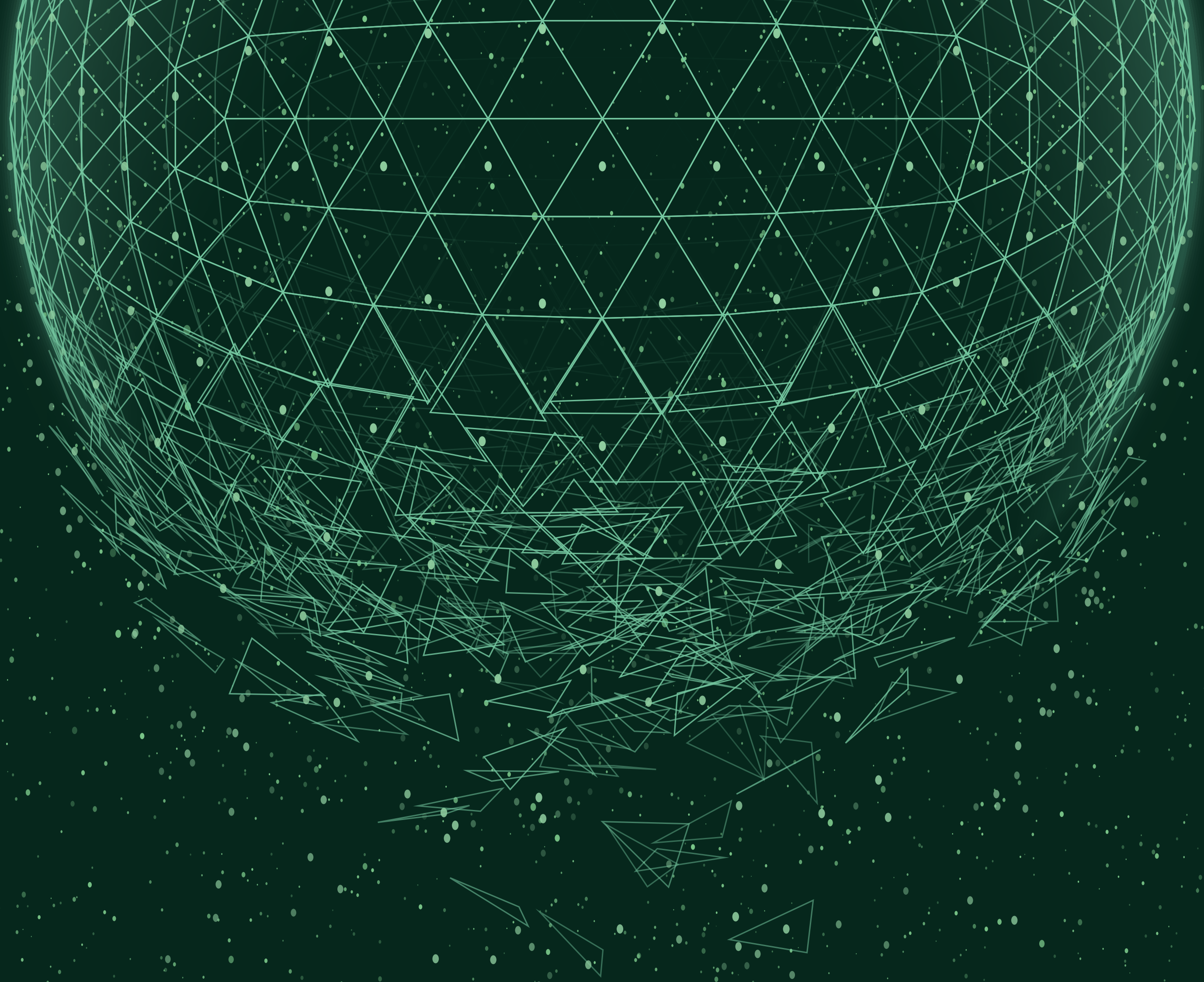


KFUPM

Journal of Undergraduate Research International

Volume 1, Issue 2, December 2025
ISSN 1532-458X

Journal of Undergraduate Research International

Volume 1, Issue 2, December 2025

Research Articles

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article *Architecture Engineering and Design/Environmental Science and Engineering*

Integrating AI into Urban Decision-Making: A Case Study of KFUPM Campus

Renada Abd Alkader and Riyad Ashmeel*

Cite https://doi.org/10.64589/juri/214394

Submitted: May 19, 2025 **Revised:** October 28, 2025 **Accepted:** November 18, 2025

ABSTRACT

King Fahd University of Petroleum and Minerals (KFUPM) currently relies on administrative judgment for urban planning, which may introduce human error and overlook environmental and social factors. This study examines the potential of Artificial Intelligence (AI) to enhance urban planning decisions by using Autodesk Forma to analyze solar exposure, wind behavior, and microclimate variations across the proposed building sites. This study utilized an AI-driven site evaluation framework to conduct multifaceted environmental analyses, including highly predictive wind movement simulations. This approach generated quantifiable metrics essential for sustainable urban design optimization, demonstrating their practical value. For example, one design option achieved 0% uncomfortable pedestrian wind conditions, which is a key measure of success unattainable through conventional evaluation methods. The findings highlight the role of AI in improving planning accuracy, reducing reliance on static datasets, and supporting data-driven decision-making. However, human expertise remains essential for interpreting AI-generated outputs, ensuring alignment with institutional goals, and validating the results. Future research should explore the integration of custom datasets for enhanced wind analysis, AI-driven site selection automation, and expert-driven refinement into planning strategies, further advancing a hybrid AI-human decision-making framework.

Keywords: artificial intelligence, urban planning, autodesk forma, campus development, sustainable design, decision-making framework

1. INTRODUCTION

Urban planning is essential for sustainable development and shaping infrastructure, environmental quality, and resource allocation. At King Fahd University of Petroleum and Minerals (KFUPM), planning decisions have traditionally relied on expert judgment and often lack predictive modeling and real-time environmental analysis. Artificial intelligence (AI) has emerged as a transformative tool that enhances efficiency through data-driven forecasting, land use optimization, and scenario simulations. Prior work in this domain has successfully applied AI and machine learning to architectural design and the initial stages of ecosensitive site assessment[1]. Despite the potential of AI, gaps remain in its application in site-specific environmental assessments and campus expansion strategies.

To address this, this study investigates AI-driven urban planning methodologies at KFUPM, focusing on site selection for new architectural buildings. This study determines how effectively AI tools (e.g., Autodesk Forma) assess microclimatic factors to identify optimal construction sites and what quantifiable differences exist in sustainable performance (e.g., wind comfort) between the proposed sites. Predictive modeling was leveraged to develop a robust and sustainable development strategy for KFUPM expansion. Specifically, the aims were to (i) evaluate the efficacy of AI-driven environmental analysis tools for site selection; (ii) quantify the differences in key sustainability metrics between the site options; (iii) provide a data-driven framework to integrate predictive simulation results into the planning process; and (iv) determine the role of human oversight when translating AI-generated metrics into decisions.

2. LITERATURE REVIEW

KFUPM is at a pivotal point in urban planning, where traditional expert-driven approaches lack predictive modeling, real-time environmental assessment, and scenario simulation. With Saudi Arabia's Vision 2030 prioritizing smart city development[2], integrating AI-driven tools presents an opportunity for sustainable data-informed campus growth. This review examines AI-assisted urban planning, focusing on its role in decision making, site analysis, and spatial strategies at KFUPM.

2.1. AI-Assisted Site Analysis with Neural Networks.

AI enhances urban planning by streamlining site analysis, with tools such as Autodesk Forma providing automated evaluations of solar exposure, wind behavior, and terrain conditions to support data-driven decision-making[1]. Although AI improves efficiency, it lacks autonomous functionality and requires expert oversight. Neural networks structure data inputs for interpreta-

Table 1. *Comparative evaluation of AI applications for environmental site analysis*

Existing AI application	Application description	Strength	Weaknesses
Urban FootPrint	A geospatial intelligent tool offering climate, community, and built environment data insights across the U.S.	Supports in decision making	Specified to only USA & above the research budget
TensorFlow	Machine learning framework for building various AI models using Python. It enables developers to create dataflow graphs with computational nodes representing mathematical operations.	widely adopted in the AI community	Requires prior coding knowledge
CityEngine	an advanced 3D modeling software that creates large, realistic interactive urban environments quickly using GIS. AI in CityEngine enhances procedural modeling, boosts efficiency, analyzes urban data, automates tasks, and provides real-time feedback for better design.	Real-time feedback and automation capabilities	No efficient data about the cost of the subscription after the end of the free trial & requires GIS data
UrbanSim	A virtual training application for mission command in counterinsurgency and stability operations.	Game-based learning environment	requires huge amount of CSV data
ENVI-met	A microscale 3D software model for simulating urban environments, incorporating principles of fluid mechanics and thermodynamics.	detailed simulation to environmental interaction, great for sustainable urban planning	requires a substantial amount of data & interface is not very user-friendly.
UrbanForm	AI technology delivering quick zoning information through maps and analysis tools.	efficient in zoning discovery	has zoning regulations for the U.S. only
Big Query	A fully managed data platform for AI-ready data management and analysis, supporting languages like SQL and Python.	Supports machine learning	Requires prior coding knowledge
Autodesk Forma	cloud-based AI software for architects, urban planners, and developers to create 3D massing models, test site proposals, and analyze factors like noise, wind, sun, and microclimate, streamlining early-stage design with intuitive, generative tools.	User-friendly and fits the research aim and objectives	None

tion, while tools such as ChatGPT assist with cultural context assessments, and MidJourney enhances spatial visualization. AI applications in transportation modeling, infrastructure forecasting, and environmental analysis accelerate timeline planning and optimize land use[3]. However, concerns about algorithmic bias, explainability, and ethical transparency remain, emphasizing the need for hybrid AI-human frameworks[4]. AI also supports sustainable urban design and evaluates environmental impacts such as waste, energy use, carbon emissions, and water needs, although human oversight remains essential for cultural, ecological, and accessibility considerations[5].

2.2. AI Applications for Environmental and Site Analysis.

2.2.1. Evaluation of AI tools for site analysis. A comparative evaluation was conducted to identify the most suitable AI platform for environmental analyses on campuses. The criteria included

1. Functionality and Scope: Types of analyses supported
2. Data Accuracy & Output Format: Visual vs. numerical data
3. Ease of Use: Accessibility for non-specialist users
4. Built-in Data Availability: The platform should offer ready-to-use datasets (e.g., terrain, climate, solar exposure) without requiring custom data uploads.
5. Platform Accessibility: Software/hardware requirements
6. Source Reliability: Documentation and developer support

Table 1 shows a comparative evaluation of AI applications. Based on this assessment, Autodesk Forma was selected as the most balanced platform because of its predictive features, real-time environmental modeling, and user-friendly interface. Its built-in datasets eliminate manual data uploads, making it well suited for AI-driven site analysis at KFUPM.

2.2.2. Autodesk Forma in sustainable site design. Autodesk Forma is categorized as an Analytical AI, offering machine-learning-powered tools for terrain analysis, solar exposure modeling, and wind behavior prediction[1]. These capabilities are particularly useful for sustainable infrastructure planning and layout optimization. However, Forma relies heavily on visual (rather than numerical) outputs and requires planners to validate insights using supplementary tools[6].

3. METHODOLOGY

To evaluate potential locations for the new Architecture and City Design Department building, Autodesk Forma was employed to conduct environmental analyses of sun exposure, wind conditions, and microclimate. The resulting data were processed and visualized using Microsoft Excel to support comparative site selection. Figure 1 illustrates a methodology flowchart outlining the sequential steps undertaken in this study.

3.1. Site Selection and Building Parameters. Four locations within the KFUPM campus were analyzed: the current

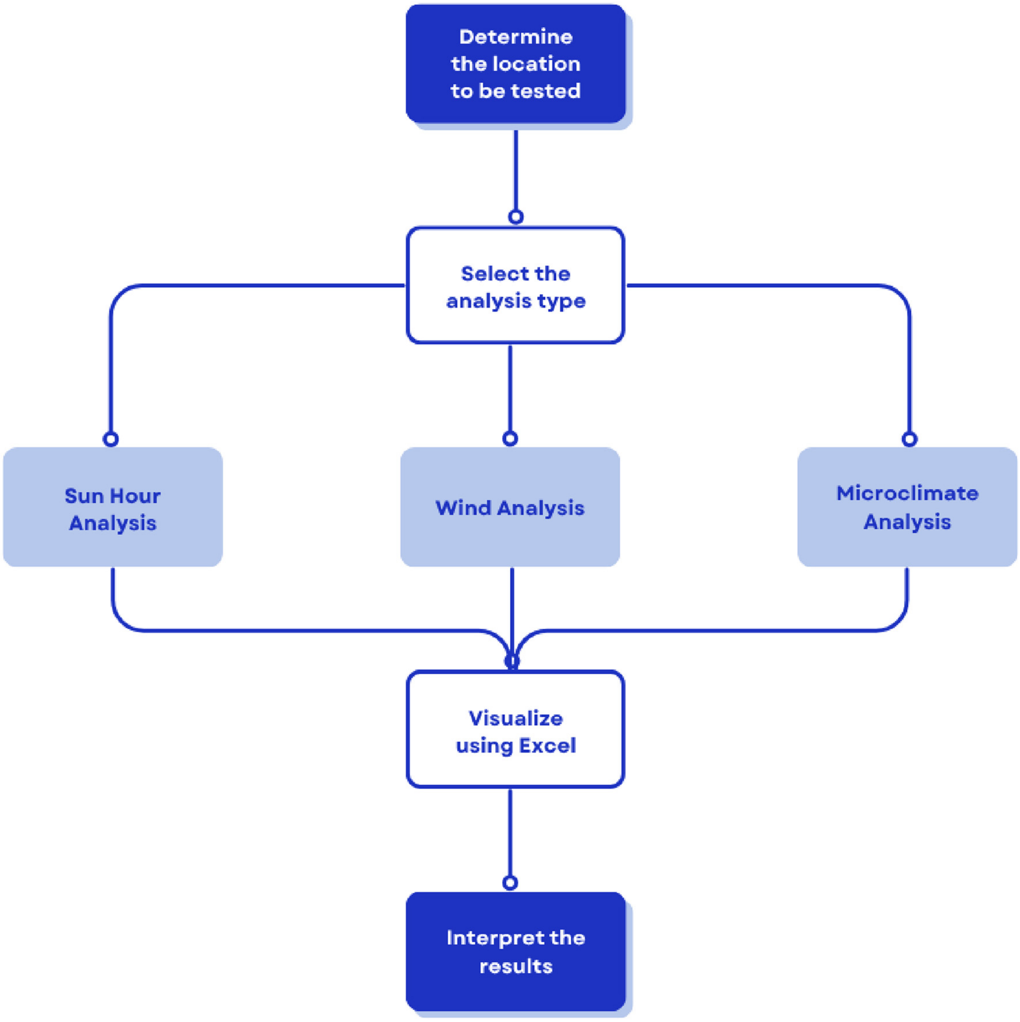


Figure 1. Methodology flowchart

department site, parking lot 19 (proposed for the new building), and parking lot 18 and 11 as alternative sites owing to their proximity. Each site was assessed within a 150-m radius to ensure consistent environmental conditions.

The proposed building design features three aboveground floors (15 m) and an underground parking level (2.5 m), totaling a structural height of 17.5 m. Owing to the absence of specific width and length data, the Bioengineering Department's rectangular footprint was used as a reference to ensure comparability across site analyses.

The AI capabilities of Autodesk Forma facilitate real-time predictive analyses of wind, noise, and operational energy, thereby enhancing the efficiency and accuracy of environmental assessments. By automating complex calculations and providing instant feedback, these tools improve data-driven decision-making for evaluating site suitability[7]. Figure 2 presents the 3D model representations of each site, illustrating the spatial characteristics of Autodesk Forma.

3.2. Analytical Tools and Procedures.

3.2.1. Sun hours analysis. The total annual sun exposure on building rooftops was assessed. Analyses were conducted on the first day of each month in 2023 to capture seasonal variations. This method aligns with the standard practices for evaluating solar exposure over time.

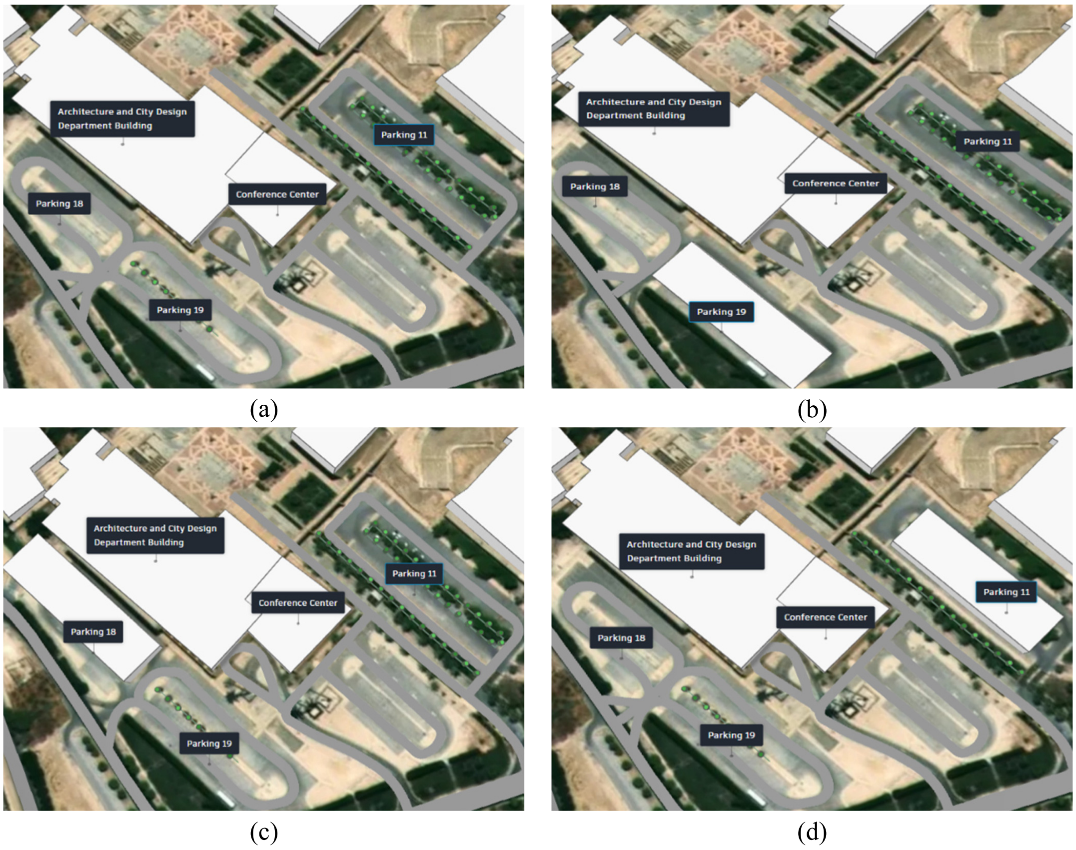
3.2.2. Wind analysis.

1. Pedestrian Wind Comfort: Comfort levels were evaluated using the default Lawson LDDC comfort scale, which categorizes conditions as uncomfortable, strolling/walking, or sitting/standing.
2. Wind Speed Classification: Wind conditions were classified into low, intermediate, and high wind categories.

Wind data were sourced from the Global Wind Atlas 3.0, using a roughness length of 0.1 and a measurement height of 1.75 m above ground level. Owing to software constraints, the analyses were conducted in real time between 6:00 PM and 7:00 PM on the 9th of May 2025.

3.2.3. Microclimate analysis.

1. Temperature exceedance: The rate at which temperatures exceeded 40°C during the hottest months (June, July, and August) at 3:00, 4:00, and 5:00 PM.
2. Air Temperature: Recorded ambient air temperatures during the same periods.



(a) (b)

(c) (d)

Figure 2. Autodesk Forma 3D models of analyzed sites: (a) model of Current Campus; (b) model of parking 19 (proposed site); (c) model of parking 18; and (d) model of parking 11

Analyses for colder months were not feasible because of software limitations restricting data collection to 8:00 AM–8:00 PM, thereby excluding pre-sunrise periods, when minimum temperatures typically occur.

3.3. Exclusions and Limitations.

Several analyses available in Autodesk Forma were excluded from this study:

1. Noise Analysis: All sites were presumed to have minimal noise variation.
2. Daylight potential: All buildings were expected to receive adequate sunlight, rendering the analysis redundant.
3. Solar Energy Analysis: All sites yielded identical results of 2171 kWh/m^2, offering no differentiation.
4. Embodied Carbon Analysis: Technical issues encountered during execution.
5. Shadow Study: Primarily illustrates shadow movement throughout the day without providing additional insights.

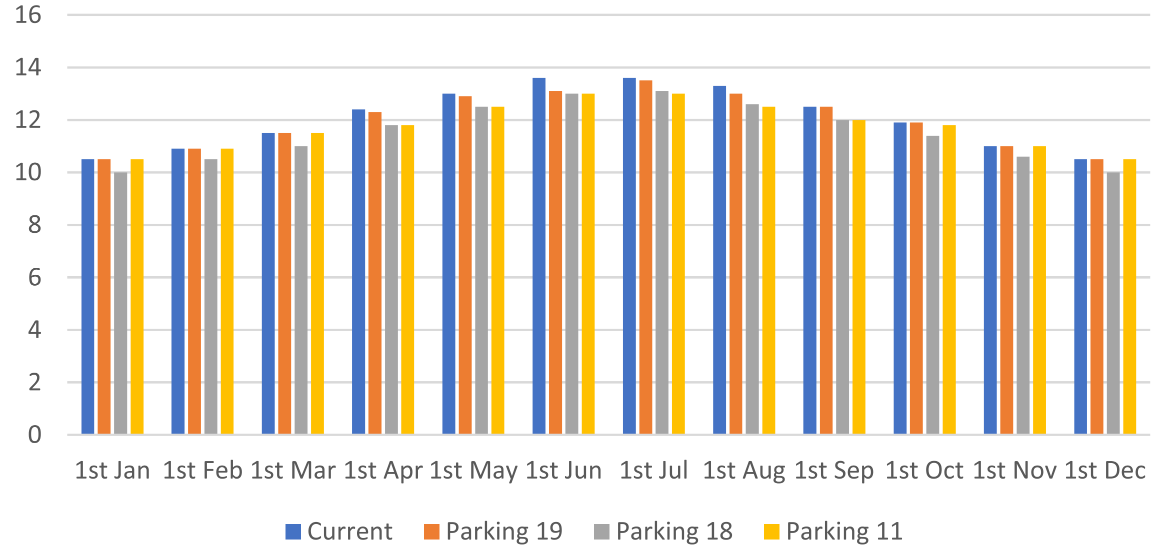


Figure 3. Annual sun exposure comparison across proposed sites

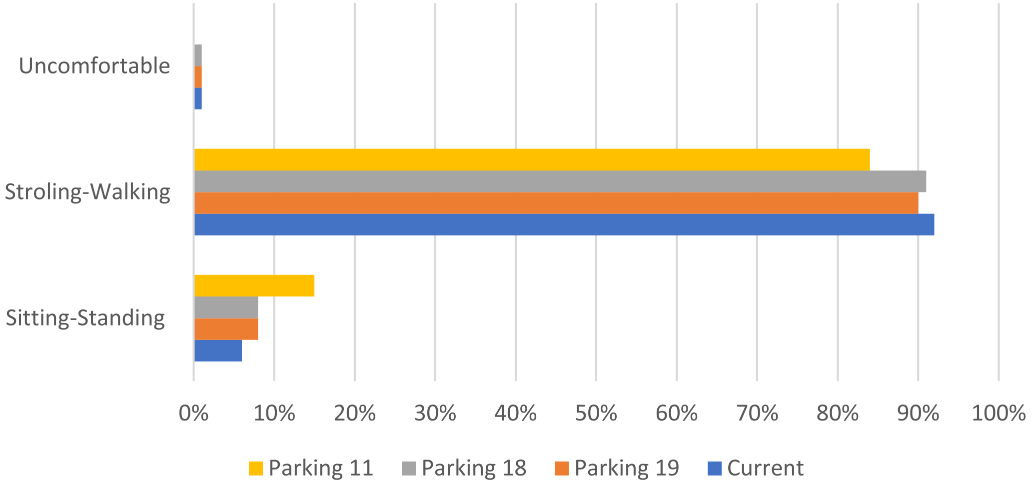


Figure 4. Pedestrian wind comfort levels at each site

4. RESULTS

4.1. Sun Hours Analysis. Figure 3 shows the annual sun exposure data for each site. Parking area 18 consistently exhibited the lowest or equal sun exposure compared with Parking area 11, suggesting that it benefits from the shading provided by the existing Architecture and City Design Department building. This reduced exposure contributes to improved thermal comfort and energy efficiency.

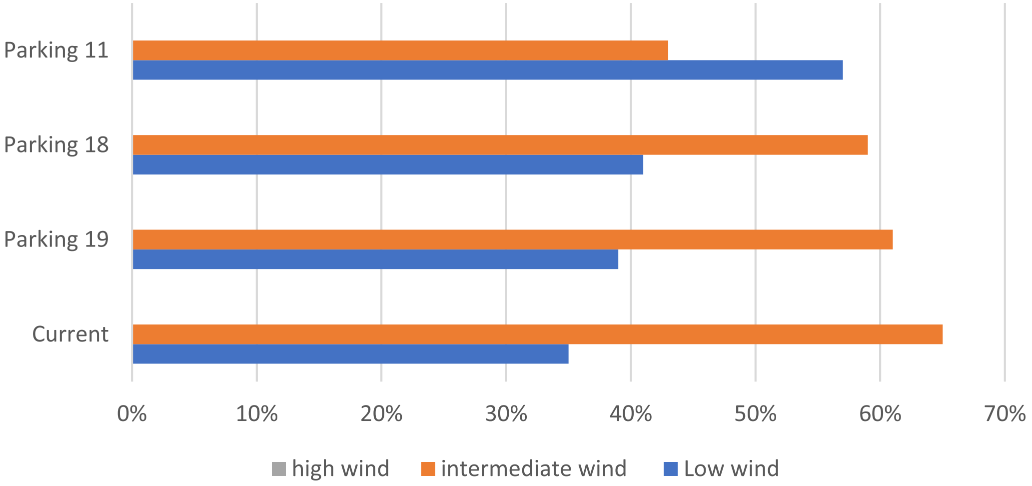


Figure 5. Wind speed classification distribution by location

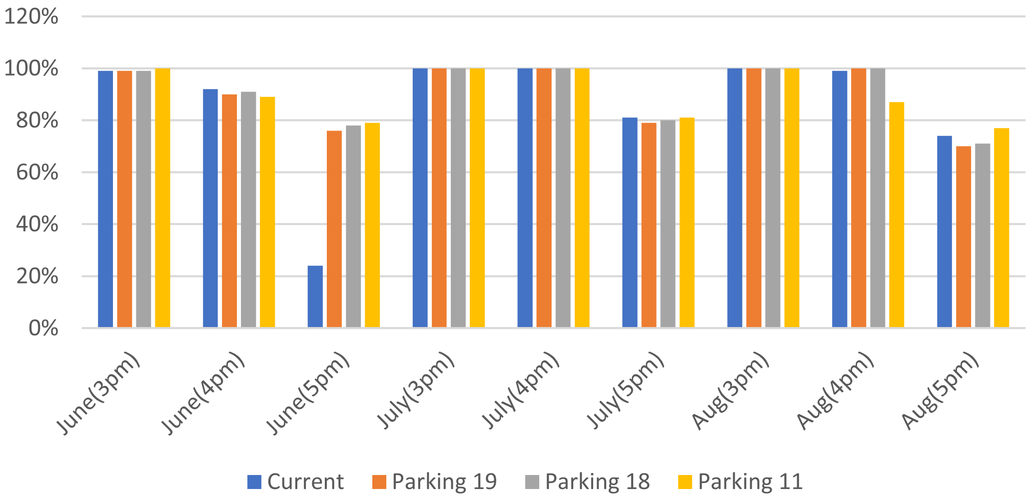


Figure 6. Wind temperature exceedance rates above 40°C in extremely hot weather

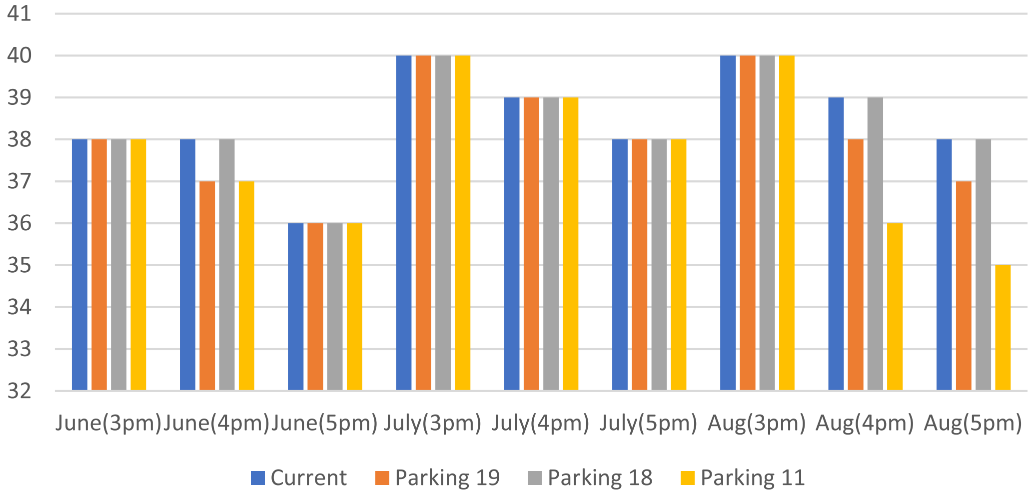


Figure 7. Average air temperature in extreme hot weather across sites

4.2. Wind Analysis.

4.2.1. Pedestrian wind comfort. As shown in Figure 4, parking lot 11 experienced the lowest percentage (0%) of uncomfortable wind conditions. The current campus and parking lot 18 demonstrated the highest percentages of strolling/walking comfort. However, an anomaly was observed across all the location data except parking lot 18, where the total comfort percentages did not sum to 100%. This discrepancy suggests potential data inconsistencies.

4.2.2. Wind speed classification. Figure 5 illustrates that all sites recorded 0% for high wind conditions, which is favorable. The current location had the highest percentage of intermediate wind speeds, followed by parking lot 19. Unlike the comfort analysis, the wind speed classification data for all sites totaled 100%, suggesting more reliable data for this aspect.

