal attention strategies. Across nine experimental runs on a balanced 7,200-image dataset, all multimodal configurations outperform unimodal baselines with gated fusion achieving the best accuracy of 96.13%.**

Keywords: Brain tumor classification; radiomic features; multimodal learning; feature fusion; medical imaging

## Introduction

In clinical practice, medical practitioners diagnose brain tumor by integrating multiple information sources i.e., symptoms, history, laboratory results, and imaging such as MRI. However, the majority of deep learning models for brain tumor classification operate on image data alone[1]. This unimodal approach to deep learning fails to replicate the multimodal reasoning in the real diagnostic workflows. To address this, we explore a two-branch multimodal framework that combines raw MRI scans with 91 radiomic features extracted per image to classify brain tumors into four categories: glioma, meningioma, pituitary, and no-tumor. A pre-trained CNN backbone (EfficientNet-B3, ResNet-50, or DenseNet-121) encodes the image stream whereas a dedicated MLP encodes the radiomic stream, with both projected to a shared 256-dimensional embedding space and fused via one of three strategies: concatenation, gated fusion with learned per-dimension modality weights, and bidirectional cross-modal attention. We performed a grid of nine experimental runs to investigate how backbone capacity and fusion strategy impact the classification performance.

## Dataset

The publicly available dataset comprises 7,200 brain MRI images distributed across four tumor categories: glioma, meningioma, pituitary tumor, and no-tumor. The dataset is perfectly balanced, with 1,600 images per class in the training set and 400 per class in the test set. The dataset was compiled from multiple publicly available MRI repositories, including the Sartaj dataset, BRaTS MRI images, and other open access sources [2,3], introducing diversity in scanner protocols, acquisition settings, and patient populations. The predefined training and testing split provided with the dataset was used in this study. Since patient identifiers were not available split should be considered imagel evel rather than patient level, and possible patient overlap between training and testing subsets cannot be completely excluded.This limitation is acknowledged in the present preliminary study.

## Experimental Setup

In the preliminary setup, radiomic features are extracted from each MRI scan and processed via a dedicated MLP to produce a radiomic embedding, and the raw MRI scan is passed through a pre-trained CNN backbone to produce a visual embedding; both are subsequently fused for classification. To extract the radiomic features, the MRI scan is resized to 224×224 pixels and passed through a three-stage unsupervised segmentation pipeline to extract three spatial regions: a brain mask extracted via CLAHE contrast enhancement and a double Otsu threshold, a tumor ROI identified by combining a global top-20% intensity threshold with a local adaptive threshold followed by morphological refinement, and a 15-pixel peritumoral ring surrounding the tumor core. From these regions, 91 radiomic features are extracted across ten families: first-order intensity statistics (15),

https://doi.org/10.18745/pb.27307

shape descriptors (12), GLCM texture (12), GLRLM run-length (8), NGTDM neighborhood tone (5), GLDM grey-level dependence (8), gradient and HOG features (12), peritumoral first-order (8), peritumoral GLCM (6), and targeted boundary/heterogeneity/fractal features (5) [4]. These numerical features are exported per image with a mask reliability flag into a CSV file.

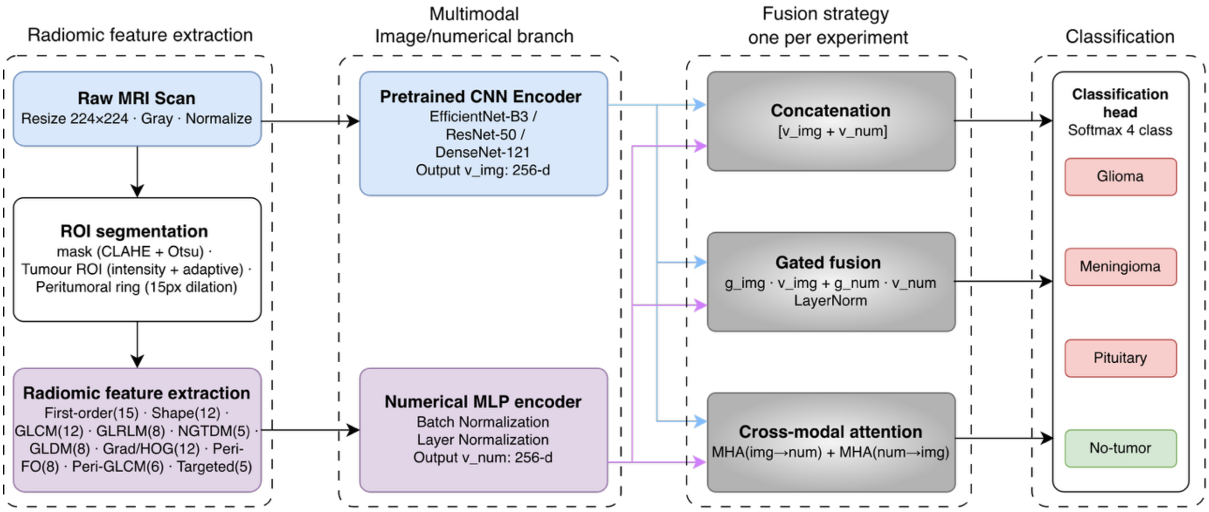



*Fig. 1. Preliminary experiment on multimodal brain tumor classification. First ROI is segmented and 91 radiomic features are extracted. Then multimodal image/numerical branches are fused via concatenation, gated, or cross-modal attention for four-class tumor prediction.*

The obtained radiomic feature vectors are processed via a numerical MLP encoder whereas the original MRI images pass through the CNN encoder. One of three pre-trained CNN backbones: EfficientNet-B3, ResNet-50, or DenseNet-121 are evaluated based on their complementary architectural designs and proven performance in medical image classification tasks [5]. The CNN outputs a 256-dimensional embedding via a dropout, linear, Layer Norm head. The radiomic features are independently encoded by a three-layer MLP (Batch Norm, Linear(128) , Linear(256), Layer Norm) to a matching 256-d embedding. The two embeddings are then fused using one of three strategies evaluated across separate runs: simple concatenation, gated fusion with dual learned sigmoid gates, or bidirectional cross-modal attention with residual connections (Fig. 1). During training the backbone is frozen for 8 epochs then fine-tuning the last three blocks for 12 epochs with weighted sampling, label smoothing, gradient clipping, and mixed precision. We performed an experiment grid of 9 runs (3 backbones × 3 fusion strategies) to evaluate the utility of multimodal fusion.

## Results and discussion

The individual CNNs established a baseline, with DenseNet-121 achieving the highest single-modality accuracy of 95.50%, slightly outperforming ResNet-50 (95.37%) and EfficientNet-B3 (94.94%), with meningioma consistently being the most challenging class across all three backbones due to its lower recall relative to the other classes (Table 1). When the radiomic branch was introduced, all multimodal configurations consistently outperformed their corresponding unimodal baselines across nine experimental runs demonstrating the robustness of multimodal feature fusion across different backbone architectures and fusion strategies. The best performing configurations were ResNet-50 + Gated fusion and DenseNet-121 + Gated fusion both achieving 96.13% followed closely by DenseNet-121 + Attention (96.06%) and DenseNet-121 + Concatenation (95.81%). Across fusion strategies gated fusion demonstrated the most consistent gains (Table 1). Moreover no-tumour and pituitary classes benefited most from multimodal fusion with near perfect recall across nearly all configurations. Meningioma remained the hardest class to classify in both unimodal and multimodal settings due to its heterogeneous appearance and morphological overlap with other tumour types. The proposed framework achieves competitive classification performance while integrating handcrafted radiomic descriptors with deep CNN image embeddings through multimodal fusion. Comparison with recent literature further

https://doi.org/10.18745/pb.27307

highlights the effectiveness of the proposed approach. Aziz et al. (2024) reported an accuracy of 96.9% using a DenseNet-based MRI classification framework, whereas the proposed multimodal CNN + radiomics fusion framework achieved 96.13% accuracy while incorporating both deep image embeddings and handcrafted radiomic descriptors.

*Table 1. Per-class F1-scores and overall accuracy for unimodal (MRI only) and multimodal (image + 91 radiomic features) across three CNN backbones and three fusion strategies.*

| Model | Type | Glioma | Meningioma | No-tumour | Pituitary | Accuracy |
|---|---|---|---|---|---|---|
| **EfficientNet-B3** | Uni | 0.897 | 0.933 | 0.976 | 0.988 | 94.94% |
| **+ Concatenation** | Multi | 0.913 | 0.932 | 0.993 | 0.988 | 95.69% |
| **+ Gated** | Multi | 0.903 | 0.932 | 0.99 | 0.982 | 95.25% |
| **+ Attention** | Multi | 0.91 | 0.932 | 0.999 | 0.986 | 95.75% |
| **ResNet-50** | Uni | 0.91 | 0.943 | 0.962 | 0.996 | 95.37% |
| **+ Concatenation** | Multi | 0.913 | 0.941 | 0.979 | 0.991 | 95.69% |
| **+ Gated** | Multi | 0.924 | 0.942 | 0.986 | 0.991 | 96.13% |
| **+ Attention** | Multi | 0.915 | 0.935 | 0.994 | 0.984 | 95.75% |
| **DenseNet-121** | Uni | 0.906 | 0.944 | 0.971 | 0.995 | 95.5% |
| **+ Concatenation** | Multi | 0.915 | 0.931 | 1 | 0.985 | 95.81% |
| **+ Gated** | Multi | 0.922 | 0.946 | 0.98 | 0.995 | 96.13% |
| **+ Attention** | Multi | 0.918 | 0.932 | 0.998 | 0.994 | 96.06% |

*Table 2. Comparison of the proposed multimodal framework with recent brain tumor MRI classification studies*.

| Related work | Model | Accuracy |
|---|---|---|
| **Rasheed et al. (2023)** | CNN with image enhancement | 95.56% |
| **Aziz et al. (2024)** | Dense Net MRI classification | 96.9% |
| **Proposed work** | CNN + radiomics fusion | 96.13% |

## Conclusion

This preliminary work proposes a multimodal framework for brain tumor classification that combines CNN-based image encoding with handcrafted radiomic features across three backbone architectures and three fusion strategies. We demonstrate that multimodal fusion consistently outperformed unimodal image only baselines

https://doi.org/10.18745/pb.27307

across all nine experimental configurations. The best performing models were ResNet-50 and DenseNet-121 with gated fusion, both achieving an accuracy of 96.13%. The current multimodal framework uses two complementary representations derived from the same MRI source: deep image embeddings and handcrafted radiomic descriptors. However, it is important to acknowledge that the observed improvement over the unimodal baselines is modest, which is largely attributable to the fact that both branches derive their information from the same MRI image. Since the radiomic features are extracted from the same source signal as the CNN input, the degree of true complementarity between modalities is limited. Future work will include a radiomics only baseline and feature family ablation studies to quantify the contribution of individual radiomic descriptors. In the future, we plan to evaluate the multimodal deep models by replacing the radiomic branch with independent clinical data sources, e.g. structured clinical reports, laboratory results, or patient history, which would introduce information entirely absent from the image domain and more accurately replicate the multimodal reasoning of clinical diagnosis.

## References

[1]. BAlsabti, S.M., Al-Gburi, R.M., Mustafa, A., AHMED, S.K., Issa, A.M., Al-Naimi, T.M., AlSaad, R. and Elhenidy, A.M., 2025. Advances in Deep Learning for Multimodal Brain Imaging: A Comprehensive Survey. *Neuroscience Informatics*, p.100252.

[2]. Rasheed, Z., Ma, Y.K., Ullah, I., Ghadi, Y.Y., Khan, M.Z., Khan, M.A., Abdusalomov, A., Alqahtani, F. and Shehata, A.M., 2023. Brain tumor classification from MRI using image enhancement and convolutional neural network techniques. *Brain Sciences*, *13*(9), p.1320.

[3]. Gómez-Guzmán, M.A., Jiménez-Beristaín, L., García-Guerrero, E.E., López-Bonilla, O.R., Tamayo-Perez, U.J., Esqueda-Elizondo, J.J., Palomino-Vizcaino, K. and Inzunza-González, E., 2023. Classifying brain tumors on magnetic resonance imaging by using convolutional neural networks. *Electronics*, *12*(4), p.955.

[4]. Zwanenburg, A., Vallières, M., Abdalah, M.A., Aerts, H.J., Andrearczyk, V., Apte, A., Ashrafinia, S., Bakas, S., Beukinga, R.J., Boellaard, R. and Bogowicz, M., 2020. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, *295*(2), pp.328-338.

[5]. Aziz, N., Minallah, N., Frnda, J., Sher, M., Zeeshan, M. and Durrani, A.H., 2024. Precision meets generalization: Enhancing brain tumor classification via pretrained DenseNet with global average pooling and hyperparameter tuning. *PloS one*, *19*(9), p.e0307825

https://doi.org/10.18745/pb.27307

# Improving Clinical Monitoring Sensitivity in Stroke Rehabilitation via Hardware-Based Signal Conditioning of a Robotic Orthosis

**Callium Dawes, Vignesh Velmurugan, Luke Wood, Farshid Amirabdollahian**
*Corresponding author: c.dawes@herts.ac.uk

## Abstract

This study presents a hardware-based signal conditioning upgrade for the WiGlove, a home-based robotic rehabilitation orthosis. To address signal noise in joint-angle measurements, a custom PCB featuring an active Sallen–Key low-pass filter was integrated into the data acquisition system. Experimental validation using 3D-printed cylinders and a custom Python logging tool demonstrated that the hardware filter reduced signal variability by approximately 41% for the index finger and 27% for the middle finger compared to software smoothing. These improvements provide the stable measurements necessary for reliable monitoring of grasp consistency and sensitive detection of functional changes during stroke recovery.

Keywords: WiGlove; Stroke Rehabilitation; Sallen–Key Filter; Signal Conditioning; Data Acquisition (DAQ)

## Introduction

The WiGlove[1] is a home-based robotic orthosis designed to assist stroke survivors in rehabilitating hand and wrist function. The device measured joint angles during training using rotary potentiometers to provide objective performance feedback. While its Data Acquisition (DAQ) system that uses a software based moving average filter facilitates reliable feedback, it is still prone to signal noise, which can negatively impact the repeatability of the measurements. Improving this is critical for the clinical utility of the device by enabling more sensitive detection of gradual functional changes in the user. This study aims to address this by implementing a hardware-based filtering solution on a custom data acquisition unit.

## Experimental/Simulation

To enhance signal stability, a custom PCB [fig.1b] was designed and integrated into the WiGlove's data acquisition electronics. The core of this upgrade is an active Sallen–Key low-pass filter [2] configuration implemented for each sensor channel using operational amplifiers. The Sallen–Key topology [3] [fig.1a] was chosen because its high input impedance isolates the potentiometers, while its low output impedance matches the sample-and-hold circuit of the microcontroller's analog-to-digital converter (ADC), preventing signal loading effects. The filter was engineered utilizing an equal-component design with resistor values of R1 = R2 = 22Kohms and capacitor values of C1 = C2 = 1uf. This establishes a cut-off frequency $f_c$ of approximately 7.23Hz given by the formula:

$$f_c = \frac{1}{2\pi RC}$$

This cut-off frequency provides a sharp 40dB/decade roll-off that attenuates high-frequency contact noise, electromagnetic interference, and structural artifacts while keeping the low-frequency DC voltage shifts associated with voluntary finger flexion and extension.

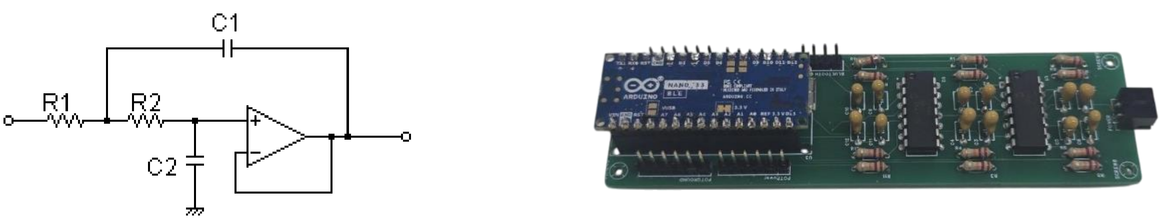



*Fig.1 (a)The design of a singular filter (b)The PCB design*

To validate this hardware, a Python-based data logging and visualization tool was developed. This tool guides users through a timed sequence of neutral, flexion, and extension wrist movements while recording data from

https://doi.org/10.18745/pb.27307

five sensors representing the thumb, fingers, and wrist. An experiment was designed using three 3D-printed cylinders of varying diameters to simulate different grasp sizes [Fig.2]:

- Small: 50 mm
- Medium: 60 mm
- Large: 70 mm

*Fig.2. Experiment configuration showing the different size*

## Results and discussion

This abstract presents the Index and Middle finger channel data to represent the performance of the overall system. Figure 3 compares the unfiltered raw signal after software moving-average smoothing with the hardware-filtered signal across the four steady grasp-hold conditions. The proposed hardware-filtered implementation resulted in lower variability than the former [Fig.3.]. An average reduction in standard deviation of approximately 41% for the Index channel and 27% for the Middle channel was observed. Applying a causal moving average to the unfiltered raw signal produced only minimal improvement, and increasing the window size failed to match hardware-filter performance without introducing substantial delay, suggesting that the observed benefit likely arose from analogue pre-digitization filtering rather than simple post hoc smoothing. These results support and validate the performance of the proposed Sallen–Key low-pass filtering upgrade in WiGlove's new data acquisition system.

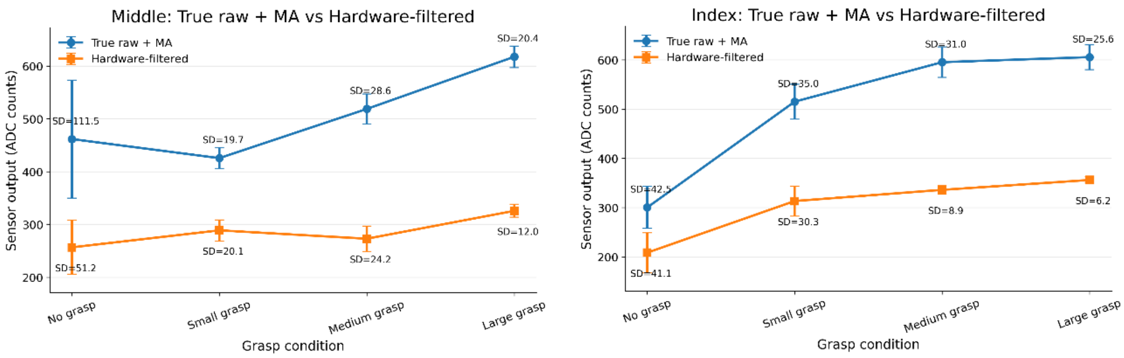



*Fig.3. (a) Variability of sensor readings corresponding to the middle finger (b) Variability of sensor readings corresponding to the index finger*

## Conclusion

This study shows that the upgrade can provide meaningful signal stability for the WiGlove's DAQ. Future work will evaluate its impact on gestures and grasp classification. This improvement will contribute to the use of the WiGlove as a home-based training progress assessment and game-play-based therapy tool for tracking rehabilitation performance over time.

https://doi.org/10.18745/pb.27307

## References

**[1]** V. Velmurugan, L. Wood, and F. Amirabdollahian, "A User-centred Design and Feasibility Analysis of the WiGlove - A Home-based Rehabilitation Device for Hand and Wrist Therapy after Stroke," University of Hertfordshire, ACHI_acceptedPaper_vignesh.pdf, 2023.

**[2]** R. P. Sallen and E. L. Key, "A Practical Method of Designing RC Active Filters," IRE Transactions on Circuit Theory, vol. 2, no. 1, pp. 74–85, Mar. 1955.

**[3]** Texas Instruments, "Single-supply, 2nd-order, Sallen-Key low-pass filter circuit," Application Report SBOA226, June 2021. [Online]. Available: https://www.ti.com/lit/pdf/sboa226

https://doi.org/10.18745/pb.27307

# Soccer Mind: Bold Hearts' Approach to Adaptive Decision-Making

**Oliver Samuel[1*], Alessandra Rossi[1], Daniel Barry[1], Bente Riegler [1] and Daniel Polani[1]**
*[1]RoboCup Bold Hearts, University of Hertfordshire*
*Corresponding author: ollydrummer@gmail.com

*Autonomous soccer-playing robots must successfully perform in complex, unpredictable environments that demand both rapid reactive responses and strategic decision-making, making the design of robust and flexible cognitive architectures a critical research challenge. To address these challenges, we present Soccer Mind - a high-level decision-making architecture for soccer playing robots. It utilises a* behaviour *tree with a blackboard architecture and uses the Belief-Desire-Intent cognitive model. The behaviour tree provides excellent reactive and deliberative control and the cognitive model permits very effective decoupling from the robot's other modules. Our framework is suitable for use on robots with different hardware and software implementations. Its modular and hardware-agnostic design furthermore lays a strong foundation for extension beyond well-defined contexts, such as soccer, making it well suited for deploying robust decision-making in less structured and more unpredictable environments.*

Keywords: RoboCup; Blackboard; Belief-Desire-Intent; Behaviour Trees; Robotics;

## Introduction

Autonomous soccer playing robots, such as those seen in RoboCup, operate in highly dynamic, non-deterministic environments. During a soccer game, robots must proactively plan, manage, and execute both team-oriented and individual objectives, while continuously modelling a dynamic environment that includes other players. To achieve this, they need to form purposeful, context-aware observations, build and update their understanding of the environment and the agents within it, whether human or machine, and translate that knowledge into coherent, adaptive actions [1]. These requirements present challenges on multiple levels. At the cognitive level, one of the primary difficulties lies in developing systems that balance high-level strategic deliberation with low-level reactive control, such as fall recovery and handling catastrophic failures in real time. There is also the need for software systems flexible enough to operate across heterogeneous hardware implementations; the Bold Hearts team now maintains multiple humanoid platforms simultaneously (i.e., Bold Bots, BoldRbots). To this extent, this work aims at creating a hardware-agnostic, modular decision-making framework that decouples the high-level decision-making logic from the robot's sensor and actuator implementations while providing both reactive and deliberative control.

## Experimental Architecture

The framework is implemented within the ROS 2 ecosystem, combining a blackboard architecture with a behaviour tree (BT) and the Belief-Desire-Intent (BDI) model. Figure 1 gives a high-level overview of the general structure. Separate nodes for Blackboard Management and the BT provide effective decoupling between state storage from decision-making logic.

## Blackboard Architecture

In this context a blackboard architecture refers to a centralised data structure/hub (the blackboard) that can collate the state of several otherwise unrelated systems. It is a well-established architecture [2] that has been used for many applications where multiple systems need to be coordinated by a central decision-making process. In Soccer Mind, each module is treated as a functionally independent agent, each providing periodic updates on their state to the blackboard node via a service. The blackboard node manages the state by providing a range of services and topics that allow it to be updated and monitored. It includes validation logic to prevent things like conflicting states from occurring (e.g. a robot cannot be simultaneously laying on its back and its front).

https://doi.org/10.18745/pb.27307

## Belief-Desire-Intent Model

The BDI Model [3] of cognition permits the decoupling of the decisions made by the robot and its low-level behaviours by posting *intents* to the blackboard instead of directly interacting with any other parts of the system. The BDI model treats the blackboard states as *beliefs*. The list of goals it has available to carry out are its desires. A *desire* becomes an intent when it has been selected by the decision-making logic.

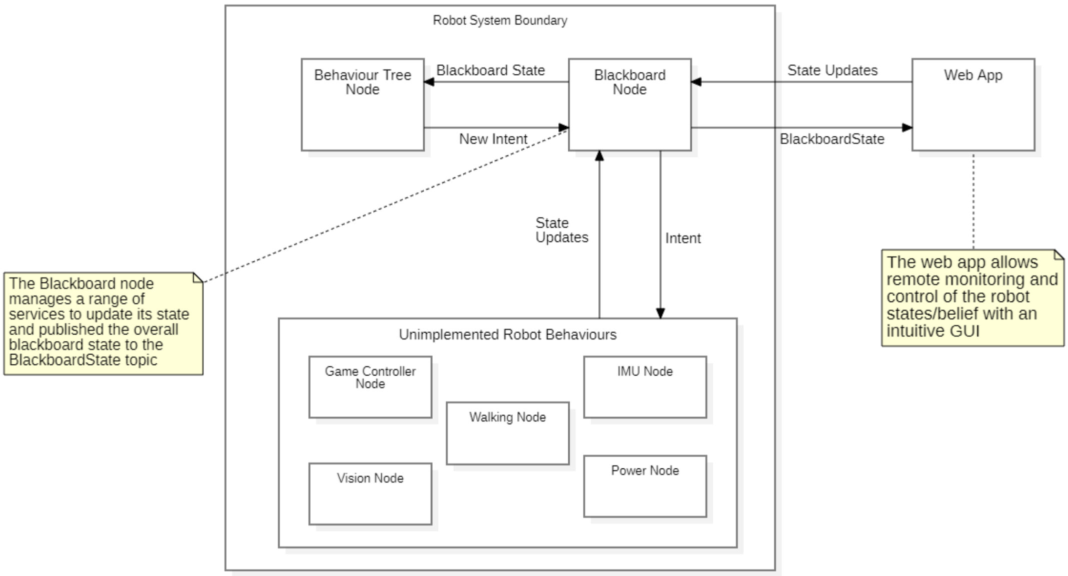



*Figure 1. High-Level Box Diagram of the Soccer Mind Architecture*

## Behaviour Tree

A behaviour tree serves here as the desire selection logic. Composed of a tree of nodes, a root node sends out a signal called a *tick* [4]. This travels down the tree at a set frequency (e.g. 20Hz) evaluating beliefs and setting intents by calling a service provided by the blackboard node. The Soccer Mind BT functions like any other behaviour tree except that its action nodes only return a status of *running*, never *success* or *failure*, and there are only two executive behaviours which set the intents on the blackboard. For example, if the BT sets the "move to ball" intent on the blackboard, other behaviour modules will see this and then make their own lower-level decisions about how to behave. To prevent the data from being updated during a tick, the blackboard data is first copied to the BT's own blackboard using a pre-tick handler method. A BT's hierarchical structure (see Figure 2) implements a Subsumption Architecture [5] which enables high priority behaviours to interrupt non-critical soccer playing branches. For example, a low power or fall detection branch can instantly interrupt the "dribble" sequence, ensuring a level of safety and control.

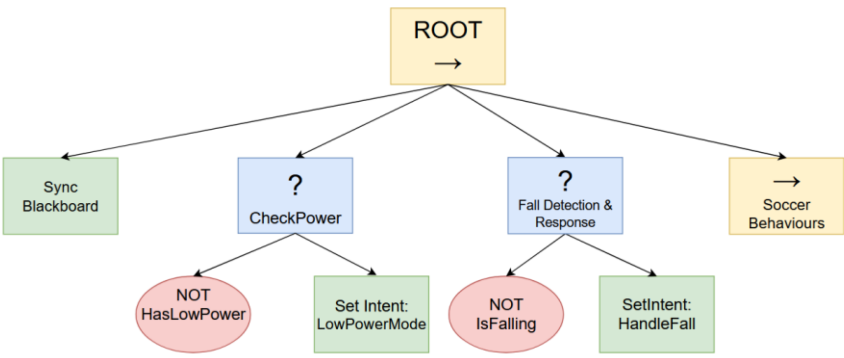



*Figure 2. Simplified Soccer Mind Behaviour Tree*

https://doi.org/10.18745/pb.27307

## Preliminary Evaluation of the Framework

The system was tested on a Raspberry Pi 5 with 4 GB of RAM. The reaction time for a backwards fall was recorded over 10 trials at both a tick rate of 10 Hz and 20 Hz. Reaction times of the backwards fall seen in Table 1 compare favourably to human benchmarks, particular for deliberative control where American football players had a decision reaction time of between 284ms and 462ms [6].

*Table 1. Reaction Times of Backwards Fall Response: Tick Rate of 10Hz vs 20Hz*

| Tick Rate | Mean reaction time | Standard Deviation | CPU Usage | Memory Usage |
|---|---|---|---|---|
| **10Hz (100ms)** | 195.4 ms | 32.3 ms | ~3.9% | 3.16% (~125.9MB) |
| **20Hz (50ms)** | 87.8 ms | 21.6 ms | ~5.9% | 3.16% (~125.9MB) |

A tick rate of 20 Hz more than halved the mean reaction time and lowered the standard deviation, although it increased the CPU load by 2%, from 3.9% TO 5.9%. This is expected as the behaviour tree can detect a change in state more quickly. At speeds faster than 20 Hz, the system became less reliable with signs of “jitter” where the blackboard intent state had not been updated before the next tick began.

## Conclusion & Future Works

This work presented Soccer Mind, a modular, hardware-agnostic decision-making framework designed for autonomous soccer-playing robots. By combining a blackboard architecture, a behaviour tree, and the Belief-Desire-Intent cognitive model, the framework successfully decouples high-level decision-making logic from low-level sensor and actuator implementations, addressing one of the key challenges faced by teams such as Bold Hearts that operate across multiple heterogeneous robotic platforms. The experimental results demonstrate that Soccer Mind achieves competitive reaction times, particularly at a tick rate of 20 Hz, where mean response times of 87.8 ms compare favourably against human deliberative benchmarks. The hierarchical structure of the behaviour tree further ensures that critical behaviours (e.g., fall recovery) can reliably interrupt non-critical actions, providing both safety and robustness during play.

Future work will focus on extending the framework to support decision making in more dynamic and less well-known environments.

## References

[1] Rossi, A., Paetzel-Prüsmann, M., Keijsers, M. et al. The human in the loop Perspectives and challenges for RoboCup 2050. Auton. Robot 48, 8 (2024). https://doi.org/10.1007/s10514-024-10159-3

[2] A. S. Vinze and A. Sen, "Blackboard Architecture and Applications: Introduction to Session," in *Proceedings of the Twenty-Sixth Hawaii International Conference on System Sciences*, Kauai, HI, USA, Jan. 5-8, 1993, pp. 554-555.

[3] M. Bratman, *Intention, Plans, and Practical Reason*. Cambridge, MA, USA: Harvard University Press, 1987.

[4] M. Colledanchise and P. Ögren, *Behavior Trees in Robotics and AI: An Introduction*. Boca Raton, FL, USA: CRC Press, 2018. doi: 10.1201/9780429489105.

[5] R. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal on Robotics and Automation*, vol. 2, no. 1, pp. 14–23, Mar. 1986, doi: 10.1109/JRA.1986.1087032.

[6] D. D. Larish, "Influence of stimulus-response translations on response programming: Examining the relationship of arm, direction, and extent of movement," *Acta Psychologica*, vol. 61, no. 1, pp. 53–70, 1986. doi: 10.1016/0001-6918(86)90021-1.

University of Hertfordshire UH

https://doi.org/10.18745/pb.27307

# The Characteristics of Flax Fibre Reinforced Composites under Impact Loading

[1]*Chizaram Ohadomere, [1]Opukuro David-West

[1]*School of Physics, Engineering and Computer Science, University of Hertfordshire, AL10 9AB*

*Corresponding author: co22abm@herts.ac.uk

**Abstract**

Although the strength properties of natural fibres do not match the synthetic ones, they have the advantage of the environmental benefits and hence reduced carbon footprint. Consequently, natural fibre reinforced composites have gained attention in recent years. In this study the impact behaviour of composites reinforced with flax fibres was investigated using LS-DYNA explicit finite element tool. The results revealed that the stacking sequence has effect on impact response. The $[30/60]_{2S}$ laminate absorbed the most energy, indicating a prolonged contact duration and improved energy dissipation. In general, the laminate architecture can be tailored to specific design requirements.

Keywords: composites; flax; impact, energy absorbed, finite element analysis.

## Introduction

The composite materials have consolidated their place in various industries such as the aerospace, automotive, and defence companies, because of the good mechanical characteristics, flexibility, and capability of satisfying the growing need of lightweight and high-strength structures. The carbon fibre reinforced composites are among the materials that have received attention because of their high strength to weight and stiffness to weight ratios. The impact on composites usually occurs during manufacturing or in service such as runway debris hurled against composite body, road debris kicking up against composite part and tool drop on composite during production.

The study by Ouyang *et al* [1] focused on the relationship between matrix cracking and delamination in composites subjected to low-velocity impact and the results showed that matrix cracks are initiated first and subsequently triggered delamination. This supports the report by Li *et al* [2, 3] on the quasi-static indentation behaviour of composite laminates to understand the damage mechanisms associated with object impacts where they observed that natural composites exhibited progressive damage and that natural composites rarely failed through a single damage mechanism. Cordin *et al* [4] observed that the mechanical property of natural composite depends on the directions of the fibre and the tensile strength. In the investigation by Zhang *et al* [5] on the relationship between impact resistance and tensile strength they concluded that during impact, tensile loading can cause cracks in a material. Therefore, higher tensile strength correlates with better impact resistance.

The demand for composites in the various industries and the challenges associated with the disposal and recyclability of synthetic fibres have led to the growing demand for more eco-friendly substitutes such as flax. This study has investigated the impact load effect on composites reinforced with flax fibres considering laminates of different configurations. The composites reinforced with flax fibre are usually applicable for the non-load bearing components in the automotive and aviation industries such as the dashboard and cabin panels. This helps with weight saving and sustainability.

## Simulation

The impact on the composites was modelled using the commercial finite element code LS-DYNA. The striker was represented as a spherical rigid body with the MAT 020-RIGID material card, with the impact velocity and mass set at 20 mm/ms and 20 kg respectively. The composites investigated in this simulation comprised of eight plies

https://doi.org/10.18745/pb.27307

of flax fibre reinforced laminate, and each ply was 0.3 mm thick. The stacking configurations were $[0/90]_{2S}$, $[15/-75]_{2S}$, $[30/60]_{2S}$, $[45/-45]_{2S}$ and $[0/90/-45/45]_S$, with plate dimensions of 100 x 100 mm.

The material properties of the composites were taken from Mahboob *et al* [6] and loaded in the MAT_058_SOLID-LAMINATED_COMPOSITE_FABRIC_SOLID material card, and the specific layer direction directions with the properties keyed into the PART_COMPOSITE card. This was then moved on to the shell elements to represent the composite, the Belytschko-Tsay shell elements formulation was adopted. To capture the interaction between the impactor and the laminate the AUTOMATIC-SURFACE-TO-SURFACE penalty-based contact algorithm was used.

## Results and discussion

The composite plates considered in this study, with the peak load and energy absorbed from the impact response are shown in ***Table 1***. At the structural level high stiffness will lead to high forces as they are directly related, and the peak load will usually connect to the maximum force before failure initiation. While the energy absorbed connects to the various damage processes in the structure resulting in the dissipation of the energy in the composites, this is usually through matrix crack, fibre breakage, fibre pull-out and delamination. ***Fig, 1*** illustrates the representation of the striker on the target and some stress contours. The primary impact response follows spikes, where the force increases rapidly with time, after which an oscillatory decay is observed due to the loss of stiffness in the structure, which continues as the force approaches zero. This characteristic is shown on ***Fig. 2***. The sudden drops in the load are usually attributed to damage such as fibre breakage, matrix crack, delamination etc. (Bandaru *et al* [7]).

***Table 1***: Laminate impact properties.

| Composite plate | Peak load (kN) | Absorbed energy (J) |
|---|---|---|
| $[0/90]_{2S}$ | 31 | 24 |
| $[15/-75]_{2S}$ | 14.6 | 4.8 |
| $[30/60]_{2S}$ | 35.3 | 49 |
| $[45/-45]_{2S}$ | 47.2 | 31 |
| $[0/90/-45/45]_S$ | 31.7 | 28 |

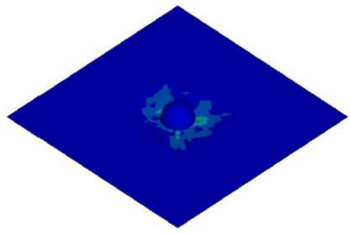


***Fig. 1***: Illustration of stress intensity.

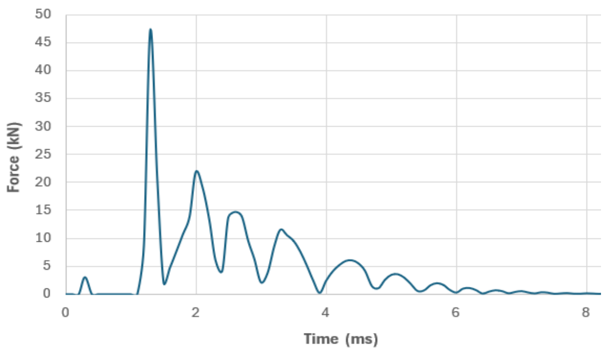



***Fig. 2***: Representative force history response.

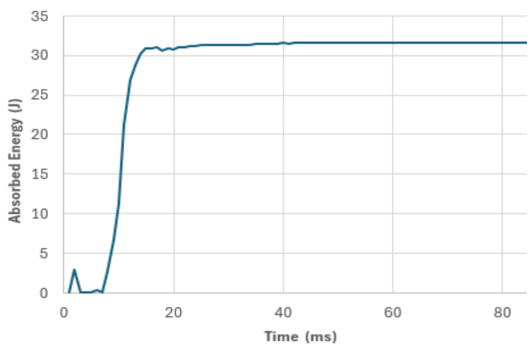



***Fig. 3***: Illustrative profile of the energy absorbed.

https://doi.org/10.18745/pb.27307

The $[45/-45]_{2S}$ sequence exhibited the highest peak force of approximately 47 kN and a few sharp rises and drops during its short oscillatory response. By contrast, the $[15/-75]_{2S}$ sequence recorded the lowest peak force with a short oscillatory period. However, a small spike was observed between the two largest peaks in this sequence. The $[30/60]_{2S}$ laminate showed a peak force of approximately 35 kN and a proportionally longer oscillatory response, indicating a prolonged contact period between the projectile and laminate. This implies a more gradual distribution of the shock load.

The $[0/90]_{2S}$ and $[0/90/-45/45]_{2S}$ laminate exhibited similar force-time behaviours. Both sequences showed comparable peak forces of 31 kN and 32 kN with a sustained oscillatory period, indicating prolonged contact between the projectile and the target and a gradual load distribution. Where torsional or multi directional loading is involved, it is necessary to have the ±45° plies in the laminates [8]. The energy absorbed as the various damage due to the impact loading is shown in ***Table 1*** and a representative profile shown in ***Fig. 3***. The $[30/60]_{2S}$ laminate absorbed the highest energy with an approximate value of 49 J suggesting superior energy absorption capability compared to others and $[15/-75]_{2S}$ recorded the lowest energy absorbed.

## Conclusion

This study investigated the impact response of flax fibre reinforced composites with an emphasis on the influence of stacking sequence. The models were built using LSDYNA commercial finite elements code. The results showed that the fibre orientations in the laminate plays a critical role in governing the impact response. The [0/90/-45/45]s composite presented the most balanced impact performance with the characteristics of moderate peak force and energy absorption relative to the others, hence the hybrid combination of fibre directions in laminate is recommended for a mixed loading environment. This work was limited to numerical analysis, and no experiment was conducted as currently there is no instrumented drop test system within the school laboratory facilities.

## References

[1] Ouyang, T., Sun, W., Bao, R. and Tan, R., 2021. Effects of matrix cracks on delamination of composite laminates subjected to low-velocity impact. *Composite Structures*, *262*, p.113354.

[2] Li, J., Hunt, J.F., Gong, S. and Cai, Z., 2017. Quasi-static compression and low-velocity impact behavior of tri-axial bio-composite structural panels using a spherical head. *Materials*, *10*(2), p.185.

[3] Li, J., Hunt, J.F., Gong, S. and Cai, Z., 2019. Low-velocity impact and compressive behavior for shifted-tri-axial composite panels. *Journal of Sandwich Structures & Materials*, *21*(2), pp.670-688.

[4] Cordin, M., Bechtold, T. and Pham, T., 2018. Effect of fibre orientation on the mechanical properties of polypropylene–lyocell composites. Cellulose, 25(12), pp.7197-7210.