Overall, the predominant wind direction was from the north with a consistent extreme wind speed of 4.5 m/s across all sites.

4.3. Microclimate Analysis.

4.3.1. Temperature exceedance. Figure 6 displays the percentage of time that temperatures exceeded 40°C during June, July, and August from 3 PM to 5 PM. In June, the current campus location showed an outlier with a 92% exceedance at 4:00 PM, dropping to 24% at 5 PM, suggesting possible data inconsistencies in Autodesk Forma's real-time temperature modeling. Parking lot 19 had the next lowest exceedance in June and had

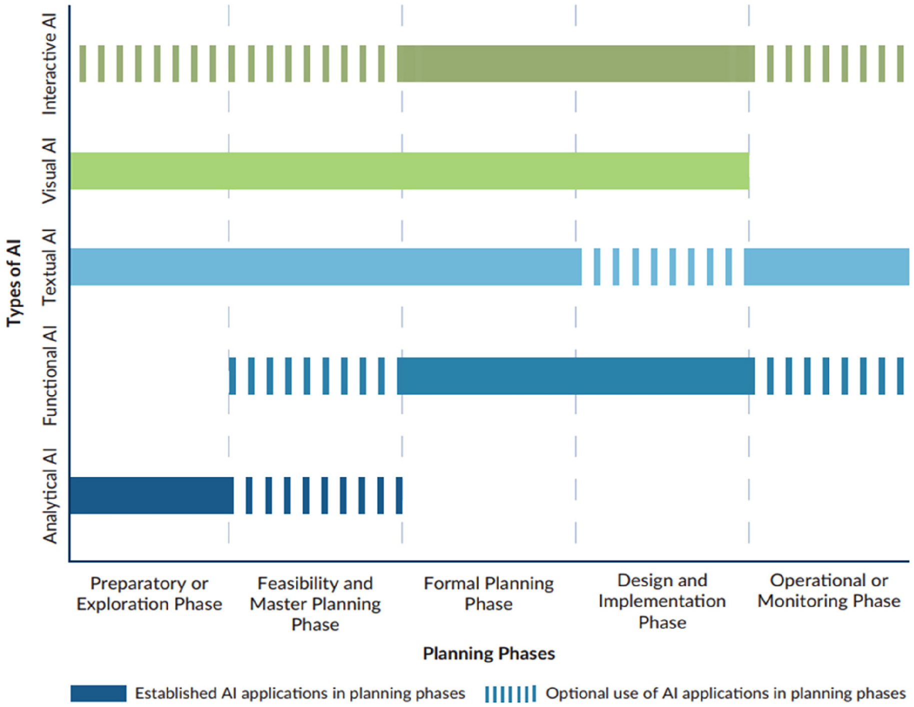


Figure 8. Use of AI tools in different planning phases. A proposal for the use of AI tools related to urban planning in different planning phases. Source: Own illustration based on Diller et al. (2017) and Urban Learning (2024). From Othengrafen et al.[4]

the lowest exceedance in July. In August, parking lot 11 had the lowest exceedance.

4.3.2. Air temperature.

As depicted in Figure 7, the June temperatures were slightly lower at parking lots 19 and 11 compared to the other sites. In July, all sites exhibited similar temperatures. In August, parking lot 11 recorded the lowest air temperature, followed by parking lot 19.

5. DISCUSSION

This study strongly confirms the utility of AI-driven tools, such as Autodesk Forma, in providing quantifiable metrics for optimal site selection, thereby significantly moving beyond traditional expert judgment. The successful generation of comparative environmental metrics across the proposed sites validated the predictive capability of the AI model for complex microclimatic analyses. This successful application aligns directly with the framework proposed in Figure 8 by Othengrafen[4], demonstrating the value of Analytical AI in the early assessment phase. This approach allows KFUPM planning cycles to support the resource-conscious goals of Vision 2030. Furthermore, the efficacy is enhanced by the necessity of importing custom datasets and assigning specific weight thresholds, which allows the tailoring of AI outputs to specific institutional priorities.

Despite its demonstrated utility, the investigation revealed critical constraints that reinforced the need for expert consultation. Autodesk Forma's wind analysis, which is limited to real-time data, restricts its ability to capture long-term trends; therefore, future research must incorporate historical meteorological data. Crucially, observed data inconsistencies, such as wind comfort percentages not totaling 100% and AI's inability to independently select the optimal site, emphasize the necessity of human oversight. Expert interpretation is essential for validating AI outputs, interpreting anomalies, and aligning development initiatives with strategic goals. While AI streamlines processes, the final strategic decision must be made by human experts to ensure that the predictions are balanced with practical knowledge. Future studies should focus on validating simulations using onsite sensor data and enhancing data integration for a truly hybrid planning mode.

6. CONCLUSIONS

This study confirms the potential of AI-driven tools, specifically Autodesk Forma, in enhancing sustainable urban planning at KFUPM by providing real-time, quantifiable environmental insights (solar exposure, wind, and microclimate). AI-powered analysis improved planning accuracy and adaptability, particularly through the use of custom datasets and the ability to assign weight thresholds. However, critical constraints were identified. Autodesk Forma's wind analysis was limited to real-time data, hindering long-term trend capture. In addition, AI currently lacks the capability to suggest optimal locations independently. Ultimately, although AI streamlines processes and optimizes resource allocation, expert interpretation is essential for validating outputs and aligning development initiatives with strategic goals. Future work must focus on enhanced data integration, including historical weather data and expert-driven refinement processes, to further improve the AI-assisted campus planning at KFUPM.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Renada Abd Alkader – *Department of Architecture and City Design, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;*
0009-0000-1400-852X
Email: s202370710@kfupm.edu.sa

Corresponding Author

Riyad Ashmeel – *Research Mentor, Assistant Professor, Department of Architecture and City Design, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;*
0009-0007-1583-463X
Email: rashmeel@kfupm.edu.sa

ACKNOWLEDGMENTS

The authors acknowledge support provided by the College of Design and Built Environment, Department of Architecture and City Design, King Fahd University of Petroleum and Minerals (KFUPM). This research was funded by the Undergraduate Research Department of KFUPM. The authors extend their gratitude to the faculty advisors and the JURI editorial team for their valuable guidance and support throughout the project.

REFERENCES

(1) Zatsepina, A., Bardina, G. & Shindina, P. Eco-sensitive site assessment: Integrating neural networks for environmentally conscious pre-project planning. *E3S Web Conf.* **614**, 05003 (2025).

(2) Aldegheishem, A. Urban Growth Management in Riyadh, Saudi Arabia: An Assessment of Technical Policy Instruments and Institutional Practices. *Sustainability* **15**, 10616 (2023).

(3) Sajjad Naseri. AI in Architecture and Urban Design and Planning: Case studies on three AI applications. *GSC Adv. Res. Rev.* **21**, 565–577 (2024).

(4) Othengrafen, F., Sievers, L. & Reinecke, E. From Vision to Reality: The Use of Artificial Intelligence in Different Urban Planning Phases. *UP* **10**, 8576 (2025).

(5) Toyyibah, W. *et al.* Human natural and artificial intelligence collaboration in urban design: a case study of Indonesia's new capital city. *IOP Conf. Ser.: Earth Environ. Sci.* **1394**, 012030 (2024).

(6) Gür, M. & Karadag, I. Machine Learning for Pedestrian-Level Wind Comfort Analysis. *Buildings* **14**, 1845 (2024).

(7) Autodesk AI |Artificial Intelligence. https://www.autodesk.com/ae/solutions/autodesk-ai.

KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Article *Bioengineering/Material Science and Engineering*

Thermal and Mechanical Quality Analysis of Briquettes of Arabica Coffee and Rice Husk Composites as an Alternative Bio-fuel

Indri Winarti, Rizky Safarina, Ryan Nur Iman, Asti Sawitri, Ridwan Ramdani, and Hasniah Aliah*

Cite https://doi.org/10.64589/juri/209735

Submitted: May 25, 2025 **Revised:** August 3, 2025 **Accepted:** November 18, 2025

ABSTRACT

Coal remains a primary energy source for fuel and electricity generation; however, it is a major contributor to global carbon emissions. Therefore, researchers are exploring strategies to reduce coal usage in fuel briquettes. A promising alternative is biomass-based briquettes, which offer similar performance to traditional charcoal while being more sustainable. This study investigated the thermal and mechanical properties of briquettes made from rice husk (RH) and coffee husk (CH), focusing on how the particle sizes of the raw materials affect the heating value. RH and CH are rich in lignin, cellulose, and hemicellulose, which are crucial for briquette structure, whereas tapioca flour, used as a binder, provides amylose and amylopectin to enhance bonding. Carbonization was performed at 400 °C for 60 min. A 25:75 (wt/wt%) CH: RH blend was used and three mesh sizes (60, 80, and 100) were evaluated by thermal and mechanical characterizations. The briquettes were assessed on the basis of their calorific value, moisture content, shatter index, volatile matter, fixed carbon content, density, ash content, and combustion rate. The 149 µm (100 mesh) sample had the highest density (662.81 kg/m^3), moisture content (6.7%), volatile-matter content (7.07%), and burning rate (0.39 gr/min). In contrast, the 250 µm (60 mesh) sample had the highest ash content (5.01%) and the lowest fixed carbon content (82.04%) and shatter index (4.36%). These findings indicate that controlling the particle size of rice husk and coffee husk can improve the performance of biomass composite briquettes, highlighting their potential as a sustainable alternative to coal for domestic use and small-scale energy applications.

Keywords: rice husk, carbonization, briquettes, coffee husk, mesh

1. INTRODUCTION

Coal remains a primary energy source and is commonly utilized for fuel and electricity production[1]. The combustion of fuel coal emits substantial quantities of CO_2 and other pollutants, resulting in local air pollution and contributing to global atmospheric CO_2[2,3]. To reduce these environmental impacts, researchers are exploring ways to minimize coal usage in briquettes by developing biomass-based briquettes offer comparable quality to traditional charcoal briquettes[4–8].

The quality, i.e., strength and durability, of briquettes is determined by the raw materials, material composition, particle size, level of compaction, and binder used during briquette molding[9,10]. Another method for improving briquette quality is to use additives to improve their physical, mechanical, and thermal properties[11]. The addition of biomass can enhance the calorific value per unit volume, improve particle size consistency, and boost overall quality[9].

Coffee husk (CH) and rice husk (RH) are currently underutilized biomass resources with the significant potential. CH is the dry outer shell of coffee cherries discarded during processing, while RH is the hard-protective layer removed from rice grains during milling. In particular, one hectare of coffee plantation can generate approximately 1.8 tons of CH waste. Currently, this byproduct is primarily used as an organic fertilizer without further processing to maximize its potential. CH is advantageous because of its high calorific value, substantial moisture content (75–80%), and low sulfur content. In their raw form, CH typically contains 2.2% moisture, 0.73% ash, 74.20% volatile matter, 25.07% fixed carbon, and is composed of 10.78% crude protein, 33.13% crude fiber, 24.67% lignin, and 20.22% cellulose[9].

RH is the outer shell of rice grains that is separated during the milling process. The West Java region of Indonesia has a harvested rice area of 1,587 hectares, yielding approximately 9.017 million tons of dry milled grain (GKG). After milling, approximately 20–30% of the rice weight (1.8–2.7 million tons) remains as waste RH[12]. The advantages of RH include its good biochemical and physicochemical properties, making it a suitable carbon source for fabricating activated carbon, adsorbents, supercapacitors, and electrodes[13]. RH typically contains 18–29% ash, 6.0–10.0% moisture, and 35.0–42.0% fixed carbon. It also comprises approximately 40% lignocellulose, 40% cellulose, and approximately 5% hemicellulose; the remaining components include mineral carbon in the form of silica and alkali elements[14]. RH is typically used as an organic fertilizer, animal feed, and substrate in a Fenton catalyst applications[15–19]. Hence, there is broad potential for using CH and RH wastes as alternative fuel sources. These materials are rich in cellulose and lignin, both of which are crucial components for briquettes[20].

Several studies have investigated the effects of adjusting certain parameters on briquette quality. Nagarajan and Prakash reported a composite briquette made of sugarcane bagasse, RH, and corncobs, with the highest calorific value (27.07 MJ/kg) obtained for the 1:1:2 mixture[21]. Setters et al. characterized the solid and liquid fractions of CH briquettes and the effect of particle size using the pyrolysis method, but no improvement in the volatile-matter content[22]. Lubwama and Yiga examined the effect of the binder concentration on the quality of CH/RH composite briquettes; they showed that adding cassava flour as a binder results in high heating values[20]. Similarly, Saeed et al. studied the moisture content of RH briquettes and reported the highest calorific value (17.688 MJ/kg) at 14% water content[14]. Although several studies have focused on improving briquette quality by modifying factors such as the raw material ratio, binder type, and moisture content, there is a lack of research on the influence of the particle size on the performance of composite briquettes made from a mixture of multiple biomass types. Although Setters et al. analyzed the effects of particle size of a single biomass in a pyrolysis study[22], no studies have specifically examined the impact of variations in particle size on the properties of composite briquettes.

To address the aforementioned gap in the existing literature, this study investigated the effect of the particle sizes of Arabica CH and RH waste on the physical and thermal properties of composite briquettes. Three types of briquettes were prepared with CH and RH raw materials with different sieve sizes. A proximate analysis based on ASTM standards, physical properties, and burning capability was performed. The characteristics of the briquettes were compared to the standard quality of usable briquettes.

2. RESEARCH METHODOLOGY

2.1. Briquetting.

Briquettes were prepared from CH waste and RH obtained from Pangalengan Bandung and Pebayuran Kab. Bekasi, West Java, Indonesia. Tapioca flour was used as the binder for the briquettes. CH and RH were sieved to obtain 100 mesh (<149 μm), 80 mesh (<177 μm), and 60 mesh (<250 μm) fractions. The composite briquettes contained 25% CH and 75% RH fractions (by weight).

The husk materials for the briquettes were dried using an oven at 100 °C for 60 min under air atmosphere operation. Then, the material was carbonized in a furnace at 400 °C for 60 min, with both heating and cooling rate set at 20°C/min. The carbonized samples were ground and sieved through 60, 80, and 100 mesh sieves. The sifted material was mixed with a binder derived from tapioca flour. The binder was made by mixing the tapioca flour and water. The mixture between the carbonized material and binder were then molded at a pressure of 4.4 tons using manual pressure for five minutes to form a briquette measuring 6 cm × 6,5 cm. The resulting briquette samples were dried in an oven at 100 °C for 60 min. The procedure for fabricating the composite briquettes is shown in Figure 1.

2.2. Characterization Methods.

2.2.1. Proximate testing.

The moisture content was determined according to the ASTM 1959 standard. Density measurements were conducted according to the ASTM D 2395 standard by calculating the mass and volume of the briquettes using the tube method. The ash content was determined according to the ASTM 1959 standard by heating the samples in a furnace at 900 °C for 6 min. The fixed carbon content was defined as the carbon content of the briquettes after excluding volatile matter, ash, and moisture. The shatter index test was performed according to ASTM D 440, which involved measuring the initial mass of the briquettes and then dropping them from a height of 1.8 m to evaluate their durability.

2.2.2. Burning rate.

The burning rate was examined by measuring the change in briquette mass during combustion. The initial briquette mass before burning was recorded as m_0. To maintain the burning process, a portable fan was used to blow air over the briquettes from a distance of approximately 15–20 cm. An observation period of $t = 5$ min was used, after which the remaining mass m_1 was recorded. The burning rate was then calculated using Equation 1:

$$Burning\ Rate = \frac{m_0 - m_1}{t} \tag{1}$$

3. RESULTS AND DISCUSSION

The density of briquette is influenced by the size and distribution of charcoal particles within the briquette and plays a crucial role in determining how long briquettes burn and how fast they combust. Generally, briquettes with a higher density exhibit a higher combustion rate, whereas those with a lower density burn less vigorously[23]. Figure 2 (a) shows the densities of briquettes made with different particle sizes of CH and RH. The particle size significantly influences briquette density, with the highest density of 662.81 kg/m^3 obtained for the smallest particle size (<149 μm) and the lowest density of 570.44 kg/m^3 obtained for the largest particle size (<250 μm). The briquette density increases with decreasing particle size because finer particles enable the formation of a more compact structure during molding, which reduces the number of internal voids and limits oxygen diffusion.

The moisture content of the composite CH:RH briquettes meets the quality standards set for briquettes. According to SNI 01-6235-2000, the acceptable moisture content for briquette production is less than 8%. Lower moisture levels help reducing energy consumption, and moisture plays a crucial role in combustion because a high moisture content decreases the calorific value of the briquette. Figure 2 (b) shows that the moisture content tends to increase as the particle size decreases, because larger particles absorb less water than smaller particles for the same mass of material[24].

The volatile-matter content characterizes the mass lost during combustion. Lower volatile-matter levels are preferable for high-quality briquettes to avoid excessive smoke production[25]. Figure 2 (c) shows similar volatile-matter contents for all briquettes produced using different particle sizes. The content is 6.58% for the 250 μm sample, which decreases to 6.12% for the 180 μm sample, and then increases to 7.07% for the 149 μm

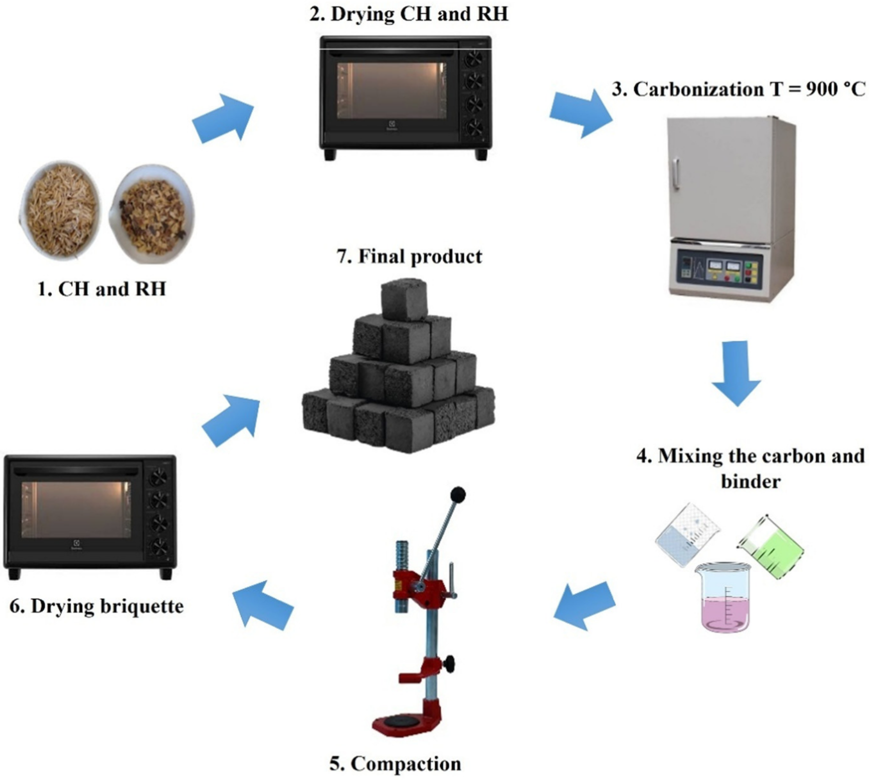


Figure 1. Schematic of the process for fabricating CH:RH composite briquettes

sample. Elevated volatile matter levels are often linked to inefficient carbonization resulting from an insufficient carbonization temperature and/or time. Specifically, higher temperatures and longer carbonization times typically remove more volatile matter, reducing the residual content in the briquettes[25,26]. Smaller particle sizes tend to exhibit enhanced reactivity due to their higher surface area, promoting faster reactions, but they can also intensify secondary reactions and alter volatile components, particularly in composite briquettes[27]. These effects may explain the lower volatile-matter content measured for the 177 µm sample.

Ash in solid fuel refers to the mineral residue that remains after combustion and cannot be burned[28]. Ash primarily consists of minerals containing silicon, calcium, magnesium, and phosphorus. Among these, the silica content has the most significant impact on the ash value and negatively affects the calorific value of briquettes, thereby reducing their quality[29]. Figure 2 (d) shows that the ash content in the briquettes decreases slightly with decreasing particle size. These findings are consistent with those reported by Bhattarai et al.[30]; they showed that larger particles were more difficult to compact and had weaker bonds, enhancing oxygen flow, leading to more complete combustion and a lower ash content. In addition, the ash content depends on the silica content of the raw materials. The addition of tapioca flour binder can also affect the ash content[31] because the binder contains inorganic components that increase the ash content of the briquettes.

Figure 2 (e) shows the fixed carbon contents for the briquette prepared with different particle sizes of CH and RH. The highest fixed carbon percentage (82.57%) was observed at a particle size of 177 µm. Briquettes with lower volatile matter, ash, and moisture contents tend to have a higher fixed carbon content as low levels of the former components indicate efficient carbonization, resulting in a greater proportion of stable carbon remaining in the briquette. A higher fixed carbon content generally enhances the energy density and combustion efficiency of briquettes, making them more effective fuel sources. Thus, controlling these parameters during briquette production is essential for optimizing the fixed carbon content and overall fuel quality[32].

The shatter index is used to assess the mechanical strength and robustness of briquettes by examining their ability to withstand impacts. In this test, the briquettes were dropped from a height of 1.8 m, and the extent of fragmentation or the number of fragments resulting from the drop was measured to gauge their durability. Figure 2 (f) presents the shatter indexes for the briquettes prepared with different particle sizes of CH and RH. The sample prepared with 250 µm particles has the lowest shatter index, with fewer particles lost during the test. These results suggest that the 250 µm briquette has better durability (higher quality) than the other samples.

The burning rate describes the amount of briquette mass consumed over a specific period during combustion. Burning rate tests were conducted using the briquettes prepared with different feedstock particle sizes to assess the efficiency of the briquettes as a fuel source, and the results are shown in Figure 2 (g). The sample with the smallest particle size (149 µm) shows the highest burning rate of 0.39 g/min. This is because increasing the particle size decreases the rate of briquette burning and ash formation. Similarly, when the briquette particle size is small, the combustion rate increases.

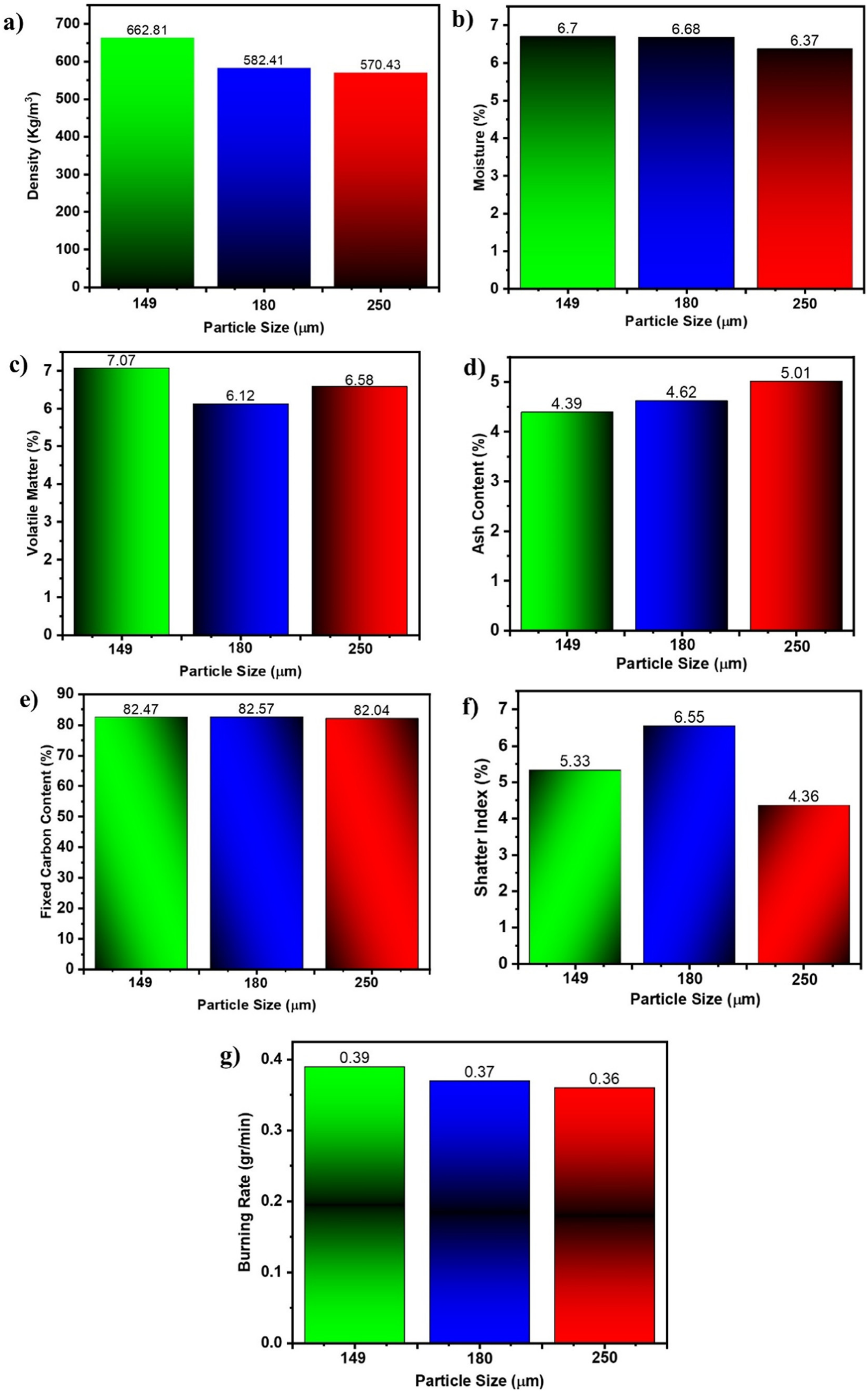


Figure 2. Effect of the particle size on the (a) density, (b) moisture content, (c) volatile-matter content, (d) ash content, (e) fixed carbon content, (f) shatter index, and (g) burning rate

4. CONCLUSIONS

This study demonstrated the potential of utilizing CH and RH biomass to produce sustainable and eco-friendly charcoal briquettes as a viable alternative fuel to coal. A comprehensive evaluation of briquettes prepared with different biomass particle sizes showed that the particle size significantly influences the physical and thermal properties of the briquettes. The briquette prepared with the smallest particles (149 µm) yielded the highest density (662.81 kg/m^3) and burning rate (0.39 g/min), while meeting moisture content standards (<8%). The highest fixed carbon content (82.57%) was observed for the 180 µm briquettes, indicating a good balance between compaction and carbonization. Although smaller particles enhanced the density and burning rate, they slightly increased the moisture and volatile matter content. Conversely, the ash content decreased with smaller particle sizes, indicating higher fuel quality. Overall, the composite briquettes made from 149 µm CH and RH showed the highest density, moisture content, volatile matter content, and burning rate. In contrast, the 250 µm sample had the highest ash content, along with the lowest fixed carbon content and shatter index. Future studies incorporating visual analyses, such as SEM, element analysis, and rheological testing, are recommended to further verify the influence of the particle size on the composition and rheological properties of briquettes.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Authors

Indri Winarti – *Department of Physics, Universitas Islam Negeri Sunan Gunung Djati Bandung, Jalan A.H. Nasution No. 105, 40614, Kota Bandung, Indonesia;* 0009-0009-4284-0394
Email: indriwinarti00@gmail.com

Rizky Safarina – *Department of Physics, Universitas Islam Negeri Sunan Gunung Djati Bandung, Indonesia;* 0009-0002-2054-7655
Email: rizkysafarina81@gmail.com

Authors

Ryan Nur Iman – *Study Program of Electrical Engineering, Universitas Islam Nusantara, Indonesia;* 0000-0001-6654-0737
Email: ryan.iman2650@alum.kfupm.edu.sa
Email: ryannuriman@uninus.ac.id

Asti Sawitri – *Department of Physics, Universitas Halim Sanusi, Indonesia;* 0009-0008-5742-5698
Email: astisawitri@uhs.ac.id;

Ridwan Ramdani – *Department of Physics, Universitas Islam Negeri Sunan Gunung Djati Bandung, Indonesia;* 0009-0001-5882-6900
Email: ridwanr@uinsgd.ac.id

Corresponding Author

Hasniah Aliah – *Research Mentor, Department of Physics, Universitas Islam Negeri Sunan Gunung Djati Bandung, Indonesia;* 0000-0002-4527-0253
Email: hasniahaliah@uinsgd.ac.id

ACKNOWLEDGEMENTS

The authors acknowledge the support provided by the Ministry of Religious Affairs of Indonesia through UIN Sunan Gunung Djati Bandung (No: 117/Un.05/V.2/KP.01.1/01/2022) and the Materials Physics Laboratory for sample testing.

REFERENCES

(1) Li, Q. The view of technological innovation in coal industry under the vision of carbon neutralization. *Int. J. Coal Sci. Technol.* **8**, 1197–1207 (2021).

(2) Kumar, J. A., Kumar, K. V., Petchimuthu, M., Iyahraja, S. & Kumar, D. V. Comparative analysis of briquettes obtained from biomass and charcoal. *Mater. Today Proc.* **45**, 857–861 (2021).

(3) Cheng, F. *et al.* Eliminating environmental impact of coal mining wastes and coal processing by-products by high temperature oxy-fuel CFB combustion for clean power Generation: A review. *Fuel* **373**, 132341 (2024).

(4) Narzary, A., Brahma, J. & Das, A. K. Utilization of waste rice straw for charcoal briquette production using three different binder. *Clean. Energy Syst.* **5**, 100072 (2023).

(5) Espuelas, S., Marcelino, S., Echeverría, A. M., del Castillo, J. M. & Seco, A. Low energy spent coffee grounds briquetting with organic binders for biomass fuel manufacturing. *Fuel* **278**, 118310 (2020).

(6) Lukas, A. G., Lombok, J. Z. & Anom, I. D. K. Briquettes made with mixtures of salak seed (Salacca zalacca) charcoal and coconut shell charcoal and the potential as an alternative Energy Source. *Int. J. Appl. Eng. Res.* **13**, 10588–10592 (2018).

(7) Wirabuana, A. D. & Alwi, R. S. Influence of starch binders composition on properties of biomomass briquettes from Durian peel (Durio kutejensis Becc). *AIP Conf. Proc.* **2349**, (2021).