[5]- Zhang, J., Bai, P., Yan, C., Liu, S. and Wang, X., 2019. Experimental investigation on relations between impact resistance and tensile properties of cement-based materials reinforced by polyvinyl alcohol fibers. *Applied Sciences*, *9*(20), p.4434.

[6] Mahboob, Z., El Sawi, I., Zdero, R., Fawaz, Z. and Bougherara, H., 2017. Tensile and compressive damaged response in Flax fibre reinforced epoxy composites. *Composites Part A: Applied Science and Manufacturing*, *92*, pp.118-133.

[7] Bandaru, A.K., Patel, S., Ahmad, S. and Bhatnagar, N., 2018. An experimental and numerical investigation on the low velocity impact response of thermoplastic hybrid composites. Journal of Composite Materials, 52(7), pp.877-889.

[8] 'Li, X., Wang, L., Zhao, W., Wu, H. and Xu, B., 2022. Selection and analysis of optimum stacking sequence and geometry dimensions on bismaleimides composite joints under tensile loading. *Polymer Testing*, *105*, p.107421.

https://doi.org/10.18745/pb.27307

# Bayfordbury Observatory: going beyond "one of the best equipped and largest teaching observatories in the country"

**Samantha M Rolfe[1*]**

[1]*Bayfordbury Observatory, TSG, SPECS, University of Hertfordshire, UK*

*Corresponding author: s.rolfe2@herts.ac.uk

**Bayfordbury Observatory is notable for being one of the largest and best equipped observatories in the country. Its primary purpose has been to mimic a professional astrophysics observatory to prepare students for a future career in research. Beyond teaching, the observatory has hosted undergraduate, postgraduate, and academic and industry research projects. Beyond astrophysics, there have been multidisciplinary uses of the observatory and surrounding Bayfordbury campus for its 59-year history, and this poster aims to showcase the facilities to continue to form collaborations and opportunities for teaching and research within SPECS, UH and beyond.**

Keywords: astronomy; astrophysics; atmospheric physics; datasets; multidisciplinary

## Introduction and Aim

Bayfordbury Observatory is notably one of the largest and best equipped observatories in the country [1]. With one of the largest optical telescopes available to undergraduates and postgraduates in the UK [2], the 60 cm CDK24 Planewave is one of the many instruments at the observatory. The observatory also boasts the only known three-dish radio interferometer, and longest baseline at 115 m, built for teaching in the UK. A suite of atmospheric instrumentation has been collecting data for 15+ years, with datasets available for teaching and research. Though teaching has always been the primary purpose of the observatory, beyond this, all of the instrumentation and datasets are available for use by researchers, with many past examples over the 56-year history. The aim of this poster will be to showcase the multidisciplinary facilities at the observatory and Bayfordbury campus, with the intention of sparking ideas for projects and collaborations for broadening the use of the observatory in teaching and research throughout all the disciplines within SPECS, UH and beyond.

## Instrumentation

Present instrumentation at the observatory includes, but is not limited to, eight optical telescope domes, with one 60 cm telescope, one 50 cm telescope (historical instrument custom built in 1970, eye-piece only), four 40 cm telescopes, one 35 cm telescope, and Solar Observing Platform composed of four telescopes. The sensors that are typically mounted are CMOS cameras, which can detect blue wavelengths through to the near infrared using astronomical broad- and narrow-band filters R, V, B, I, H-alpha, S-II and O-III, though there are also some other specialist filters available, including methane. There is a spectrometer that could be mounted on the telescope in place of the CMOS camera. Since 2012, several of the optical telescopes are able to run robotically, which have created an archive dataset of over 112,000 astronomical images via automated observing plans submitted by students and astrophysics technical staff.

There are four radio dishes, presently receiving at 1420 MHz (21 cm neutral Hydrogen line), but could be modified to receive other frequencies. Presently, the interferometer adds the signal from radio sources in real time, but building a correlator could be an interesting undergraduate project – just one of many possibilities for projects at the observatory.

The atmospheric gantry instrumentation includes a day and night allsky camera, taking images every 20 to 60 seconds with a dataset of over 15 years, a solar photometer on the NASA AERONET network, a pyranometer and pyrgeometer, with space for further instrumentation to be mounted.

https://doi.org/10.18745/pb.27307

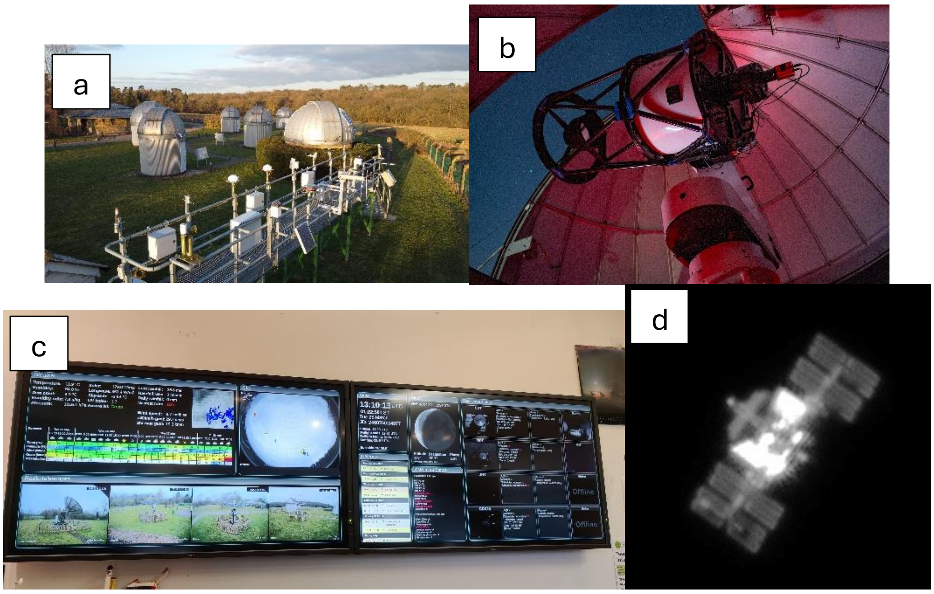



*Fig. 1. Bayfordbury Observatory and some of the facilities. (a) aerial photograph of the atmospheric gantry and the optical telescope domes. (b) 60 cm CDK24 installed in September 2021. (c) displays of some of the data available from the observatory. (d) example of an image of the International Space Station taken with the CDK24.*

## Examples

Other than its primary use case of teaching astrophysics students, the observatory and the surrounding Bayfordbury campus is used for many other purposes, which could be expanded. PAM has a prominent outreach and public engagement programme that sees ~5000 visitors to the observatory annually. Civil engineering uses the campus to complete surveying practicals. Film and TV students often use the observatory and wider campus for projects. HML has had a field station since the then Hatfield Polytechnic acquired the campus in 1967, each year there are many ecological and environmental teaching practicals and ongoing research projects. There was a recent collaboration between the Engineering department, UCL and an external company which utilised part of the observatory grounds to set up complex, proof-of-concept equipment.

## References

[1] University of Hertfordshire, "Bayfordbury Observatory," 2026. [Online]. Available: https://www.herts.ac.uk/bayfordbury/bayfordbury-observatory. [Accessed 03 05 2026].
[2] University of Hertfordshire, "Bayfordbury Observatory telescope," 2021. [Online]. Available: https://www.herts.ac.uk/about-us/campus-investment/new-telescopetransformingteaching. [Accessed 03 05 2026].

https://doi.org/10.18745/pb.27307

# Evaluating AI-Driven Alert Triage Accuracy in Security Operations Centres

**Amil Saman[1*], Alexios Mylonas[1], Javed Khan[1]**

*University of Hertfordshire*

*Corresponding author: A.Saman@herts.ac.uk

**This early-stage PhD proposal examines AI-driven alert triage accuracy in Security Operations Centres (SOCs). The study has not yet been implemented and no simulations or results are reported. The current phase uses literature review and gap mapping to determine whether a focused future evaluation of AI-assisted triage is feasible, original and methodologically defensible. Possible future work may compare AI-assisted and rule-based triage using controlled adversary-emulation data, SOC telemetry and measurable classification metrics such as precision, recall, F1-score and false-positive rate.**

Keywords: Security Operations Centre; Alert Triage; Machine Learning; False Positives; Research Gap Mapping

## Introduction

Security Operations Centres (SOCs) play a central role in monitoring, analysing and responding to cyber threats. However, SOC teams continue to face high alert volumes, false-positive investigations and analyst workload challenges [1], [2]. The false-alarm problem is also well established in intrusion detection research, where even a small false-positive rate can create a large operational burden when benign activity is frequent [3].

The abstract title focuses on evaluating AI-driven alert triage accuracy in SOCs. Work is currently at the early proposal and gap-mapping stage. The study has not yet been carried out, simulations have not been completed, and no empirical results are being claimed. The current objective is to examine the literature, identify unresolved research gaps and decide whether a future controlled evaluation of AI-assisted SOC alert triage would be feasible and original.

Machine learning has been widely studied for intrusion detection and cyber security analytics [4], [5]. More recent work also considers alert fatigue and automation in SOC environments [1], [2], [6]. However, the key research issue is not simply whether AI can be applied to SOC triage. The stronger research problem is whether AI-assisted triage can be evaluated under realistic conditions, using defensible datasets, clear ground truth, interpretable outputs and operationally meaningful metrics. This proposal therefore treats AI triage as a possible research direction rather than a fixed conclusion.

## Experimental/Simulation

No experimental work or simulations have been completed at this stage. This section describes a possible future methodology that may be used if the ongoing literature review confirms that controlled SOC alert triage evaluation is the most defensible PhD direction.

The current research-phase methodology is literature-informed gap mapping. Papers are being reviewed to extract the research problem, dataset type, method, evaluation metrics, stated limitations and future-work recommendations. These findings are then grouped into candidate gap themes, including operational realism, dataset representativeness, labelled ground truth, explainability, privacy risk, model robustness and analyst adoption.

If a future simulation-based study is selected, a controlled SOC lab may be designed using open-source tooling. Adversary behaviour could be generated using Atomic Red Team and MITRE Caldera, mapped to MITRE ATT&CK techniques [7]-[9]. Endpoint and network telemetry may be collected using tools such as Sysmon, Zeek and Suricata, and alerts may be generated through a SIEM or rule-based workflow such as Wazuh [10].

https://doi.org/10.18745/pb.27307

A future AI-assisted triage model could use structured alert features such as event type, rule identifier, severity, host context, process or network attributes, frequency, temporal patterns and ATT&CK technique mapping. Supervised learning models could then be compared with rule-based triage outputs using accuracy, precision, recall, F1-score and false-positive rate. Public datasets such as UNSW-NB15 and CICIDS2017 may be used as methodological reference points or secondary benchmarks, but they would not automatically replace the need for SOC-relevant alert-triage data [11], [12].

## Results and discussion

As this is proposed early-stage PhD work, no results are reported. The discussion at this stage focuses on the research gap and on how a future study could avoid making broad or premature claims about AI effectiveness in SOC environments.

Initial literature mapping suggests that the important research question may not be whether AI can improve triage in general, because AI-enabled security tooling is already being explored and implemented. Instead, the unresolved question is how AI-assisted alert triage should be evaluated in a way that is realistic, reproducible, explainable and operationally useful. Table 1 summarises the current candidate gap themes that may shape the final PhD topic.

*Table 1. Candidate gap themes identified for further investigation.*

| **Candidate gap theme** | **Why it matters** | **Relevance to possible future study** |
|---|---|---|
| Operational realism | Many studies do not fully reproduce SOC workflows, alert queues, analyst pressure or noisy enterprise telemetry. | Supports a lab or hybrid evaluation design rather than a purely benchmark-based ML comparison. |
| Dataset and ground-truth quality | Labelled SOC alert data is difficult to obtain, and public datasets may not represent real alert triage decisions. | Shapes whether the study uses public datasets, controlled simulation, or a hybrid strategy. |
| Explainability and trust | Analysts may not act on AI triage outputs if the reasoning is opaque or difficult to validate. | Suggests that model interpretability and analyst-facing explanation should be considered. |
| Privacy and security risk | SOC logs and alerts may contain sensitive operational data, and AI models may introduce leakage or misuse risks. | May influence feature selection, data handling and model-governance requirements. |
| Adoption and human factors | AI can increase rather than reduce workload if it does not fit SOC processes or analyst decision-making. | Keeps the research focused on augmentation rather than replacement of analysts or rule-based detection. |

The expected contribution at this stage is therefore not a claim that AI improves triage. The contribution is a clearer research direction: identifying where existing AI-assisted SOC triage research remains weak and defining a possible evaluation framework that could be tested later. If the final PhD topic remains focused on alert triage accuracy, the future study would need to justify its dataset strategy, ground-truth labelling, baseline comparison and operational relevance.

https://doi.org/10.18745/pb.27307

## Conclusion

This revised abstract clarifies that the work is a proposed, early-stage PhD research direction rather than a completed empirical study. Focus remains on AI-driven alert triage accuracy, but the current phase is literature review, gap mapping and topic refinement. No simulations have yet been completed and no results are claimed.

The proposed direction is to investigate whether a focused and reproducible evaluation of AI-assisted SOC alert triage can form a defensible PhD contribution. The likely value of the work lies in identifying realistic evaluation gaps around datasets, ground truth, explainability, privacy, robustness and analyst adoption, and then using those gaps to shape a future methodology.

## References

[1] B. A. Alahmadi, L. Axon and I. Martinovic, “99% False Positives: A Qualitative Study of SOC Analysts’ Perspectives on Security Alarms,” in Proc. 31st USENIX Security Symposium, 2022.
[2] S. Tariq, M. B. Chhetri, S. Nepal and C. Paris, “Alert Fatigue in Security Operations Centres: Research Challenges and Opportunities,” ACM Computing Surveys, vol. 57, no. 9, 2025, doi: 10.1145/3723158.
[3] S. Axelsson, “The Base-Rate Fallacy and the Difficulty of Intrusion Detection,” ACM Transactions on Information and System Security, vol. 3, no. 3, pp. 186-205, 2000, doi: 10.1145/357830.357849.
[4] R. Sommer and V. Paxson, “Outside the Closed World: On Using Machine Learning for Network Intrusion Detection,” in Proc. IEEE Symposium on Security and Privacy, pp. 305-316, 2010, doi: 10.1109/SP.2010.25.
[5] A. L. Buczak and E. Guven, “A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection,” IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, 2016, doi: 10.1109/COMST.2015.2494502.
`Lecture Notes in Computer Science, Springer, 2024, doi: 10.1007/978-3-031-54129-2_36.
[7] MITRE, “ATT&CK Framework,” MITRE Corporation. [Online]. Available: https://attack.mitre.org
[8] Red Canary, “Atomic Red Team Documentation.” [Online]. Available: https://docs.redcanary.com/docs/atomic-red-team
[9] MITRE, “Caldera Documentation.” [Online]. Available: https://caldera.readthedocs.io
[10] Wazuh, “Wazuh Documentation.” [Online]. Available: https://documentation.wazuh.com
[11] N. Moustafa and J. Slay, “UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems,” in Proc. Military Communications and Information Systems Conference, 2015, doi: 10.1109/MilCIS.2015.7348942.
[12] I. Sharafaldin, A. H. Lashkari and A. A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization,” in Proc. ICISSP, pp. 108-116, 2018, doi: 10.5220/0006639801080116.

https://doi.org/10.18745/pb.27307

# Inclusive hydrogen economy in Kenya: Institutional Readiness, Policy Sequencing, and Pathways for Equitable Transition

**Yasser Sorouraddin[1]*, Chamu Kuppuswamy[2], Mahmoud Chizari[1],**

[1] *School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK*

[2] *School of Law and Education, University of Hertfordshire, Hatfield, UK*

*Corresponding author: y.sorouraddin@herts.ac.uk

**This paper examines whether Kenya can build an inclusive hydrogen economy that supports green industrialisation rather than reproducing enclave, export-led transition pathways. Using a Political, Economic, Technological and Environmental (PETE) framework weighted through Analytic Hierarchy Process (AHP), Kenya is compared with South Africa as an analytical contrast. The analysis indicates that Kenya's main advantage lies in its renewable resource base and increasingly low-carbon electricity system, while its principal constraints are institutional coordination, finance mobilisation, and domestic demand creation. South Africa offers stronger industrial depth and hydrogen-policy visibility but also demonstrates the risks of justice tensions and policy fragmentation. For Kenya, credible sequencing, not technological enthusiasm alone, will determine equitable transition outcomes.**

**Keywords**: *green hydrogen; Kenya; South Africa; institutional readiness; equitable transition*

## Introduction

Green transition strategies in Africa increasingly need to connect decarbonisation with industrial development and social inclusion. For Kenya, hydrogen is promising not only because of its renewable energy endowment, but because it could support energy production, industrial decarbonisation, clean transport, and export-oriented manufacturing if accompanied by the right institutions and policy sequencing [1, 2]. Yet early-stage hydrogen economies are rarely constrained by technology alone. They are shaped by governance quality, financing capability, infrastructure coordination, and the extent to which benefits are localised through jobs, skills, and domestic value chains. In this regard, this paper reframes Kenya's hydrogen transition as an institutional readiness and industrial policy question. It applies a Political, Economic, Technological and Environmental (PETE) framework, weighted through the Analytic Hierarchy Process (AHP), to assess readiness for an inclusive hydrogen economy. South Africa is used as an analytical contrast rather than a normative model. This comparison is useful because South Africa has a more advanced hydrogen commercialisation discourse, stronger industrial depth, and an explicit localisation agenda, while also facing major just-transition challenges linked to coal dependence and inequality [3, 4]. The paper asks two questions: first, what are the main binding constraints on Kenya's inclusive hydrogen transition; and second, what policy sequence is most appropriate for moving from opportunity to implementation?

## Method

The study employs a Multi-Criteria Decision Analysis (MCDA) design using the PETE framework. Political indicators cover governance effectiveness, regulatory quality, and policy stability. Economic indicators capture investment capacity, market scale, and industrialisation potential. Technological indicators focus on renewable-energy capacity, electricity-system readiness, and absorptive capability. Environmental indicators address emissions reduction and sustainability co-benefits. Criterion weights are derived using AHP [5, 6]. A pairwise comparison matrix was constructed based on a consistent importance scale ranging from 1 (equal importance) to 5 (extreme importance) between the four PETE criteria. This matrix was then used to calculate the relative priority weights shown in Table 1 through the standard AHP eigenvalue method. These weights reflect the proposition that in early-stage hydrogen transitions, economic viability and technological readiness are usually the most binding conditions, while political and environmental criteria remain necessary enabling dimensions.

https://doi.org/10.18745/pb.27307

*Table 1. PETE–AHP analytical structure and criterion weights used for the Kenya–South Africa comparison. Weights were derived through a pairwise comparison matrix based on a 1–5 importance scale using AHP.*

| Criterion | Illustrative focus | Weight |
|---|---|---|
| Political | Governance, regulation, coordination | 0.0909 |
| Economic | Finance, demand, local value addition | 0.4545 |
| Technological | Renewables, infrastructure, skills | 0.2727 |
| Environmental | Emissions, efficiency, climate gains | 0.1818 |

## Results and discussion

Kenya's principal constraint is institutional and economic readiness rather than renewable-resource availability. Latest IRENA country profiles show that renewable sources account for 83% of Kenya's installed electricity capacity, leaving 17% non-renewable, whereas in South Africa renewables account for 21% and non-renewables 79% [7]. Figures 1 (A, B) present these shares separately for clarity. This comparison is important, but it should be interpreted cautiously. It does not demonstrate that Kenya is already outperforming South Africa in the hydrogen economy overall. Rather, it indicates that Kenya enters the hydrogen transition with a stronger renewable-electricity base for potential green hydrogen production, while South Africa enters with greater industrial depth, commercialisation experience, and localisation capacity [3, 4]. The implication is that Kenya may possess an upstream advantage in renewable-energy foundations, whereas South Africa currently exhibits stronger readiness in the downstream institutional and industrial dimensions required to convert hydrogen potential into large-scale green industrialisation.