(8) Wang, C., Zhang, S., Wu, S., Sun, M. & Lyu, J. Multi-purpose production with valorization of wood vinegar and briquette fuels from wood sawdust by hydrothermal process. *Fuel* **282**, 118775 (2020).

(9) Harsono, S. S., Fauzi, M., Koekoeh, R. & Wibowo, K. Biomass Stove with Low Carbon Monoxide Emission Fueled by Solid Fuel Coffee-Husk Biopellet. (2022).

(10) Khan, A. U. *et al.* Evaluating Coal–Jatropha curcas Composite Briquettes as Sustainable Fuel Alternatives. *Int. J. Energy Res.* **2025**, 8847392 (2025).

(11) Setter, C. & Oliveira, T. J. P. Evaluation of the physical-mechanical and energy properties of coffee husk briquettes with kraft lignin during slow pyrolysis. *Renew. Energy* **189**, 1007–1019 (2022).

(12) Badan Pusat Statistik. Luas Panen, Produksi, dan Produktivitas Padi Menurut Provinsi 2018, 2019, 2021. *Luas Panen, Produksi, dan Produktivitas Padi Menurut Provinsi 2018, 2019, 2021* (2021).

(13) Anggraeni, S., Girsang, G. C. S., Nandiyanto, A. B. D. & Bilad, M. R. Effects of particle size and composition of sawdust/carbon from rice husk on the briquette performance. *J. Eng. Sci. Technol.* **16**, 2298–2311 (2021).

(14) Saeed, A. A. H. *et al.* Moisture content impact on properties of briquette produced from rice husk waste. *Sustain.* **13**, (2021).

(15) Vadiveloo, J., Nurfariza, B. & Fadel, J. G. Nutritional improvement of rice husks. *Anim. Feed Sci. Technol.* **151**, 299–305 (2009).

(16) Dere, I. *et al.* Preparation of slow-release fertilizer derived from rice husk silica, hydroxypropyl methylcellulose, polyvinyl alcohol and paper composite coated urea. *Heliyon* **11**, e42036 (2025).

(17) Selvarajh, G. *et al.* Enriched rice husk biochar superior to commercial biochar in ameliorating ammonia loss from urea fertilizer and improving plant uptake. *Heliyon* **10**, e32080 (2024).

(18) de Oliveira Simas, A. L. *et al.* Production of Phytase, protease and xylanase by aspergillus niveus with Rice Husk as a Carbon source and

application of the enzymes in animal feed. *Waste and Biomass Valorization* **15**, 3939–3951 (2024).

(19) Chu, J.-H., Kang, J.-K., Park, S.-J. & Lee, C.-G. Enhanced sonocatalytic degradation of bisphenol A with a magnetically recoverable biochar composite using rice husk and rice bran as substrate. *J. Environ. Chem. Eng.* **9**, 105284 (2021).

(20) Lubwama, M. & Yiga, V. A. Characteristics of briquettes developed from rice and coffee husks for domestic cooking applications in Uganda. *Renew. Energy* **118**, 43–55 (2018).

(21) Nagarajan, J. & Prakash, L. Preparation and characterization of biomass briquettes using sugarcane bagasse, corncob and rice husk. *Mater. Today Proc.* **47**, 4194–4198 (2021).

(22) Setter, C. *et al.* Slow pyrolysis of coffee husk briquettes: Characterization of the solid and liquid fractions. *Fuel* **261**, (2020).

(23) Aliah, H. *et al.* Physical properties of briquettes composite from coffee husks (Coffea arabica L) and corncob (Zea mays) for alternative fuel. AIP Conf. Proc. 2646, 60004 (2023).

(24) Abineno, J. C., Dethan, J. J. S., Bunga, F. J. H. & Bunga, E. Z. H. Characterization and performance analysis of Kesambi branch biomass briquettes: A study on particle size effects. *J. Ecol. Eng.* **26**, (2025).

(25) Inegbedion, F. Estimation of the moisture content, volatile matter, ash content, fixed carbon and calorific values of saw dust briquettes. *MANAS J. Eng.* **10**, 17–20 (2022).

(26) Gündüz, F., Akbulut, Y., Koyunoğlu, C., Önal, Y. & Karaca, H. Obtaining the best temperature parameters for co-carbonization of lignite (yatağan)-biomass (peach seed shell) by structural characterization. *Heliyon* **8**, e10636 (2022).

(27) Zhong, S., Baitalow, F. & Meyer, B. Experimental investigation on fragmentation initiation of mm-sized coal particles in a drop-tube furnace. *Fuel* **234**, 473–481 (2018).

(28) Staničić, I., Brorsson, J., Hellman, A., Rydén, M. & Mattisson, T. Thermodynamic analysis on the fate of ash elements in chemical looping combustion of solid fuels – Manganese-Based oxygen carriers. *Fuel* **369**, 131676 (2024).

(29) Liu, Y. *et al.* Ash chemistry in chemical looping process for biomass valorization: A review. *Chem. Eng. J.* **478**, 147429 (2023).

(30) Bhattarai, P., Sapkota, R. & Ghimire, R. M. Effects of Binder and Charcoal Particle Size on the Physical and Thermal Properties of Beehive Briquettes. *Proc. IOE Grad. Conf.* **1**, 57–63 (2016).

(31) Yuliah, Y., Kartawidjaja, M., Suryaningsih, S. & Ulfi, K. Fabrication and characterization of rice husk and coconut shell charcoal based bio-briquettes as alternative energy source. in *IOP Conference Series: Earth and Environmental Science* **65**, (2017).

(32) Jahiding, M., Ngkoimani, L.O., Erzam, S.H., Ratnawati, W.O., & Maymanah, S. Pengembangan briket hybrid berbasis sekam padi dan batubara muda (brown coal) sebagai bahan bakar alternatif. Jurnal Aplikasi Fisika, 7(1), 12–21 (2011).

KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Article *Chemistry/Material Science and Engineering*

Nitrogen-Doped Carbon Dots from Pomegranate Peels as Sustainable Corrosion Inhibitors for Carbon Steel in Acidic Medium

Ali M. Al Nasser, Hifsa Kurdshid, and Saviour A. Umoren*

Cite https://doi.org/10.64589/juri/214906

Submitted: July 28, 2025 **Revised:** October 29, 2025 **Accepted:** November 30, 2025

ABSTRACT

Carbon steel suffers from rapid corrosion in acidic environments such as those encountered during acid cleaning and petroleum-well acidification, leading to significant economic losses and environmental challenges. Conventional corrosion inhibitors are often toxic, nonbiodegradable, and expensive, motivating the search for sustainable biomass-derived alternatives. In this study, carbon dots (CDs) synthesized from pomegranate peel waste (PPCDs) and their nitrogen-doped analogs (N-PPCDs) were evaluated as ecofriendly corrosion inhibitors for carbon steel in 5 wt.% HCl. The structural, morphological, and chemical properties of the materials were characterized using UV-Vis, Fourier-transform infrared, transmission electron microscopy, and energy-dispersive X-ray analyses. Electrochemical impedance spectroscopy and potentiodynamic polarization measurements demonstrated that both PPCDs and N-PPCDs acted as mixed-type inhibitors, effectively inhibiting both anodic and cathodic corrosion reactions. Nitrogen doping enhanced inhibition efficiency from 70% to 80% at CD concentration of 200 ppm owing to enhanced adsorption and chemisorption through N- and O-donor functional groups, as confirmed by surface analysis. These findings establish the nitrogen-doped pomegranate peel CDs as promising and sustainable alternatives to conventional synthetic inhibitors for corrosion control in acidic media.

Keywords: carbon dots, pomegranate peel, carbon steel, acid corrosion, corrosion inhibition

1. INTRODUCTION

Corrosion has been a significant and long-standing challenge, as it affects a wide spectrum of sectors, such as transportation, construction, and energy sectors, leading to structural failures, operational inefficiencies, and high maintenance costs[1]. Carbon steel (CS), despite its widespread use, is particularly vulnerable to corrosion in acidic environments, for example, when exposed to acidic cleaning agents or harsh environments[2], necessitating the development of effective corrosion inhibitors. While many corrosion inhibitors have been shown to be effective, they are often toxic and expensive[3]. This has intensified research into greener alternatives capable of offering comparable or superior inhibition performance while minimizing ecological impact[4]. Hence, there is a demand for robust and environmentally friendly corrosion inhibitors that can perform in extreme environments. In meeting this demand, green corrosion inhibitors offer several advantages, such as a lack of toxicity and the utilization of inexpensive waste resources[5]. Recently, carbon dots (CDs) synthesized from agricultural waste have attracted significant attention[6]. These CDs are nanosized and may be derived from a variety of low-cost materials, such as carbohydrates and organic solvents[7]. CDs synthesized from renewable biomass and agricultural waste precursors, such as fruits, vegetables, and dairy byproducts, offer advantages over traditional inhibitors, because these precursors are non-toxic, cost-effective, and environmentally friendly[8]. Reports have shown that CDs can effectively suppress corrosion in acidic media by blocking metal's active sites, hindering electron transfer, and establishing barrier films that restrict the ingress of aggressive species[9]. In addition, CDs can serve as electron donors and acceptors, enhancing its corrosion-inhibition properties[10]. Incorporation of nitrogen into CDs, in particular, has been found to enhance their inhibition capability due to improved electron-donor characteristics and stronger metal–inhibitor interactions[4], improving their ability to donate electrons, suppress electrochemical reactions, and provide corrosion protection. Several studies have demonstrated that N-doped CDs can reduce the corrosion rate of CS in hydrochloric acid, with the nitrogenated functional groups strengthening the adsorption process[11]. For instance, carbon-dot-based complexes modified with sulfosalicylic acid served as corrosion inhibitors for CS in an HCl solution, markedly decreasing the corrosion rate[12]. Biomass-derived CDs were employed as corrosion inhibitors for Q235 steel in a 1 M HCl solution, mitigating corrosion by self-aggregation and adsorption onto the metal surface[13]. Additionally, the efficacy of CDs as corrosion inhibitors for CS has been evidenced in various acidic solutions, with the diminished corrosion rate ascribed to the development of a protective coating on the metal surface[14–16]. While CDs derived from various natural and synthetic sources have demonstrated considerable potential as corrosion inhibitors for various metals under diverse environmental conditions[4,8], the use of CDs derived from pomegranate peels for corrosion protection of CS has not been reported to date. Furthermore, nitrogen-doped CDs are an emerging class of functional nanomaterials that have received limited attention in the context of corrosion inhibition. This study addresses this gap by investigating CDs synthesized from pomegranate peel waste (PPCDs) and their nitrogen-doped analogs (N-PPCDs) obtained via hydrothermal carbonization of pomegranate peel waste using glycine for N doping. The inhibition of CS in HCl was evaluated through electrochemical and surface analyses, highlighting their potential as sustainable high-performance corrosion inhibitors.

2. EXPERIMENTAL

2.1. Materials and Materials Preparation. Pomegranates were purchased from a local market and glycine was obtained from Sigma-Aldrich. X60 CS was used as the metal substrate, and its chemical composition matched that previously documented in the literature[17]. The corrosive medium used in the research work was 5% solution of hydrochloric acid (HCl). It was prepared by diluting analytical-grade concentrated HCl (Merck) with distilled water. The CS samples were polished by wet mechanical abrasion using sandpapers of different grades (#120, #320, #600, #800, and #1000). Prior to each experimental run, the polished samples were thoroughly rinsed with distilled water, followed by ethanol, and subsequently dried using warm air to ensure a clean, uniform surface.

2.2. Synthesis of Undoped and N-doped PPCDs. The pomegranate peel waste was sun-dried, and the dried biomass was ground into a fine powder that served as the carbon source. Glycine was used as the nitrogen source. The pomegranate peel (1.5 g) and glycine (0.5 g) precursors were mixed with distilled water (100 mL) and placed in a Teflon-lined stainless-steel autoclave. To facilitate the hydrothermal process, the autoclave was kept in a drying oven at 200 °C for a full day. Once the reaction finished, the autoclave was cooled to room temperature, and the mixture was filtered using Whatman #1 filter paper. To purify the CDs, the supernatant was filtered. The dialyzed filtrate was then dried in an oven (70 °)C to produce the solid-state product (N-PPCDs). A solid product obtained from the powdered pomegranate peel waste material without blending with glycine is referred to as the undoped pomegranate peel CDs (PPCDs).

2.3. Characterization of Synthesized PPCDs. The optical properties of PPCDs and N-PPCDs were examined using a JASCO 770 UV–vis spectrophotometer (Tokyo, Japan) operated at a scanning speed of 200 nm min^{-1} and a spectral resolution of 1 nm. Distilled water served as the reference for baseline correction. Fourier-transform infrared (FTIR) spectra were collected with an ATR-FTIR system (Nicolet iS5, Thermo Scientific, Waltham, MA, USA) over the wavenumber range of 4000–400 cm^{-1}. Morphological observations were performed using a JEOL JSM-6610LV scanning electron microscope (JEOL, Japan). For SEM analysis, a drop of the ethanolic dispersion of PPCDs or N-PPCDs was deposited onto an aluminum stub, allowed to dry at ambient conditions to form a thin film, and subsequently coated with a thin layer of gold. Imaging was conducted at an accelerating voltage of 20 kV. Elemental composition was evaluated using an energy-dispersive X-ray spectroscopy (EDAX) detector integrated with the SEM system. Transmission electron microscopy (TEM) was used to determine particle size and morphology. Samples were prepared by placing a small amount of the colloidal PPCD or N-PPCD suspension onto a carbon-coated copper grid, followed by drying at room temperature. TEM images were obtained using a JEOL JEM-2100F microscope (Tokyo, Japan) operated at 200 kV[15].

2.4. Corrosion Inhibition Evaluation. The inhibitory effectiveness of PPCDs and N-PPCDs on CS in 5 wt.% HCl was evaluated using weight loss and electrochemical measurements. The equations (1 and 2) and a detail description of the experimental procedures used for weight loss evaluation are provided in Section S5 and Table S2 of the Supplementary Materials. A Gamry 1010E system was employed to conduct the electrochemical tests in accordance with the protocol detailed in Section S6. Linear polarization resistance (LPR) quantities were estimated by scanning ±10 mV around the corrosion potential (E_{corr}) at a rate of 0.167 mV/s, whereas potentiodynamic polarization (PDP) measurements were conducted over a range of ±250 mV using a scan rate of 0.5 mV/s. Data analysis was conducted by Echem Analyst v6.2 software. The inhibition efficiencies (IE) were calculated using Equations (3) and (4) in Table S2, corresponding to the PDP, EIS, and LPR methods.

2.5. Surface Analysis. The surface morphology of the carbon steel samples was examined using a scanning electron microscope (SEM; JEOL JSM-6610LV, Japan) equipped with an energy-dispersive X-ray spectroscopy (EDAX) system. This combination enabled both imaging and elemental analysis of the sample surfaces after exposure to the corrosive environment, with and without the presence of inhibitors. The method for the preparation of the sample for SEM analysis is described in Section S7 of the Supplementary Materials. In addition, a 3D optical profiler (Profilm 3D, Filmetrics, USA) was used to evaluate alterations in the surface topography of the steel specimens. This analysis was performed on samples exposed to 5% HCl under inhibited and uninhibited conditions to assess changes in roughness and damage features. AFM measurements were conducted using an MFP-3D Infinity Oxford Instruments (UK), operated under ambient conditions, to provide nanoscale insight into surface modifications resulting from corrosion and inhibitor adsorption.

3. RESULTS AND DISCUSSION

3.1. Characterization of PPCDs and N-PPCDs. Prior to the examination of the corrosion inhibition by the CDs, characterization techniques were used to verify the successful synthesis of PPCDs and N-PPCDs. TEM analysis (Figure 1a) revealed that PPCDs exhibit an amorphous spherical morphology, with particle sizes ranging from 5 to 10 nm. The UV–vis spectrum (Figure 1b) of PPCDs exhibited two absorption peaks at approximately 256 and 367 nm, corresponding to the $\pi-\pi^*$ transitions in conjugated systems and the n–π^* transitions associated with the surface C=O and C–O functionalities[2,18]. The inset shows that the

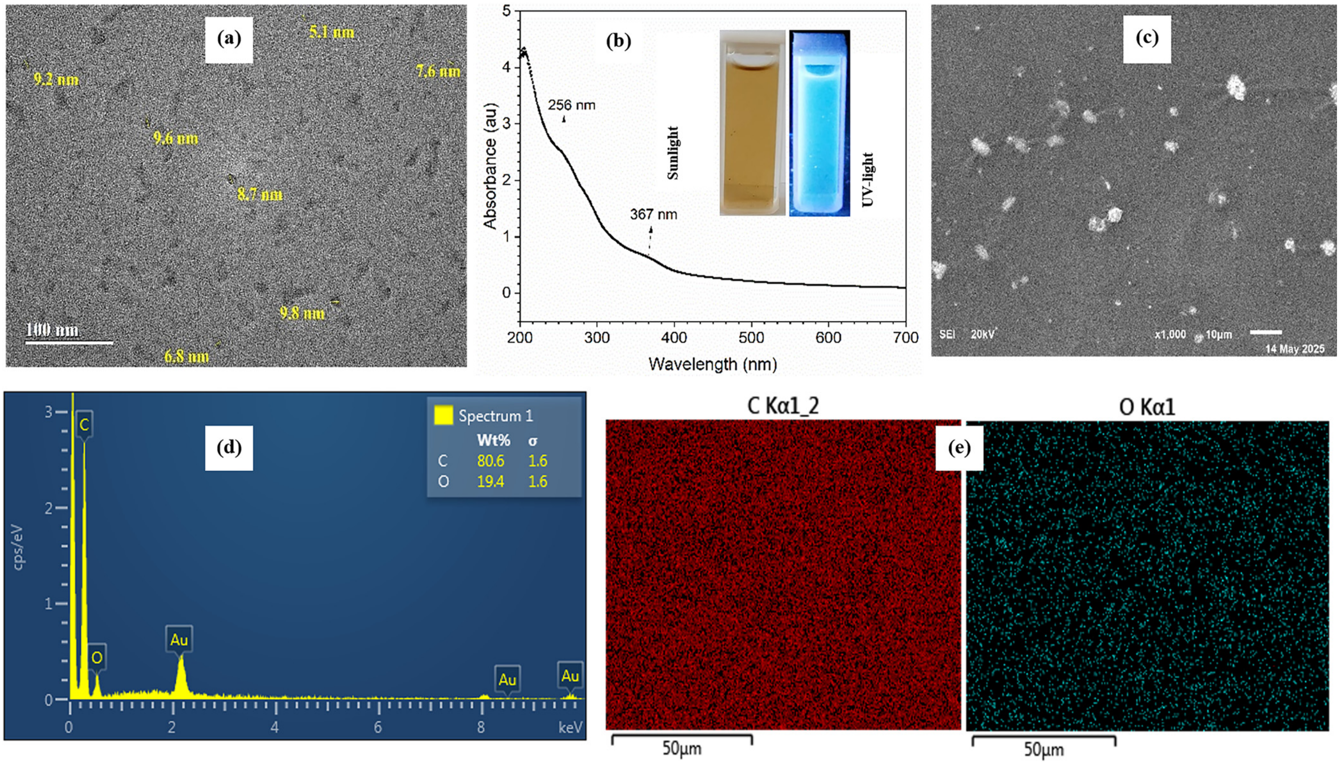


Figure 1. Results of (a) TEM, (b) UV-vis, (c) SEM, (d) EDAX and (e) elemental mapping characterizations of undoped PPCDs

PPCD solution appeared orange under visible light and emitted deep-blue fluorescence under UV light, confirming its fluorescent nature. The peaks at 1000–1275 cm^{-1} correspond to the C-H in-plane bending vibration and C–O stretching vibration, and the peaks at 1000–650 cm^{-1} correspond to the C–H out-of-plane bending vibration[13]. These results confirm that PPCDs include abundant C and O functional groups and unsaturated bonds which enhance their potential for use as corrosion inhibitors. SEM imaging and EDAX mapping (Figures 1(c–e)) further confirmed that carbon and oxygen were the primary elements in the PPCDs. Elemental analysis (Figure 1(c–e)) confirmed the absence of nitrogen in the undoped PPCDs, indicating that the observed n–π^* transitions originated solely from oxygenated groups rather than from N-containing moieties.

For N-PPCDs, TEM analysis (Figure 2a) revealed an amorphous spherical morphology similar to that observed for PPCDs, but with reduced particle sizes in the range of 2–5 nm due to nitrogen doping. The UV–vis spectrum (Figure 2b) presented a single peak at approximately 325 nm which is attributed to the n–π^* transitions involving heterocyclic structures[19,2], and the solution displayed the same orange color and blue fluorescence under UV illumination, indicating that the fluorescence properties were retained upon N doping. Figure 2(c–e) displays the results of the SEM analysis and EDAX mapping, confirming that N, C, and O are the main components of the N-PPCDs. As described above, the UV–Vis spectrum of undoped PPCDs (Figure 1b) exhibits two distinct absorption bands: a strong $\pi \rightarrow \pi$ transition near 274 nm, assigned to the aromatic C=C bonds within the conjugated carbon core, and a weaker n$\rightarrow \pi^*$ transition at approximately 367 nm corresponding to the transitions involving the C=O and C–O groups. After nitrogen doping (Figure 2b), the $\pi \rightarrow \pi^*$ transition became less distinct and merged with the n$\rightarrow \pi^*$ band to yield a single broad absorption feature centered at approximately 325 nm. This apparent disappearance of the $\pi \rightarrow \pi^*$ band is attributed to the perturbation of the π-electron system caused by nitrogen incorporation[20,6]. Nitrogen atoms introduce new electronic states (C–N, C=N, pyridinic N) within the carbon framework, disrupting the planarity and delocalization of π orbitals and enhancing charge-transfer interactions. The resulting hybridization between π and n orbitals gives rise to a blue-shifted, broadened absorption band characteristic of N-doped carbon materials. This observation confirms successful nitrogen doping and its effect on the optical and electronic properties of the CDs.

FTIR analysis of PPCDs (Figure 3(a)) revealed the presence of several functional groups; for example a broad peak at 3333 cm^{-1} can be attributed to intramolecular hydrogen bonds such as O-H[10], while the absorption peaks at 2940 and 2065 cm^{-1} can be attributed to the methyl or methylene groups[10], and the peak at 1574 cm^{-1} is attributed to the C=C vibration of the benzene ring skeleton[13]. The absorption peak at 1208 cm^{-1} corresponds to the C–O stretching vibration of carboxylic acid. The spectrum of N-PPCDs presented in Figure 3(b) reveals the presence of C=C, C–N, and C=N bonds, indicating the successful transformation of the precursors into N-PPCDs. The peak at 1597 cm^{-1} is assigned to the C—N stretching vibration[10]. The edges of the CDs are formed by chemical structures such as –CONR, which shows an absorption peak at 1651 cm^{-1}, confirming the doping process[6]. The N-PPCD shows the absorption peak of conjugated double bonds (N=C=O and C=C=N) at 2353 cm^{-1},

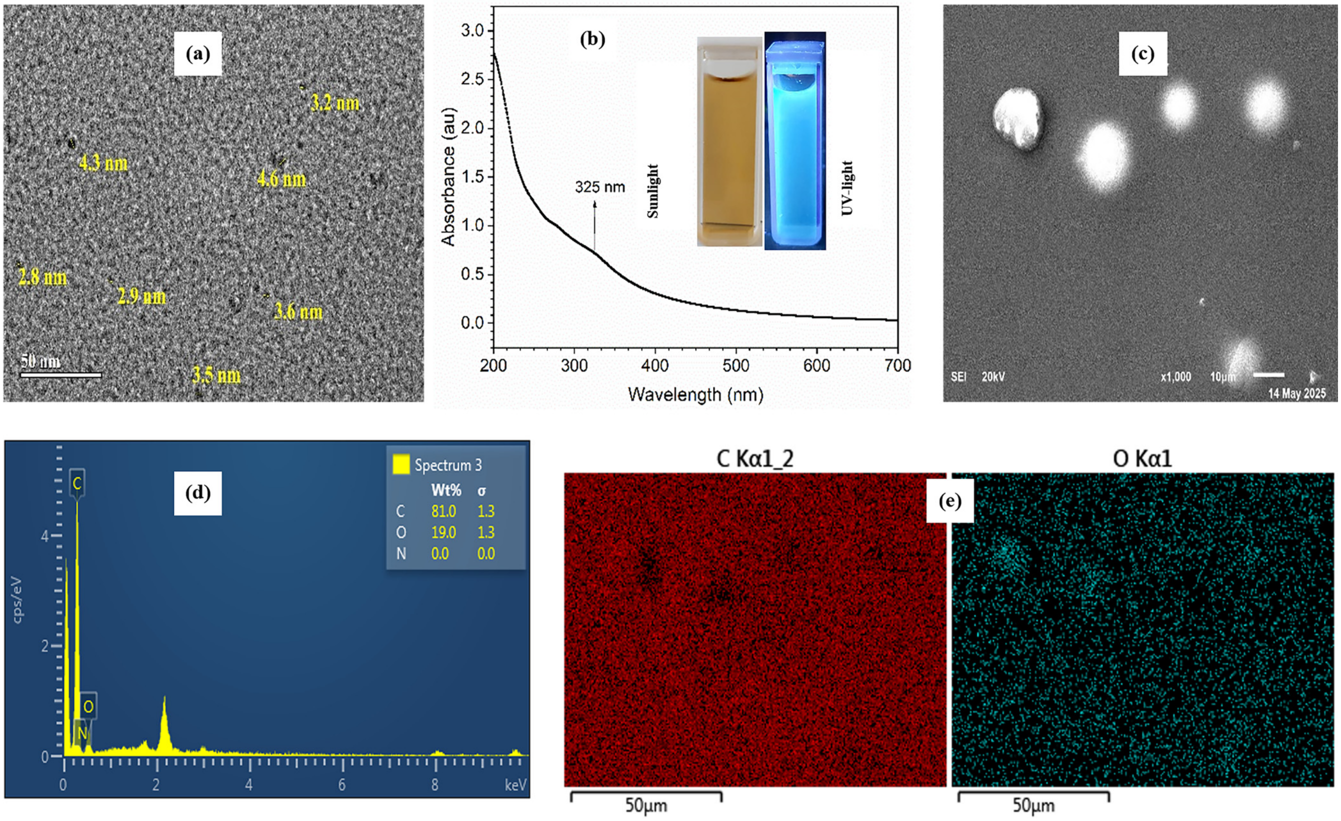


Figure 2. Results of (a) TEM, (b) UV-vis, (c) SEM, (d) EDAX and (e) elemental mapping characterizations of doped N- PPCDs

distinguishing it from its PPCD precursor[6]. The peaks centered at 1385 and 1302 cm^{-1} indicate the presence of C–N–C bonds[3]. Hence, the findings indicate that the PPCDs were successfully modified through nitrogen doping.

3.2. Corrosion Inhibition Performance.

3.2.1. Weight loss measurements. The corrosion rate and corresponding IE obtained from weight loss tests for PPCDs and N-PPCDs at two different temperatures are illustrated in Figure 4. Both the corrosion rate and IE show a clear dependence on the PPCD and N-PPCD concentrations, with increasing inhibitor concentration leading to higher IE and a corresponding decrease in the corrosion rate. This trend was attributed to the greater surface coverage of CS by the inhibitor molecules which hindered metal dissolution[21]. However, for PPCD, IE declined as the temperature was increased from 30 to 60 °C, as shown in Figure 4(a, b), likely due to the desorption of the inhibitor molecules from the CS surface and the increased molecular mobility; both of these effects reduce surface protection at elevated temperatures. The N-PPCDs showed better performance at 60 °C than at 30 °C in terms of IE. These results demonstrate the thermal stability of N-PPCDs and their ability to function at higher temperatures.

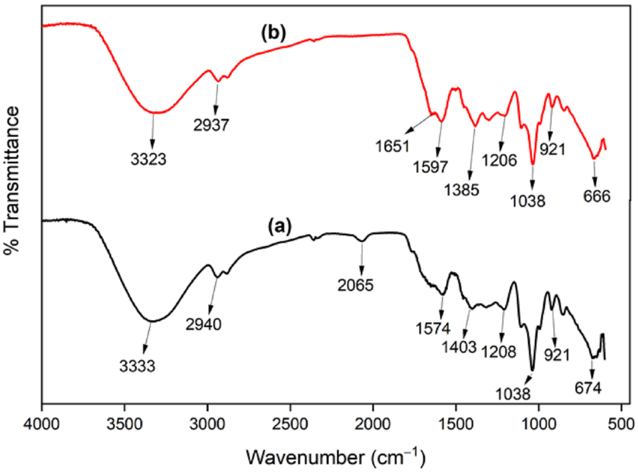


Figure 3. ATR-FTIR spectra of (a) undoped PPCDs and (b) doped N-PPCDs

3.2.2. Electrochemical measurements. PDP measurements were used to investigate the effect of the CDs on the anodic and cathodic redox reactions of CS in a 5% HCl solution. The polarization curves obtained for the CS with and without different concentrations of PPCDs and N-PPCD are shown in Figure 5. Figure 5(a-b) shows the concentration-dependent behavior of steel at a temperature of 25 °C, evaluated across the PPCD and N-PPCD concentration range of 25–200 ppm. The IE, corrosion potential (E_{corr}), corrosion current density (I_{corr}), and the anodic and cathodic slopes (β_a and β_c) were calculated by extrapolation of the anodic and cathodic current to the corrosion potentials, and the obtained results are presented in Table 1. These results indicate that the addition of CDs resulted in a noticeable decrease in both β_a and β_c accompanied by a significant reduction in the I_{corr} and E_{corr} values. These findings strongly suggest that PPCDs and N-PPCDs act as mixed-type inhibitors, effectively blocking both the cathodic and anodic reactions. The E_{corr} value

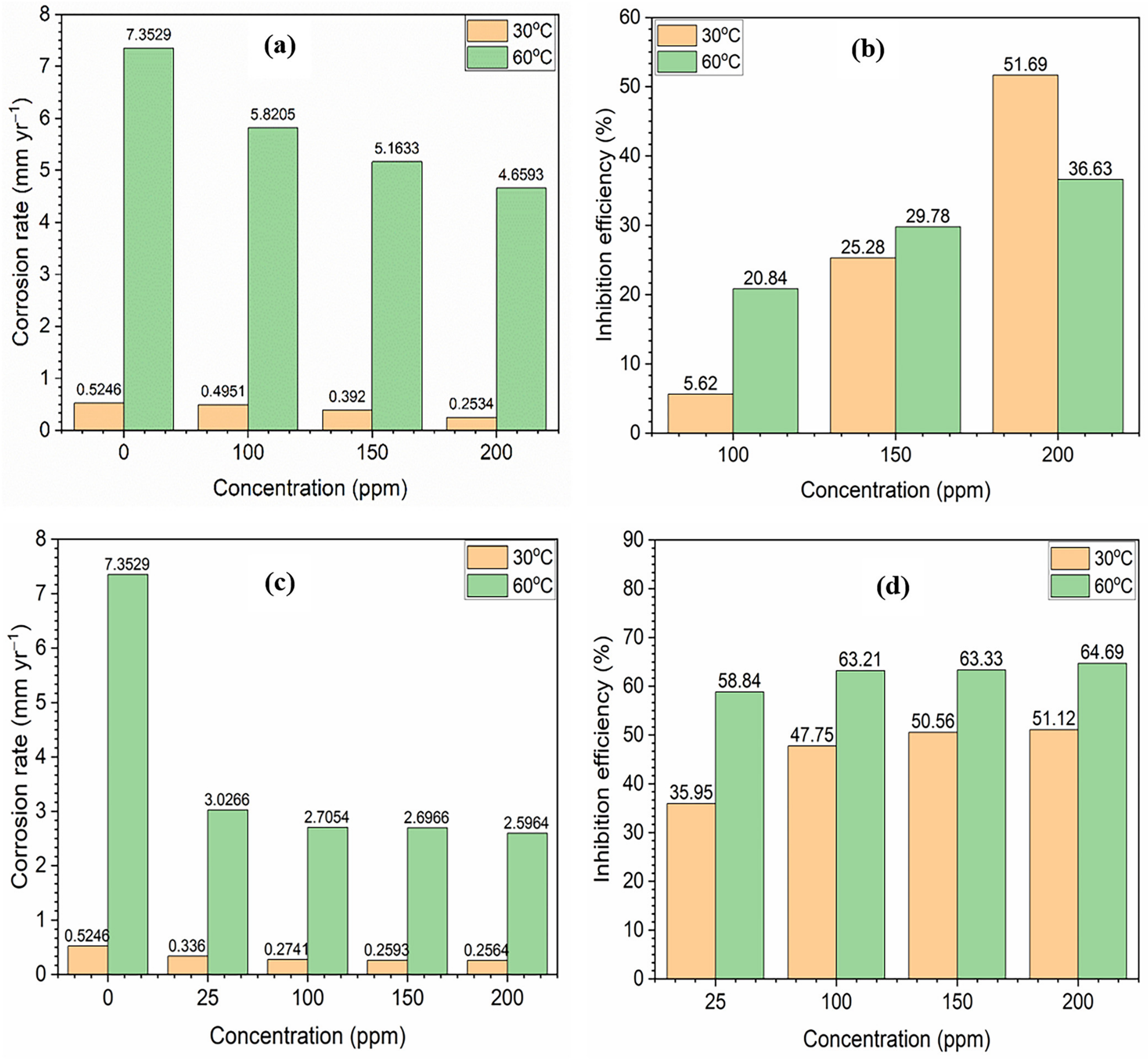


Figure 4. Corrosion rate and IE *obtained by weight loss measurements* for CS in 5% HCl solution at different concentrations of (a, b) undoped PPCDs and (c, d) doped N-PPCDs at 30 and 60 °C

became more positive after the addition of CDs, indicating that the inhibitors adsorbed onto the steel surface and blocked the active sites for corrosion.

Examination of the results presented in Table 1 shows that incorporation of PPCDs and N-PPCDs at higher concentrations increased the IE values as observed using both the PDP and LPR techniques. The highest IE values were 70.09% for PPCD and 79.58% for N-PPCD at 200 ppm obtained using the PDP technique, and 65.17% for PPCD and 75.88% for N-PPCD obtained with the LPR technique. These results showed that 200 ppm was the optimum concentration for further testing. At this concentration, the corrosion rate was 0.97 mm/y of N-PPCD, compared to 1.41 mm/y for PPCD. The higher IE and lower corrosion rate observed for N-PPCD suggested that the nitrogen-containing functional groups served as active adsorption sites, forming a protective film on the metal surface that inhibited corrosion.

The open-circuit potential (OCP) measurements at room temperature are described in the Supplementary Material, Figure S1, and section S1. The results show that the potential values shifted toward more positive potentials with increasing concentrations of PPCDs and N-PPCDs, with the most significant shift observed at 200 ppm. All inhibition curves stabilized over time, indicating the development of a protective layer on the metal surface.

To investigate the electrochemical behavior, Nyquist and Bode plots were recorded for CS in 5% HCl with varying concentrations (25–200 ppm) of PPCDs and N-PPCDs at 25 °C. Figures 6(a) and 6(c) show the Nyquist plots, and Figures 6(b) and 6(d) display the corresponding Bode plots. The Nyquist plots revealed a notable increase in the semicircle diameter upon inhibitor addition, indicating enhanced charge transfer resistance. The diameter further increased at higher inhibitor concentrations, with 200 ppm exhibiting the most significant surface coverage. N-PPCD exhibited larger diameters than PPCD at all concentrations, confirming its superior inhibition efficiency.

The Bode plots show the impedance modulus and phase angle over a frequency range between 10^5 and 10^{-2} Hz. Both PPCD

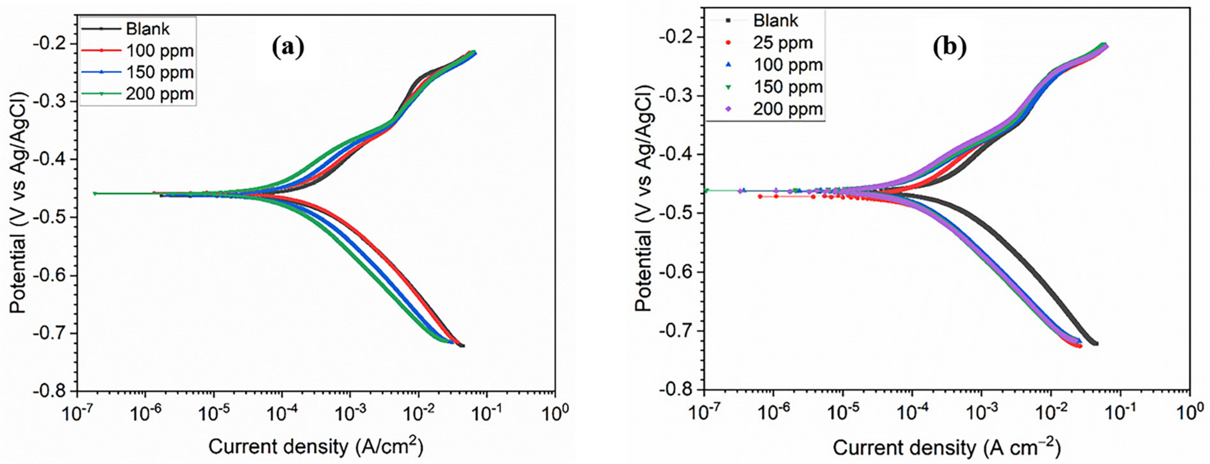


Figure 5. PDP profiles of CS immersed in 5% HCl at 25 °C, comparing the uninhibited solution with different concentrations of (a) undoped PPCDs and (b) doped N-PPCDs

and N-PPCD increased the impedance modulus, with higher values obtained with increasing concentration, suggesting effective barrier film formation on the CS surface. Additionally, the increasing peak heights in the phase angle plots with higher inhibitor concentrations indicate enhanced capacitive behavior at the metal-solution interface, which is attributed to the greater adsorption of inhibitor molecules[22].

The electrical equivalent circuit shown in the Supplementary Materials, Figure S2, and Section S2 was used to fit the experimental EIS data, and the obtained parameters are listed in Table 2. The results presented in Table 2 indicate that compared to the blank solution, the R_p values in the acidic solutions containing PPCDs and N-PPCDs increased. Specifically the R_p value increased from 51.91 $\Omega\cdot cm^2$ (uninhibited solution) to 151.09 $\Omega\cdot cm^2$ for the solution including 200 ppm PPCD, while it increased to 217.38 $\Omega\cdot cm^2$ for the solution with 200 ppm N-PPCD. This demonstrates that the PPCDs and N-PCDs formed a dense adsorbed layer, minimizing the number of active corrosion sites. The decrease in Y_{odl} suggests a reduced double-layer capacitance owing to the adsorbed inhibitor molecules forming a less permeable film. N-doped PPCDs consistently showed better performance than undoped PPCDs at the same concentrations, most likely due to their enhanced electron-donating capability, improved surface affinity, and denser film formation attributed to the nitrogen functional groups. At 200 ppm, the N-PPCDs provide R_{ct} and R_p of 168.3 $\Omega\cdot cm^2$ and 217.38 $\Omega\cdot cm^2$, respectively, with IE = 76.12%, outperforming undoped PPCDs. The sharp improvement in the inhibition efficiency, even at 25 ppm, of N-PPCDs emphasizes their potential as highly efficient green inhibitors at low dosages.

3.3. Surface Analysis. SEM-EDS analysis was carried out on the CS samples immersed in 5% HCl for 24 h, both with and without the addition of 100 ppm CDs, to assess surface morphology and elemental composition. The steel surface exposed to HCl without CDs exhibited significant damage owing to metal dissolution and degradation (Figure 7(a)). Exposure to HCl resulted in notable deformation and deep pits on the surface. Conversely, the micrographs of the steel surface treated with 200 ppm PPCD (Figure 7(c)) and N-PPCD (Figure 7(e)) display less surface damage with fewer imperfections compared to the untreated sample. EDS analysis of the CS sample (Figure 7(b)) revealed peaks consistent with the chemical composition of the corroded steel. In the EDS spectrum of the untreated sample, the chloride peaks indicated the adsorption of chloride ions on the steel surface, which can accelerate corrosion[10]. The reduction of the chlorine peak in the EDAX results for the PPCD-inhibited

Table 1. *PDP and LPR parameters for CS in 5% HCl without and with different concentrations of undoped PPCDs and doped N-PPCDs at 30 °C*

			PDP Technique					LPR Technique		
CDs	Conc. (ppm)	E_{corr} (mV vs Ag/AgCl)	I_{corr} ($\mu A/cm^2$)	β_a (mV/ dec)	β_c (mV/ dec)	Corrosion rate (mm/yr)	IE (%)	R_p (Ω cm^2)	Corrosion rate (mm/yr)	IE (%)
	Blank	−463.0	408.0	178.4	114.4	4.74	–	52.46	5.76	–
PPCDs	100	−458.0	301.0	137.5	102.1	3.49	26.23	67.59	4.47	22.38
	150	−462.0	199.0	137.3	105.9	2.31	51.23	102.3	2.96	48.72
	200	−459.0	122.0	120.9	106.2	1.41	70.09	150.6	2.01	65.17
N-PPCDs	25	−472.0	156.0	125.5	112.3	1.81	61.76	145.4	2.08	63.92
	100	−462.0	99.0	98.1	96.3	1.15	75.74	152.5	1.98	65.60
	150	−461.0	90.1	97.5	100.7	1.05	77.92	202.1	1.49	74.04
	200	−463.0	83.3	108.2	92.7	0.97	79.58	217.5	1.39	75.88

Table 2. *Impedance parameters for CS in 5% HCl without and with different concentrations of undoped PPCDs and N-PPCDs at 25 °C*

CDs	Conc. (ppm)	R_s (Ω cm^2)	Yo_f (mF cm^{-2} s^{n-1})	n_f	R_f (Ω cm^2)	Y_{odl} (μF cm^{-2} s^{n-1})	n_{dl}	R_{ct} (Ω cm^2)	R_p (R_f+ R_{ct}) (Ω cm^2)	$\chi^2 \times 10^{-4}$	IE (%)
	Blank	0.81	1.25	0.78	14.43	433.7	0.97	37.48	51.91	4.23	–
PPCDs	100	0.84	0.51	1.0	29.45	810.9	0.78	45.72	75.17	8.76	30.94
	150	0.65	0.23	0.99	30.11	523.7	0.79	78.74	108.85	8.54	52.31
	200	0.67	0.18	0.97	25.29	664.9	0. 79	125.8	151.09	7.41	65.64
N-PPCDs	25	0.76	1.16	0.87	3.03	321.6	0.83	141.2	144.23	1.86	64.01
	100	0.74	0.76	0.76	26.29	252.4	0.98	127.6	153.89	7.31	66.27
	150	0.76	0.87	0.76	18.29	229.3	0.92	177.9	196.19	5.84	73.54
	200	0.67	0.77	0.75	49.08	365.2	0.92	168.3	217.38	1.10	76.12

and N-PPCD-inhibited CS samples indicates that the inhibitor molecules effectively blocked the adsorption of chloride ions on the surface (Figures 7(d),(f)). Additionally, the nitrogen peak in the EDS spectrum of the N-PPCD-inhibited CS sample suggests that N-PPCD was adsorbed onto the steel surface (Figure 7(f)).

AFM surface analysis was used to examine the changes in the CS surface topography after immersion in 5% HCl solution, both with and without the inhibitor. Figure S4(a–d) show 2D and 3D micrographs of the CS submerged in the corrosive medium with and without CDs. The polished sample surface shows no visible damage (Figure S4(a)). By contrast, the sample immersed in the corrosive media exhibited clear pits and fissures, with the surface roughness increasing dramatically from 0.5166 to 68.468 nm, indicating significant corrosion. Conversely, the surface of the sample in the inhibitor-containing solution displayed minimal imperfections (Figure S4(c, d) in Supplementary Materials), with a roughness of only 2.433 nm for the PPCD-inhibited CS sample and of 1.121 nm for the N-PPCD-inhibited CS sample, which is much lower than the 68.468 nm observed for the sample without the inhibitor. These findings implied that PPCDs and N-PPCDs effectively protected steel from corrosion in harsh HCl environments. The N-doped PPCDs demonstrated superior inhibition performance, creating a more effective barrier layer that reduces metal dissolution and pitting.

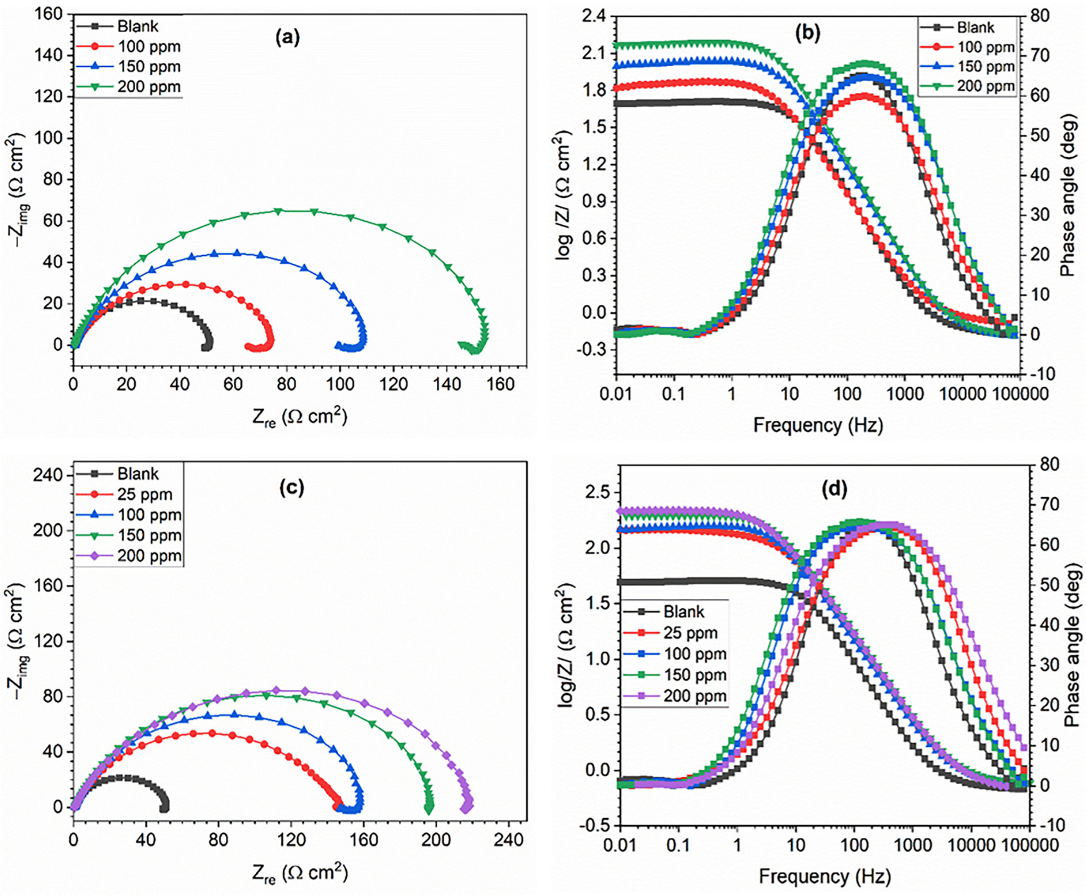


Figure 6. Impedance plots for CS in 5% HCl solution in the absence and presence of different concentrations of (a, b) undoped PPCDs and (c, d) doped N-PPCDs at 25 °C in Nyquist and Bode representations

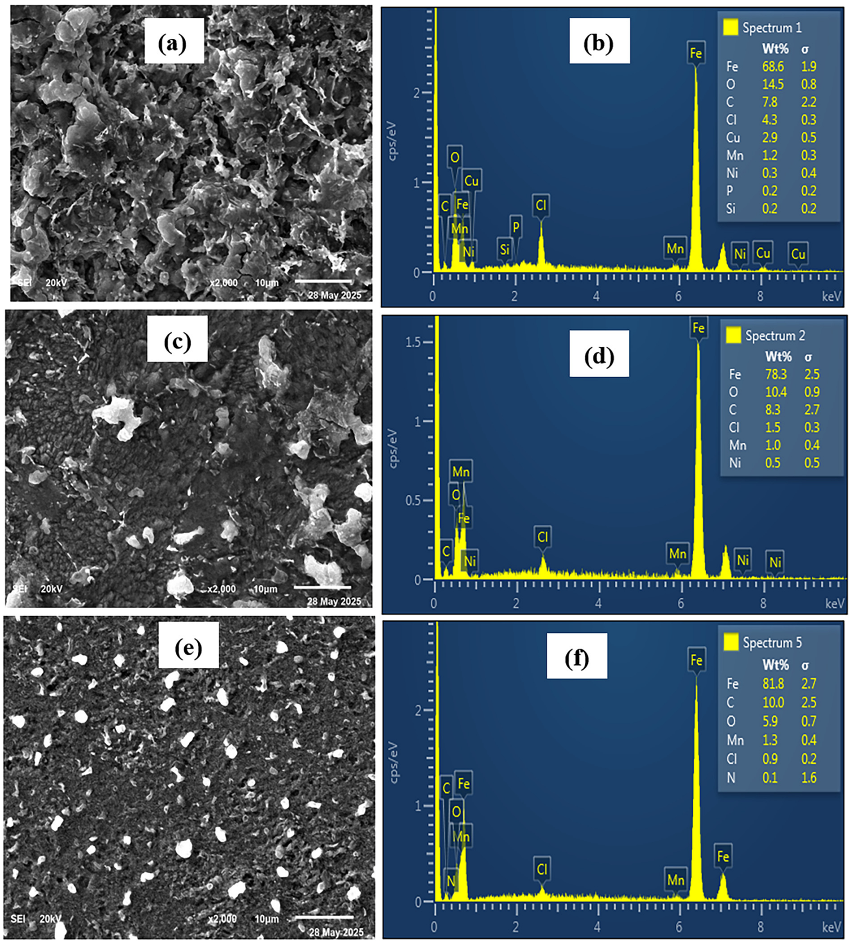


Figure 7. SEM images and corresponding EDAX spectra of CS immersed in 5% HCl (a and b) without inhibitor, (c and d) containing 100 ppm of undoped PPCDs and (e and f) containing 100 ppm of doped N-PPCDs at 30 °C after 24 h immersion

Optical profilometry was performed to further validate the microscale surface protection behavior and Figure S3 in the Supplementary Materials, Section S3 shows the 2D and 3D profilometer images obtained under identical test conditions. Table S1 summarizes the surface roughness parameters obtained from optical profilometry for CS samples exposed to 5% HCl solution, both in the absence and presence of 200 ppm of undoped PPCDs and N-doped PPCDs at 30 °C. As discussed in Section S4 of the Supplementary Materials, the sample treated with N-doped PPCDs demonstrated markedly improved surface protection. These roughness values were significantly lower than those of the blank and undoped PPCD-treated samples.

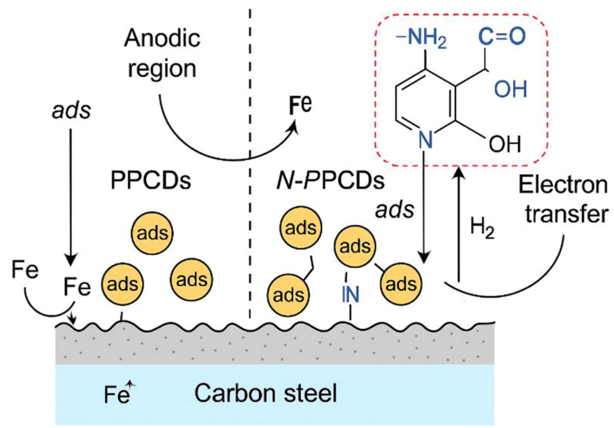


Figure 8. Proposed mechanism of corrosion inhibition of carbon steel in 5% HCl by undoped (PPCDs) and nitrogen-doped (N-PPCDs) pomegranate-peel-derived carbon dot

3.4. Possible Mechanism and Comparison to Previous Studies.

Taken together, the electrochemical and surface characterization results indicate that both PPCDs and, N-PPCDs inhibit CS corrosion in 5% HCl via competitive adsorption and compact film formation at the metal/electrolyte interface with a more pronounced effect obtained for N-PPCDs. In 5% HCl, the steel surface rapidly accumulates specifically adsorbed Cl^-, while the oxygen- and nitrogen-bearing groups (carboxylates, hydroxyls; and for N-PPCDs: amines/amide/pyridinic N) on the CDs in the solution are partially protonated. The protonated functional groups in the CDs are attracted to the Cl^--decorated surface regions, displacing water and anions and binding to Fe through O-donor (PPCDs) and additional N-donor (N-PPCDs) sites via electrostatic attraction and van der Waals/$\pi-\pi$ interactions. The adsorbed CD layer functions as a barrier to charge transfer (raising R_{ct}) and as a diffusion barrier for H^+ and Cl^-. For the cathodic reaction (H_2 evolution), the inhibitor film reduces access of H^+ to catalytic sites. For the anodic reaction ($Fe \rightarrow Fe^{2+} + 2e^-$), coverage of active Fe sites and partial

Table 3. *Comparison of corrosion inhibition performance of biomass-derived CDs for carbon steel in an acid medium*

S/N	CD inhibitor types	Metal substrate	Corrosive medium	Concentration	Inhibition efficiency (%)	Ref.
1	Acerola seed-derived CDs	Mild steel	1 M HCl	200 ppm	94.1	[23]
2	Pumpkin seed-derived CDs	carbon steel	1 M HCl	10 ppm	94.6	[24]
3	Sugarcane bagasse-derived CDs	carbon steel	1 M HCl	100 ppm	94.0	[25]
4	Lichi leaves derived CDs	Q235 steel	1 M HCl	200 ppm	86.1	[26]
5	Peanut shell-derived CDs	Cold-rolled steel	1 M HCl	50 mg/L	92.4	[27]
6	Grapefruit-derived CDs	Q235 steel	1 M HCl	200 ppm	94.1	[13]
7	*Z. bungeanum* leaves-derived CDs	Q235 steel	1 M HCl	200 ppm	95.9	[28]
8	Lupine seed-derived CDs	Carbon steel	1 M HCl	175 ppm	89.3	[29]
9	Pomegranate peel-derived CDs	Carbon steel	5% HCl	200 ppm	70.0	This study
10	N-doped-pomegranate peel-derived CDs	Carbon steel	5% HCl	200 ppm	80.0	This study

complexation of Fe^{2+} by O/N groups at the interface reduce effective dissolution. The organic layer formed by the CDs blocks both anodic dissolution and cathodic hydrogen evolution, consistent with the mixed-type behavior observed in the PDP experiments (simultaneous reductions in β_a and β_c with modest E_{corr} shifts), the increased R_p/R_{ct} and decreased double-layer admittance found by EIS measurements, and the markedly reduced roughness and suppressed chloride signals observed by AFM/SEM-EDAX. The enhanced performance of the N-PPCDs was attributed to stronger chemisorption and tighter packing due to the nitrogen functional groups, yielding a denser and less permeable film. The observed retention or improvement of IE at 60 °C for N-PPCDs suggests a chemisorption-dominated component for which the binding is less strongly affected by temperature, whereas PPCD protection shows greater decreases with higher temperature, suggesting a stronger physisorption contribution to the binding.

Hence, overall the data show (i) mixed-type inhibition (PDP: simultaneous decreases in β_a and β_c with modest, <85 mV, E_{corr} shifts), (ii) substantial increases in R_p/R_{ct} with rising inhibitor dose (EIS), (iii) marked decreases in double-layer admittance/capacitance (CPE_{dl}/Y_{odl}), and (iv) dramatically smoother surfaces with suppressed Cl signals (AFM/SEM-EDAX). This mechanism is schematically depicted in Figure 8. Together, these features support an adsorption/film formation mechanism enhanced by the nitrogen functionalities in the N-PPCDs for the inhibition of corrosion.

Table 3 compares the corrosion inhibition efficiencies of various biomass-derived CDs tested on mild or carbon steels in acidic HCl environments. The inhibition efficiencies reported in previous studies range from 86% to 96% and were typically achieved at concentrations between 10 and 200 ppm in 1 M HCl (≈3.65%). Notably, all previous studies used milder acid strengths and similar or higher doses of inhibitors. Although most reported CDs achieved 86–96% efficiency in 1 M HCl, the present study N-doped CDs exhibited 80% inhibition even in the substantially more aggressive 5% HCl medium. The 10% improvement of inhibition upon nitrogen doping confirmed the beneficial role of N-donor functionalities in enhancing chemisorption and compact film formation on the steel surface. Considering the higher acid strength and carbon steel substrate used in this study, the observed performance was comparable to or better than that of the previously reported multi-heteroatom-doped CDs, highlighting the chemical economy and green synthetic advantages of the present formulation. Additionally, the present study is the first report on the use of nitrogen-doped pomegranate peels CDs for corrosion inhibition.

4. CONCLUSIONS

This study demonstrates that CDs synthesized from pomegranate peel waste, and particularly nitrogen-doped PPCDs (N-PPCDs), can serve as effective green corrosion inhibitors of CS in acidic environments. N-PPCDs achieved superior inhibition efficiency (up to 79.58%) owing to enhanced surface interactions and film formation. Electrochemical and surface analyses confirmed the reduced corrosion rates and smoother steel surfaces in the presence of these inhibitors. Overall, N-PPCDs offer a sustainable and high-performance alternative to conventional inhibitors, supporting the use of biomass-derived nanomaterials for corrosion protection.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Ali M. Al Nasser – *Department of Materials Science and Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0009-0008-2138-4011
Email: s202012520@kfupm.edu.sa

Authors

Hifsa Kurdshid – *Postdoctoral Fellow, Interdisciplinary Research Center for Advanced Materials, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0000-0001-5295-1332
Email: hifsa.khurshid@kfupm.edu.sa

Corresponding Author

Saviour A. Umoren – *Research Mentor, Principal Research Scientist, Interdisciplinary Research Center for Advanced Materials, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia.*

Department of Materials Science and Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia; 0000-0002-8564-4897
Email: umoren@kfupm.edu.sa

ACKNOWLEDGEMENTS

The authors acknowledge the support provided by the Interdisciplinary Research Center for Advanced Materials at King Fahd University of Petroleum and Minerals.

REFERENCES

(1) He, H. J. *et al.* Exploiting machine learning for controlled synthesis of carbon dots-based corrosion inhibitors. *J. Clean. Prod.* **419**, (2023).

(2) Yang, Y., Lu, R., Chen, W., Mei, P. & Lai, L. Amphiphilic carbon dots as high-efficiency corrosion inhibitor for N80 steel in HCl solution: Performance and mechanism investigation. *Colloids Surfaces A Physicochem. Eng. Asp.* **649**, 129457 (2022).

(3) Aslam, R., Sun, X., Zhao, J., Wang, Q. & Yan, Z. Assessment of Anticorrosive Application of Biomass-Derived Carbon Quantum Dots for Mild Steel in 15% HCl Solution: Chemical and Electrochemical Studies. *Waste and Biomass Valorization* (2025) doi:10.1007/s12649-025-03090-w.

(4) Cui, M., Ren, S., Zhao, H., Wang, L. & Xue, Q. Novel nitrogen doped carbon dots for corrosion inhibition of carbon steel in 1 M HCl solution. *Appl. Surf. Sci.* **443**, 145–156 (2018).

(5) Gupta, D., Priyadarshi, R., Tammina, S. K., Rhim, J.-W. & Agrawal, G. Fruit Processing Wastes as Sustainable Sources to Produce Multifunctional Carbon Quantum Dots for Application in Active Food Packaging. *Food Bioprocess Technol.* (2024) doi:10.1007/s11947-024-03578-8.

(6) Matter, E. A., El-Naggar, G. A., Nasr, F. & Ahmed, G. H. G. Facile synthesis of N-doped carbon dots (N-CDs) for effective corrosion inhibition of mild steel in 1 M HCl solution. *J. Appl. Electrochem.* **53**, 2057–2075 (2023).

(7) Moradi, M., Molaei, R., Kousheh, S. A., T. Guimarães, J. & McClements, D. J. Carbon dots synthesized from microorganisms and food by-products: active and smart food packaging applications. *Crit. Rev. Food Sci. Nutr.* **63**, 1943–1959 (2023).

(8) Long, W. J., Li, X. Q., Zheng, S. Y. & He, C. A novel effective carbon dots-based inhibitor for carbon steel against chloride corrosion: From inhibition behavior to mechanism. *Carbon N. Y.* **218**, 118708 (2024).