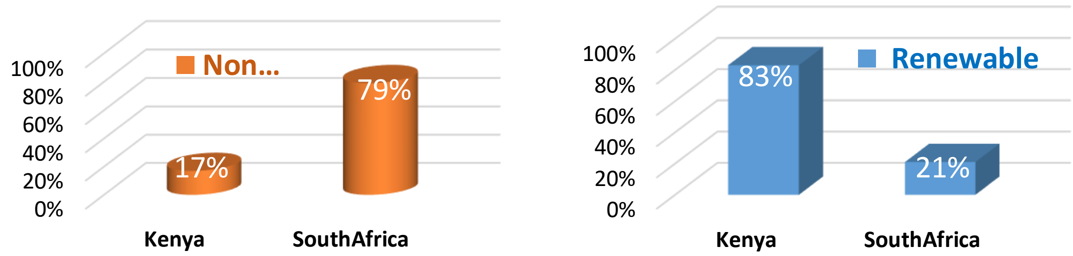



*Fig. 1. Installed electricity capacity mix in 2025: (A) Renewable share of total installed capacity – Kenya 83% versus South Africa 21%; (B) Non-renewable share – Kenya 17% versus South Africa 79% (Source: IRENA Statistical Profile, 2025).*

This distinction has direct implications for policy sequencing. Kenya does not need first to prove that renewable resources exist; rather, it must strengthen the governance and market architecture that would allow hydrogen projects to become bankable, coordinated, and development oriented. Kenya already has a Green Hydrogen Strategy and Roadmap, Guidelines on Green Hydrogen and its Derivatives, and formal recognition of hydrogen within the National Energy Policy 2025–2034 [2, 4]. However, the policy documents also identify high capital costs, limited infrastructure, insufficient regulatory frameworks, and inadequate local technical capacity as immediate barriers [2,3]. By contrast, South Africa's hydrogen policy architecture is more explicitly commercialisation-oriented: the Hydrogen Society Roadmap and Green Hydrogen Commercialisation Strategy emphasis market support, localisation, secure financing, domestic demand creation, and socio-economic development [5, 7]. They also emphasise domestic demand creation in green steel, fertiliser, mobility, and sustainable aviation fuel, while linking hydrogen to jobs, gender equality, community participation, and industrial localisation [7].

*Table 2. Comparative institutional readiness, key bottlenecks, and policy sequencing priorities for an inclusive green hydrogen economy: Kenya vs. South Africa (renewable power base, policy architecture, main constraints, anchor demand, equity/localisation mechanisms).*

| Dimension | Kenya | South Africa | Implication for Kenya |
|---|---|---|---|
| Renewable power base | High renewable share in installed capacity; strong geothermal, hydro, solar and wind platform [7]. | Lower renewable share overall, but strong solar/wind resource potential and large industrial system [7]. | Kenya can prioritise pilot hydrogen demand without waiting for complete power-sector decarbonisation. |
| Hydrogen policy architecture | Strategy and Roadmap (2023), Guidelines (2024), and policy recognition in National Energy Policy 2025–2034 [2,8]. | Hydrogen Society Roadmap (2021) and Green Hydrogen Commercialisation Strategy (2024) [3,10]. | Kenya should move from strategy formation to implementable regulation and project coordination. |
| Main bottlenecks | High capital costs, limited infrastructure, regulatory gaps, and weak local technical capacity [1,2,9]. | Financing, logistics and infrastructure needs, plus just-transition tensions around coal and inequality [4,9]. | Institutional coordination and blended finance are higher priorities than export ambition. |
| Anchor demand / use cases | Agriculture, industry, transport, power generation, and export markets are recognised policy use cases [9]. | Export hydrogen/derivatives, green steel, fertiliser, hydrogen mobility and sustainable aviation fuel [10]. | Kenya should start with domestic and regional demand where development benefits are easier to localise. |
| Equity and localisation | Local capacity, innovation, and jobs are highlighted, but operational mechanisms remain emerging [9]. | Skills, job creation, BBBEE, gender equality, and community participation are explicitly emphasised [10]. | Embed youth, women and local-value-chain goals early in project design, procurement, and training. |

Accordingly, Kenya should not mechanically imitate South Africa. A more appropriate pathway is to sequence reform through: (1) regulatory clarification and inter-agency coordination, (2) common infrastructure and blended finance for early projects, (3) targeted pilot demand in fertiliser, industrial heat, transport, and ammonia-linked value chains, and (4) localisation through skills, supplier development, and inclusive procurement. Under the PETE–AHP weighting, this confirms that economic and technological factors are decisive, but also that political coordination determines whether these advantages translate into equitable transition outcomes.

## Conclusion

Our analysis confirms that Kenya's hydrogen opportunity is genuine, but its realization depends less on resource endowment than on policy sequencing and institutional readiness. Compared to South Africa, Kenya is better positioned in terms of renewable-electricity structure, yet less mature in its commercialisation architecture. For an equitable transition, Kenya should first prioritize regulation, coordination, and finance, then scale domestic hydrogen demand in ways that strengthen employment, local capabilities, and industrial value creation. Future work could extend the PETE–AHP framework by incorporating explicit country scores, sensitivity testing, and broader inclusion indicators such as employment, regional distribution, and gender outcomes.

## References

[1] M. Apergi, L. Eicke, A. Goldthau, J. Kurniawan, E. Schuch, and S. Weko, "Pathways to a sustainable electricity sector in Kenya: Challenges and transformational factors," *Utilities Policy*, vol. 91, art. 101854, 2024. https://doi.org/10.1016/j.jup.2024.101854

[2] Energy and Petroleum Regulatory Authority (EPRA), *Kenya's Guidelines on Green Hydrogen and its Derivatives*, 2024. https://www.epra.go.ke/sites/default/files/Documents/Guidelines-on-Green-Hydrogen-and-its-Derivatives.pdf (Accessed: 15 Apr. 2026).

[3] Department of Science and Innovation, South Africa, *Hydrogen Society Roadmap for South Africa*, 2021. https://www.dsti.gov.za/images/South_African_Hydrogen_Society_RoadmapV1.pdf (Accessed: 15 Apr. 2026).
[4] H. R. Bohlmann, J. A. Bohlmann, M. Chitiga-Mabugu, and R. Inglesi-Lotz, "Just Energy Transition of South Africa in a Post-COVID Era," *Sustainability*, vol. 15, no. 14, art. 10854, 2023. https://doi.org/10.3390/su151410854
[5] T. L. Saaty, *The Analytic Hierarchy Process*. New York, NY, USA: McGraw-Hill, 1980.
[6] O. M. Vaidya and S. Kumar, "Analytic hierarchy process: An overview of applications," *European Journal of Operational Research*, vol. 169, no. 1, pp. 1–29, 2006. https://doi.org/10.1016/j.ejor.2004.04.028
[7] International Renewable Energy Agency (IRENA), *Statistical Profile: Kenya*, 2025. https://www.irena.org/IRENADocuments/Statistical_Profiles/Africa/Kenya_Africa_RE_SP.pdf (Accessed: 15 Apr. 2026).
[8] Ministry of Energy and Petroleum, Kenya, *Green Hydrogen Strategy and Roadmap for Kenya*, 2023. https://energy.go.ke/sites/default/files/KAWI/Publication/GHS_15_11_2023_COMP.pdf (Accessed: 15 Apr. 2026).
[9] Ministry of Energy and Petroleum, Kenya, *National Energy Policy 2025–2034 (Final Draft)*, 2025. https://www.energy.go.ke/sites/default/files/Final%20Draft_National%20Energy%20Policy%202025_November%202025.pdf (Accessed: 15 Apr. 2026).
[10] Department of Trade, Industry and Competition, South Africa, *Revised Green Hydrogen Commercialisation Strategy*, 2024. https://www.thedtic.gov.za/wp-content/uploads/Revised-Green-Hydrogen-Commercialisation-Strategy.pdf (Accessed: 15 Apr. 2026).
[11] P. Mirzania, J. A. Gordon, N. Balta-Ozkan, R. C. Sayan, and L. Marais, "Barriers to powering past coal: Implications for a just energy transition in South Africa," *Energy Research & Social Science*, vol. 101, art. 103122, 2023. https://doi.org/10.1016/j.erss.2023.103122

https://doi.org/10.18745/pb.27307

# Simulation Study to Leverage Bio-Inspired Enhancement of Uniaxial Mechanical Properties of Polymeric Structures

**Arwa Mohsin[1*], Sikiru O. Ismail [2], Liam O'Connor [3], Tom de Vuyst [4], and Andreas Chrysanthou[5]**

*University of Hertfordshire*

*Corresponding author: am25aev@herts.ac.uk

**Biomimicry has been gaining attraction from last few decades as it offers advantage of achieving desirable pursuits. This study examined the use of diatom inspired structures to improve the mechanical properties of polymers. The mechanical behaviours of these structures were evaluated using finite element analysis (FEA). Polylactic acid (PLA) and acrylonitrile butadiene styrene (ABS) are used because these are affordable, multipurpose and easy to process. Results revealed that the strength-to-weight ratio can be improved by up to 36% when using diatom inspired geometries, which yields a promising solution to lightweight structure. Results showed that the material and geometry work synchronously.**

Keywords: Biomimicry, Diatom Structures, Polymeric Materials, Finite Element Simulation, Mechanical Properties.

## Introduction

Structural innovations, inspired by many biological structures have been achieved for anti-fogging and anti-fouling, hydrophobic surfaces, super wettability, and extraordinary adhesion and wetting and optical behaviours [1, 2]. Even with these breakthrough improvement, there is still needed to improve the mechanical properties of materials. This is because previous studies focused on changing the material composition, which is costly and ineffective. Improvement of the mechanical properties such as strength, flexibility, impact resistance and fatigue life of material is still one of the important scientific challenges. In industries like aerospace, automotive, etc., light weight structures are desirable. Changing chemical composition of material can be advantageous but is not effective especially when we want to use a specific material that must hold desired properties. Bio-inspired design helps to achieve this. Researchers look for organisms that hold that desired property at exceptional level and then mimic it to achieve similar results in materials. One of the organisms in nature is diatoms which are microscopic algae and hold exceptional tensile properties.

Gutierrez et al. [3] demonstrated an FEA of Coscinodiscus diatom, suggesting that the middle layer is structurally dominant and the upper layer is structurally minor due to its thinness. Linnemann et al. [4] optimized CAD models based on diatoms, and obtained as much as 93% displacement reduction, with thickness-optimized models being best. FEA confirms these results. The studies mentioned above do not study the thickness of the layers as a variable which is responsible for the stiffness of a diatom and is a key factor in this research.

## Simulation

To attain desired results, accurate dimensions of Consinodiscus Asteromophalus are mimicked as shown in Fig. 1. To gain an understanding of the mechanical property of lattice geometry and for comparison, three different CAD designs including biomimicry are designed. In real-life engineering applications, failure criteria begin at yield; therefore, this research is limited to the yield strength of polymer materials. Models are simulated in ANSYS. Traditional tensile and compressive loading and boundary conditions are applied. Pressure of 31MPa and 65MPa for PLA and 38MPa 50MPa for ABS was applied for tensile and compressive tests, respectively. Material properties are extracted from datasheet of Bambu Filament [5] and RS-Pro [6].

https://doi.org/10.18745/pb.27307

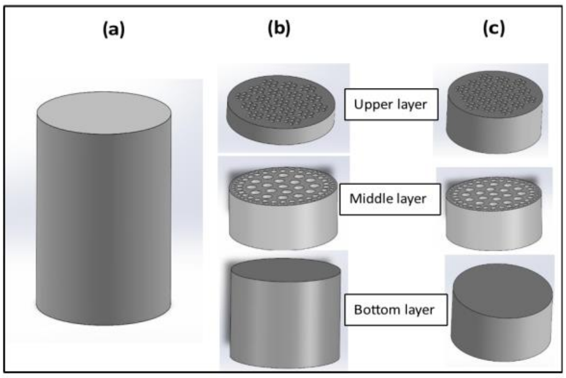



*Fig. 1. Diatom-inspired model on SolidWorks.*

## Results and discussion

### 1. Tensile Behaviour

The solid models show maximum absolute stiffness and load carrying capacity as their geometry is fully dense.

However, in terms of mass, strength-to-weight performance is better with diatom inspired biomimetic structures. This is in line with the results obtained by Aitken et al. [7] which demonstrated the possibility of simultaneously obtaining lightness, strength, and structural integrity through diatom biomimicry. In the case of PLA, bio-inspired design enhanced 11.5% strength-to-weight ratio. In the case of ABS, there is a much larger enhancement of the strength-to-weight ratio, up to ~47.8% (with a variation model recording a 36% enhancement) (Fig 2a and 2b).

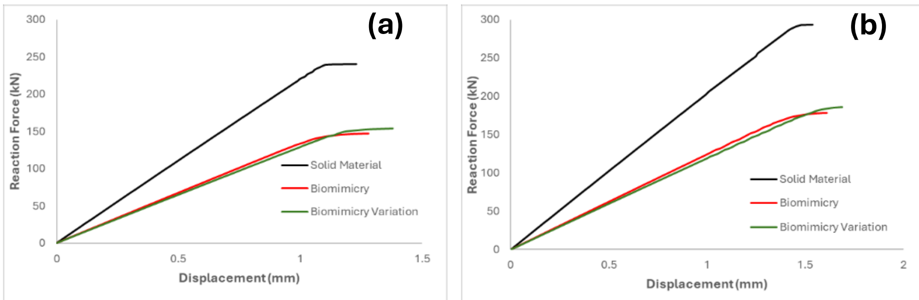



*Fig. 2. Reaction force vs displacement for, tensile loading (PLA) in the left (a), tensile loading (ABS) right (b)*

The biomimetic designs are found to have reduced axial rigidity, experiencing a lesser stress and reaction force at given strains as compared to solid models. This validates that load in biomimetic structures is absorbed by regulated deformation and redistribution of loads but not stress concentration [8].

### 2. Compressive Behaviour

The load-carrying capacity is highest in the solid models. This is consistent with the known cell solid theory that the reduction of the relative density caused by internal voids leads to a reduction of the ability to carry load and decrease of stiffness in the axial direction [9].

In case of PLA structures, although the mass of biomimetic structures reduces by about 11-15%, increase in specific strength is observed (Fig. 3). This can be related to more structural efficiency of diatom-inspired geometry, which can carry same amount of load using less material. In case of ABS, similar trend is observed.

Nevertheless, the enhancement was smaller in ABS than PLA, because of its stress distribution. That is because ABS is generally softer than PLA.

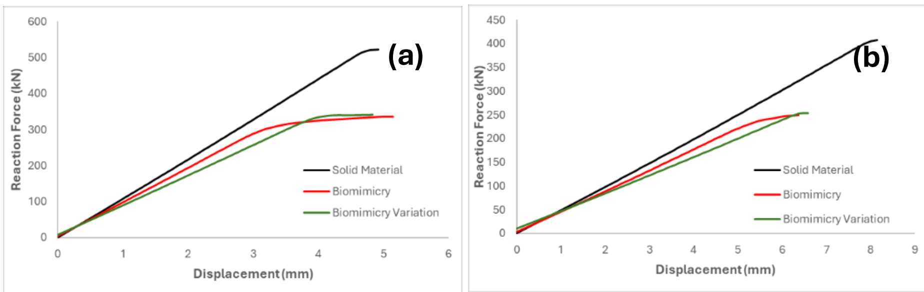



*Fig. 3. Reaction force vs displacement for, compressive loading (PLA) in the left (a), compressive loading (ABS) right (b)*

Overall, the compressive tests showed that biomimetic designs are capable to achieve lightweight designs while maintaining a compressive load carrying capacity.

## Conclusion

The mechanical behaviors of diatom-inspired PLA and ABS biomimetic structures were analysed and compared with traditional solid structures in tensile and compression loading conditions. It is concluded that the biomimetic structures exhibited low mass and efficient load distribution. These models showed noticeable rise in strength-to-weight ratio which increased up to 47.8%. These findings are similar to those reported by Mehdikhani et al. [9] which suggested that voids such as internal geometric discontinuities decrease stiffness but increase the ability to rotate and deform locally. These results confirmed that diatom-inspired designs can offer a promising solution to create lightweight and structurally efficient components, with a strength-to-weight ratio and less material consumption as a criterion. These bioinspired structures have significant applications in aerospace structures, automotive engineering (for weight reduction), biomedical devices (lightweight implants), and protective systems.

## References

[1] Y. Duan et al., “Robust Antifogging and Antifouling Coating Tailored with Zwitterionic Nanocellulose for Multi-Functional Applications,” Advanced Science, vol. 12, no. 39, Sep. 2025.
[2] C. Lee et al., “Bioinspired nanoscale hierarchical pillars for extreme superhydrophilicity and wide angular transmittance,” Nanoscale Advances, vol. 4, no. 3, pp. 761–771, Dec. 2021.
[3] H. Z. Hassan and N. M. Saeed, “Advancements and applications of lightweight structures: a comprehensive review,” Discover Civil Engineering, vol. 1, no. 1, Aug. 2024.
[4] A. Ramos, V. G. Angel, M. Siqueiros, T. Sahagun, L. Gonzalez, and R. Ballesteros, “Reviewing Additive Manufacturing Techniques: Material Trends and Weight Optimization Possibilities Through Innovative Printing Patterns,” Materials, vol. 18, no. 6, p. 1377, Mar. 2025.
[5] Bambu Filament Chemicals. (n.d.), November 11, 2025, “PLA technical data sheet,” Source. [Online]. Available: https://files.bbystatic.com/G8TU2klversusR6MVeDE4zgPkw%3D%3D/Bambu%2BFilament%2B-Technical%2BData%2BSheet
[6] RSPro, November 11, 2025, “ABS technical data sheet,” Source. [Online]. Available: A700000006781685.pdf
[7] Z. H. Aitken, S. Luo, S. N. Reynolds, C. Thaulow, and J. R. Greer, “Microstructure provides insights into evolutionary design and resilience of *Coscinodiscus* sp. frustule,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 8, pp. 2017–2022, Feb. 2016.
[8] D. A. Chen and M. M. Gibbons, “Bio-Inspired Sutures: Localizing Damage by Isolating Strain Energy,” *Biomimetics*, vol. 10, no. 2, p. 102, Feb. 2025.
[9] M. Mehdikhani, L. Gorbatikh, I. Verpoest, and S. V. Lomov, “Voids in fiber-reinforced polymer composites: A review on their formation, characteristics, and effects on mechanical performance,” *Journal of Composite Materials*, vol. 53, no. 12, pp. 1579–1669, Nov. 2018.

https://doi.org/10.18745/pb.27307

# A Computational Investigation into Carbon Nanotube Reinforced Composite Bone Plates Used for Internal Fixation

**Kaan Buyuktas[1], Dizem Arifler[2], and Volkan Esat[3*]**

*[1]Middle East Technical University – Northern Cyprus Campus, Sustainable Environment & Energy Systems, Kalkanli, Guzelyurt, Mersin 10, Turkiye*

*[2]Middle East Technical University – Northern Cyprus Campus, Physics, Kalkanli, Guzelyurt, Mersin 10, Turkiye*

*[3]Automotive Engineering, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK*

*Corresponding author: v.esat@herts.ac.uk

## Abstract

**Titanium alloy plates, screws, and implants are widely used to treat orthopedic fractures and severe bone cracks caused by trauma or surgery. This study investigates the mechanical performance of carbon nanotube (CNT)-reinforced titanium and PEEK (polyether ether ketone) composites for internal fixation of a transversely fractured femur using finite element analysis. Results show that adding CNTs to Ti-6Al-4V increases Young's modulus by up to 54% while reducing weight by up to 8% compared with the conventional alloy. All evaluated bone plate designs achieved an acceptable factor of safety above 1.5.**

Keywords: carbon nanotube reinforced composite; titanium alloy; PEEK; multiscale finite element modelling, femur fixture

## Introduction

This research investigates the mechanical performance of carbon nanotube (CNT)-reinforced composites and their implications for biomechanical applications and design optimization. By integrating materials science principles with finite element analysis (FEA), the study evaluates the potential of CNT reinforcements in commonly used biomechanical implants, including plates and screws. Computer-aided design (CAD) is employed to model biomechanical components, human bone structures, and compact fixation designs, which are subsequently assessed using FEA. The numerical results are validated through comparisons with existing literature. The inclusion of CNTs in conventional engineering materials has been shown to significantly enhance properties such as Young's modulus and ultimate tensile strength, leading to broad adoption of CNT-reinforced composites across industries including aerospace, automotive, electronics, and biomedicine [1]. The primary objective of this work is to examine how CNT reinforcement in widely used biomechanical materials can improve reliability, durability, sustainability, and failure resistance mechanisms. This study demonstrates that carbon nanotube–reinforced materials can achieve a high level of sustainability while serving as effective and promising candidates for biomechanical applications, particularly bone plate fixtures.

Prior research indicates that the toxicity of CNTs is influenced by several factors, including surface charge, morphology, length, diameter, degree of agglomeration, functionalization, and material purity [2]. Animal-based experimental studies involving both soft tissue and bone regeneration have shown that CNT-based polymer and metal composites do not induce significant toxicity or severe inflammatory responses [3].

## Methodology

Multiscale modelling approaches for carbon nanotube-reinforced materials have been developed, with FEA carried out using Hexagon's Digimat-FE software [4]. Within this framework, methods such as the Monte Carlo approach, STAT (statistical discretization), and Hybrid-IA (integral approximation) are employed to determine the spatial distribution of individual inclusions. The software assigns appropriate orientations and positions for each inclusion within the representative volume element (RVE), enabling effective reinforcement of the matrix material. Some of the generated CNT inclusions in a unit cubic RVE, and the material properties of the multi-walled CNT (MWNT) and matrices used in FEA simulations can be seen in Fig. 1 and Table 1 below, respectively.

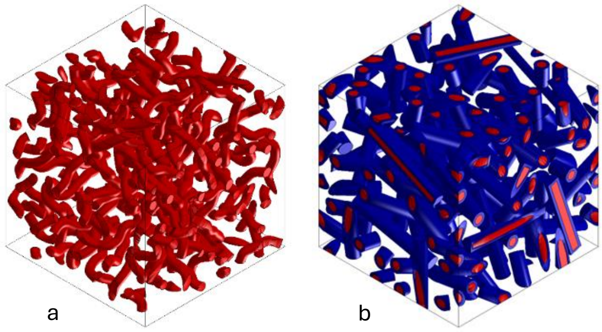



*Fig. 1. Some of the generated CNT inclusions in a unit cubic RVE; (a) 3D randomly aligned curved, and (b) 3D randomly aligned straight with*

*Table 1. Material properties of the MWNT and*

| | $\rho\ (kg/m^3)$ | $E\ (GPa)$ | $v$ | $\varepsilon$ (%) |
|---|---|---|---|---|
| MWCNT [5] | 2600 | 1200 | 0.35 | 12 |
| Ti-6Al-4V [6] | 4430 | 114 | 0.342 | 10 |
| PEEK [7] | 1320 | 3.76 | 0.4 | 50 |

The FEA case study for femur evaluation was conducted under defined assumptions and boundary conditions to assess the performance of carbon nanotube-reinforced titanium and PEEK materials. Two design configurations were examined: one involving a 5 mm gap representing a transverse fracture, and another with the same fracture gap filled with bone cement. All applied loads and constraints on the femur were modeled as surface effects (Fig. 2). Specifically, forces were distributed across 79 selected faces, with a load of 700 N applied at the femoral head to replicate hip joint loading. This applied force corresponds to the full body weight of a 70 kg individual under a 100% loading condition.

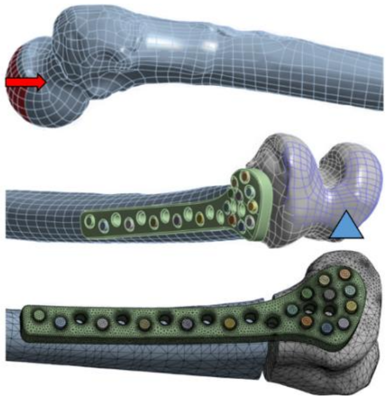


*Fig. 2. Femur FEA study boundary conditions and*

## Results and Discussion

The influence of carbon nanotube volume fraction, orientation, interphase characteristics, and geometric properties on CNT-reinforced composites was examined for both Ti-CNT and PEEK-CNT materials. The femur model containing a 5 mm transverse fracture on the distal side was analysed using distal femoral locking plates made from different composite materials. The fixation approach also permitted the fracture gap to be filled with PMMA bone cement to assess its effect on overall biomechanical behaviour. Figure 3 presents the finite element analysis results, illustrating the maximum displacement (top) and von Mises stress distribution (bottom) in the bone, plate, and screws for a complete femur with a 5 mm transverse fracture. FEA results for the maximum displacement and von Mises stress in the bone, plate, and screws with a 5-mm PMMA filled transverse fracture are provided in Table 2.

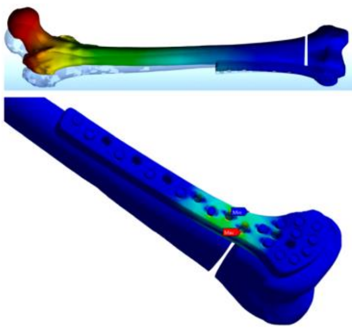


*Fig. 3. FEA results for the femur*

Table 2. Maximum displacement and von Mises stress results with PMMA

| | $\Delta x_{max}\ (mm)$ | | | $\sigma_{V,max}\ (MPa)$ | | |
|---|---|---|---|---|---|---|
| Materials | Bone | Plate | Screws | Bone | Plate | Screws |
| Ti-6Al-4V | 4.057 | 0.139 | 0.137 | 51.649 | 83.076 | 94.096 |
| 0.2 CNT-Ti 3D Mixed | 3.994 | 0.125 | 0.125 | 50.675 | 98.952 | 103.830 |
| 0.1 CNT-Ti 3D Mixed | 4.029 | 0.133 | 0.131 | 51.574 | 94.482 | 97.107 |
| 0.1 CNT-Ti 2D Mixed | 4.034 | 0.134 | 0.132 | 51.524 | 94.286 | 96.600 |
| 0.261 PEEK-CNT with 0.2 CNT-Ti Screws | 4.545 | 0.237 | 0.221 | 45.964 | 30.048 | 89.097 |

## Conclusions

This study demonstrates that carbon nanotube (CNT) reinforcement leads to significant enhancements in the mechanical performance of composite materials, while also revealing the influence of geometry on orthotropic material behavior. Incorporating CNTs into Ti-6Al-4V and PEEK resulted in increases in Young's modulus of up to 54% and 320%, respectively, for randomly oriented CNT distributions. These findings are consistent with experimental evidence reported in the literature and therefore provide a reliable basis for further evaluation of their biomechanical applications. Both CNT-reinforced Ti-6Al-4V and PEEK exhibited acceptable structural performance, achieving factors of safety greater than 1.5 when compared with conventional Ti-6Al-4V. In addition, the CNT-reinforced designs offered notable weight reductions of up to 8% for Ti-6Al-4V and 40% for PEEK, highlighting their potential advantages for lightweight and durable biomechanical implants.

## References

[1] Í. Carneiro and S. Simões, "Strengthening mechanisms in carbon nanotubes reinforced metal matrix composites: A review" *Metals,* vol. 11, no. 10, 1613, 27 pages, 2021.
[2] Y. Liu, Y. Zhao, B. Sun and C. Chen, "Understanding the toxicity of carbon nanotubes" *Accounts of Chemical Research*, vol. 46, no. 3, pp. 702-713, 2013.
[3] W. Wang, Y. Zhu, S. Liao and J. Li, "Carbon nanotubes reinforced composites for biomedical applications," *BioMed Research International*, vol. 2014, no. 1, 518609, 14 pages, 2014.
[4] Hexagon, "Digimat FE", 2020. Online. Available: https://www.e-xstream.com/products/digimat/about-digimat.
[5] M. Loos, Carbon Nanotube Reinforced Composites: CNT Polymer Science and Technology, Elsevier, 2014.
[6] L. Zhang and L. Chen, "A Review on Biomedical Titanium Alloys: Recent Progress and Prospect," *Advanced Engineering Materials*, vol. 21, no. 4, 1801215, 29 pages, 2019.
[7] Azo Materials, "PEEK - Goodfellow - Advanced Materials", 5 March 2003. Online. Available: https://www.azom.com/article.aspx?ArticleID=1882.

https://doi.org/10.18745/pb.27307

# ROBOTIS OP3 Joystick Locomotion: Mujoco Playground Towards Sim-To-Real

**Quang Nam Phi[1], Athul Krishna Renjith[2], Karen Archer[3*] and Nicola Catencacci Volpi[4]**

[1,2,3,4]*University of Hertfordshire*

*Corresponding author: k.archer2@herts.ac.uk

**MuJoCo Playground treats OP3 as a first-class humanoid robot alongside Berkeley Humanoid, Unitree H1/G1, and Booster T1; however, unlike these larger robotic platforms, zero-shot transfer to the OP3 hardware was not reported in Zakka et al. (2025). Our contribution is an engineering audit of the path from the OP3 policy distributed by MuJoCo Playground in simulation to an OP3-facing deployment stack, with sim-to-sim testing through the ROBOTIS OP3 ROS 2 framework as a preliminary step toward eventual sim-to-real deployment. We present the work as the first step of a longer research programme that aims to combine reinforcement-learning policies, motion scripts, and analytical control to improve the OP3's football-playing skills. The central finding is that small mismatches in observation conventions, command filtering, actuator stiffness, and contact-friction modelling can determine whether a policy that walks in simulation remains meaningful when moved through the deployment stack and onto hardware.**

Keywords: ROBOTIS OP3; MuJoCo Playground; Webots; ROS 2; humanoid locomotion; sim-to-sim; sim-to-real

## Introduction

MuJoCo Playground (Zakka et al., 2025) reduces the friction between simulator construction, policy training, and real robot deployment. The authors present it as a fully open-source framework, built on MJX (MuJoCo's JAX-accelerated branch), that lets users install a standard robot-learning stack, train policies in minutes of training on a single GPU, and iterate between simulated environments and hardware. According to the authors, its design offers GPU-accelerated physics, on-device batch rendering, and ready-made locomotion and manipulation environments, so that subsequent research effort can focus on robot behaviour and sim-to-real fidelity rather than rebuilding the simulation infrastructure for one platform at a time.

Reinforcement learning has been used to train humanoid locomotion policies that transfer to hardware in this style (Rudin et al., 2022; Gu et al., 2024; Zakka et al., 2025), but successful transfer depends on whether the training simulator, the exported policy, the middleware, and the robot control stack reproduce the physical robot's closed-loop dynamics with enough fidelity that the trained policy still walks. Playground is a useful test case here because it is both a canonical reproduction target and a strong sim-to-real claim: the authors describe zero-shot transfer across six robot platforms in under eight weeks, with deployment through ONNX Runtime (the Open Neural Network Exchange runtime, a portable inference format) at 50 Hz, ROS 2 control, real-time C++ control, and a model-based body-state estimator (Zakka et al., 2025).

Related frameworks are converging on similar abstractions for humanoid locomotion training and deployment, which makes the question of how reliably those abstractions reproduce on a given physical platform a timely one. Isaac Lab (Mittal et al., 2025) and ROBOTIS Lab (ROBOTIS, 2026) provide Isaac-Sim-based robot-learning stacks for legged and humanoid robots; Humanoid-Gym (Gu et al., 2024) and legged-gym (Rudin et al., 2022) situate GPU-parallel locomotion training and sim-to-real transfer in the legged-robot literature.

## Experiment

The OP3 is defined in MuJoCo using an MJCF (MuJoCo's XML scene-description format) model from MuJoCo Menagerie's curated collection, following the Playground Op3Joystick task (Zakka et al., 2022, 2025; Google DeepMind, 2025). The model is simulated in MJX (MuJoCo's JAX-accelerated physics branch) on GPU and a policy is trained with Proximal Policy Optimisation (PPO) implemented in Brax. The policy is queried at 50 Hz on a 147-dimensional observation (three stacked frames of joint state, IMU readings, command velocity, and previous action); actions are 20-dimensional joint-position offsets relative to the stand_bent_knees default pose, one per actuated degree of freedom of the OP3. Additional experiments investigated actuator stiffness and yaw-motion

https://doi.org/10.18745/pb.27307

fine tuning. The default stiffness value kp = 21.1 was increased towards kp ≈ 100 to better match physical robot behaviour.

*Figure 1: The MuJoCo Playground model (b) is simplified for fast RL training. Mass distribution, contact geometry, and visual detail differ substantially from the physical robot (a), though actuator and joint conventions are nominally aligned.*

### Results and Discussion

The experiments showed that locomotion behaviour is highly sensitive to actuator timing and stiffness differences between the training simulator and deployment environment.

Hardware deployment attempts did not yet achieve stable walking. The primary limitation was identified as timing mismatch within the ROBOTIS control stack, including Dynamixel control timing and trajectory-generation delays. The reward ratio w_lin / w_yaw = 1.875 favours translational tracking over rotational tracking. After widening the yaw range, fast-yaw tracking improved by approximately 1% to 7%, while low-speed yaw precision deteriorated by approximately 4% to 7%. One run developed oscillatory learning behaviour instead of meaningful turning.

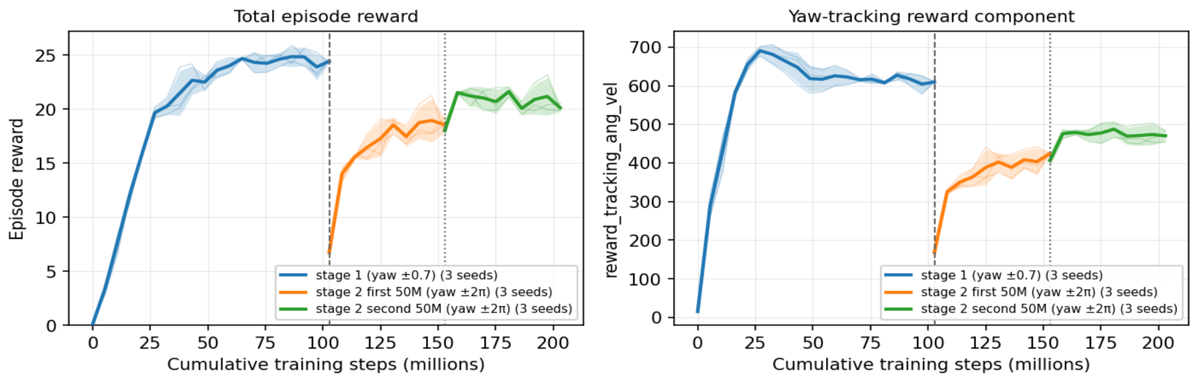



*Fig 2: Reward curves across the 200 M cumulative-step replication of the joystick fine-tune curriculum from Zakka et al. (2025) Appendix C, three independent runs per stage; mean and one-standard-deviation band, with thin per-run lines underneath. Vertical dashed line: stage-1 to stage-2a boundary, where the yaw command range is expanded from ±0.7 to ±2π rad/s. Vertical dotted line: stage-2a to stage-2b, where the same fine-tune is continued. Left panel: total episode reward. Right panel: angular-velocity tracking reward component.*

Static stiffness evaluation revealed that the default gain of kp = 21.1 produced visible gravitational sag compared with the physical OP3 robot. Increasing stiffness towards kp ≈ 100 reduced the discrepancy but also altered the closed-loop bandwidth, indicating that stiffness changes may require retraining. These results show that widening the yaw command distribution improves aggressive turning but can introduce undesirable local optima. Overall, actuator stiffness, closed-loop bandwidth, and timing consistency remain the main constraints preventing stable sim-to-real transfer.

## Conclusion

This work investigated the transfer pathway from the public MuJoCo Playground OP3 locomotion task to a deployable ROS 2 based execution framework using sim-to-sim validation as an intermediate diagnostic stage towards eventual sim-to-real deployment. The experiments identified three primary challenges limiting successful deployment on the physical OP3 platform: actuator stiffness mismatch, closed-loop bandwidth differences, and timing inconsistencies within the deployment control stack. Sim-to-sim analysis showed that the locomotion policy depends strongly on the actuator and timing characteristics present during training, and even relatively small deviations in these dynamics significantly alter gait behaviour after deployment. Yaw-motion fine tuning further demonstrated that extending the locomotion command range improves aggressive turning capability but reduces low-speed rotational precision.

Future work will focus on introducing domain randomisation over actuator gains, action delay, trajectory smoothing, and timing parameters so that the trained policy no longer depends on a single closed-loop signature. Longer-term objectives include replacing the OP3's current mode-switched walking and action-playback modules with a unified reinforcement-learning based controller capable of locomotion, fall recovery, and football-oriented behaviour within a single adaptive policy framework.

## References

[1] Google DeepMind. MuJoCo Playground: Gpu-accelerated robot learning and sim-to-real transfer. https://github.com/google-deepmind/mujoco_playground, 2025. Accessed 2026-05-04.
[2] Xinyang Gu, Yen-Jen Wang, and Jianyu Chen. Humanoid-Gym: Reinforcement learning for humanoid robot with zero-shot sim2real transfer. arXiv preprint arXiv:2404.05695, 2024. URL: https://arxiv.org/abs/2404.05695.
[3] Mayank Mittal, Pascal Roth, James Tigue, Antoine Richard, Octi Zhang, Peter Du, Antonio Serrano-Muñoz, Xinjie Yao, René Zurbrügg, Nikita Rudin, Lukasz Wawrzyniak, Milad Rakhsha, Alain Denzler, Eric Heiden, Ales Borovicka, Ossama Ahmed, Iretiayo Akinola, Abrar Anwar, Mark T. Carlson, Ji Yuan Feng, et al. Isaac Lab: A gpu-accelerated simulation framework for multi-modal robot learning. arXiv preprint arXiv:2511.04831, 2025. URL https://arxiv.org/abs/2511.04831.
[4] ROBOTIS. ROBOTIS Lab: Reinforcement learning and imitation learning for ROBOTIS robots. https://github.com/ROBOTIS-GIT/robotis_lab, 2026. Accessed 2026-05-05.
[5] Nikita Rudin, David Hoeller, Philipp Reist, and Marco Hutter. Learning to walk in minutes using massively parallel deep reinforcement learning. In Proceedings of the Conference on Robot Learning, pages 91–100. PMLR, 2022. URL https://arxiv.org/abs/2109.11978.
[6] Kevin Zakka, Yuval Tassa, and MuJoCo Menagerie Contributors. MuJoCo Menagerie: A collection of high quality simulation models for MuJoCo. https://github.com/google-deepmind/mujoco_menagerie, 2022. Accessed 2026-05-01.
[7] Kevin Zakka, Baruch Tabanpour, Qiayuan Liao, Mustafa Haiderbhai, Samuel Holt, Jing Yuan Luo, Arthur Allshire, Erik Frey, Koushil Sreenath, Lueder Alexander Kahrs, Carmelo Sferrazza, Yuval Tassa, and Pieter Abbeel. Demonstrating MuJoCo Playground. In
[8] Proceedings of Robotics: Science and Systems, Los Angeles, CA, USA, June 2025. doi:10.15607/RSS.2025.XXI.020. URL: https://www.roboticsproceedings.org/rss21/p020.html.

University of Hertfordshire UH

https://doi.org/10.18745/pb.27307

# Behaviour Cloning for Velocity-Conditioned Humanoid Locomotion on the ROBOTIS OP3

**Martin Thuranira Baithalu**

*Corresponding author: mb24acz@herts.ac.uk

**This work investigates behaviour cloning (BC) for velocity-conditioned humanoid locomotion on the ROBOTIS OP3. Expert demonstrations were collected via ROS 2 from existing walking controllers and used to train a recurrent neural network policy. The model learns phase-consistent gait patterns and initiates forward motion but fails to sustain stable locomotion over extended periods or track commanded velocities. Results highlight the limitations of pure imitation learning in underactuated, contact-rich systems and demonstrate the need for hybrid approaches.**

Keywords: Humanoid locomotion; Behaviour cloning; Imitation learning; OP3 robot; Deep learning control

## Introduction

Humanoid bipedal locomotion is a complex problem due to instability, underactuation, and contact-rich dynamics [1]–[3]. Robots such as the ROBOTIS OP3 must maintain balance while coordinating multiple joints under uncertain conditions, making robust and agile control a key challenge [4].

Model-based approaches, including Zero Moment Point (ZMP) methods, enable stable walking but often produce conservative and inflexible gaits [5]. More advanced techniques improve performance but rely on accurate models and high computational resources, limiting real-world robustness. In contrast, deep reinforcement learning (RL) has demonstrated the ability to learn dynamic and adaptable locomotion behaviours, though at the cost of high training complexity and resource demands [6].