(9) Ganesan, K. *et al.* Candle soot derived carbon dots as potential corrosion inhibitor for stainless steel in HCl medium. *J. Appl. Electrochem.* **54**, 89–102 (2024).

(10) Aslam, R. *et al.* Adsorption and Corrosion Inhibition Evaluation of Agro-Industrial Waste-Derived Sustainable Carbon Dots for Corrosion Protection of Q235 Steel in 5% HCl. *Langmuir* (2025) doi:10.1021/acs.langmuir.4c04571.

(11) Tengfei Xiang, Jiaqi Wang, Yanli Liang, Walid Daoudi, Wei Dong, Ruiqian Li, Xuxin Chen, Shengjun Liu, Shunli Zheng, and K. Z. Carbon Dots for Anti-Corrosion. *Adv. Funct. Mater.* **34**, 1–26 (2024).

(12) Wu, S., Wang, J., Liu, T., Guo, X. & Ma, L. Sulfosalicylic acid modified carbon dots as effective corrosion inhibitor and fluorescent corrosion indicator for carbon steel in HCl solution. *Colloids Surfaces A Physicochem. Eng. Asp.* **661**, 130951 (2023).

(13) Zheng, S. *et al.* Study on the corrosion inhibition of biomass carbon quantum dot self- aggregation on Q235 steel in hydrochloric acid. *Arab. J. Chem.* **16**, 104605 (2023).

(14) Saraswat, V. & Yadav, M. Improved corrosion resistant performance of mild steel under acid environment by novel carbon dots as green corrosion inhibitor. *Colloids Surfaces A Physicochem. Eng. Asp.* **627**, 127172 (2021).

(15) Umoren, P. S., Kavaz, D., Nzila, A., Sankaran, S. S. & Umoren, S. A. Biogenic Synthesis and Characterization of Chitosan-CuO Nanocomposite and Evaluation of Antibacterial Activity against Gram-Positive and -Negative Bacteria. *Polymers (Basel).* **14**, (2022).

(16) Xu, X. *et al.* Nitrogen-doped carbon quantum dots for effective corrosion inhibition of Q235 steel in concentrated sulphuric acid solution. *Mater. Today Commun.* **29**, 102872 (2021).

(17) Usman, B. J., Umoren, S. A. & Gasem, Z. M. Inhibition of API 5L X60 steel corrosion in CO_2-saturated 3.5% NaCl solution by tannic acid and synergistic effect of KI additive. *J. Mol. Liq.* **237**, 146–156 (2017).

(18) Zhu, M. *et al.* Insights into the newly synthesized N-doped carbon dots for Q235 steel corrosion retardation in acidizing media: A detailed multidimensional study. *J. Colloid Interface Sci.* **608**, 2039–2049 (2022).

(19) Zhu, M. *et al.* Insights into the newly synthesized N-doped carbon dots for Q235 steel corrosion retardation in acidizing media: A detailed multidimensional study. *J. Colloid Interface Sci.* **608**, 2039–2049 (2022).

(20) Xu, Z. *et al.* The n–π^* electronic transition induced by nitrogen vacancies enhances photocatalytic hydrogen production in carbon nitride. *Chem. Eng. J.* **501**, 157670 (2024).

(21) Cen, H., Chen, Z. & Guo, X. N, S co-doped carbon dots as effective corrosion inhibitor for carbon steel in CO_2-saturated 3.5% NaCl solution. *J. Taiwan Inst. Chem. Eng.* **99**, 224–238 (2019).

(22) Mourya, P., Banerjee, S. & Singh, M. M. Corrosion inhibition of mild steel in acidic solution by Tagetes erecta (Marigold flower) extract as a green inhibitor. *Corros. Sci.* **85**, 352–363 (2014).

(23) Paiva, V. M. *et al.* A new carbon quantum dot obtained from acerola (Malpighia glabra) seed waste and its application as a corrosion inhibitor of mild steel in hydrochloric acid. *Biomass and Bioenergy* **203**, 108264 (2025).

(24) Magno Paiva, V. *et al.* Pumpkin (Cucurbita maxima) seed-derived nitrogen, phosphorus, and sulfur carbon quantum dot as an inhibitor of corrosion for mild steel in HCl solution. *J. Mater. Res. Technol.* **28**, 2504–2515 (2024).

(25) Nunes, R. da S. *et al.* Sugar Cane (Saccharum officinarum L.) Waste Synthesized Si,N,S-Carbon Quantum Dots as High-Performance Corrosion Inhibitors for Mild Steel in Hydrochloric Acid. *ACS Omega* **9**, 50246–50259 (2024).

(26) Long, W.-J., Li, X.-Q., Yu, Y. & He, C. Green synthesis of biomass-derived carbon dots as an efficient corrosion inhibitor. *J. Mol. Liq.* **360**, 119522 (2022).

(27) Guan, Y., Li, X., Tan, B., Yang, S. & Deng, S. Hydrothermal synthesis of peanut shell-derived carbon dots as the novel efficient inhibitor for the corrosion of steel. *Colloids Surfaces A Physicochem. Eng. Asp.* **728**, 138517 (2026).

(28) Dong, L. *et al.* High-Efficiency Corrosion Inhibitor of Biomass-Derived High-Yield Carbon Quantum Dots for Q235 Carbon Steel in 1 M HCl Solution. *ACS Omega* **8**, 46934–46945 (2023).

(29) Abd-El-Nabey, B. A., Mahmoud, M. E., Abdelrahman, A. & Abd-El-Fatah, M. A. Effective Performance of Derived Sustainable Lupine Carbon Quantum Dots as a Superior Inhibitor Material for Carbon Steel Corrosion in a Hydrochloric Acidic Environment (1.0 Molar). *Ind. Eng. Chem. Res.* **63**, 2901–2915 (2024).

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article *Civil Engineering*

Evaluating the Sustainability of Ultra-high-performance Concrete Mixes: A Multi-criteria Life Cycle Assessment Incorporating Embodied Carbon and Energy

Hussain Abdulatif Karam, Ashraf A. Bahraq*, Mohammed A. Al-Osta, and Shamsad Ahmad

Cite https://doi.org/10.64589/juri/214907

Submitted: September 11, 2025 **Revised:** September 21, 2025 **Accepted:** November 30, 2025

ABSTRACT

Ultra-high-performance concrete (UHPC) represents a modern class of cementitious materials characterized by superior strength and durability. However, the high embodied carbon and energy of UHPC relative to those of normal concrete (NC) pose major sustainability challenges. This study presents a comprehensive evaluation of 40 UHPC mixes through an integrated life cycle assessment that incorporated embodied carbon, embodied energy, and compressive strength. The assessment focused on the raw material product, transportation, and maintenance stages and examined both local and global material-sourcing scenarios. The findings showed that UHPC produced locally in Saudi Arabia exhibited 3.28% to 8.21% higher embodied carbon and 1.33% to 4.34% higher embodied energy than its global equivalents, highlighting the need to optimize regional production practices. To evaluate sustainability from a performance-based perspective, an eco-strength index (ESI) was employed to capture the tradeoff between embodied carbon and compressive strength. The ESI values ranged from 7.3 to 14.9 kg·CO_2/MPa. A case study involving a single-girder bridge constructed with both UHPC and NC revealed that although UHPC exhibited a higher impact per cubic meter, its superior mechanical performance enabled material reduction, thereby lowering the total emissions per meter of bridge length. These results emphasize that when structural efficiency is considered, UHPC represents a sustainable solution for high-performance infrastructure applications.

Keywords: ultra-high-performance concrete, life-cycle impact assessment, embodied carbon, embodied energy, sustainable performance index, sustainability goals

1. INTRODUCTION

Construction industry growth in recent years has greatly increased the need for concrete with superior performance for various construction projects. Ultra-high-performance concrete (UHPC) is a novel class of concrete that exhibits high mechanical strength and durability and thus represents a promising option to meet this need[1]. UHPC is characterized by compressive and tensile strengths exceeding 120 MPa and 6.9 MPa, respectively, along with resistance to chloride attack and freeze–thaw cycles, compared to conventional concrete[2–4]. These superior properties are achieved via high amounts of binders and fine aggregates that increase the packing density, a low water-to-binder ratio, and the inclusion of specialized chemical admixtures and steel fibers[5].

Despite its superior mechanical properties, concerns regarding the environmental impact of UHPC have increased because of its high cement content compared with that of normal concrete (NC). Cement accounts for approximately 8% of all greenhouse gas emissions worldwide because of the exorbitant energy required to heat the raw materials used to produce cement[6]. Therefore, numerous researchers have investigated various mix designs to reduce the embodied carbon of UHPC. These designs include substituting cement with other binders such as fly ash, silica fume, and ground-granulated blast furnace slag to reduce the carbon footprint without compromising the mechanical properties[7].

The existing literature has primarily focused on optimizing the strength and enhancing the durability of UHPC blends, with few studies having investigated comprehensive sustainability assessment frameworks that account for environmental burdens alongside mechanical performance. Moreover, most life cycle assessment (LCA) studies on UHPC are confined to cradle-to-gate analyses and often overlook critical phases, such as transportation, maintenance, and structural efficiency. In addition, the influence of regional variability in material sourcing, particularly the environmental implications of locally manufactured materials, has rarely been addressed.

Therefore, in this study, we performed a multicriteria sustainability evaluation of 40 UHPC mixes that integrates embodied carbon, embodied energy, and compressive strength into a unified framework. The environmental sustainability of the UHPC formulations was evaluated by analyzing their embodied carbon and energy using the LCA method. A selection of UHPC mix designs from the existing literature was used to estimate their environmental footprints relative to that of NC to identify more sustainable alternatives. In addition, an eco-strength index (ESI) of the UHPC and NC mixes was used to assess both the environmental and mechanical performance, thus providing a balanced sustainability perspective that integrates structural efficiency and carbon footprint. To further capture the combined effects of environmental impact and structural efficiency, we propose a sustainable performance index. This index enables a more holistic comparison of UHPC mixes and supports the classification of their sustainability profiles based on performance trade-offs.

The remainder of this paper is structured as follows. Section 2 presents the research methodology adopted in this study; describes the UHPC criteria, life cycle inventory (LCI) analysis development and assumptions, life cycle impact assessment (LCIA) categories and methods, and ESI calculations; and presents a case study comparing UHPC and NC bridges. Section 3 presents a comparative analysis of the UHPC mixes that showcase the trends and influence of the raw materials, a comparison of the mixes with NC, a detailed analysis of the key contributors to environmental impact, a comparison of the eco-strength evaluation of UHPC and its tradeoffs, and a comparison of the environmental impact and long-term sustainability of both UHPC and NC bridges as part of the case study. Finally, Section 4 concludes the paper by summarizing the key findings, showcasing the implications and recommendations, and establishing a path for future research.

2. RESEARCH METHODOLOGY

2.1. Ultra-High-Performance Concrete Mixes. Criteria for source selection were used to ensure a fair and consistent comparison of the UHPC mixes. First, there were no restrictions on the year of publication. Although some sources were older, they contain mixed designs comparable to those in more recent literature. Second, only mixes composed of the following materials were used: Ordinary Portland Cement (OPC) as the primary cementitious material and silica fume (SF), natural pozzolan (NP), fly ash (FA), and ground granulated blast-furnace slag (GGBFS) as secondary cementitious materials. Quartz powder (QP) and fine aggregates were considered as filler materials, and superplasticizer (SP), steel fibers, and water were also included. These criteria ensured a fair and consistent application of environmental impact comparisons using the available EPDs. Finally, all selected UHPC mixes were required to have a minimum 28-day compressive strength of 120 MPa, as specified by ASTM C1856[8], which is commonly used to define UHPC[2–4].

Based on the above conditions, 40 UHPC and 3 NC mix designs were collected from the literature, as presented in Table 1 and Table 2. The selected mixes were established in various geographic regions, including North America, Europe, and Asia, and captured a broad spectrum of UHPC compositions. Each mix included detailed information on the raw material proportions facilitate consistent analysis of the environmental effects based on the raw material quantities. All mixes incorporated one or more supplementary cementitious materials (SCMs) and varying steel fiber and SP contents. The compressive strengths at 28 days ranged from 120 MPa to over 160 MPa, supporting broad high-stress structural applications. This dataset forms the basis for the LCA and ESI comparisons in the subsequent sections.

To evaluate the impact of material sourcing, this study distinguished between locally produced materials within the Kingdom of Saudi Arabia and imported materials that were sourced internationally. Internationally sourced materials generally exhibit lower environmental impacts at the raw material product stage (A1–A3) due to more efficient manufacturing practices (Figure 1). However, these benefits are partially offset by the additional environmental burden associated with long-distance transportation (A4). In contrast, local materials offer advantages in terms of logistics, availability, and supply chain resilience, particularly in responding to site-level demands or shortages. Of all the raw materials, only two had local EPDs: OPC and NP. Although some materials may have local manufacturing facilities, such as fine aggregate, local materials were considered the average of global materials and assigned zero transportation emissions. Therefore, in the comparisons of local and global UHPC mixes, the only factors that impacted the outcome of the LCA were OPC, NP, and transportation.

2.2. Life Cycle Assessment Framework. To evaluate the environmental performance of the proposed materials and mix designs, a structured LCA approach was developed. The following framework outlines the overall structure and boundaries adopted in this study's LCA, describing the stages, inclusions, and exclusions considered throughout the analysis. The LCA primarily followed a cradle-to-gate system boundary (raw material product stage, A1–A3), although downstream modules were also included, namely, the transportation of raw materials (A4) and maintenance-related activities (B2–B5) during the service life of the structure. Other life cycle phases, such as construction and installation (A5), use-phase operations (B1, B6, and B7), and end-of-life stages (C1–C4), were explicitly excluded from the analysis because their contributions were considered irrelevant or negligible.

2.3. Life Cycle Inventory Analysis. LCA measures environmental impact via a standardized framework through various steps of an item's life cycle, starting with the extraction of raw materials and ending with the demolition of structures. The LCA process includes several key phases according to ISO 14040[32]: defining the assessment's goals and scope, conducting a comprehensive LCI to gather data on resource use and emissions, performing an LCIA to evaluate the environmental consequences of these activities, and interpreting the results to reach meaningful conclusions[7]. Applying these standardized methods will help to estimate the emissions associated with UHPCs, thereby enabling the discovery of more viable options for enhancing sustainability.

LCI involves collecting data on all material inputs, specifically the associated carbon emissions and energy use of these inputs, and the outputs for the defined product system. This study's LCI focused exclusively on the constituent materials of the UHPC mixtures. The scope included raw material extraction and manufacturing as well as transportation to Dammam Port, Saudi Arabia for imported materials. Downstream processes such as mixing, placement, and end-of-life treatment were excluded. EPDs were used to obtain standardized, third-party verified environmental

Table 1. *UHPC mix designs, including their various proportions, 28-day compressive strength, and flow*

		Quantity (kg/m^3)										Compressive	
Ref.	Mix No.	OPC	SF	NP	FA	GGBFS	Fine Agg.	QP	SP	Water	Steel fiber	strength (MPa)	Flow (mm)
Li et al.[1]	1	856.5	214.1	-	-	-	1177.5	-	19.7	184	120	157.0	n/a
Ha et al.[9]	2	625	250	-	-	375	1100	-	70	170	156	151.3	n/a
	3	1000	250	-	-	-	1100	-	70	170	-	159.3	n/a
Regalla and Kumar[10]	4	712	231	-	-	-	1020	211	24.92	178	-	144.0	240
Huang et al.[11]	5	488.8	135.2	-		416	1114	-	26	187.2	-	125.0	225
Du et al.[12]	6	470.7	-	-	108.3	433.3	990	-	4.2	232.4	156	120.0	n/a
Abdolpour et al.[13]	7	697	190	-	150	-	1068	-	25	182	79	153.6	220
	8	820	190	-	150	-	1068	-	25	182		168.7	290
Nassar et al.[14]	9	780	95	-	-	35	1185	-	18.2	195	22.75	143.2	219
Shafieifar et al.[15]	10	712	231	-	-	-	1020	211	30.7	109	156	138.0	n/a
Alsalman et al.[16]	11	1044	261	-	-	-	1212	-	113.7	261	-	158.0	200
Dingqiang et al.[17]	12	900	125	-	-	-	1060	-	36	180	-	130.0	212
Hasnat and Ghafoori[18]	13	991	70	-	-	-	1174	-	13.6	216	156	143.0	244
	14	881	31	-	122	-	1174	-	18.3	205	234	136.0	248
	15	660	47	-	81	120	1174	-	13.7	184	156	131.0	245
Shi et al.[19]	16	744	159	-	159		1169	-	19.1	181	120	124.0	240
	17	641	160	-	-	267	1175	-	14.3	182	120	138.0	240
Li et al.[20]	18	872	261	-	-	131	1046	-	51	240	156	158.6	n/a
Ahmad et al.[21]	19	900	165	55	-	-	1021	-	40	163	157	154.8	198
	20	810	220	90	-	-	1002	-	40	163	157	157.2	201
Al-Osta et al.[22]	21	900	220	-	-	-	1005	-	40	163	157	149.5	200
Ahmad et al.[23]	22	1000	250	-	-	-	818	-	44	188	157	137.0	200
	23	1000	250	-	-	-	784	-	25	219	157	135.0	200
Ahmad et al.[24]	24	900	220	-	-	-	1005	-	40	162	157	161.0	230
	25	900	88	132	-	-	1042	-	40	162	157	150.0	195
	26	900	88	-	132	-	1005	-	40	162	157	158.0	210
Wu et al.[25]	27	792	264	-	-	-	1056	-	21.1	173	156	158.8	n/a
Mousavinezhad et al.[2]	28	839	104	-	83	21	920	-	62.3	147	119	134.5	370
	29	839	104	-	62	42	921	-	65.3	147	119	132.5	370
Randl et al.[26]	30	401	124	-	-	328	833	211	30	170	155	139.4	285
	31	401	124	-	328	-	833	397	30	170	155	124.7	290
Wu et al.[27]	32	729	124	-	-	-	833	397	30	170	155	166.1	280
	33	863	216	-	-	-	923	-	92.79	194	156	153.6	240-250
Wu et al.[28]	34	472	262	-	-	315	1049	-	21	178	157	125.8	227
	35	792	264	-	-	-	1056	-	21.1	173	156	149.8	n/a
	36	475	264	-	-	317	1056	-	21.1	173	156	144	n/a
	37	317	264	-	475	-	1056	-	21.1	173	156	142.5	n/a

(Continued)

Table 1. *Continued*

Ref.	Mix No.	Quantity (kg/m^3) OPC	SF	NP	FA	GGBFS	Fine Agg.	QP	SP	Water	Steel fiber	Compressive strength (MPa)	Flow (mm)
Feng et al.[29]	38	833.3	208.3	-	-	-	1041.6	-	31.3	193.7	152.9	127.61	200
Tafraoui et al.[30]	39	691	172	-	-	-	759	276	22	187.3	138	155	n/a
Guo et al.[30]	40	565.1		-	-	780.4	780.4	-	40.5	225.1	117	134.2	280

Table 2. *NC mix designs*

Ref.	Class	**Quantity (kg/m^3)** OPC	SF	Fine aggregates	Gravel	SP	Water
Han et al.[31]	C40	400	0	700	1125	0.116	152
	C60	400	0	700	1125	0.166	120
	C80	600	67	550	800	0.266	173.42

data for each material. These documents were sourced from recognized platforms, including the International EPD System, EPD Hub, and manufacturer websites depending on the material type and regional availability[33]. Transport distances were treated as a key assumption in this analysis. For locally sourced materials, transport distances were assumed to be negligible. For globally sourced materials, only port-to-port maritime shipping was considered. Saudi Arabia was selected as the case study location because of its prominent position in the Middle East and ambitious Vision 2030 development agenda. The Kingdom is undertaking numerous giga-projects that demand high-performance concrete solutions, including UHPC, making it a strategically relevant context for environmental assessment.

The analysis included both local and global sources to represent real-world sourcing scenarios. Internationally sourced materials frequently offer greater consistency and purity, whereas locally sourced materials support regional supply chains and exhibit lower transportation-related emissions. Evaluating both options allows for a more thorough understanding of the environmental tradeoffs in the UHPC formulations. Table 3 lists the average embodied carbon and energy of the materials collected from the respective EPDs (A1–A3 scope). Figure 1 presents a comparative analysis of the average environmental impacts of OPC production from local and global sources. The results indicate that local OPC production in Saudi Arabia exhibits a higher global warming potential (GWP) (approximately 900 kg·CO$_2$/t) and total energy use (1600 MJ/t) than its global counterparts, which include factories from the USA, UK, and Europe. Although global factories also have considerable environmental loads, their mean GWP and energy use are lower, suggesting more efficient production technologies or cleaner energy inputs. This highlights the need for targeted improvements in regional cement manufacturing to align with the global sustainability benchmarks. Furthermore, this disparity in the impact of OPC production plays a critical role in the LCA of UHPC mixes, as OPC is a primary contributor to the overall embodied carbon and energy. This influence is discussed further in the subsequent sections of this study.

Table 3. *Average embodied carbon and energy of the materials collected from respective EPDs (A1–A3 scope)*

Material	GWP (Kg·CO$_2$/kg)	Total energy* (MJ/kg)
OPC	0.8983	5.272
SF	0.3435	5.773
NP	0.205	3.097
FA	0.0348	0.415
GGBFS	0.2045	2.549
Fine aggregate	0.00201	0.043
QP	0.0232	0.817
SP	0.8045	17.773
Steel fibers	2.275	27.308

*Total energy is the sum of renewable primary, renewable secondary, non-renewable primary, and non-renewable secondary.

2.4. Life Cycle Impact Assessment.

A typical LCIA includes a range of environmental indicators such as GWP, energy use, ozone depletion potential, acidification potential, and water deprivation potential. In this study, GWP and energy use were selected as the primary LCIA indicators and determined across the entire life cycle of the UHPC mixes, thereby

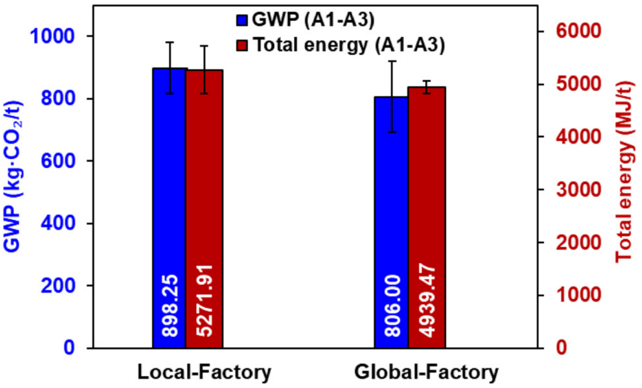
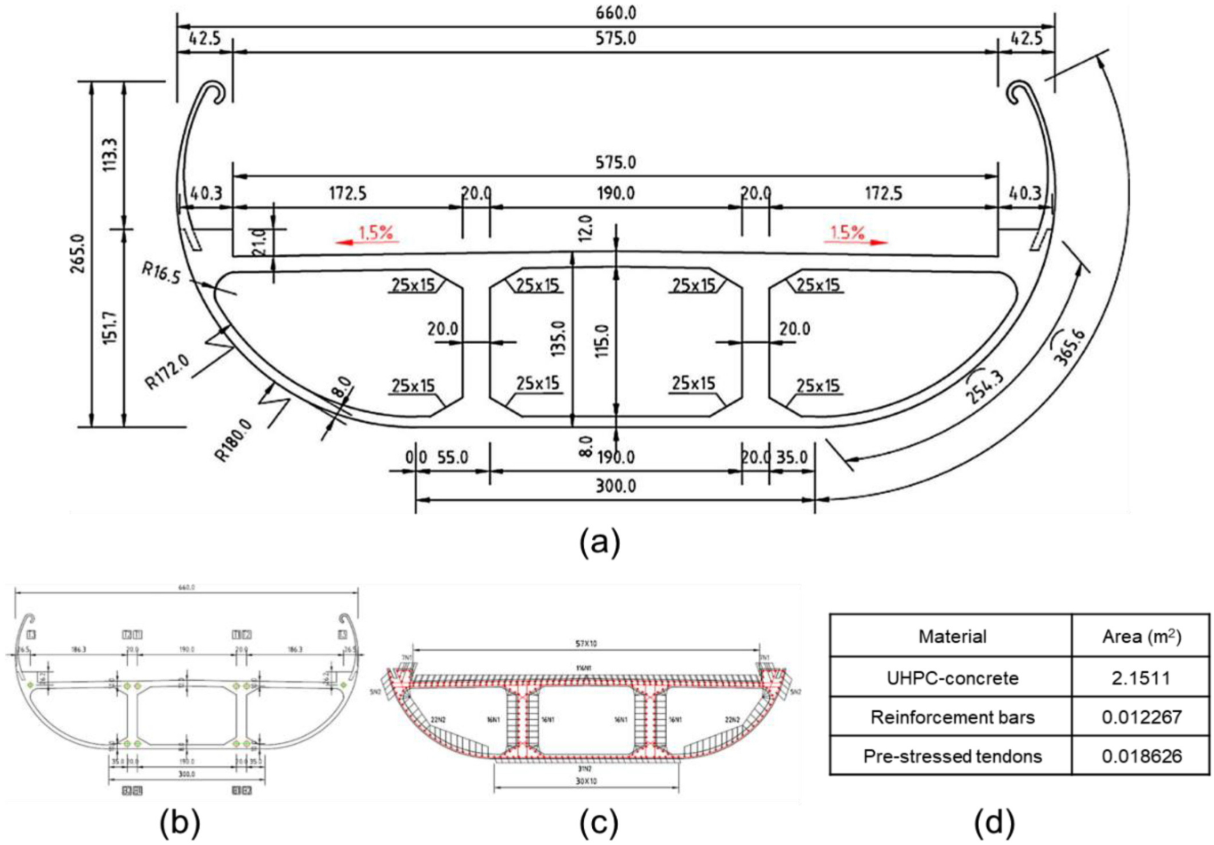


Figure 1. Comparison of average GWP and total energy consumption (A1–A3 scope) for OPC produced by local and global factories



Material	Area (m^2)
UHPC-concrete	2.1511
Reinforcement bars	0.012267
Pre-stressed tendons	0.018626

Figure 2. UHPC bridge: (a) Cross-section; (b) Configuration of pre-stressed tendons; (c) Details on the reinforcement bars; and (c) Cross-sectional areas of materials used in the bridge design (Adopted from Dong[34] with modification)

encompassing the manufacturing, transportation, and maintenance phases. A functional unit of 1 m^3 of concrete was adopted for all LCIA calculations. The results were subsequently compared to identify the concrete mix with the lowest environmental footprint based on the ESI.

The GWP and embodied energy of the manufacturing phase were determined by calculating the infrastructure's cumulative CO_2 emissions and energy use for each mix. The embodied carbon (kg·CO_2/m^3) and embodied energy (MJ/m^3) associated with concrete (EC_{manu} and EE_{manu}, respectively) were calculated as follows[34]:

$$EC_{manu} = \sum_{i=1}^{n} m_i \cdot cf_i \tag{1}$$

$$EE_{manu} = \sum_{i=1}^{n} m_i \cdot ef_i \tag{2}$$

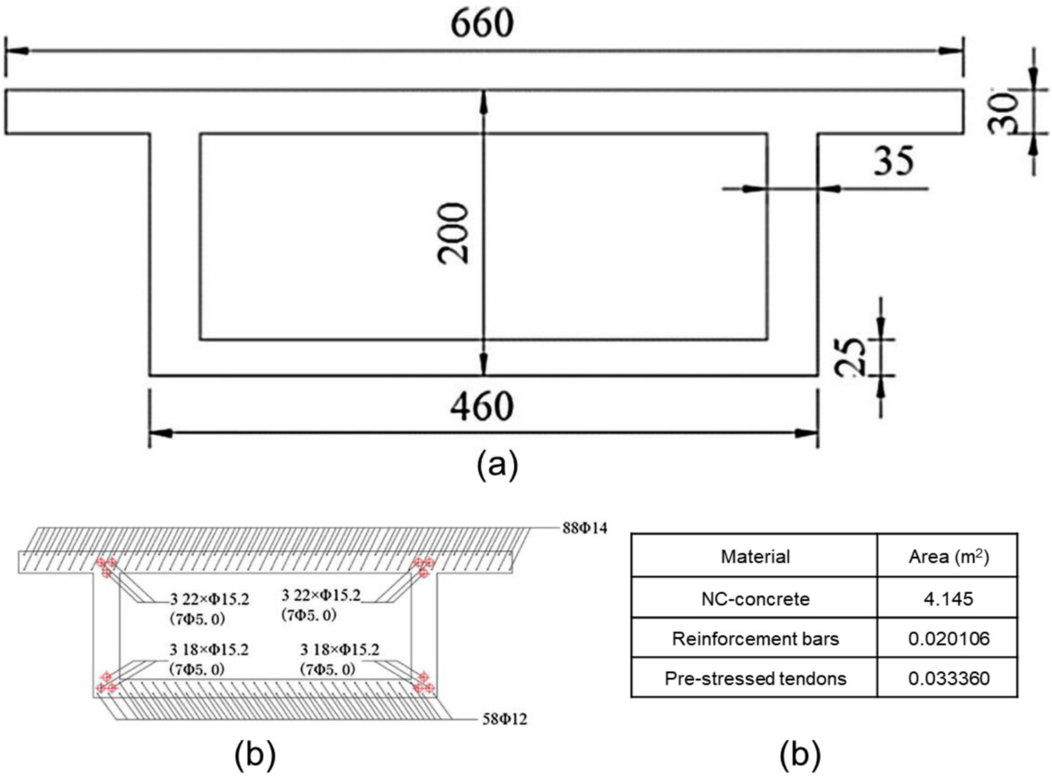
where m_i is the amount of constituent materials in the concrete mix (kg/m^3); cf_i and ef_i are the CO_2 emissions and energy use factors associated with the manufacturing of constituent material I, respectively (Table 3); and n is the number of constituent materials within the concrete. For a typical UHPC mix, the total CO_2 emissions and energy expenditure associated with the production phase of the concrete can be calculated as long as the CO_2 emissions and energy expenditure factors are known.