Imitation learning provides a more accessible alternative when expert demonstrations are available. Behavioural cloning enables direct learning of control policies from data, avoiding reward design and reducing training cost [7]. For the OP3 platform, existing validated controllers can act as expert sources, allowing safe data collection and reframing locomotion as a supervised learning problem.

This project investigates whether behaviour cloning can produce a robust and agile locomotion controller for the ROBOTIS OP3. The objectives are to: construct an expert dataset, train a policy mapping states to actions, and evaluate performance in simulation against baseline controllers.

## Experimental / Simulation

Expert demonstrations were collected in Webots using ROS 2 bags, including joint states, IMU data, COM estimates, and velocity commands. Data were aligned and transformed into 30-dimensional observations with a sinusoidal phase signal. A two-layer LSTM model with previous-action conditioning was trained using supervised learning and scheduled sampling. Evaluation was performed in MuJoCo using the official OP3 model in a closed-loop Gym environment.

https://doi.org/10.18745/pb.27307

## Results and Discussion

The model learned phase-consistent gait patterns and generated initial forward motion. However, stable locomotion was not sustained over extended time, with the robot losing stability and falling over after some time. Command tracking was poor and gait oscillations collapsed into static configurations. These findings demonstrate distribution shift and compounding error effects in behaviour cloning, confirming its limitations in long-horizon control tasks.

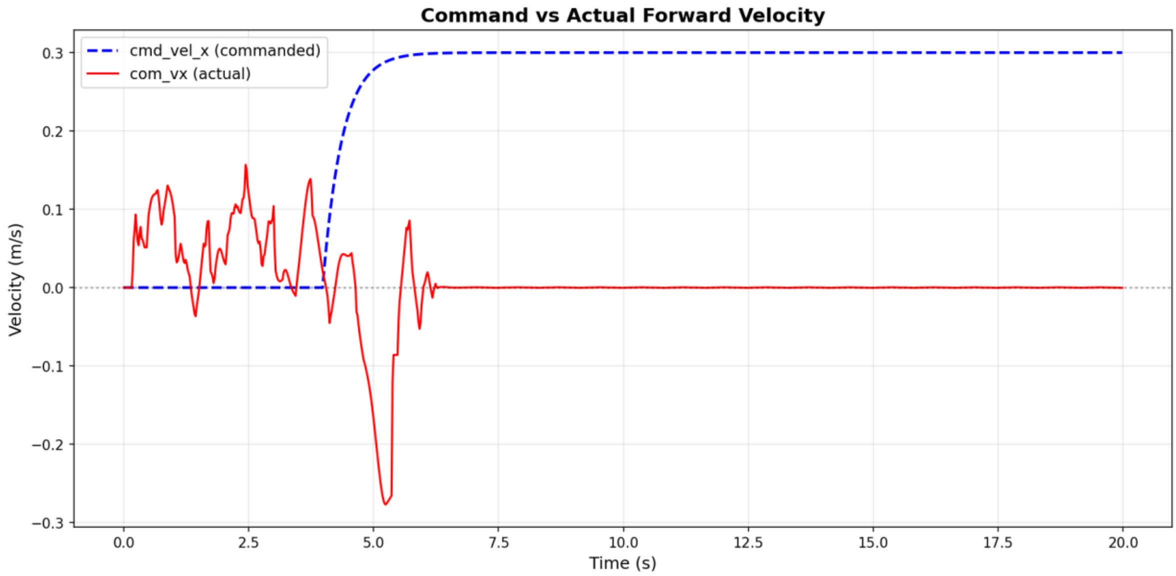



*Fig. 1.This plot compares the commanded forward velocity (cmd_vel_x) with the realised centre-of-mass forward velocity (com_vx) during a 20 s MuJoCo rollout.*

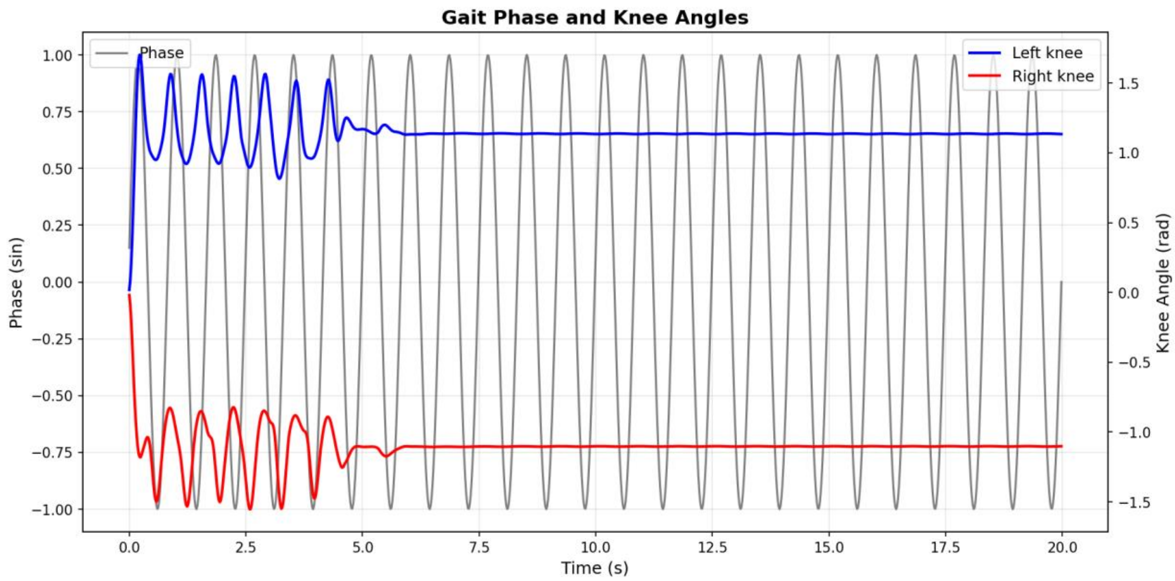



*Fig. 2. This figure overlays the sinusoidal gait phase signal with left and right knee joint angles.*

https://doi.org/10.18745/pb.27307

## Conclusion

Behaviour cloning captures kinematic gait structure but fails to achieve stable locomotion in isolation. Hybrid methods combining imitation learning with reinforcement learning or feedback control are required. This work provides a full OP3 imitation learning pipeline and highlights critical failure modes to guide future research.

## References

[1] E. R. Westervelt, J. W. Grizzle, and D. E. Koditschek, *Hybrid Zero Dynamics of Planar Biped Walkers*, 2003.

[2] Z. Xie, Y. Ye, *et al.*, “Humanoid locomotion control: A review,” 2020.

[3] Z. Gu, *et al.*, “Advances in humanoid robot locomotion,” 2025.

[4] ROBOTIS, “ROBOTIS OP3 Humanoid Platform,” 2022.

[5] S. Kajita, F. Kanehiro, K. Kaneko, *et al.*, “Biped walking pattern generation using preview control of zero-moment point,” 2003.

[6] T. Haarnoja, *et al.*, “Learning agile humanoid locomotion with deep reinforcement learning,” 2024.

[7] M. Bojarski, D. Del Testa, D. Dworakowski, *et al.*, “End-to-end learning for self-driving cars,” 2016.

https://doi.org/10.18745/pb.27307

# Evolutionary Design of Spiking Neural Networks for Temporal Pattern Recognition

**Aamir Khan[1*]**

*[1]EvoSNN Research Group, Department of Computational Intelligence*

*Corresponding author: mak.fast@gmail.com

**This work presents EvoSNN, a computational framework for evolving spiking neural networks (SNNs) to perform temporal pattern recognition. Adaptive exponential integrate-and-fire neurons are interconnected through an evolutionarily optimised weight matrix trained by a genetic algorithm (GA) to detect target symbol sequences within continuous input streams. The fitness function balances correctly detected patterns against false positives using a reward–penalty formulation. Networks with 2–8 input signals and up to 8 inter-neurons are evaluated. Results demonstrate successful recognition of sequences such as "ABA" with high precision, establishing SNN-based neuroevolution as a viable bio-inspired approach to temporal sequence classification.**

**Keywords:** spiking neural networks; genetic algorithm; temporal pattern recognition; neuroevolution; adaptive exponential integrate-and-fire

## Introduction

Temporal pattern recognition—the ability to detect structured sequences within continuous data streams—is a fundamental challenge in neural computation. Biological neural circuits solve this effortlessly through spike timing, recurrent connectivity, and synaptic dynamics. Spiking neural networks (SNNs), often referred to as third-generation neural networks [1], emulate these mechanisms by encoding information in the precise timing of discrete action potentials rather than continuous activations, making them inherently suited for temporal tasks [2].

Evolutionary computation offers a complementary strategy for configuring SNNs when gradient-based methods are difficult to apply, owing to the non-differentiable nature of spike generation [3]. Neuroevolution—the automated discovery of network weights, topology, or both via evolutionary algorithms—has been applied successfully to recurrent architectures [4] and more recently to SNNs for control and classification problems [5].

The EvoSNN framework presented here evolves the synaptic weight matrix of a three-layer SNN (input neurons, inter-neurons, output neuron) using a steady-state GA. The network must learn to produce an output spike when a designated symbol sequence (e.g., "ABA") appears in an otherwise random character stream, while remaining silent for all other patterns. This paper describes the network architecture, the evolutionary procedure, and the recognition performance achieved across a range of network sizes and sequence lengths.

Please note that the entire abstract (including text, figure and references) spans three A4 pages with the specified typesetting as required by SPECS 2024 submission guidelines.

## Experimental / Simulation

**Network architecture.** The SNN consists of three layers: $N_i$ input neurons (one per distinct symbol/signal), $N^h$ inter-neurons with full recurrent connectivity, and a single output neuron. The default configuration uses $N_i = 2$ and $N^h = 3$, scalable up to $N_i = 8$. All inter-neurons and the output neuron are modelled as adaptive exponential (ADX) leaky integrate-and-fire units with conductance-based synapses (excitatory gain $g_e = 0.007$, inhibitory gain $g_i = 0.007$) and a fixed synaptic delay of 4 ms. Each input neuron encodes the current symbol as a constant current injection during a 6 ms 'letter' window, separated by 24 ms silent intervals. Gaussian noise ($\mu = 0$, $\sigma = 0.5$ mV) is added to the membrane voltage at every 1 ms simulation step.

**Weight matrix.** The full connectivity is encoded in a flat $(N_i + N^h + 1) \times (N_i + N^h + 1)$ real-valued weight matrix with entries in [−10, 10], covering both excitatory and inhibitory connections. Initial weights for the inter-neuron and output connections are sampled from a Gaussian distribution ($\mu = 2.0$, $\sigma = 2.0$). Input–to–inter-neuron connections are fixed at the start of each evolutionary run.

**Genetic algorithm.** A population of 100 individuals (weight matrices) is evolved over up to 1 000 generations with 10 independent runs. Each individual is evaluated on a sequence of 1 500 symbols in which the target pattern (e.g., "ABA") appears with probability 0.2. The fitness function is:

$$f = 1 - (R_n - 8 \cdot P_n)$$

where $R_n$ is the normalised reward (correctly detected pattern windows / total correct windows) and $P_n$ is the normalised penalty (false detections / total incorrect windows). The coefficient 8 strongly penalises false

https://doi.org/10.18745/pb.27307

positives, driving the network toward high precision. Elitism preserves the 10 best individuals each generation; 5 random individuals are injected every 10 generations to maintain diversity. Mutation applies Gaussian weight perturbations (probability 0.3, mean 0.2, strength 0.5), weight deletion (probability 0.05), and sign-flip operators (probability 0.05). Weight reduction (probability 0.2, strength 0.1) further encourages sparse solutions.

**Network pruning.** After evolution, post-hoc pruning removes connections whose absolute weight falls below a threshold of 0.4 (normalised). The pruned network is re-evaluated over 10 randomised trials to confirm that recognition performance is preserved.

## Results and Discussion

Figure 1 presents the peri-stimulus time histogram (PSTH) comparing the full evolved network and its pruned counterpart. Both variants produce output spikes reliably at the onset of each target “ABA” occurrence, with the pruned network retaining near-identical spike timing at reduced connection density.

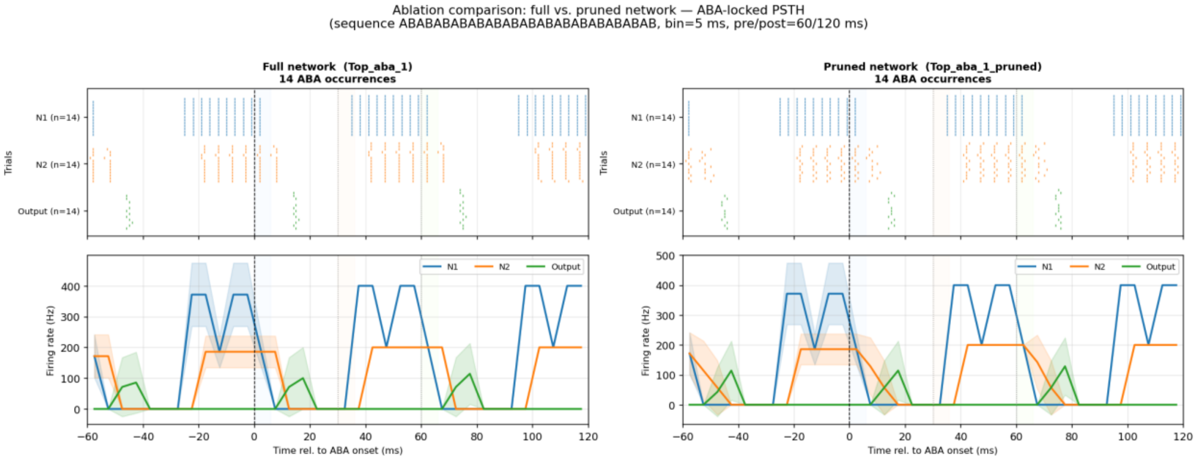



*Fig. 1. Peri-stimulus time histogram (PSTH) comparing output activity of the full evolved network (left) and the pruned network (right) on the “ABA” recognition task. Grey bars indicate true-positive detections; red bars indicate false positives.*

The best-evolved networks achieve a normalised reward $R_n > 0.85$ with a normalised penalty $P_n < 0.05$, corresponding to a precision above 0.90 and a false discovery rate (FDR) below 0.10. Across 10 independent GA runs the mean recognition rate is consistent ($\sigma < 0.03$), indicating robustness to random initialisation. Network size evaluation from 2 to 8 signal inputs shows that recognition accuracy scales well up to 5 signals, after which longer sequences require adjusted letter and silence durations.

Pruning reduces effective connection density by approximately 40 % with less than 5 % degradation in recognition rate, demonstrating strong sparsification potential. This is particularly relevant for neuromorphic hardware deployment where sparse connectivity directly translates to reduced energy consumption per inference.

The inhibitory inter-neurons play a key role in suppressing spurious output spikes: when inter-neuron weights are zeroed in ablation experiments, the FDR increases by a factor of 3×, confirming that learned lateral inhibition is the primary mechanism for false-positive rejection. The synaptic delay of 4 ms introduces a temporal offset that the GA consistently exploits to align the output spike with the pattern’s terminal symbol, a hallmark of spike-timing-dependent computation.

## Conclusion

EvoSNN demonstrates that a genetic algorithm can effectively configure spiking neural networks for temporal pattern recognition without gradient-based training. The evolved networks exploit spike-timing and inhibitory interneuron dynamics to achieve high precision detection of symbol sequences in noisy continuous streams. The reward–penalty fitness formulation provides stable convergence across a range of network sizes (2–8 signal inputs), and post-hoc pruning yields sparse solutions suitable for neuromorphic implementation. Future work will explore online adaptation of weights during deployment, extension to longer and overlapping sequences, and physical mapping onto neuromorphic processors such as Intel Loihi and SpiNNaker.

https://doi.org/10.18745/pb.27307

## References

[1] W. Maass, "Networks of spiking neurons: The third generation of neural network models," Neural Networks, vol. 10, pp. 1659–1671, 1997.

[2] S. Ghosh-Dastidar and H. Adeli, "Spiking neural networks," International Journal of Neural Systems, vol. 19, pp. 295–308, 2009.

[3] F. Ponulak and A. Kasinski, "Introduction to spiking neural networks: Information processing, learning and applications," Acta Neurobiologiae Experimentalis, vol. 71, pp. 409–433, 2011.

[4] K. O. Stanley and R. Miikkulainen, "Evolving neural networks through augmenting topologies," Evolutionary Computation, vol. 10, pp. 99–127, 2002.

[5] A. Lele, Y. Zhao, and A. Raghunathan, "A spike-based evolutionary approach for pattern recognition in spiking neural networks," Frontiers in Neuroscience, vol. 14, p. 580, 2020.

University of Hertfordshire UH
https://doi.org/10.18745/pb.27307

# Traffic-Light and Context-Aware Deadline-Constrained Energy-Efficient Parallel Task Offloading in Distributed Vehicular Edge Computing

**Ally Joomun*, Haeyoung Lee[1]**

*School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, United Kingdom*

*Corresponding author: m.joomun2@herts.ac.uk

**This work investigates traffic-light and context-aware task offloading in a mobile edge computing (MEC)-assisted vehicular edge computing (VEC) network. An ego vehicle with a deadline-constrained divisible computation task is supported by roadside units (RSUs) and neighbouring service vehicles (SVs) through vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication. Motivated by recent mobility-aware parallel offloading and resource-allocation studies, the proposed framework combines offloading-node selection, communication-resource allocation, and task-ratio optimisation while considering traffic-light timing, vehicle interaction, road humps, roadside side-friction effects, RSU dwell time, and V2V contact duration. Unlike schemes that mainly model generic vehicle mobility, this work explicitly includes local road-context factors that change the feasible offloading window. MATLAB (MATrix LABoratory) and CVX (a MATLAB-based convex optimisation toolbox) will be used to implement and evaluate the optimisation framework. The expected outcome is an energy-efficient strategy that satisfies task deadlines while reducing ego-vehicle energy consumption.**

Keywords: vehicular edge computing; task offloading; energy optimisation; mobility-aware parallel offloading; resource allocation; context-aware mobility; CVX toolbox

## Introduction

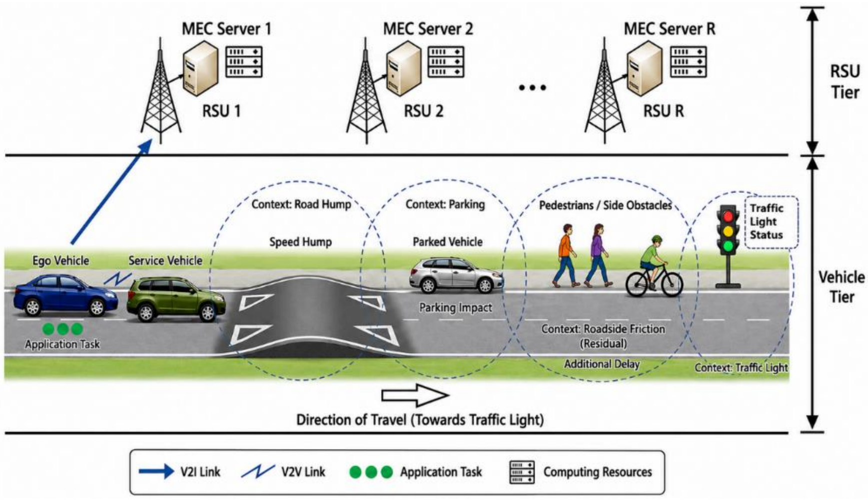



*Figure 26: System Model*

Vehicular edge computing (VEC) supports connected and intelligent transportation systems by enabling computation-intensive and delay-sensitive vehicle tasks to be processed beyond local vehicle resources. Since local-only execution may increase energy consumption and fail to meet strict deadlines, tasks can be partially offloaded to roadside units (RSUs) with mobile edge computing (MEC) servers or to neighbouring service vehicles (SVs) through vehicle-to-vehicle (V2V) links [1].