The CO_2 emissions of the transportation phase were determined based on the assumptions that all global materials are shipped by large container vessels from nearby ports and emissions within a country are negligible. The CO_2 emissions associated with transportation were calculated using the following formula:

$$EC_{trans} = \sum_{i=1}^{n} m_i \cdot d_i \cdot T \tag{3}$$

where m_i is the amount of constituent materials in the concrete mix (kg/m^3); d_i is the average distance from ports close to multiple manufacturing facilities for material i (km) to Dammam Port in Saudi Arabia; and T is the CO_2 emissions factor associated with very large container vessels, which is equivalent to 0.003 kg·CO_2/ton·km[35].

Over time, all buildings deteriorate and corrode and thus require maintenance to prevent their structural performance from deteriorating. This type of maintenance can be minor or major. Minor maintenance can halt or slow down a structure's degeneration, whereas major maintenance is performed when the structure's corrosion reaches a certain threshold at which performance is compromised and requires immediate attention[36]. For example, if a bridge's corrosion reaches the point at which the structural performance is affected, then essential maintenance is applied to reverse this process. The embodied carbon linked to



(a)

(b)

Material	Area (m^2)
NC-concrete	4.145
Reinforcement bars	0.020106
Pre-stressed tendons	0.033360

(b)

Figure 3. NC bridge: (a) Cross-section; (b) Details of pre-stressed tendons and reinforcement bars; and (c) Cross-sectional areas of materials used in bridge design (Adopted from Dong[34] with modification)

maintenance was expressed using the following formula[34]:

$$EC_{main} = \sum_{i=1}^{n} EC_{ini} \cdot r_{m,i} \tag{4}$$

$$EE_{main} = \sum_{i=1}^{n} EE_{ini} \cdot r_{m,i} \tag{5}$$

where EC_{ini} is the initial embodied carbon associated with manufacturing and transport and $r_{m,i}$ is the ratio between the embodied carbon of maintenance action i and the embodied carbon of EC_{ini}. In this study, the maintenance period was 80 years, which is considered the average life expectancy of an NC bridge[37]. The maintenance interval was set at 20 years for UHPC bridges and 10 years for NC bridges, based on previous studies[34]. The maintenance ratio $r_{m,i}$, is defined as 0.1 for minor maintenance and 0.5 for major maintenance activities[34]. Given the superior durability and longer service life of UHPC, these structures were assumed to require only minor maintenance throughout the analysis period. In contrast, NC bridges are expected to undergo major maintenance every fourth cycle[34,38], thus reflecting their higher vulnerability to deterioration over time.

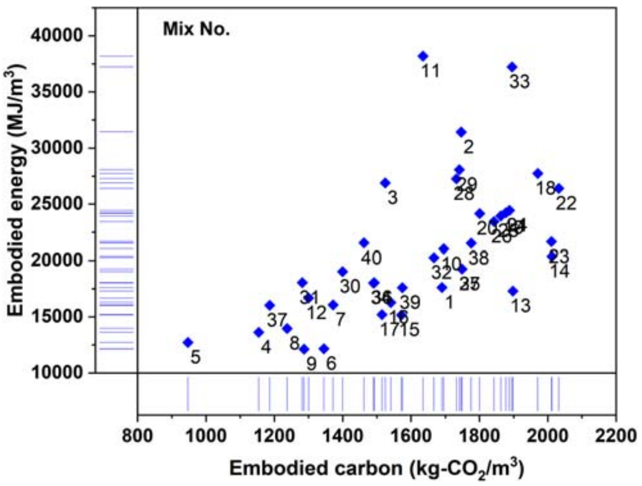


Figure 4. LCA plot of UHPC mixes (each point is labeled by the mix number)

2.5. Eco-Strength Index Calculation.

Another metric used in this study to compare the UHPC mixes was the ESI, which measures the total embodied carbon over the strength in MPa at 28 days. The ESI is calculated as follows[39]:

$$ESI = \frac{EC}{f'_c} \tag{6}$$

where EC is the total embodied carbon over the concrete's lifecycle and f'_c is the 28-day compressive strength of the concrete, expressed in MPa. This index offers a combined perspective on environmental and mechanical performance. A higher ESI indicates that a mix emits a larger amount of CO_2 per unit strength, suggesting lower environmental efficiency. Conversely, a lower ESI reflects a more sustainable mix that achieves higher strength with a smaller carbon footprint. When comparing concrete mixes, a lower ESI is preferred because it signifies the potential to construct high-strength environmentally efficient structures with reduced carbon emissions. In addition to the ESI, this study analyzed the relative contributions of embodied carbon, embodied

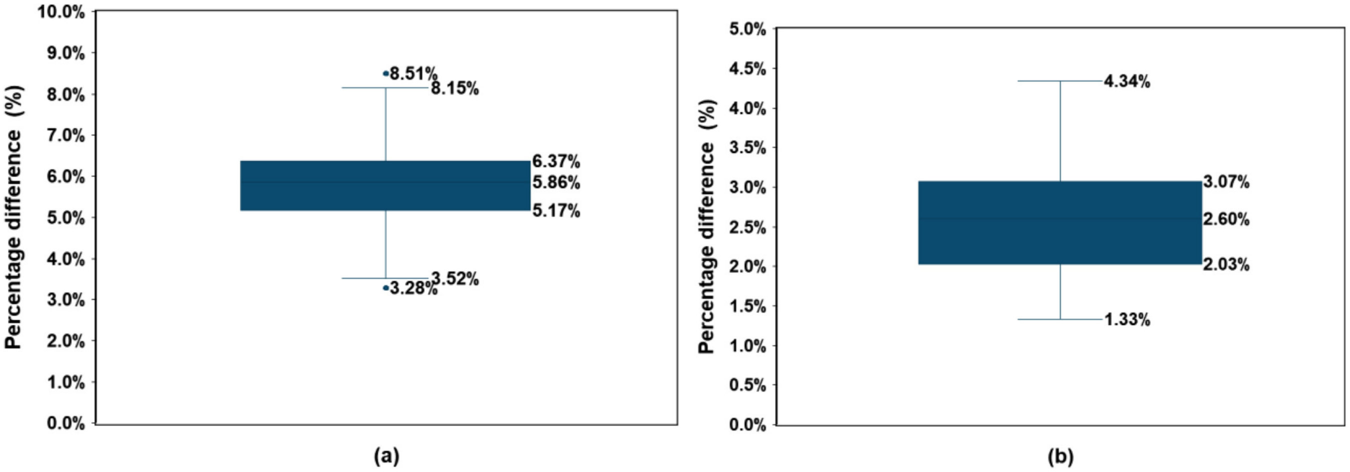


Figure 5. Percentage difference in the sustainability performance of UHPC mixes using locally sourced materials compared to global sources: (a) Embodied carbon and (b) Embodied energy

energy, and compressive strength to provide a comprehensive perspective on the sustainability of UHPC mixes.

2.6. Case Study: Ultra-High-Performance Concrete vs. Conventional Bridge.

The case study examined the UHPC and NC bridges reported by Dong[34]. The previously described LCA techniques were applied to both bridges to evaluate and compare their environmental performance. The UHPC bridge was designed as a single-girder structure with a cross-sectional width of 6.6 m (Fig. 2), and its prestressed tendons and reinforcement layout are detailed in Fig. 2(b–c). In contrast, the NC bridge features a different cross-sectional geometry and reinforcement configuration (Fig. 3). Owing to the lower mechanical performance of NC compared to UHPC, the NC bridge required a larger cross-sectional area to carry equivalent loads.

Using the previously defined equations, the embodied carbon and energy of both the UHPC and NC bridges can be calculated. The cross-sectional areas of the materials, including concrete, prestressed tendons, and reinforcement, in each bridge are shown in Fig. 2(d) and 3(b). These areas were multiplied by the corresponding unit LCA values for UHPC and NC to compute the total environmental impact per unit length of the bridge. This method enables a direct comparison of the two materials and facilitates the assessment of the environmental performance across any bridge length.

3. RESULTS AND DISCUSSION

3.1. Comparative Analysis of the Ultra-High-Performance Concrete Mixes.

The indices presented in Section 2 were calculated, and the results were analyzed to identify key trends. Figure 4 shows the outcomes of the embodied carbon and energy of all 40 UHPC mixes considered in this study. The marginal distributions along the axes represent the density of the mixes in terms of their respective embodied environmental burdens. A notable observation was the variation in

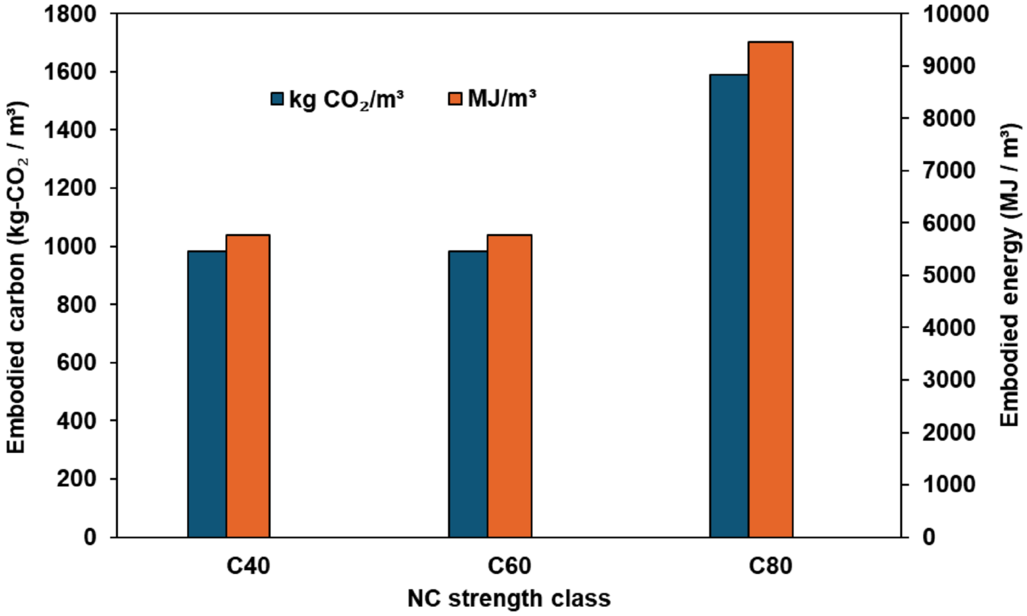


Figure 6. Total embodied carbon and energy for NC mixes at different strength classes

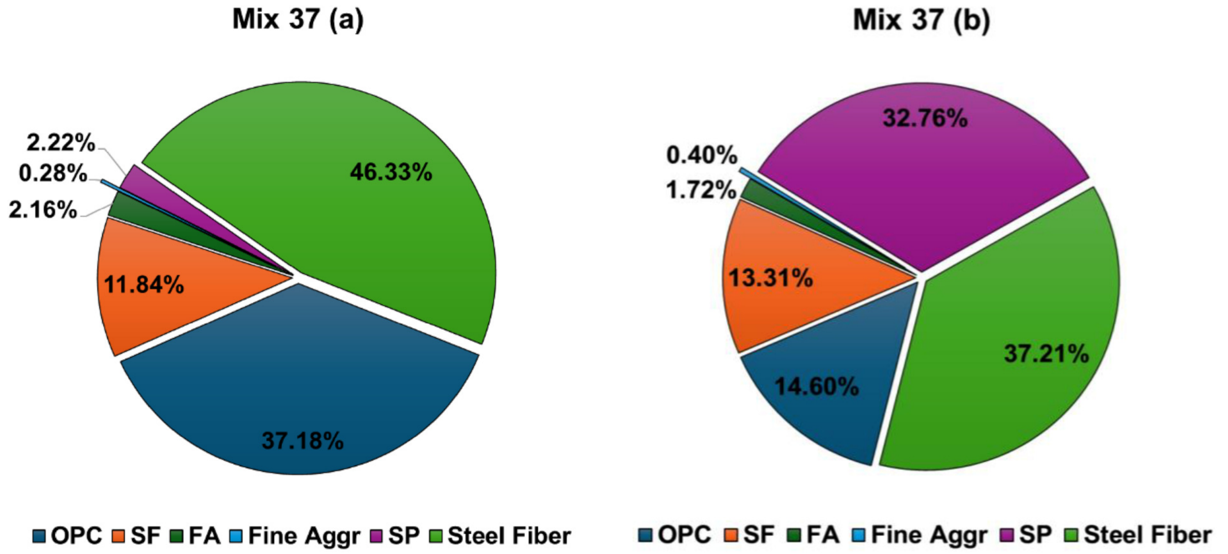


Figure 7. Proportional contribution of each material to the environmental impact of UHPC Mix-37: (a) Embodied carbon and (b) Embodied energy

embodied carbon among the UHPC mix designs, which ranged from 947.5 to 2032.2 kg CO2/m^3. Moreover, the embodied energy values ranged from 12,120.5 to 38,185.9 MJ/m^3. Mixes 4, 5, and 37 exhibited the lowest embodied carbon values, whereas Mixes 14, 22, and 23 exhibited the highest. Mixes 5, 6, and 9 (located in the lower-left quadrant) presented the lowest embodied energy values, whereas Mixes 2, 11, and 33 (located in the upper-right quadrant) presented the highest values.

The LCA conducted on the UHPC mixes revealed that OPC and steel fibers had the highest impact on embodied carbon. Mixes that contained high amounts of OPC and steel fibers exhibited the highest carbon emissions. Meanwhile, mixes produced with a lower OPC content but higher SCM content achieved significant results, even outperforming some mixes with higher embodied carbon.

A comparison of the local and global UHPC mixes showed that only OPC and NP were locally available; therefore, only these materials affected the LCA results. Although local fine aggregate was available, local EPDs were not found; consequently, global EPDs were used to estimate its embodied carbon and energy. Thus, both local and global fine aggregate were comparable. A comparative analysis of the local and global UHPC mixes indicated that local mixes exhibited 3.28–8.51% higher embodied carbon and 1.33–4.34% higher embodied energy throughout the LCA (Fig. 5). These differences in embodied carbon and energy are substantial but may not represent the best option when considering costs or material availability. As such, improving the carbon efficiency of OPC factories within the Kingdom is the best course of action for creating readily available carbon-efficient UHPC mixes.

To identify the most environmentally friendly concrete, the LCAs of both UHPC and NC were compared. The C40 and C60 grades of NC exhibited nearly identical LCA results, with values of approximately 981.0 kg·CO_2/m^3 for embodied carbon

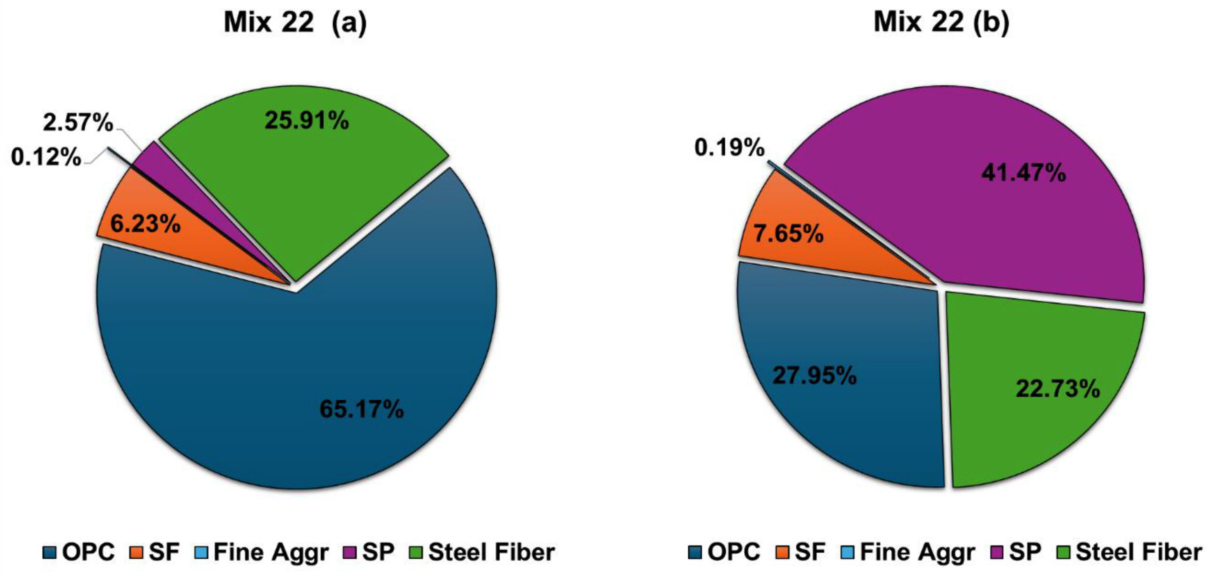


Figure 8. Proportional contribution of each material to the environmental impact of UHPC Mix-22: (a) Embodied carbon and (b) Embodied energy

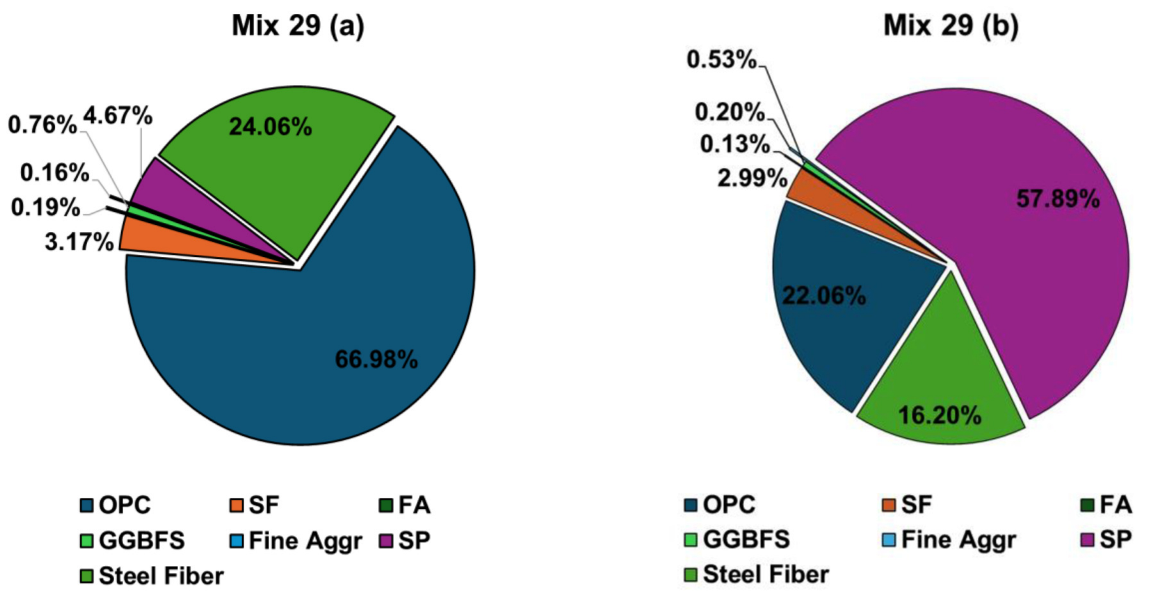


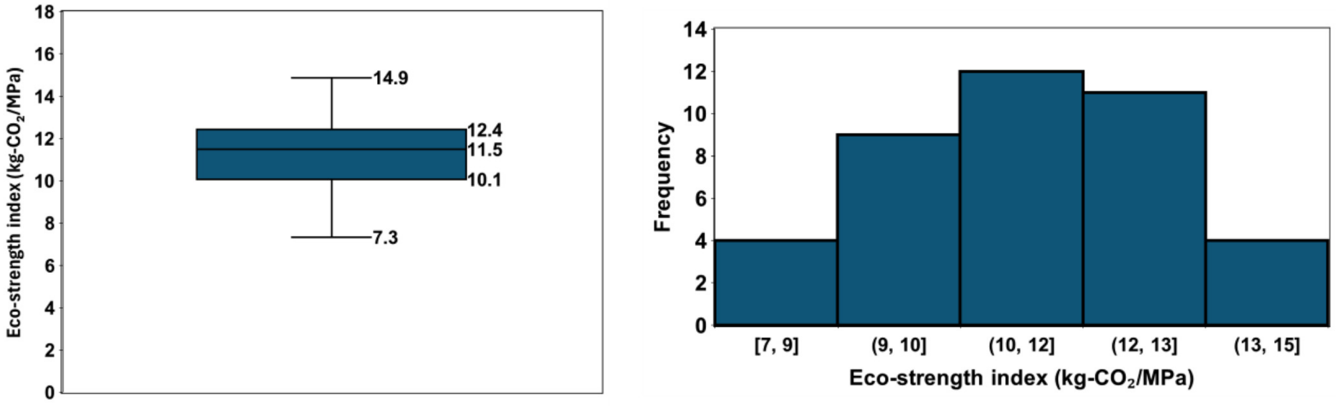
Figure 9. Proportional contribution of each material to the environmental impact of UHPC Mix-22: (a) Embodied carbon and (b) Embodied energy

and 5780 MJ/m^3 for embodied energy (Figure 6). In contrast, the C80 mix showed a significantly higher environmental impact, with 1590.9 kg·CO$_2$/m^3 and 9452.8 MJ/m^3. In terms of embodied energy, all three NC grades exhibited lower values than the UHPC mixes. In terms of embodied carbon, both C40 and C60 exhibited lower emissions than most of the UHPC mixes, with the exception of Mix 5. The C80 mix fell within the range of UHPC values, outperforming some UHPC mixes while being outperformed by others. Of note, this comparison was based on equal volumes of concrete. However, in practice, a UHPC structure requires less material than an equivalent NC structure owing to its superior strength. Therefore, these results provide only an initial comparison, while the case study offers a more accurate evaluation of the relative environmental efficiencies of UHPC and NC.

3.2. Key Contributors to Environmental Impact.

Identifying the materials that contribute most significantly to the embodied carbon and energy of the UHPC mixtures is critical for improving environmental performance without compromising strength. The targeted reduction of high-impact components and use of alternative materials are essential steps toward achieving more sustainable UHPC mix designs. A major contributor to the material's environmental impact is the production stage within the A1–A3 scope.

Among the UHPC components, OPC exhibited the highest carbon emissions at 0.8983 kg·CO$_2$/kg. Owing to the large demand for OPC and its substantial carbon footprint, OPC accounts for 5–7% of global CO$_2$ emissions[40]. The production of OPC is energy-intensive and requires kiln temperatures exceeding 1300 °C, typically achieved by burning fossil fuels[41]. In contrast, SCMs, including SF, NP, FA, and GGBFS, exhibit considerably lower embodied carbon, making them suitable substitutes for OPC. In contrast, filler materials such as fine aggregate and QP have very low emissions and thus do not require substitution. Superplasticizers have relatively high embodied carbon.



(a) (b)

Figure 10. Analysis of the eco-strength index values of the UHPC mixes: (a) Median of the index values and (b) Histogram representing the frequency distribution of the index values

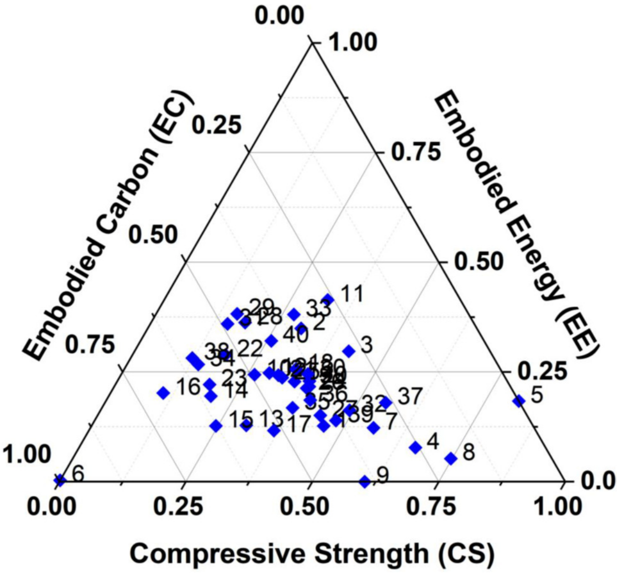


Figure 11. Ternary plot showing the normalized contributions of embodied carbon, embodied energy, and compressive strength for the UHPC mixes

However, given the small dosage required for UHPC and the limited availability of alternatives, the replacement of superplasticizer is not a priority. Steel fibers are the most environmentally intensive components in terms of both embodied carbon and energy because their manufacture requires high thermal energy to form and shape the material, typically through fossil fuel combustion[42].

Based on these data, the most effective strategy for reducing the environmental impact of UHPC is to minimize the use of OPC and steel fibers. This aligns with current UHPC mix design research trends that focus on substituting OPC with low-carbon alternatives. A representative example is Mix 37, which uses 317 kg/m^3 of OPC, 264 kg/m^3 of SF, and 475 kg/m^3 of FA, yielding an OPC-to-SCM ratio of approximately 3:7 (Table 1). Despite its high SCM content, Mix 37 achieved a compressive strength of 142.5 MPa. This mix achieved one of the lowest proportions of CO_2 and energy attributed to OPC among all the collected mixes, making it a strong candidate for eco-efficient design (Figure 7). Although the SCMs investigated in this study are limited, other viable options include limestone powder, rice husk ash, metakaolin, and recycled glass powder[9,30,43–52]. Similarly, steel fibers can be replaced with more sustainable alternatives, such recycled tire-based materials, to further reduce the environmental impact[4,7,53]. With the environmental contributions of the individual materials identified, the next step was to assess which components most strongly influenced the overall footprint of the UHPC mixes. Figure 8 illustrates a high-impact UHPC mix (Mix 22), where OPC contributes the most to embodied carbon, followed by steel fibers. In terms of embodied energy, SP was the leading contributor, followed by OPC and steel fibers. Figure 9 presents a different UHPC mixture for comparison. The figures reveal that OPC and steel fibers are the dominant contributors to embodied carbon while SP, OPC, and steel fibers are the dominant contributors to embodied energy. These insights further support the strategy of material substitution and optimization in UHPC mix design.

3.3. Sustainability Evaluation of the Ultra-High-Performance Concrete Mixes. This section compares the UHPC mixes based on their ESI to understand their overall performance in terms of environmental and engineering aspects. As a lower value indicates better eco-efficiency, the mixes can be ranked based on their ESI to identify the mixes that offer the best balance between strength and environmental impacts. Figure 10 presents an analysis of the calculated ESI values, which range from 7.3 to 14.9 kg·CO_2/MPa, with a median of 11.5 and an interquartile range between 10.1 and 12.4. The histogram shown in Fig. 10(b) further reveals that most mixes fall within the range of 10–13 kg·CO_2/MPa. Only a few mixes exhibited considerably higher or lower index values, which reflected exceptional or poor eco-efficiency. These insights can guide the identification of candidate mixes for optimization or substitution strategies based on their environmental-to-strength performance ratios.

Of the 40 mixes, Mix 8 had the lowest index value of 7.3, making it the most eco-efficient, and it was followed by Mixes 6 (7.6), 4 (8.0) and 37 (8.3). Although these latter mixes were not necessarily the weakest or strongest in performance, they had a low environmental impact relative to their strength. Thus, they can be considered ideal options for sustainable construction, where reducing carbon emissions is a priority. In contrast, several mixes displayed high ESI values, indicating a lower eco-efficiency. For example, Mix 23 and Mix 24 have index values of 14.9 and 14.7, respectively, which are the highest in the dataset. Despite their moderate strengths (135 MPa and 141 MPa), the emissions associated with each strength unit were relatively high. Mixes 13 (13.3) and 28 (12.9) also fell into this category.

Notably, some mixes with high compressive strengths were also relatively eco-efficient, whereas others were not. For example, Mix 3 had one of the highest compressive strengths (159.3 MPa) and a good ESI (9.6), indicating that it delivers substantial performance with reasonable environmental efficiency. In contrast, Mix 2 had a slightly lower strength (151.3 MPa) but a worse ESI (11.5), indicating that it was less efficient in terms of emissions per unit of strength. Meanwhile, Mix 6 and Mix 8 had lower compressive strengths (125.3 MPa and 120.6 MPa, respectively) and the lowest ESI values (7.6 and 7.3, respectively). Therefore, these mixes represent the most sustainable option overall, although they would not necessarily be ideal when strength is also required. These comparisons revealed that the most sustainable UHPC mixtures were not always the strongest. Instead, the best performers achieved a balance, offering reasonable compressive strength while minimizing environmental impact.

To enable a comprehensive assessment of the sustainability performance of UHPC mixes, this study proposes a sustainable performance index (SPI), which serves as a tool to classify and compare different formulations. A ternary plot was produced using the normalized contributions of embodied carbon, embodied energy, and compressive strength for the 40 mixes analyzed (Fig. 11). Each data point in the plot represents a mix, with its position reflecting the relative balance of environmental impact and mechanical performance, which was constrained such that the sum of the three variables equals unity. This visualization enables a simultaneous comparison of trade-offs and supports the classification of mixes into three distinct zones:

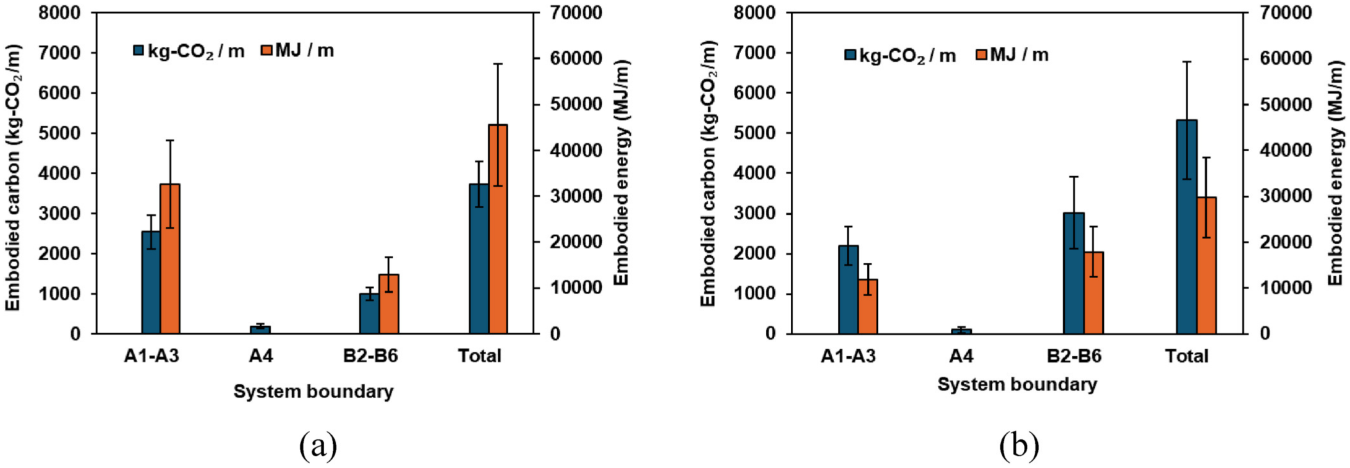


Figure 12. LCA results in terms of embodied carbon and embodied energy across different stages for the two bridge types per meter of span: (a) UHPC and (b) NC

Zone-I: optimized mixes (characterized by high strength and low environmental impact); Zone-II: balanced trade-offs (including mixes with either high strength and high environmental cost or moderate performance across all metrics); and Zone-III: environmentally intensive mixes (marked by low strength and/or high environmental burden).

The distribution of data points shows a concentration of mixes near the center of the ternary diagram, suggesting that many formulations achieved a balance among strength, carbon, and energy metrics. Exceptions were observed in the lower-impact region of the ternary plot, where Mixes 4, 5, and 8 demonstrate moderate to high compressive strength combined with comparatively low embodied carbon and energy. These mixes reflect a favorable trade-off, thus offering environmentally efficient performance without a significant compromise in strength. As such, they align with the criteria for Zone-I, which reflects balanced sustainability, and illustrate how well-optimized UHPC formulations can achieve both reduced environmental impact and adequate structural capacity. In contrast, Mix 6 was situated closer to the embodied carbon vertex, indicating a high carbon footprint relative to its strength. This mix also recorded the lowest compressive strength among all the formulations, which indicates inefficiencies in material utilization and reinforces its classification within Zone-III. Among the most promising formulations was Mix 37, which was located toward the compressive strength-dominant region but also maintained moderate embodied carbon and embodied emission values. Although this mix contained the lowest OPC content (317 kg/m^3), it achieved impressive compressive strength (142.5 MPa). This demonstrates the potential of the optimized binder systems to deliver environmentally efficient and structurally robust UHPC. Accordingly, Mix 37 is representative of the Zone-I category and exemplifies a well-balanced and sustainable concrete mix.

3.4. Case Study Insights: Ultra-High-Performance Concrete vs. Normal Concrete Bridge. The case study compared the environmental impacts of UHPC and NC bridges across different life cycle stages. The UHPC bridge demonstrated an average impact of 2541.5 kg·CO$_2$/m and 32,615.3 MJ/m for the A1–A3 phase (material production), 186.3 kg·CO$_2$/m for

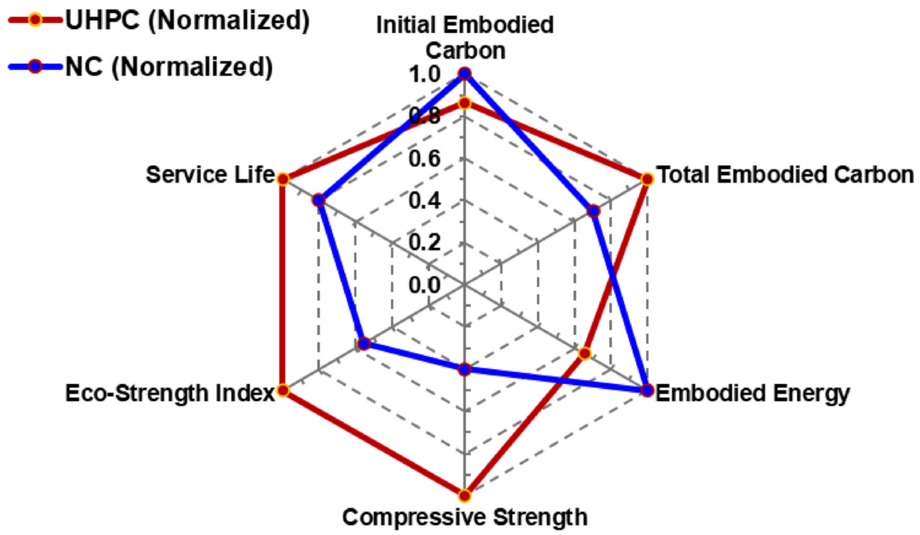


Figure 13. Performance-based comparison of UHPC and NC across six sustainability and structural metrics. Values are normalized relative to the best-performing value in each category

Table 4. *Performance-based indicators for UHPC vs. NC bridges*

Indicator	UHPC bridge	NC bridge	Interpretation
Initial embodied carbon (A1–A3)	High (~2,500 kg·CO_2/m)	Lower (~2,200 kg·CO_2/m)	NC has lower production emissions
Total embodied carbon (A1–B6)	Lower (~3,724 kg·CO_2/m)	Higher (~5,320 kg·CO_2/m)	UHPC is better over full life cycle
Total embodied energy	Higher (~45,500 MJ/m)	Lower (~30000 MJ/m)	UHPC more energy-intensive upfront
Compressive strength	High (120–170 MPa)	Moderate (40–80 MPa)	UHPC structurally superior
Eco-strength index	Low-favorable (11.3 kg·CO_2/MPa)	Moderate (20.3 kg·CO_2/MPa)	UHPC is more carbon-efficient per MPa
Service life	Excellent (100+ years)	Moderate (80 years max)	UHPC expected to last longer

A4 (transportation), and 996.3 kg·CO_2/m and 3,794.7 MJ/m for B2–B6 (use and maintenance) (Fig. 12). This resulted in a total life cycle impact of 3,724.1 kg·CO_2/m and 45,528.1 MJ/m. In contrast, the NC bridge exhibited lower initial values during production, with 2,197.8 kg·CO_2/m and 11,914.1 MJ/m for A1–A3 and 100.6 kg·CO_2/m for A4. However, due to more frequent maintenance requirements, it accumulated 3,021.3 kg·CO_2/m and 17,861.4 MJ/m during B2–B6, leading to a total impact of 5,319.8 kg·CO_2/m and 29,776.1 MJ/m. Although UHPC had a higher environmental impact during the initial stages (A1–A4), it outperformed NC over the entire life cycle, as shown in Fig. 12.

Based on this analysis, the UHPC bridge is considered more environmentally friendly than its NC counterpart for several key reasons. First, UHPC requires considerably less maintenance over its lifespan, which translates to lower long-term environmental impacts associated with its repair activities. Second, owing to its superior strength, UHPC enables a reduced structural cross-section, which means that less material is needed for the UHPC bridge compared to the NC bridge to achieve an equivalent load-bearing capacity. Third, UHPC generally has a longer service life. Although this LCA assumed an 80-year lifespan, which is aligned with the average for NC bridges, UHPC structures typically exceed this duration. Consequently, NC bridges may require reconstruction or extensive strengthening over extended periods, further increasing their cumulative environmental footprint.

To facilitate a comparative evaluation of UHPC and NC across multiple sustainability and performance metrics, a multicriteria analysis was included in the sustainability assessment. Different aspects were considered, including the initial and total embodied carbon, total embodied energy, compressive strength, ESI, and service life. Table 4 lists the performance-based indicators for the UHPC and NC bridges. Based on this, the relative performance ratio (RPR) method was employed to normalize each metric to a unitless scale based on its best-performing value. The direction of preference was first determined for each metric, namely, whether lower is better (e.g., embodied carbon, embodied energy, and ESI) or higher is better (e.g., compressive strength and service life). For "lower is better" indicators, the RPR was calculated as the ratio of the best (lowest) value to the current value. For "higher is better" indicators, the RPR was calculated as the ratio of the current value to the best (highest) value. An RPR of 1.0 represented the best performance for a given metric, whereas values closer to zero indicated lower relative performance. To ensure a fair comparison, the RPR calculations were based on the average values of each metric for UHPC and NC. For example, the metric for compressive strength was calculated using average values of 150 MPa for UHPC and 60 MPa for NC, as reported in Section 2.1. Thus, the RPR for compressive strength was calculated as follows: $RPR_{NC} = 60/150 = 0.40$ and $RPR_{UHPC} = 150/150 = 1.0$.

Figure 13 illustrates the performance-based comparison of UHPC and NC across six sustainability and structural metrics. The data indicated that NC exhibited superior performance in terms of initial embodied carbon (1.0) and embodied energy (1.0), reflecting its lower material intensity and simpler mix composition. However, its performance declined significantly in terms of structural and service life-related criteria. For instance, its compressive strength (0.4), ESI (0.6), and service life (0.8) indicated substantial trade-offs when strength and longevity are prioritized. In contrast, UHPC performed strongly in all structural and durability metrics, achieving perfect scores (1.0) in compressive strength, ESI, service life, and total embodied carbon. These results highlight the efficiency of UHPC in delivering high mechanical performance with lower long-term environmental burden. While its initial embodied carbon and energy use were relatively higher (normalized scores of 0.9 and 0.7, respectively), they were offset by gains in lifecycle performance.

4. CONCLUSIONS AND RECOMMENDATIONS

This study presents a comprehensive sustainability evaluation of 40 UHPC mixes by integrating embodied carbon, embodied energy, and compressive strength within a multicriteria assessment framework. Based on the analysis and discussion, the key findings, applications, and recommendations are further highlighted.

4.1. Key Findings.

By analyzing the UHPC mix designs, a deeper understanding of their environmental and structural performance, such as embodied carbon and energy use, is provided. UHPC mixes generally exhibit higher embodied emissions than NC due to the intensive use of materials such as cement and steel fibers. However, their superior strength enables material efficiency at the structural level, thereby reducing the total life cycle impact.

The ESI (range of 7.3–14.9 kg·CO_2/MPa) effectively captured the balance between carbon emissions and mechanical performance; thus, it represents a practical tool for evaluating the sustainability of UHPC formulations. In addition, the proposed sustainable performance index facilitates classification into optimized, balanced, and environmentally intensive zones by integrating embodied carbon, embodied energy, and compressive strength. Mixes 3, 4, 5, 7, 8, 9, and 37 demonstrated strong sustainability profiles, thus combining high strength with low environmental burdens through the strategic use of low-carbon materials.

A comparative analysis of bridge applications shows that although UHPC had higher initial embodied carbon (2,500 kg·CO_2/m) and energy (45,500 MJ/m) than NC, its total life cycle carbon was lower (3,724 vs. 5,320 kg·CO_2/m) due to its reduced maintenance and longer service life (>100 years vs. 80 years). It also achieved a more favorable ESI (11.3 vs. 20.3 kg·CO_2/MPa), which confirmed its efficiency in delivering structural performance with a reduced carbon footprint.

In conclusion, the optimization of UHPC through mix design and material sourcing produces a highly sustainable material for long-lasting, high-performance infrastructure. Its use is especially justified when durability, strength, and long-term environmental performance are critical, whereas NC may remain suitable for applications with lower structural demands and shorter service lives.

4.2. Implementation and Future Recommendations.

To support the wider adoption of UHPC in sustainable infrastructure, a strategic shift in material design, sourcing, and assessment methodologies is essential. Despite its high initial environmental footprint, UHPC's superior mechanical performance and extended service life position it as a promising material when optimally designed and implemented.

One of the most pressing challenges is the carbon and energy intensity of locally produced OPC, which are higher than those of its international counterparts. This underscores the urgency of regional process improvements, such as cleaner production technologies, clinker substitution, and alternative fuels. These efforts would not only reduce emissions but also create a clearer understanding of how local environmental conditions, production practices, and material compositions influence overall sustainability performance.

To enhance the sustainability profile of UHPC, mix designs should increasingly incorporate low-carbon materials that can effectively lower the embodied carbon and energy without compromising the mechanical strength. Equally important is the need to address the environmental burden of other key components, particularly steel fibers and chemical admixtures. Thus, future research should explore low-carbon alternatives, such as industrial waste/byproducts and recycled steel fibers, to further reduce the environmental intensity of UHPC mixes.

From a performance standpoint, tools such as the ESI and SPI framework proposed in this study offer practical means of evaluating and selecting UHPC mixes that balance structural and environmental priorities. Their adoption, alongside region-specific environmental databases and verified EPDs, will support more localized and informed sustainability assessments, thereby leading to better decision-making in future infrastructure development.

Future research should also incorporate life cycle cost analyses along with environmental assessments to support more robust decision-making. Simultaneously, evaluating the cost and environmental trade-offs will enable the identification of UHPC formulations that not only offer lower emissions but also deliver long-term economic value. This integrated approach can guide both the broader adoption of UHPC and the selection of specific blends with the most favorable life cycle performance, both environmentally and economically.

It is worth mentioning that this study has certain limitations that may influence the interpretation of its LCA results. The system boundary excludes domestic transportation within Saudi Arabia and on-site construction emissions, potentially underestimating the total environmental burden. Additionally, a fixed maintenance schedule and a minimum strength threshold of 120 MPa were used for all UHPC mixes, which may not fully capture real-world variability in durability and performance.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Hussain Abdulatif Karam – *Department of Civil & Environmental Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0009-0002-8185-4843
Email: s202352030@kfupm.edu.sa

Corresponding Author

Ashraf A. Bahraq – *Research Mentor, Post Doctoral Fellow, IRC for Construction and Building Materials, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0000-0002-0583-5002
Email: ashraf.bahraq@kfupm.edu.sa

Authors

Mohammed A. Al-Osta – *Professor, Department of Civil & Environmental Engineering and IRC for Construction and Building Materials, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0000-0002-3678-2695
Email: malosta@kfupm.edu.sa

Shamsad Ahmad – *Professor, Department of Civil & Environmental Engineering and IRC for Construction and Building Materials, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0000-0001-7939-7319
Email: shamsad@kfupm.edu.sa

ACKNOWLEDGEMENTS

The authors would like to thank the Undergraduate Research Office (URO) and the Interdisciplinary Research Center for Construction & Building Materials at King Fahd University of Petroleum & Minerals (KFUPM), Saudi Arabia.

CONFLICTS OF INTEREST AND FINANCIAL DISCLOSURE

The authors declare no conflicts of interest and no financial disclosures.

REFERENCES

(1) Li, Y. *et al.* A comparative study on mechanical properties and environmental impact of UHPC with belite cement and portland cement. *J Clean Prod* **380**, (2022).

(2) Mousavinezhad, S. *et al.* A Comprehensive Study on Non-Proprietary Ultra-High-Performance Concrete Containing Supplementary Cementitious Materials. *Materials* **16**, (2023).

(3) Guo, P., Meng, W., Du, J., Han, B. & Bao, Y. Lightweight ultra-high-performance concrete (UHPC) with expanded glass aggregate: Development, characterization, and life-cycle assessment. *Constr Build Mater* **371**, 130441 (2023).

(4) Zhuo, K. X. *et al.* Axial compressive behavior of environmentally friendly high-strength concrete: Effects of recycled tire steel fiber and rubber powder. *Journal of Building Engineering* **76**, 107092 (2023).

(5) Hasnat, A. & Ghafoori, N. Properties of ultra-high performance concrete using optimization of traditional aggregates and pozzolans. *Constr Build Mater* **299**, 123907 (2021).

(6) Sameer, H. *et al.* Environmental assessment of ultra-high-performance concrete using carbon, material, and water footprint. *Materials* **12**, (2019).

(7) Farahzadi, L., Tellnes, L. G. F., Shafei, B. & Kioumarsi, M. Life-cycle environmental assessment of ultra-high-performance concrete with sustainable materials and fiber substitutions. *Clean Eng Technol* **23**, 100846 (2024).

(8) ASTM International. ASTM C1856: Standard Practice for Fabricating and Testing Specimens of Ultra-High Performance Concrete. 1–4 Preprint at https://doi.org/10.1520/C1856_C1856M-17 (2017).

(9) Ha, N. S. *et al.* Effect of grounded blast furnace slag and rice husk ash on performance of ultra-high-performance concrete (UHPC) subjected to impact loading. *Constr Build Mater* **329**, (2022).

(10) Regalla, S. S. & Kumar, N. S. Optimizing ultra high-performance concrete (UHPC) mix proportions using technique for order preference by similarity to ideal solution (TOPSIS) for enhanced performance: Precision towards optimal concrete engineering. *Case Studies in Construction Materials* **22**, (2025).

(11) Huang, H., Ren, X., Liu, Z. & Wang, F. Development of low-carbon and cost-effective ultra-high performance concrete using carbonated recycled fine aggregate. *Constr Build Mater* **399**, (2023).

(12) Du, J. *et al.* Utilization of off-specification fly ash in preparing ultra-high-performance concrete (UHPC): Mixture design, characterization, and life-cycle assessment. *Resour Conserv Recycl* **180**, (2022).

(13) Abdolpour, H. *et al.* Performance and life cycle of ultra-high performance concrete mixes containing oil refinery waste catalyst and steel fibre recovered from scrap tyre. *Journal of Building Engineering* **79**, (2023).

(14) Nassar, R.-U.-D., Zaid, O. & Elhadi, K. M. Utilizing carbonated recycled concrete fines to develop sustainable ultra-high-performance fiber-reinforced concrete. *Journal of Building Engineering* **99**, (2025).

(15) Shafieifar, M., Farzad, M. & Azizinamini, A. Experimental and numerical study on mechanical properties of Ultra High Performance Concrete (UHPC). *Constr Build Mater* **156**, 402–411 (2017).

(16) Alsalman, A., Dang, C. N., Martí-Vargas, J. R. & Micah Hale, W. Mixture-proportioning of economical UHPC mixtures. *Journal of Building Engineering* **27**, (2020).

(17) Dingqiang, F. *et al.* A novel approach for developing a green Ultra-High Performance Concrete (UHPC) with advanced particles packing meso-structure. *Constr Build Mater* **265**, 120339 (2020).

(18) Hasnat, A. & Ghafoori, N. Properties of ultra-high performance concrete using optimization of traditional aggregates and pozzolans. *Constr Build Mater* **299**, (2021).

(19) Shi, Y., Long, G., Ma, C., Xie, Y. & He, J. Design and preparation of ultra-high performance concrete with low environmental impact. *J Clean Prod* **214**, 633–643 (2019).

(20) Li, S., Han, G., Du, B., Jiang, Z. & Sun, L. Carbon emissions assessment of ultra-high performance concrete in construction industry: Calculation method and case study. *Energy Build* **329**, (2025).

(21) Ahmad, S., Mohaisen, K. O., Adekunle, S. K., Al-Dulaijan, S. U. & Maslehuddin, M. Influence of admixing natural pozzolan as partial replacement of cement and microsilica in UHPC mixtures. *Constr Build Mater* **198**, 437–444 (2019).

(22) Al-Osta, M. A. *et al.* Effect of hybridization of straight and hooked steel fibers and curing methods on the key mechanical properties of UHPC. *Journal of Materials Research and Technology* **15**, 3222–3239 (2021).

(23) Ahmad, S., Al-Fakih, A., Bahraq, A. A. & Yusuf, M. O. Fracture toughness of UHPC mixtures: Effects of w/b ratio, cement and silica fume contents. *Constr Build Mater* **417**, 135327 (2024).

(24) Ahmad, S., Hakeem, I. & Maslehuddin, M. Development of UHPC mixtures utilizing natural and industrial waste materials as partial replacements of silica fume and sand. *Scientific World Journal* **2014**, (2014).

(25) Wu, Z., Shi, C., He, W. & Wu, L. Effects of steel fiber content and shape on mechanical properties of ultra high performance concrete. *Constr Build Mater* **103**, 8–14 (2016).

(26) Randl, N., Steiner, T., Ofner, S., Baumgartner, E. & Mészöly, T. Development of UHPC mixtures from an ecological point of view. *Constr Build Mater* **67**, 373–378 (2014).

(27) Wu, Z., Shi, C. & Khayat, K. H. Investigation of mechanical properties and shrinkage of ultra-high performance concrete: Influence of steel fiber content and shape. *Compos B Eng* **174**, 107021 (2019).

(28) Wu, Z., Shi, C., He, W. & Wang, D. Static and dynamic compressive properties of ultra-high performance concrete (UHPC) with hybrid steel fiber reinforcements. *Cem Concr Compos* **79**, 148–157 (2017).

(29) Feng, S., Xiao, H. & Li, H. Comparative studies of the effect of ultrahigh-performance concrete and normal concrete as repair materials on interfacial bond properties and microstructure. *Eng Struct* **222**, 111122 (2020).

(30) Tafraoui, A., Escadeillas, G. & Vidal, T. Durability of the Ultra High Performances Concrete containing metakaolin. *Constr Build Mater* **112**, 980–987 (2016).

(31) Han, X., Wang, B. & Feng, J. Relationship between fractal feature and compressive strength of concrete based on MIP. *Constr Build Mater* **322**, 126504 (2022).

(32) ISO. *Environmental Management - Life Cycle Assessment, Principles and Framework: 14040.* (International Organization for Standardization, Geneva, 2006).

(33) EPD International. International EPD System (IES). https://www.environdec.com/ (2025).

(34) Dong, Y. Performance assessment and design of ultra-high performance concrete (UHPC) structures incorporating life-cycle cost and environmental impacts. *Constr Build Mater* **167**, 414–425 (2018).

(35) International Chamber of Shipping. *Shipping, World Trade and the Reduction of CO2 Emissions.* (2014).

(36) Kong, J. S. & Frangopol, D. M. Life-cycle reliability-based maintenance cost optimization of deteriorating structures with emphasis on bridges. *Journal of Structural Engineering* **129**, 818–828 (2003).

(37) Srikanth, I. & Arockiasamy, M. Deterioration models for prediction of remaining useful life of timber and concrete bridges: A review. *Journal of Traffic and Transportation Engineering (English Edition)* **7**, 152–173 (2020).

(38) Müller, H. S., Haist, M. & Vogel, M. Assessment of the sustainability potential of concrete and concrete structures considering their

environmental impact, performance and lifetime. *Constr Build Mater* **67**, 321–337 (2014).

(39) Momshad, A. M. *et al.* Assessing the engineering properties and environmental impact with explainable machine learning analysis of sustainable concrete utilizing waste banana leaf ash as a partial cement replacement. *Clean Eng Technol* **24**, (2025).

(40) He, Z., Zhu, X., Wang, J., Mu, M. & Wang, Y. Comparison of CO2 emissions from OPC and recycled cement production. *Constr Build Mater* **211**, 965–973 (2019).

(41) Lee, J., Lee, T., Jeong, J. & Jeong, J. Mix design optimization and environmental impact assessment of low-carbon materials containing alkali-activated slag and high CaO fly ash. *Constr Build Mater* **267**, 120932 (2021).

(42) Stengel, T. & Schießl, P. *Life Cycle Assessment (LCA) of Ultra High Performance Concrete (UHPC) Structures. Eco-Efficient Construction and Building Materials: Life Cycle Assessment (LCA), Eco-Labelling and Case Studies* (Woodhead Publishing Limited, 2013). doi:10.1533/9780857097729.3.528.

(43) Amin, M., Agwa, I. S., Mashaan, N., Mahmood, S. & Abd-Elrahman, M. H. Investigation of the Physical Mechanical Properties and Durability of Sustainable Ultra-High Performance Concrete with Recycled Waste Glass. *Sustainability (Switzerland)* **15**, (2023).

(44) Khan, M. I., Abbas, Y. M., Abellan-Garcia, J. & Castro-Cabeza, A. Eco-efficient ultra-high-performance concrete formulation utilizing electric arc furnace slag and recycled glass powder–advanced analytics and lifecycle perspectives. *Journal of Materials Research and Technology* **32**, 362–377 (2024).

(45) Mehta, A. & Ashish, D. K. Silica fume and waste glass in cement concrete production: A review. *Journal of Building Engineering* **29**, 100888 (2020).

(46) Li, P. P., Brouwers, H. J. H., Chen, W. & Yu, Q. Optimization and characterization of high-volume limestone powder in sustainable ultra-high performance concrete. *Constr Build Mater* **242**, (2020).

(47) Zhuang, W., Li, S., Wang, Z., Zhang, Z. & Yu, Q. Impact of micromechanics on dynamic compressive behavior of ultra-high performance concrete containing limestone powder. *Compos B Eng* **243**, (2022).

(48) Mo, Z., Gao, X. & Su, A. Mechanical performances and microstructures of metakaolin contained UHPC matrix under steam curing conditions. *Constr Build Mater* **268**, (2021).

(49) Mo, Z., Wang, R. & Gao, X. Hydration and mechanical properties of UHPC matrix containing limestone and different levels of metakaolin. *Constr Build Mater* **256**, (2020).

(50) Abdellatief, M. *et al.* Development of ultra-high-performance concrete with low environmental impact integrated with metakaolin and industrial wastes. *Case Studies in Construction Materials* **18**, (2023).

(51) Xia, Y. *et al.* Designing low-carbon ultra-high performance concrete with co-combustion ash of sewage sludge and rice husk. *Materials and Structures/Materiaux et Constructions* **58**, (2025).

(52) Ha, N. S. *et al.* Effect of grounded blast furnace slag and rice husk ash on performance of ultra-high-performance concrete (UHPC) subjected to impact loading. *Constr Build Mater* **329**, (2022).

(53) Isa, M. N., Pilakoutas, K., Guadagnini, M. & Angelakopoulos, H. Mechanical performance of affordable and eco-efficient ultra-high performance concrete (UHPC) containing recycled tyre steel fibres. *Constr Build Mater* **255**, (2020).

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article

Computer Engineering

Identification of Diabetic Retinopathy Using Deep Learning and Ensemble Model Approach

Nusrat Mahee and Md. Sabbir Ejaz*

Cite https://doi.org/10.64589/juri/214968

Submitted: June 04, 2025 **Revised:** November 06, 2025 **Accepted:** December 01, 2025

ABSTRACT

One of the most prevalent complications of diabetes and a leading cause of preventable blindness worldwide is diabetic retinopathy (DR). Although several approaches have been employed in the existing literature for DR detection and classification, several critical limitations remain owing to the use of small or single-center datasets and a lack of cross-model benchmarking. These have resulted in reduced generalizability, poor sensitivity, and unquantified performance advantages. This study examines the history of DR detection methods and transition of conventional image-processing algorithms to state-of-the-art deep learning and ensemble models. This study employed the MESSIDOR dataset, which includes retinal fundus images with labels indicating DR severity. Images were enhanced and preprocessed, and deep learning-based features were extracted using several pretrained architectures (e.g., ResNet50, InceptionV3, DenseNet variants, VGG16, Xception) and a custom deep neural network (DNN). Additionally, ensemble methods were developed by integrating CNN-based feature extractors with machine learning classifiers, such as random forest (RF), support vector machine (SVM), XGBoost, and LightGBM, to improve the generalization and classification accuracy. Experimental results revealed that Xception achieved the highest performance, with a testing accuracy of 95%. Furthermore, EfficientNetB3 integrated with XGBoost achieved the highest testing accuracy (86.83%). This study highlights the possibility of deploying deep learning and ensemble AI in DR screening devices, particularly in telemedicine. Such an integration can significantly improve early detection rates, reduce the burden on healthcare providers, and make diabetic eye care more accessible in underserved regions.

Keywords: diabetic retinopathy, ensemble method, deep learning, machine learning, MESSIDOR, telemedicine

1. INTRODUCTION

DR is a major cause of blindness, resulting from long-term elevated blood sugar levels that damage the retinal blood vessels. DR is divided into two major categories; one is proliferative DR (PDR) and non-proliferative DR (NPDR). In diabetes patients, NPDR affects the retinal blood vessels. The condition is further divided into mild, moderate, and severe and NPDR is the initial phase of the condition. These stages are very crucial to detect early enough and apply medical intervention before the disease advances to other more serious stages and loss of vision[1]. Our study leverages methodological and experimental advancements to improve DR detection. Figure 1 illustrates retinal conditions during different stages of DR.

2. RELATED WORK

There are many algorithms, strategies, and methodologies which have been used to detect and classify DR. Great progress has been achieved in applying machine and deep learning methods, but there are still a number of major limitations. Most existing models were developed using small or single-center datasets (e.g., DIARETDB1 and EyePACS subsets), which often lack sufficient representation across various severity grades of DR. This limitation leads to inefficient sensitivity to detect mild or non-proliferative cases and diminishes external validity in other populations and imaging equipment. Other models are based on individual CNN architectures or classical machine learning algorithms and not on cross-model benchmarking so the performance benefits remained undetermined. Moreover, the ensemble and hybrid methods are either under-explored or not optimized or computational inefficiencies tend to impede the ability to apply them in real-time. Table 1 (in the supplementary section S1) presents a summary of some past work and their contributions.

In this work, to mitigate these drawbacks of previous studies, we diversify the dataset and analyze the state-of-the-art models in a comparative manner, as well as optimize hybrid/ensemble methods to achieve better generalizability and computational efficiency. In this direction, we apply the publicly available MESSIDOR database and expand it to 12,000 images through data augmentation that improves model generalization when applied to various imaging conditions. After that, we perform a comprehensive comparative analysis of seven deep learning models (ResNet50, InceptionV3, DenseNet121, DenseNet169, VGG16, Xception, and our custom DNN) and four ensemble

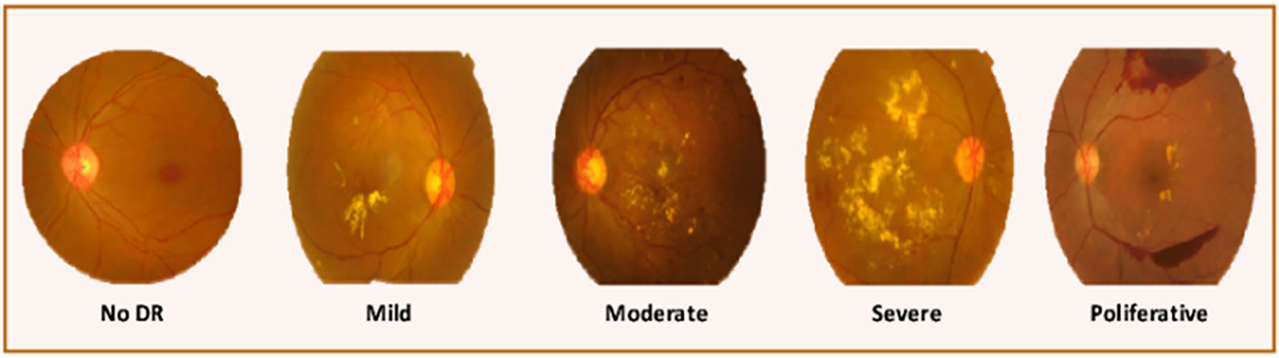


Figure 1. DR stages: (a) Normal retina; (b) mild DR; (c) moderate DR; (d) severe DR; and (e) proliferative DR[2]

models (EfficientNetB3 + XGBoost, DenseNet121 + LightGBM, ResNet50 + SVM, and ResNet50 + RF) to determine a definite hierarchy of the performance. Lastly, the hybrid and ensemble models are further optimized by using CNN-based feature extraction and advanced machine learning classifiers. Lightweight models like Xception and EfficientNetB3 are used, which are computationally efficient and clinically applicable.