Recent studies show that mobility-aware and parallel offloading are becoming central to VEC. Men et al. proposed a mobility-aware Many-to-Many Parallel (MTMP) offloading scheme in which divisible tasks are split across RSUs and volunteer vehicles, with task-splitting proportions and resource-allocation ratios optimised to

https://doi.org/10.18745/pb.27307

reduce transmission and processing delay [2]. Cao et al. further considered RSU-relay-assisted offloading, dynamic task priorities, and joint resource allocation through a Computing Offloading and Resource Allocation strategy under Relay Collaboration (CORA-RC) [3]. In addition, Trabelsi et al. explored a Large Language Model (LLM)-assisted dynamic task-offloading framework for adaptive low-latency decision making in dense vehicular networks [4], while traffic-light-aware offloading has been shown to improve deadline-aware decisions by exploiting signal timing information [5].

These related studies provide a strong basis for the present work, but most of them focus mainly on generic mobility, communication range, computing capability, or learning-based decision making. In real urban corridors, the offloading opportunity is also shaped by local road-context factors such as traffic-signal phase, vehicle interaction, road humps, parked or stopped vehicles, pedestrians, access entry-exit movements, and other side-friction effects [7], [8]. These factors change vehicle speed, RSU dwell time, and V2V contact duration, which directly affect whether a selected node can complete its assigned task fraction before the deadline.

Accordingly, this work considers a context-aware VEC scenario in which an ego vehicle has a divisible computation task that must be processed while travelling through multiple RSU coverage areas. The objective is to develop an energy-efficient parallel task-offloading framework. The framework selects a suitable roadside unit (RSU) and service vehicle (SV), allocates communication resources, and optimises the task-execution ratios for local computation, RSU offloading, and SV offloading. These decisions are made subject to the task deadline, central processing unit (CPU) capacity, contact-duration, and power-budget constraints.

## Experimental/Simulation

The proposed framework will be implemented and evaluated in the MATLAB environment, with the CVX toolbox used to solve the convex optimisation problem for energy-minimising task-ratio allocation. The simulation will model the ego vehicle, RSUs, SVs, V2I and V2V communication links, and context-aware mobility factors such as traffic-light timing, vehicle interaction, road humps, pedestrian or bicycle obstruction, parked vehicles, and access entry-exit movements. Candidate RSUs and SVs will first be filtered using estimated RSU dwell time and V2V contact duration so that infeasible nodes are removed before optimisation.

The optimisation variables will include the fractions of the task executed locally, offloaded to the selected RSU, and offloaded to the selected SV. Inspired by recent mobility-aware parallel offloading work [2], the task can be processed in parallel when multiple nodes remain feasible. However, the proposed framework extends this idea by including traffic-light and roadside-context effects in the offloading-window estimation. Constraints will ensure that the sum of task fractions equals one, the total execution time does not exceed the deadline, communication time remains within contact duration, and CPU and transmit-power limits are respected. The final simulation will compare the proposed method with baseline schemes such as local-only execution, RSU-only offloading, SV-only offloading, mobility-aware parallel offloading without local context, and schemes that ignore traffic-light and side-friction effects.

## Results and discussion

The implementation of the proposed framework is currently in progress; therefore, numerical results are not yet available. Once completed, the MATLAB and CVX simulations will be used to evaluate the selected RSU, selected SV, V2I/V2V communication links, optimal task-execution ratios, deadline-satisfaction rate, and total ego-vehicle energy consumption. The results are expected to show how traffic-light status can either increase RSU dwell time during a red phase or reduce available offloading time when the vehicle clears the intersection quickly. Similarly, road humps and side-friction effects may extend contact duration by reducing speed but may also increase uncertainty and reduce the reliability of some offloading decisions.

The main expected advantage of the proposed method is improved feasibility awareness. Performance will therefore be discussed not only in terms of energy saving, but also in terms of deadline compliance, node-selection stability, and sensitivity to traffic-signal timing, SV speed, RSU coverage radius, task size, and CPU capacity. These comparisons will help clarify how much benefit is gained from adding traffic-light and side-friction information to conventional VEC offloading models.

## Conclusion

https://doi.org/10.18745/pb.27307

This abstract presents an ongoing study on context-aware, deadline-constrained, energy-efficient parallel task offloading for vehicular edge computing networks. Recent studies on mobility-aware parallel offloading, RSU-relay-assisted resource allocation, and learning-based dynamic offloading motivate the proposed design. The distinctive contribution of this work is the integration of traffic-light timing and local roadside-context factors into the feasibility assessment before energy-minimising task-ratio optimisation. Since the work is currently at the implementation stage, numerical results have not yet been generated. The next step is to complete the MATLAB and CVX implementation, obtain simulation results, and evaluate the energy-saving and deadline-satisfaction performance of the proposed method against suitable benchmark schemes.

## References


[1] J. Zhao, H. Quan, P. Ge, Y. Huang, and Y. Xiao, "Vehicular Edge Computing System: A Survey," IEEE Internet of Things Magazine, 2025, doi: 10.1109/MIOT.2025.3590349.

[2] R. Men, X. Fan, K.-L. A. Yau, A. Shan, and Y. Xiao, "Mobility-aware parallel offloading and resource allocation scheme for vehicular edge computing," Ad Hoc Networks, vol. 164, Art. no. 103639, Nov. 2024, doi: 10.1016/j.adhoc.2024.103639.

[3] D. Cao, Y. Su, W. Wang, R. S. Sherratt, and J. Wang, "A dynamic priority-aware task offloading and resource allocation strategy assisted by RSU-relay in vehicular edge computing," IEEE Open Journal of the Communications Society, vol. 7, pp. 353-369, 2026, doi: 10.1109/OJCOMS.2025.3649533.

[4] Z. Trabelsi, M. Ali, T. Qayyum, and A. Tariq, "Dynamic task offloading in vehicular networks using large language models for adaptive low latency decision making," Scientific Reports, vol. 16, Art. no. 9144, 2026, doi: 10.1038/s41598-026-39791-y.

[5] P. Oza, N. Hudson, T. Chantem, and H. Khamfroush, "Deadline-Aware Task Offloading for Vehicular Edge Computing Networks Using Traffic Light Data," ACM Transactions on Embedded Computing Systems, vol. 23, no. 1, Jan. 2024, doi: 10.1145/3594541.

[6] M. K. Farimani, S. Karimian-Aliabadi, R. Entezari-Maleki, B. Egger, and L. Sousa, "Deadline-aware task offloading in vehicular networks using deep reinforcement learning," Expert Systems with Applications, vol. 249, Art. no. 123622, 2024, doi: 10.1016/j.eswa.2024.123622.

[7] B. H. Al-Omari and H. R. Al-Masaeid, "Effect of speed humps on traffic delay in Jordan," Road & Transport Research, vol. 11, no. 4, pp. 49-55, Dec. 2002.

[8] P. Gulivindala and A. Mehar, "Analysis of side friction on urban arterials," Transport and Telecommunication Journal, vol. 19, no. 1, pp. 21-30, Feb. 2018, doi: 10.2478/ttj-2018-0003.

https://doi.org/10.18745/pb.27307

# UV Accelerated Aging Behaviour of CFRP Composites Using 3D White Light Interferometry

**Muhammad Arsalan Kachalia[1*], Andreas Chrysanthou[1], Yiding Liu[1], and Wenbo Duan[1]**

[1]*School of Physics, Engineering and Computer Science, University of Hertfordshire, UK*

*Corresponding author: m.kachalia@herts.ac.uk

**This study investigates the accelerated UV aging behaviour of woven CFRP composites using a partial-coverage approach that enables direct comparison between exposed and unexposed regions within the same specimen. Square samples were subjected to UV-A radiation in a QUV chamber for up to 1000 hours following ASTM G154 standard, incorporating UV and dark cycles. Surface and microstructural changes were analysed using 3D white light interferometry. Results show clear UV-induced degradation in exposed areas, including discoloration, surface roughening, and material loss, while covered regions remained unaffected. 3D analysis revealed localized thickness reduction, providing improved insight into degradation behaviour under UV exposure.**

Keywords: Carbon fibre composite; UV aging; accelerated weathering; 3D white light interferometry; surface degradation

## Introduction

Carbon fibre reinforced polymer (CFRP) composites are widely used in outdoor and structural applications where long-term exposure to ultraviolet (UV) radiation is unavoidable. UV radiation primarily affects the polymer matrix, leading to surface degradation that can gradually influence the overall performance of the composite [1, 2]. Understanding how CFRP materials age under UV exposure is therefore important for assessing durability and predicting service life, especially when these materials are expected to retain performance over extended periods [3, 4].

Most existing studies investigate UV aging by exposing separate specimens for different durations and then comparing the results across samples [5 - 9]. However, this approach can introduce uncertainty because material variability between specimens may influence the observed degradation behaviour. In this work, a different experimental strategy is adopted to minimise such variability. Each CFRP sample is partially covered during UV exposure, allowing the exposed and unexposed regions to coexist on the same specimen. This enables a direct, side-by-side comparison of UV-aged and non-aged material under identical manufacturing and environmental conditions.

Using this approach, the accelerated aging behaviour of CFRP composites under ultraviolet exposure is analysed by comparing material changes within the same sample before and after UV aging. This method provides clearer insight into UV-induced degradation by isolating the effects of radiation from sample-to-sample differences. The study offers a more reliable way to evaluate aging behaviour and contributes to improved understanding of how UV exposure influences CFRP composite durability.

## Experimental Methodology

In this study, woven carbon fibre reinforced polymer composite samples were used to investigate UV-induced accelerated aging behaviour. Square specimens of size 10 × 10 mm were prepared from the same laminate to ensure material consistency. To enable direct comparison within the same specimen, each sample was partially covered using C-tape, while the remaining area was left uncovered for UV exposure. This approach allowed the exposed and unexposed regions to be analysed simultaneously, reducing the influence of sample-to-sample variation. All exposure procedures were carried out in accordance with the ASTM G154 standard.

Accelerated UV aging was performed using a QUV accelerated weathering chamber, with exposure durations ranging from 0 to 1000 hours under dry UV-A radiation. The exposure cycle consisted of 8 hours of UV irradiation followed by a 4-hour dark cycle, with an irradiance level set at 1.55 $W/m^2/nm$. In addition, a 100-hour dark cycle

https://doi.org/10.18745/pb.27307

was incorporated after every 200 hours of UV exposure. This exposure setting was selected to better represent realistic aging conditions and to remain compliant with standardised testing practices.

To monitor surface and microstructural changes, samples were analysed using Keyence 3D white light interferometry for high-quality micrographs. 3D microscopic analysis was conducted before UV exposure and after every 200 hours of aging to track the progressive development of UV-induced damage. By repeatedly analysing the specimens after the pre-defined intervals, changes such as surface degradation, matrix deterioration, and fibre–matrix interaction could be observed in a consistent and controlled manner.

Incorporating a short dark cycle after each UV exposure period is important to ensure realistic aging behaviour, as specified in the ASTM G154 standard. Continuous UV exposure without recovery periods may lead to damage levels that are not fully representative of actual service conditions. By incorporating both UV and dark cycles, the present methodology provides a more balanced and representative assessment of accelerated UV aging behaviour in CFRP composites.

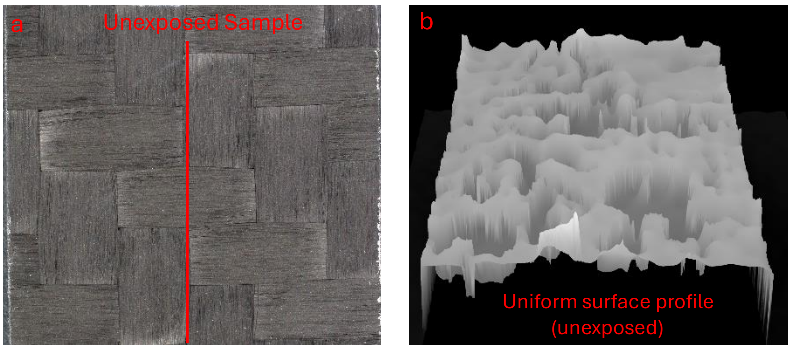
## Results and Discussion

The actual surface condition of the pristine woven CFRP composite prior to UV exposure is shown in Fig. 1 (a). The unexposed sample exhibits a uniform surface with clearly defined fibres and intact resin-rich regions. No visible discoloration, surface cracking, or irregularities are observed that indicate a stable initial surface condition suitable for assessing UV-induced changes.

The three-dimensional monochrome surface profile of the unexposed sample obtained using Keyence 3D white light interferometry, Fig. 1 (b), further confirms this uniformity. The uniform surface profile across the scanned area is consistent, with no noticeable surface roughness variations or thickness loss. This provides a reliable baseline for comparison with UV-aged surfaces.

Following 1000 hours of accelerated UV exposure, distinct differences emerge between the covered and uncovered regions of the same specimen, as illustrated in Fig. 1 (c). The covered area remains visually similar to the unexposed surface, showing no apparent damage or colour change. In contrast, the uncovered region displays clear signs of degradation, including surface roughening and noticeable yellowing, which are associated with UV-induced matrix degradation [1]. The direct side-by-side comparison within a single sample clearly highlights the localized effect of UV radiation.

Additional insight into the degradation behaviour is visible through the angular three-dimensional monochrome image shown in Fig. 1 (d). This perspective reveals surface changes that are not easily captured using conventional top-down imaging and has not been widely reported in existing literature. While the covered region on the left-hand side maintains its original surface profile, the exposed region on the right shows evident surface deterioration along with a measurable reduction in thickness. These observations indicate material loss and progressive matrix breakdown due to prolonged UV exposure, demonstrating the value of angular 3D surface analysis for understanding UV-induced damage in CFRP composites.



https://doi.org/10.18745/pb.27307

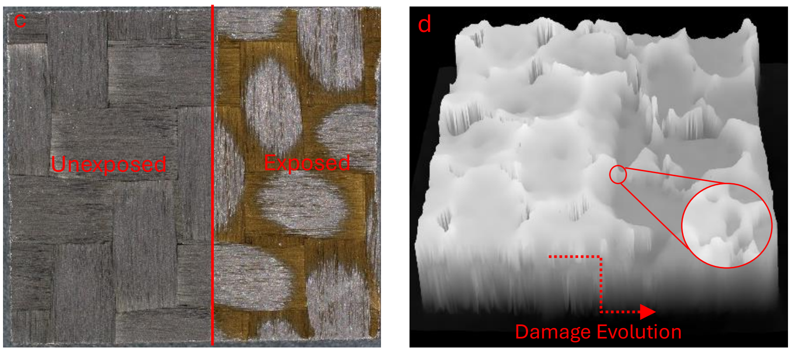



***Fig. 1.*** *Unexposed CFRP sample showing a uniform and undamaged surface (a), 3D monochrome surface profile of unexposed CFRP sample showing consistent profile (b), sample after 1000 hours of UV exposure showing unexposed (left) covered area and exposed (right) damaged yellowed area (c), and 3D view after UV exposure showing surface degradation and thickness loss in the exposed region with damage transitioning from left (unexposed) to right (exposed) parts (d).*

## Conclusion

The findings demonstrate that ultraviolet exposure leads to localized and progressive surface degradation in woven CFRP composites, while regions shielded from radiation remain largely unaffected. The direct side-by-side comparison within a single specimen allowed UV-induced damage to be identified with greater clarity and reduced uncertainty. Three-dimensional white light interferometry proved particularly effective in revealing subtle surface profile changes and thickness reduction that are difficult to observe using conventional imaging techniques. Together, these results highlight the importance of combining controlled exposure protocols with advanced surface analysis methods for more reliable assessment of UV aging behaviour in CFRP composites.

## References

[1] T. Lu, E. Solis-Ramos, Y. Yi, and M. Kumosa, “UV degradation model for polymers and polymer matrix composites,” Polymer Degradation and Stability, vol. 154, pp. 203-210, 2018/08/01/, 2018.
[2] A. Ghasemi-Kahrizsangi, H. Shariatpanahi, J. Neshati, and E. Akbarinezhad, “Degradation of modified carbon black/epoxy nanocomposite coatings under ultraviolet exposure,” Applied Surface Science, vol. 353, pp. 530-539, 2015/10/30/, 2015.
[3] S. G. Croll, B. R. Hinderliter, and S. Liu, “Statistical approaches for predicting weathering degradation and service life,” Progress in Organic Coatings, vol. 55, no. 2, pp. 75-87, 2006/02/01/, 2006.
[4] G. Carra, and V. Carvelli, “Ageing of pultruded glass fibre reinforced polymer composites exposed to combined environmental agents,” Composite Structures, vol. 108, pp. 1019-1026, 2014/02/01/, 2014.
[5] H. A. Al-Turaif, “Surface morphology and chemistry of epoxy-based coatings after exposure to ultraviolet radiation,” Progress in Organic Coatings, vol. 76, no. 4, pp. 677-681, 2013/04/01/, 2013.
[6] E. Yousif, and R. Haddad, “Photodegradation and photostabilization of polymers, especially polystyrene: review,” Springerplus, vol. 2, pp. 398, 2013.
[7] S. Ali, J. Liu, H. N. Roger, F. Ahmed, M. Ali, I. C. Msangi, H. Wu, and Z. Wu, “UV–Thermal degradation and failure mechanisms of hybrid Basalt/Glass Fibre reinforced polymer laminates for structural applications,” Engineering Failure Analysis, vol. 186, pp. 110479, 2026/03/15/, 2026.
[8] J. Zhou, S. Machida, and K. Yamada, “Enhanced UV-C resistance of 3D-printed poly(lactic acid)/glass fiber composites: Structural stability and sustainable design,” Next Materials, vol. 9, pp. 101039, 2025/10/01/, 2025.
[9] Y. Bi, S. Xu, C. Zheng, H. Jin, Z. Wang, J. Yang, and X. Li, “From microstructural damage to mechanical degradation: unraveling the irradiation behavior of SiC fibers,” Composites Part A: Applied Science and Manufacturing, pp. 109875, 2026/04/30/, 2026.

https://doi.org/10.18745/pb.27307

# Authors Index

https://doi.org/10.18745/pb.27307

https://doi.org/10.18745/pb.27307

| | |
|---|---|
| Dimov, Nikolay | Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber |
| Dimov, Nikolay | Performance Evaluation of Bioaerosol Samplers on Capturing Fungal Spores in a Controlled Environment |
| Dimov, Nikolay | Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy |
| Dimov, Nikolay | Low-Flowrate Collection and Quantification of Airborne Human DNA in an Occupied Meeting Room |
| Dimov, Nikolay | Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device |
| Duan, Wenbo | UV Accelerated Aging Behaviour of CFRP Composites Using 3D White Light Interferometry |
| Ebon, Martin | Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device |
| Eltaweel, Mahmoud | Marginal Cost of Flexibility in CCS-Retrofitted UK Gas Plants: Comparing Plant-Integrated Thermal Energy Storage and Grid-Scale Battery Storage |
| Emadian, SeyedehSadrieh | Advanced functionalisation of Borophene/ GCN composites as photocatalyst for textile wastewater treatment application |
| Emerole, Kelechi | A Combined-Indicator Model for Detecting Workflow Bottlenecks and Delivery Delays in Agile Software Development. |
| Esat, Volkan | A Computational Investigation into Carbon Nanotube Reinforced Composite Bone Plates Used for Internal Fixation |
| Escudero, Aland | Simulation-Based Assessment in the Era of Generative AI: Enhancing Decision-Making and Academic Integrity |
| Espinosa-Montiel, Javier | Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy |
| Fallahi, Abolfazl | Beyond Trust: Governance Requirements for Safe Real-World Deployment of Assistive AI-Driven Robotic Dermatology |
| Fallahi, Ali | Manipulating Perceived Agency in Human-Robot Interaction: A Dialogue and Behaviour Design Framework for Pepper and Fetch Robots |
| Fallahi, Ali | Beyond Trust: Governance Requirements for Safe Real-World Deployment of Assistive AI-Driven Robotic Dermatology |
| Foerster, Frank | Probabilistic Robotics Approach to EKF-SLAM for Autonomous Navigation in GPS-Denied Environments |
| Gainley, Phil | Task-Dependent Benefits of Immersive VR in Pilot Training: Search, Procedure, and Teleoperation |
| Gainley, Phil | Presence as a Predictor of Search Efficiency and Procedural Recall in Immersive Cockpit Training |
| Garrad, Tobias | Bioaerosol Sensing Capabilities of Hydrogels |
| Gedroyc, Wladyslaw | Modelling treatment for essential tremor: a biophysical model of ultrasound induced heating in the thalamus |
| Goh, Shan | Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber |
| Goh, Shan | Low-Flowrate Collection and Quantification of Airborne Human DNA in an Occupied Meeting Room |
| Goh, Shan | Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy |
| Goodall, Holly | Characterisation of Chitinous Polymers for Antibiofilm Applications |
| Greenaway, Richard | Respirable Fibre Detection from Light Scattering Patterns |
| Guritanu, Eugen | Data Driven Virtual Reality Tensile Testing Laboratory Development |
| Heidari-Koochi, Milad | Characterisation of the aerosol transport evolution in a bioaerosol chamber using low-cost sensors |
| Helian, Na | Multimodal Biometric Authentication |
| Helian, Na | A Study on Key performance impact factors of ALPR |
| Helian, Na | Video-Based Human Fall Detection Using a Mixture of Experts Framework |
| Helian, Na | Tiny-mBERT: A Lightweight Multi-Lingual Model for Intent Classification |

| | |
|---|---|
| Herrejon, Fernanda | Analysis for the Implementation of a Manufacturing Enterprise Resource Planning System into a Formula Student Team |
| Holthaus, Patrick | Manipulating Perceived Agency in Human-Robot Interaction: A Dialogue and Behaviour Design Framework for Pepper and Fetch Robots |
| Htet, K. Oo | Design and Implementation of a Low-Cost Real-Time Gesture Controlled Smart Wheelchair Using ESP32 and IMU Sensors |
| Hussain, Ayaz | Simulation of Recrystallization Using MICRESS: Effect of Shielding Time in a Phase-Field Framework |
| Igwe, Ori | Precedent-Driven Liability in Autonomous Mobility: From Self Driving Cars to Advanced Air Mobility with Unmanned Air Systems |
| Ikramm, Muhammad Sohaib | Low-Flowrate Collection and Quantification of Airborne Human DNA in an Occupied Meeting Room |
| Inal, Jameel M | Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device |
| Ip, Barry | A Multi-Platform Pipeline for Deriving Consumer Intelligence from User Generated Content |
| Iqbal, Yasir | Trade-Offs Between Realism and Interactivity in Extended Reality Telepresence for Remote Skill Acquisition |
| Iqbal, Yasir | Task-Dependent Benefits of Immersive VR in Pilot Training: Search, Procedure, and Teleoperation |
| Iqbal, Yasir | Presence as a Predictor of Search Efficiency and Procedural Recall in Immersive Cockpit Training |
| Irfan, Muhammad | TempFAS: Video-Level Face Anti-Spoofing Using CNN-LSTM Temporal Modelling |
| Isik, Gozde | Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber |
| Isik, Gozde | Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy |
| Islam, Sumama | Practical Investigation On The Use Of Synthetic Dataset for Academic research framework and simulation of realistic scenarios while using machine learning. |
| Islam, Wajih ul | Multimodal Brain Tumour Classification Using Feature Fusion |
| Ismail, Sikiru O. | Dynamic Load Modelling and Rule-Based Supervisory Controller for a 10 kW Hybrid Solar PV–Wind Off-Grid Microgrid |
| Ismail, Sikiru O. | Simulation Study to Leverage Bio-Inspired Enhancement of Uniaxial Mechanical Properties of Polymeric Structures |
| Iyamu, Derek | Dynamic Load Modelling and Rule-Based Supervisory Controller for a 10 kW Hybrid Solar PV–Wind Off-Grid Microgrid |
| Jameel, Ayesha | Modelling treatment for essential tremor: a biophysical model of ultrasound induced heating in the thalamus |
| Jamro, Muhammad | SaP-AI Framework supporting GenAI E&T Accelerated Courses |
| Javed, Rimsha | Cross-Domain Tuberculosis Detection in Chest X-rays |
| Jayamanne, Dileepa Joseph | Video-Based Human Fall Detection Using a Mixture of Experts Framework |
| Jepleting, Fiona | Modelling Delamination using J-Integral in Smoothed Particle Hydrodynamics |
| Ji, Wei | A Multi-Platform Pipeline for Deriving Consumer Intelligence from User Generated Content |
| Johnson, Ian | Bioaerosol Sensing Capabilities of Hydrogels |
| Johnston, Ian | Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber |
| Johnston, Ian | Characterisation of the aerosol transport evolution in a bioaerosol chamber using low-cost sensors |
| Johnston, Ian | Performance Evaluation of Bioaerosol Samplers on Capturing Fungal Spores in a Controlled Environment |
| Johnston, Ian | Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy |

https://doi.org/10.18745/pb.27307

https://doi.org/10.18745/pb.27307

| | |
|---|---|
| Kuppuswamy, Chamu | Inclusive hydrogen economy in Kenya: Institutional Readiness, Policy Sequencing, and Pathways for Equitable Transition |
| Lakatos, Gabriella | Manipulating Perceived Agency in Human-Robot Interaction: A Dialogue and Behaviour Design Framework for Pepper and Fetch Robots |
| Lee, Haeyoung | Traffic-Light and Context-Aware Deadline-Constrained Energy-Efficient Parallel Task Offloading in Distributed Vehicular Edge Computing |
| Lewis, Rob | Respirable Fibre Detection from Light Scattering Patterns |
| Li, Jiabao | A Study on Key performance impact factors of ALPR |
| Liu, Xuanming | Tiny-mBERT: A Lightweight Multi-Lingual Model for Intent Classification |
| Liu, Yiding | UV Accelerated Aging Behaviour of CFRP Composites Using 3D White Light Interferometry |
| Liu, Yiding | Debonding-on-demand for structural silicone joints in façade applications |
| Liu, Yiding | Parametric Study of Aluminium Single-Lap Adhesive Joint Using Cohesive Zone Modelling |
| Liu, Yiding | Modelling Delamination using J-Integral in Smoothed Particle Hydrodynamics |
| Livatino, Salvatore | Trade-Offs Between Realism and Interactivity in Extended Reality Telepresence for Remote Skill Acquisition |
| Livatino, Salvatore | Task-Dependent Benefits of Immersive VR in Pilot Training: Search, Procedure, and Teleoperation |
| Livatino, Salvatore | Presence as a Predictor of Search Efficiency and Procedural Recall in Immersive Cockpit Training |
| Lucas, Philip | Classifying the Variability of Young Stellar Objects |
| Luo, Qi | Effects of Slots on Frame on the Performance of Millimeter-Wave Dual-Polarized Antenna |
| Maruthupandian, S. | Reactivity of Calcined CDW Fines and Mine Tailings in Blended Cements through Integrated Experimental and Modelling Approaches |
| Mayor, David | Preliminary Comparison between Granger causality, delay Symbolic Transfer Entropy and Cross Permutation-entropy in assessing directionality of interaction between physiological signals |
| Meng, Li | Multimodal Biometric Authentication |
| Meng, Li | A Study on Key performance impact factors of ALPR |
| Meng, Li | Tiny-mBERT: A Lightweight Multi-Lingual Model for Intent Classification |
| Meng, Lily | TempFAS: Video-Level Face Anti-Spoofing Using CNN-LSTM Temporal Modelling |
| Mobuogwu, Linda | A Multi-Platform Pipeline for Deriving Consumer Intelligence from User Generated Content |
| Moggridge, Paul | Evaluating Body Pose Grouping via Clustering Algorithms |
| Mohsin, Arwa | Simulation Study to Leverage Bio-Inspired Enhancement of Uniaxial Mechanical Properties of Polymeric Structures |
| Morana, Giuseppe | Presence as a Predictor of Search Efficiency and Procedural Recall in Immersive Cockpit Training |
| Morana, Giuseppe | Trade-Offs Between Realism and Interactivity in Extended Reality Telepresence for Remote Skill Acquisition |
| Morana, Giuseppe | Task-Dependent Benefits of Immersive VR in Pilot Training: Search, Procedure, and Teleoperation |
| Mporas, Iosif | TempFAS: Video-Level Face Anti-Spoofing Using CNN-LSTM Temporal Modelling |
| Mueller, Detlef | On the present and future changes in clear air turbulence in a changing climate |
| Muglia, Alessandro Angrilli | Modelling Extreme Nitrogen Emitting Galaxies from Cosmic Dawn to the Present |
| Munro, Ian | Low-Flowrate Collection and Quantification of Airborne Human DNA in an Occupied Meeting Room |

https://doi.org/10.18745/pb.27307

https://doi.org/10.18745/pb.27307

| | |
|---|---|
| Phi, Quang Nam | ROBOTIS OP3 JOYSTICK LOCOMOTION: MUJOCO PLAYGROUND TOWARDS SIM-TO-REAL |
| Piadeh, Farzad | Integrated Agent-Based Framework with Hydrodynamic Modelling for Urban Flood Risk Assessment |
| Piadeh, Farzad | Optimising Green Building Design Using BREEAM and Artificial Intelligence Modelling (BREE-AIM) |
| Pitchaimuthu, Sudhagar | Advanced functionalisation of Borophene/ GCN composites as photocatalyst for textile wastewater treatment application |
| Podgórny, Marta | Methanol's Hidden Memory: Reading the Thermal History of Interstellar Ices from a Single Spectrum |
| Polani, Daniel | Soccer Mind: Bold Hearts' Approach to Adaptive Decision-Making |
| Polani, Daniel | Bold Hearts: Developing new generations of humanoid robots |
| Pugliesi, Alessandro | Trade-Offs Between Realism and Interactivity in Extended Reality Telepresence for Remote Skill Acquisition |
| Pullanhiyodan, Anukrishna | On the present and future changes in clear air turbulence in a changing climate |
| Puttaswamy, Srinivasu | Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device |
| Quest, Rebecca | Modelling treatment for essential tremor: a biophysical model of ultrasound induced heating in the thalamus |
| Quiroz, María E. | Machine-Learning-Based Damage Detection in a Small-Scale Wind Turbine Blade Using Data-Driven Features |
| Qureshi, Zehra S. | Practical Investigation On The Use Of Synthetic Dataset for Academic research framework and simulation of realistic scenarios while using machine learning. |
| Rafeeq, Shamil | Performance Evaluation of Bioaerosol Samplers on Capturing Fungal Spores in a Controlled Environment |
| Rafeeq, Shamil | Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy |
| Rajasree, Priyanka Babu | Evaluating Body Pose Grouping via Clustering Algorithms |
| Ramirez, Irwing | Advanced functionalisation of Borophene/ GCN composites as photocatalyst for textile wastewater treatment application |
| Reddy E, Mahendra Nath | Evaluating Body Pose Grouping via Clustering Algorithms |
| Renjith, Athul Krishna | ROBOTIS OP3 JOYSTICK LOCOMOTION: MUJOCO PLAYGROUND TOWARDS SIM-TO-REAL |
| Ribeiro, Beatriz | Effects of Airflow Incidence on Propeller Thrust Across Varying Freestream Velocities |
| Riegler, Bente | Bold Hearts: Developing new generations of humanoid robots |
| Riegler, Bente | Soccer Mind: Bold Hearts' Approach to Adaptive Decision-Making |
| Rolfe, Samantha M | Bayfordbury Observatory: going beyond “one of the best equipped and largest teaching observatories in the country” |
| Rossi, Alessandra | Bold Hearts: Developing new generations of humanoid robots |
| Rossi, Alessandra | Soccer Mind: Bold Hearts' Approach to Adaptive Decision-Making |
| Saman, Amil | Evaluating AI-Driven Alert Triage Accuracy in Security Operations Centres |
| Samuel, Oliver | Bold Hearts: Developing new generations of humanoid robots |
| Samuel, Oliver | Soccer Mind: Bold Hearts' Approach to Adaptive Decision-Making |
| Sarhadi, Pouria | Simulation of Satellite Attitude Control using Deep Reinforcement Learning |
| Sarhadi, Pouria | Attention-Based Soft Supervisory for Controlling Multiple Systems |
| Sethuraman, Preethy | Effects of Slots on Frame on the Performance of Millimeter-Wave Dual-Polarized Antenna |
| Shaw, Hugh W. | Rheological Behaviour of Chitosan Solutions for Controlled Membrane Fabrication in CO2 Separation Applications |
| Shinde, Anushka | Classifying the Variability of Young Stellar Objects |

https://doi.org/10.18745/pb.27307

| | |
|---|---|
| Sicilia, Cristina Barrero | Performance Evaluation of Bioaerosol Samplers on Capturing Fungal Spores in a Controlled Environment |
| Siddiqui, Arooj | Experimental Evaluation of LoRa Communication Performance in Enclosed Environments for HUMS Applications |
| Sikapite, Gloria | Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device |
| Sikora, Ivan | Precedent-Driven Liability in Autonomous Mobility: From Self Driving Cars to Advanced Air Mobility with Unmanned Air Systems |
| Simpson, Oluyomi | Dynamic Load Modelling and Rule-Based Supervisory Controller for a 10 kW Hybrid Solar PV–Wind Off-Grid Microgrid |
| Smith, Daniel J. B. | A Bayesian approach for classifying radio sources fitting photometry and spectroscopy |
| Smith, Joely | Modelling treatment for essential tremor: a biophysical model of ultrasound induced heating in the thalamus |
| Smith-Thorn, Ashton | Comparative Study of the Effect of Ply Orientation and Stacking Sequence on the Tensile Strength of Natural Fibre Composite for Motorsport Applications |
| Soe, Ye Lin | Practical Investigation On The Use Of Synthetic Dataset for Academic research framework and simulation of realistic scenarios while using machine learning. |
| Sorouraddin, Yasser | Inclusive hydrogen economy in Kenya: Institutional Readiness, Policy Sequencing, and Pathways for Equitable Transition |
| Stanley, Alex | Trade-Offs Between Realism and Interactivity in Extended Reality Telepresence for Remote Skill Acquisition |
| Stanley, W. | Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber |
| Steffert, Tony | Preliminary Comparison between Granger causality, delay Symbolic Transfer Entropy and Cross Permutation-entropy in assessing directionality of interaction between physiological signals |
| Steuber, Volker | Multimodal Brain Tumour Classification Using Feature Fusion |
| Steuber, Volker | Video-Based Human Fall Detection Using a Mixture of Experts Framework |
| Steuber, Volker | Cross-Domain Tuberculosis Detection in Chest X-rays |
| Stopford, Chris | Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber |
| Stopford, Chris | Respirable Fibre Detection from Light Scattering Patterns |
| Stopford, Chris | Metagenomic of Indoor Air: A Comparative Study of Sampler Efficacy |
| Subramani, Dharshini | Simulation of Satellite Attitude Control using Deep Reinforcement Learning |
| Summers, James | HYDRA – Hydrogels for Bioaerosol Capture |
| Sun, Yi | Video-Based Human Fall Detection Using a Mixture of Experts Framework |
| Sun, Yi | A Synthetic Data Pipeline for Instruction-Following Compliance in Clinical Conversational AI |
| Sun, Yi | Modelling high intensity focussed ultrasound for breast tumour ablation |
| Swanton, Eamonn | A Synthetic Data Pipeline for Instruction-Following Compliance in Clinical Conversational AI |
| Syed, Mohisin | Temperature-Dependent Loss Analysis of a PMSM for Electric Vehicle Traction Under WLTP Drive Cycle Using MATLAB/Simulink |
| Taheri, Mina | Optimising Green Building Design Using BREEAM and Artificial Intelligence Modelling (BREE-AIM) |
| Tamanna, Nusrat | Simulation of Recrystallization Using MICRESS: Effect of Shielding Time in a Phase-Field Framework |
| Tatarov, Boyan | On the present and future changes in clear air turbulence in a changing climate |
| Tiwari, P.R. | On the present and future changes in clear air turbulence in a changing climate |

https://doi.org/10.18745/pb.27307

| | |
|---|---|
| Toh, Say Meng | Video-Based Human Fall Detection Using a Mixture of Experts Framework |
| Traill, Angela | Simulation-Based Assessment in the Era of Generative AI: Enhancing Decision-Making and Academic Integrity |
| Tveretina, Olga | Visual Classification of Mechanical Linkage Trajectories Using Vector Field Overlays |
| Ummaleti, Bala Vijay Sai | Human Activity Recognition: Classifying Physical Actions Using Smartphone Inertial Sensors |
| Varagnolo, Silvia | Advanced functionalisation of Borophene/ GCN composites as photocatalyst for textile wastewater treatment application |
| Vargas, Alix | Analysis for the Implementation of a Manufacturing Enterprise Resource Planning System into a Formula Student Team |
| Vargas, Alix | Simulation-Based Assessment in the Era of Generative AI: Enhancing Decision-Making and Academic Integrity |
| Vasdev, Nikhil | Mapping Metadata from the Isolation of Extracellular Vesicles from Urine Samples Using PCA and Pressure Readings from a Mesofluidic Tangential Flow Filtration Device |
| Vaz, Neil | Simulation-Based Assessment in the Era of Generative AI: Enhancing Decision-Making and Academic Integrity |
| Velmurugan, Vignesh | Improving Clinical Monitoring Sensitivity in Stroke Rehabilitation via Hardware-Based Signal Conditioning of a Robotic Orthosis |
| Vendra, Chandra | Temperature-Dependent Loss Analysis of a PMSM for Electric Vehicle Traction Under WLTP Drive Cycle Using MATLAB/Simulink |
| Volpi, Nicola Catencacci | ROBOTIS OP3 JOYSTICK LOCOMOTION: MUJOCO PLAYGROUND TOWARDS SIM-TO-REAL |
| Vuyst, Tom De | Modelling Delamination using J-Integral in Smoothed Particle Hydrodynamics |
| Vuyst, Tom de | Simulation Study to Leverage Bio-Inspired Enhancement of Uniaxial Mechanical Properties of Polymeric Structures |
| Wang, Zhouyu | Design and Implementation of a Low-Cost Real-Time Gesture Controlled Smart Wheelchair Using ESP32 and IMU Sensors |
| Weeasiri, Lanka | Evaluation of Aerosol Sampler Performance for DNA-Based Detection of Aerosolised Biological Particles in a Controlled Chamber |
| Weerasiri, Lanka | Low-Flowrate Collection and Quantification of Airborne Human DNA in an Occupied Meeting Room |
| West, Jonathan S. | Performance Evaluation of Bioaerosol Samplers on Capturing Fungal Spores in a Controlled Environment |
| Wood, Luke | Improving Clinical Monitoring Sensitivity in Stroke Rehabilitation via Hardware-Based Signal Conditioning of a Robotic Orthosis |
| Wu, Hongwei | Marginal Cost of Flexibility in CCS-Retrofitted UK Gas Plants: Comparing Plant-Integrated Thermal Energy Storage and Grid-Scale Battery Storage |
| Yaqoob, Muhammad | Early Comparative Evaluation of Transformer Models for Multilingual Software Vulnerability Detection |
| Yaqoob, Muhammad | Cross-Domain Tuberculosis Detection in Chest X-rays |
| Yaqoob, Muhammad | Multimodal Brain Tumour Classification Using Feature Fusion |
| Yates, Rob | From stardust to us: Predicting planet habitability in the Universe from computational simulations of galaxy chemical evolution |
| Yousif, Nada | Modelling treatment for essential tremor: a biophysical model of ultrasound induced heating in the thalamus |
| Yousif, Nada | Modelling high intensity focussed ultrasound for breast tumour ablation |
| Zocco, Alessandro | Trade-Offs Between Realism and Interactivity in Extended Reality Telepresence for Remote Skill Acquisition |
| Zocco, Alessandro | Task-Dependent Benefits of Immersive VR in Pilot Training: Search, Procedure, and Teleoperation |
| Zocco, Alessandro | Presence as a Predictor of Search Efficiency and Procedural Recall in Immersive Cockpit Training |

https://doi.org/10.18745/pb.27307

University of Hertfordshire UH

# SPECS Conference 2026

## Where Technical Excellence Enables Research Culture

**Citation:** Ayub, A., Datsiou, K. C., Pournaghshband, A., & Yousif, N. (2026). Proceedings of Abstracts: School of Physics, Engineering and Computer Science Conference 2026. *University of Hertfordshire. https://doi.org/10.18745/PB.27307*