3. METHODOLOGY

This study implements a systematic approach to DR detection. It employs both deep learning and ensemble models. The approach emphasizes dataset preparation, preprocessing, model selection, and training pipelines. The overall workflow of the applied methodology is illustrated in Figure 2.

3.1. Dataset Preparation.

In this study, the publicly available MESSIDOR[19] dataset was used and it contained fully anonymized retinal images acquired with the ethical consent of the original providers. None of the patient information can be identified, and the data can be used in academic research. The data set has retinal fundus pictures with Grade 0 (546 pictures), Grade 1 (153 pictures), Grade 2 (247 pictures) and Grade 3 (254 pictures). We used data augmentation to increase the size of the data to 12,000 images. Augmentation techniques comprised random flips, rotations (up to 30°), color jitter, resizing to 224 × 224 pixels and affine transformations. The augmented images were saved in structured, class-wise folders while preserving the original class proportions to enhance model robustness. Table 2. summarizes the number of original and augmented images per class.

In the second step of model training, we implemented real-time preprocessing using the ImageDataGenerator by Keras. New images were not saved under this technique, but transformations were made dynamically when loading a batch. They consisted of shear transformations, zooming, horizontal flipping, rotation (up to 20°) and changes in brightness (between 0.8 and 1.2°), width and height changes. Moreover, a preprocessing input function was also used to normalize the input images according to the model architecture that is being trained.

Table 1. *Grade-wise image counts before and after augmentation*

DR grade	Original count	Augmented count	Final count
Grade 0	546	4914	5460
Grade 1	153	1377	1530
Grade 2	247	2223	2470
Grade 3	254	2286	2540
Total	**1200**	**10800**	**12000**

3.2. Overview and Rationale of Implemented Models.

Our study utilized seven deep learning models and four ensemble models. The deep learning models include ResNet50, InceptionV3, DenseNet121, DenseNet169, Xception, VGG16, and DNN. These models were selected because they represent diverse and well-established CNN architectures. ResNet and DenseNet address gradient flow through residual and dense connections, InceptionV3 and Xception provide efficient multi-scale and depthwise-separable feature extraction, and VGG16 offers a strong convolutional baseline. A custom DNN is included as a lightweight benchmark. Together, they allow a comprehensive evaluation of different feature-learning mechanisms for DR classification.

The ensemble models include EfficientNetB3 + XGBoost, ResNet50 + RF, ResNet50 + SVM, and DenseNet121 + LightGBM. The primary variations among the ensemble models lie in the choice of deep feature extractor and machine learning classifier. The EfficientNetB3 + XGBoost model combines a compact yet efficient CNN with gradient boosting to achieve high accuracy and interpretability. The ResNet50 + RF model combines the residual training of ResNet and the resiliency of decision-tree ensembles to enhance generalization. The ResNet50 + SVM model integrates deep residual features with a linear kernel SVM, which is particularly suitable for high-dimensional feature spaces. Finally, the DenseNet121 + LightGBM model leverages the densely connected CNN layers to generate rich features and uses them in a LightGBM model that is constrained to multiclass classification based on the log loss. These architectural variations influence how each model captures fine-grained features and balances the bias–variance trade-off, thereby affecting classification performance and error behavior. An overview of the implemented models is included in the supplementary section S2.

3.2.1. Deep learning model configuration.

In this study, we applied transfer learning employing the pretrained CNN

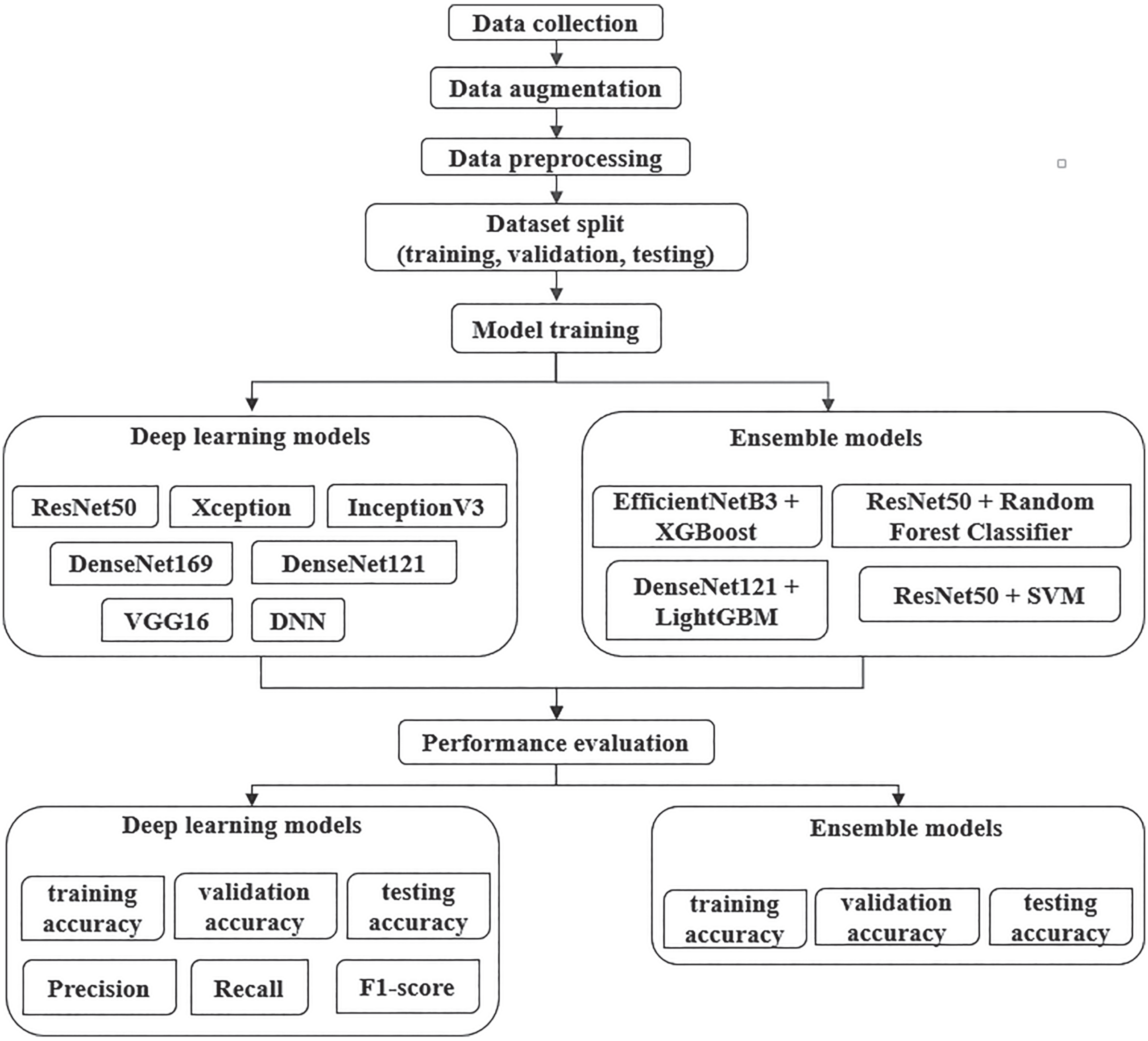


Figure 2. Schematic representation of the workflow

architectures of the implemented models. These models, originally trained on ImageNet, provide strong feature-extraction capabilities. We tailored the models for our task by selectively unfreezing 4–20 of the final layers depending on the model. These layers were fine-tuned to learn high-level retinal features, whereas low-level features from pretraining were retained.

We replaced the original classifier with a fully connected custom head, starting with a global average pooling layer to reduce the dimensions, followed by dense layers (1024 and 512 units) with rectified linear unit activation. We included dropout (rate of 0.3) and L2 regularization (0.001) to prevent overfitting. The final layer used softmax activation for multiclass classification. Training was performed using the Adam optimizer (learning rate = 0.0001) and categorical cross-entropy loss.

3.2.2. Ensemble model configuration. In our study, all four ensemble models follow a similar ensemble pipeline that combines deep learning-based feature extraction with classical machine learning classifiers. A pretrained CNN, EfficientNetB3, ResNet50, or DenseNet121, is loaded without including its final layers and treated as a feature extractor. These extracted features are then fed into a classic machine learning model (e.g., XGBoost, SVM, RF, or LightGBM), which is trained as a multiclass classifier. After training, all models are assessed using accuracy measures, and classification reports, which provide quantitative and visual measures of the predictive capacity of each model at each DR severity level.

3.2.3. Hardware and software environment. All experiments were conducted in a Windows-based system using Jupyter Notebook (Anaconda Navigator). The models were implemented in Python. The same hardware and software environment was used across all models to ensure consistency, fairness, and reproducibility.

4. RESULTS AND DISCUSSION

This section summarizes the results obtained from implementing the deep learning and ensemble models on the MESSIDOR

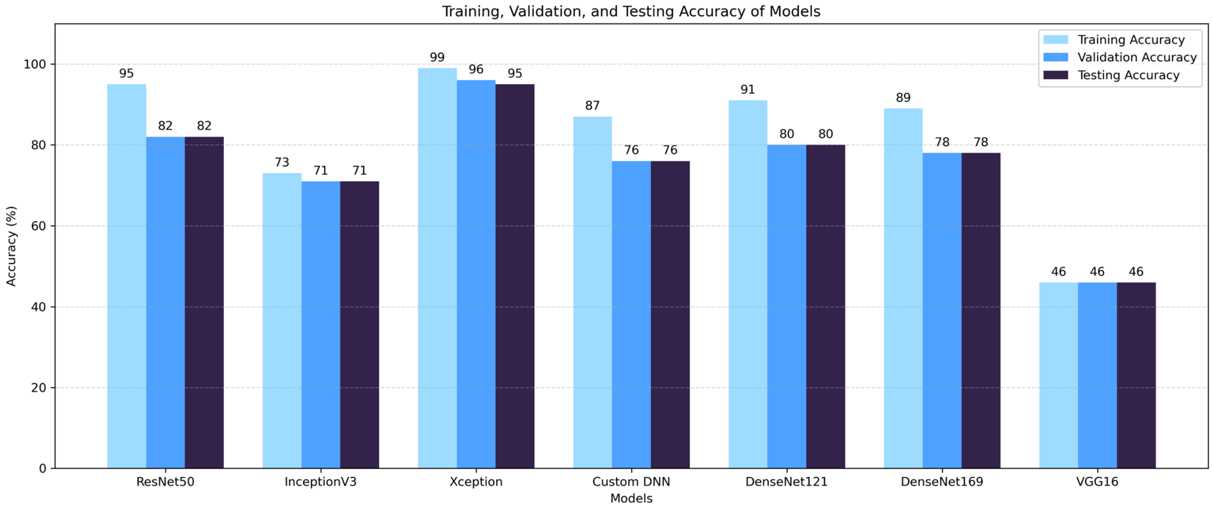


Figure 3. Bar chart of training, validation, and testing accuracies for deep learning models

dataset. The analysis is organized into four parts: Section 4.1 presents the results of the deep learning models, and Section 4.2 discusses the performance of the ensemble models. Section 4.3 provides a comparative analysis of existing studies, and Section 4.4 offers a detailed discussion of the key findings.

4.1. Deep Learning Model Results.

In this study, we evaluated seven deep learning models for DR detection. The training, validation, and testing accuracies are presented in Table 3 (provided in the supplementary section S3), and Figure 3.

Xception demonstrated the highest performance, with training, validation, and testing accuracies of 99%, 96%, and 95%, respectively, showing excellent generalization. ResNet50 trained with 95% accuracy but only 82% on validation and testing, indicating a marginal overfitting. Smaller overfitting (91% training, 80% testing, 89% training, and 78% testing) in DenseNet121 and DenseNet169 also indicated the successful extraction of features, as well as the memorization of training features. The custom DNN scored 87% training and 76% testing, indicating moderate overfitting, as the model fit some non-generalizable patterns. InceptionV3 scored worse in validation and testing (71%), likely owing to underfitting or architectural incompatibility. VGG16 achieved the lowest accuracy (46%) due to its shallow, uniform architecture and lack of residual connections, which limit gradient flow and the ability to capture fine-grained retinal features. Consequently, it struggles to distinguish subtle lesions across DR severity levels. Table 4 outlines the accuracy, recall, and F1-scores of the models for the four DR grades in the MESSIDOR dataset.

A comparative analysis of the models on the MESSIDOR dataset reveals that Xception was the most successful, with the highest precision, recall, and F1-scores across all DR grades. This excellent performance can be explained by its depthwise separable convolutions and residual connections that enable highly efficient feature extraction by decoupling channels and spatial learning. Consequently, Xception can capture fine retinal lesions while remaining computationally efficient.

ResNet50 and DenseNet169 also exhibited robust performance owing to their deep residual and dense connections, which are more effective in learning hierarchical lesion patterns. Conversely, although both InceptionV3 and VGG16 are stable and interpretable, they achieved relatively lower performance because they are both shallow and less optimized for complex feature hierarchies. Custom DNN performed poorly due to pure dense structures that do not possess the spatial feature extraction. Generally, models with an effective convolutional design and feature reusability were highly effective in differentiating DR severity levels.

4.2. Ensemble Model Results.

This section summarizes the performances of the four ensemble models. A comprehensive

Table 2. *Class-wise precision, recall, and F1-scores across DR grades*

Model	**Precision across grades**				**Recall across grades**				**F1-score across grades**			
	0	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Xception	0.94	0.93	0.95	0.99	0.98	0.89	0.92	0.96	0.96	0.91	0.93	0.97
ResNet50	0.87	0.60	0.77	0.92	0.85	0.76	0.74	0.87	0.86	0.67	0.75	0.89
DenseNet121	0.76	0.61	0.76	0.90	0.91	0.56	0.61	0.73	0.83	0.58	0.68	0.81
DenseNet169	0.80	0.68	0.67	0.94	0.90	0.45	0.77	0.75	0.85	0.54	0.72	0.84
Custom DNN	0.46	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.63	0.00	0.00	0.00
InceptionV3	0.70	0.68	0.64	0.84	0.90	0.39	0.51	0.70	0.79	0.49	0.57	0.76
VGG16	0.77	0.61	0.66	0.94	0.91	0.49	0.64	0.70	0.83	0.54	0.65	0.80

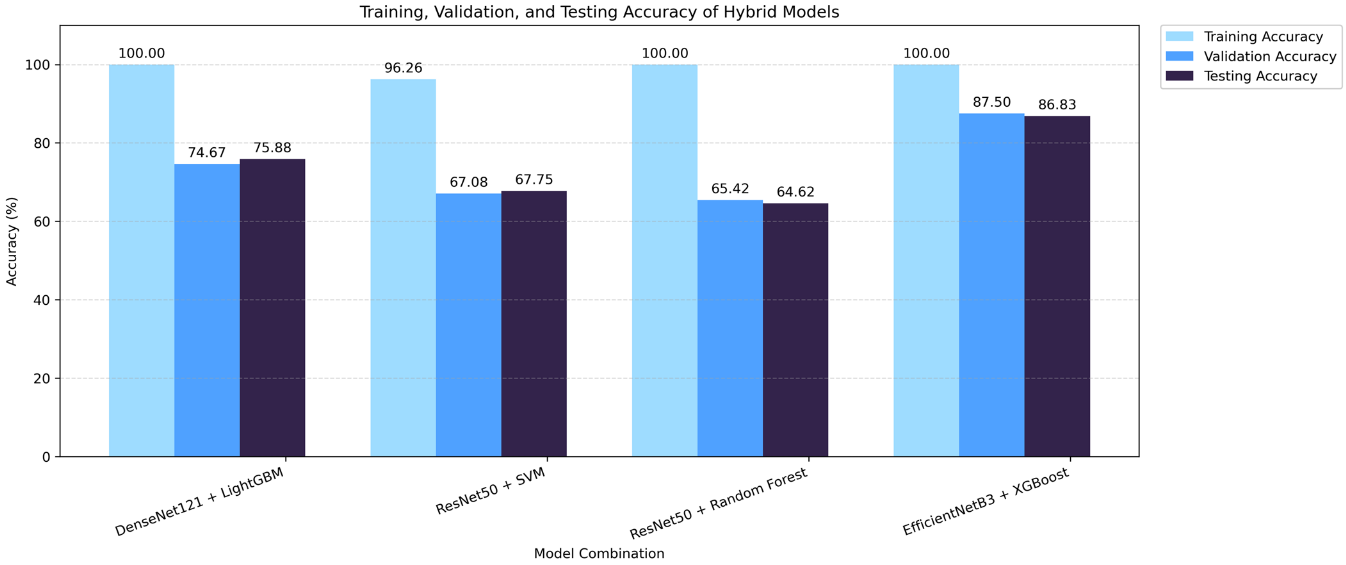


Figure 4. Bar chart of training, validation, and testing accuracies for ensemble models

summary of the experimental results is presented in Table 5 (provided in supplementary section S4) and illustrated in Figure 4.

EfficientNetB3 + XGBoost achieved training, validation, and test accuracies of 100%, 87.50%, and 86.83%, respectively. The gap between training and validation/testing indicates overfitting. The ensemble memorized the training patterns but was still adequately generalized to unseen data. The highest test accuracy can be attributed to the strong feature extraction capability of EfficientNetB3. The model is able to capture the finer retinal details. The enhanced robustness and classification capability were caused by XGBoost capacity to capture the complicated relationship between these features. The DenseNet121 + LightGBM also had 100% training and lower validation (74.67%) and testing (75.88%) performances due to overfitting caused by memorization of training features. ResNet50 + SVM and ResNet50 + RF had good training accuracy (96–100%) and poor test accuracy (67.75% and 64.62%, respectively), which were characteristics of overfitting and ineffective feature integration.

4.3. Comparison with Existing Works.

Compared with previous studies, our study on the MESSIDOR dataset demonstrates competitive and, in several cases, superior performance. Earlier research using MESSIDOR achieved notable results, such as VGG16 (94.57%), DenseNet121 (79.75%)[8], IoTDL-DRD (99.08%)[9], and ResNetGB (98.88%)[10], often relying on complex ensemble models or optimization-based techniques. By contrast, our implementation of the Xception model achieved a test accuracy of 95%, demonstrating strong performance without heavy reliance on metaheuristic algorithms. In addition, our ensemble model, EfficientNetB3 + XGBoost, outperformed several ensemble models from previous studies, such as the ensemble of ResNet50, InceptionV3, Xception, DenseNet121, and DenseNet169, which reached only 80.08%[1], and the VGG16 + XGBoost model, which reached 79.5%[11] by achieving 86.83% test accuracy. Although some studies on other datasets, such as EyePACS and DIARETDB1, have reported better results (ResNetGB: 99.73%, Modified U-Net: 98.68%), these datasets are typically larger or more diverse. Despite the moderate size and limited variability of MESSIDOR, our models (e.g., DenseNet121: 80%; ResNet50: 82%) showed strong learning capabilities. This highlights that with effective data augmentation, model fine-tuning, and thoughtful ensemble combinations, a high performance in DR detection is achievable, even on more constrained medical imaging datasets.

4.4. Discussion of the Findings.

Early training revealed diverse behaviors across architectures. DenseNet121 + LightGBM and EfficientNetB3 + XGBoost both showed 100% training accuracy with lower validation/testing (74.67%–87.5%), indicating overfitting but reasonable generalization. VGG16 underfit with 46% accuracy, while InceptionV3 and ResNet50 showed moderate training (>70%) but reduced validation/testing performance. These results suggest that although advanced deep learning architectures effectively capture complex retinal features, the performance of the ensemble models can be further improved by combining the strengths of models. Nevertheless, the findings also suggest that the success of ensemble-based methods is highly dependent on the combination of the best models and the tuning of hyperparameters to make the ensemble approach robust and generalizable.

This study has some limitations. It uses only the MESSIDOR dataset, which may limit how well the results apply to other populations or imaging conditions. Only a few ensemble combinations were tested, and important validation measures like AUC and cross-validation were not fully included. Future research should include larger and more diverse datasets and explore additional ensemble strategies. It should also focus on optimizing model architectures and incorporating supplementary validation metrics to improve robustness and diagnostic accuracy.

5. CONCLUSIONS

This study evaluated deep learning and ensemble models for DR detection using the MESSIDOR dataset. Xception achieved the highest test accuracy (95%), followed by ResNet50 (82%) and DenseNet121 (80%), confirming the strong feature-extraction capability of advanced CNN architectures. The EfficientNetB3

+ XGBoost ensemble reached a notable test accuracy of 86.83%, outperforming some standalone models and demonstrating the value of well-designed ensemble strategies. However, not all ensembles were effective; combinations such as ResNet50 + SVM and ResNet50 + RF performed poorly, underscoring the importance of selecting compatible feature extractors and classifiers. Clinically, the findings have important implications. High-performing models such as Xception and EfficientNetB3 + XGBoost could support early detection of DR, assist in telemedicine-based screening programs, and improve real-world clinical workflows by reducing the burden on ophthalmologists and expanding diagnostic services to underserved populations. Incorporating these AI models into routine practice has the potential to enhance patient outcomes and facilitate large-scale DR screening. Overall, the results align with previous findings and highlight the need for careful architectural and hyperparameter choices to maximize DR classification performance.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Nusrat Mahee – *Department of Information and Communication Engineering, Noakhali Science and Technology University, Bangladesh;* 0009-0004-6103-5175
Email: mahee1115@student.nstu.edu.bd

Corresponding Author

Md. Sabbir Ejaz – *Department of Information and Communication Engineering, Noakhali Science and Technology University, Bangladesh;* 0009-0001-1789-2258
Email: sabbirejaz.ice@nstu.edu.bd

AUTHOR CONTRIBUTION STATEMENT

Nusrat Mahee: Collection and/or assembly of data, data analysis and interpretation, writing the article. **Md. Sabbir Ejaz:** Research concept and design, critical revision of the article, final approval of the article.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Md. Sabbir Ejaz, who has helped me with this research in terms of his constant guidance, encouragement, and helpful ideas. I also acknowledge the Department of Information and Communication Engineering (ICE), Noakhali Science and Technology University (NSTU) for providing the resources and conducive environment that made this work possible.

FUNDING

This research was financially supported by the NSTU Research Cell, Noakhali Science and Technology University (NSTU), which is gratefully acknowledged.

CONFLICTS OF INTEREST

The author states that there are no financial, personal or professional conflict of interest and denies any competing interests with regards to the work.

REFERENCES

(1) Qummar, S., *et al.* A deep learning ensemble approach for DR detection. *IEEE Access.* **7**, 150530–150539 (2019).

(2) Bashir, I. *et al.* RDS-DR: An Improved Deep Learning Model for Classifying Severity Levels of DR. *Diagnostics,* **13**, 3116 (2023).

(3) Deepa, V., Kumar, C. S., & Cherian, T. Ensemble of multi-stage deep convolutional neural networks for automated grading of DR using image patches. *Journal of King Saud University- Computer and Information Sciences*, **34**(8), 6255–6265 (2022).

(4) P. R., & A. P. Diagnosis of DR using machine learning techniques. *ICTACT Journal on Soft Computing,* **3**(4), 563–575 (2013).

(5) S. Dutta, B. C. S. Manideep, S. M. Basha, R. D. Caytiles, & N. C. S. N. Iyengar. Classification of DR images by using deep learning models. *International Journal of Grid and Distributed Computing,* **11**(1), 89–106 (2018).

(6) Naveen, R., Sivakumar, S. A., Maruthi Shankar, B., & Keerthana Priyaa, A. DR detection using image processing. *International Journal of Engineering and Advanced Technology,* **8**(6 Special Issue), 937–941 (2019).

(7) Carrera, E. V., Gonzalez, A., & Carrera, R. Automated detection of DR using SVM. Proc. 2017 IEEE 24th International Congress on Electronics, Electrical Engineering and Computing (INTERCON), (2017).

(8) T. Swapna, D. Akhila, P. Srija, T. Shivani, and P. Srinidhi, DR Classification using Transfer Learning. *International Journal of Advanced Trends in Computer Science and Engineering,* **12**(3), (2023).

(9) T. Palaniswamy and M. Vellingiri, Internet of Things and Deep Learning Enabled DR Diagnosis Using Retinal Fundus Images. *IEEE Access*, **11**, 27590–27601 (2023).

(10) F. Saeed, M. Hussain, and H. A. Aboalsamh, Automatic DR Diagnosis Using Adaptive Fine-Tuned Convolutional Neural Network. *IEEE Access*, **9**, 41344–41359 (2021).

(11) Mohanty, C., *et al.* Using deep learning architectures for detection and classification of DR. *Sensors*, **23**(12), 5726, (2023).

(12) A. Erciyas and N. Barışçı, An Effective Method for Detecting and Classifying DR Lesions Based on Deep Learning. *Computational and Mathematical Methods in Medicine*, **2021**(1), 9928899 (2021).

(13) A. Skouta, A. Elmoufidi, S. Jai-Andaloussi, O. Ouchetto, Hemorrhage semantic segmentation in fundus images for the diagnosis of DR by using a convolutional neural network. *Journal of Big Data*, **9**(1), (2022).

(14) R. Reguant, S. Brunak, & S. Saha. Understanding inherent image features in CNN-based assessment of DR. *Scientific Reports,* **11**(1), 1–12 (2021).

(15) Ramasubramanian, B. & Prabhakar, G. An Early Screening System for the Detection of DR using Image Processing. *International Journal of Computers and Applications*, **61**, 6–10 (2013).

(16) Gao, Z. *et al.* Diagnosis of DR Using Deep Neural Networks. *IEEE Access* **7**, 3360–3370 (2019).

(17) Samanta, A., Saha, A., Satapathy, S. C., Fernandes, S. L. & Zhang, Y. D. Automated detection of DR using convolutional neural networks on a small dataset. *Pattern Recognition Letters*, **135**, 293–298 (2020).

(18) Zhao, Y. *et al.* Using Machine Learning Techniques to Develop Risk Prediction Models for the Risk of Incident DR Among Patients with Type 2 Diabetes Mellitus: A Cohort Study. *Frontiers in Endocrinology (Lausanne)*, **13**, 876559 (2022).

(19) E. Decencière, X. Zhang, G. Cazuguel, B. Laÿ, B. Cochener, *et al.*, Feedback on a publicly distributed image database: The Messidor database. *Image Analysis & Stereology*, **33**(3), 231–234 (2014).

KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Article
Computer Science

Benchmarking Drift-Resilient Anomaly Detection in Streaming Industrial Data: A Case Study on Turbofan Engine Failures

Rakan M. AlKhulaif

Cite https://doi.org/10.64589/juri/214614

Submitted: June 04, 2025 **Revised:** October 26, 2025 **Accepted:** November 23, 2025

ABSTRACT

Concept drift undermines the reliability of unsupervised anomaly detection in safety-critical systems, in which sensor distributions evolve over time. Using a modified version of NASA's C-MAPSS FD001 turbofan dataset, this study benchmarks the drift robustness of two widely used detectors (Isolation Forest and One-Class SVM) alongside a streaming change detector (ADWIN). To emulate gradual system degradation, we injected 5% synthetic anomalies across sensors 1–5 and applied a multiplicative drift ($1.0 \rightarrow 1.5$) from cycles 80 to 120. All models were trained on pre-drift data (<cycle 80) using min-max normalization and evaluated across the pre-, mid-, and post-drift phases. The isolation forest achieved an F1 score of 0.857 pre-drift but degraded to 0.140 mid-drift and 0.080 post-drift; retraining at cycle 120 failed to recover the performance. One-Class SVM exhibited weak detection throughout (pre-drift $F1 = 0.364$; post-drift $= 0.080$). Meanwhile ADWIN, used naïvely as an anomaly signal on aggregate sensor values, returned $F1 = 0$ throughout. These findings illustrate that static detectors collapse under moderate drift, simple retraining is ineffective when drift co-occurs with anomalies, and drift signals alone do not reliably flag anomalies. We released our simulation protocol and codebase to support future work on drift-resilient anomaly detection for streaming safety-critical environments.

Keywords: concept drift, anomaly detection, streaming data, turbofan engines, predictive maintenance

1. INTRODUCTION

Industrial cyber-physical systems (CPS), including aerospace turbofans, power-generation turbines, and manufacturing lines, increasingly rely on continuous sensor monitoring for predictive maintenance[1]. Unsupervised anomaly detectors are widely adopted because they require no labelled failure data and can theoretically identify novel fault signatures. However, these methods typically assume a stationary data distribution[2,3]. In practical deployments, sensor readings evolve over time because of component ageing, environmental conditions, or maintenance actions: a phenomenon known as concept drift[2,4]. When static detection boundaries are applied to drifting data, reliability plummets, resulting in missed fault signals or a proliferation of false alarms. Thus ensuring robust anomaly detection during drift is critical for maintaining system integrity and safety[1,5,6].

Although concept drift has been studied extensively in the context of classification tasks, there exists a marked paucity of systematic benchmarks for evaluating unsupervised anomaly detectors under realistic drift scenarios. Prior work often relies on synthetic benchmarks or focuses exclusively on classification accuracy. In the domain of anomaly detection for CPS, few studies examine the interplay between drift and rare fault occurrences, particularly in continuous, streaming settings. As a result, there is limited guidance on whether static detectors suffice, how quickly adaptive schemes respond, or how retraining strategies should be orchestrated[2,4,7–10].

This study presents a reproducible pipeline to simulate gradual drift and evaluates three approaches: (i) use of static unsupervised detectors (Isolation Forest[11] and One-Class SVM[12]) trained solely on pre-drift data, (ii) the utilisation of streaming drift detector ADWIN[13] as a naive anomaly proxy, and (iii) scheduled retraining of static detectors post-drift. By employing NASA's C-MAPSS FD001 turbofan dataset[14], the aim is to quantify the robustness of each method under drifting sensor distributions in a safety-critical streaming context and provide a full codebase and drift protocol to facilitate replication and extension.

The remainder of this paper is organized as follows:

Section 2 reviews related work on concept drift and unsupervised anomaly detection.

Section 3 describes on the dataset, drift simulation, anomaly injection, and model configurations.

Section 4 presents experimental results, including F1 scores and recovery analyses.

Section 5 discusses implications, limitations, and potential future directions.

Section 6 concludes the paper.

2. RELATED WORK

2.1. Concept Drift in Streaming Environments. Concept drift refers to changes in the statistical properties of data streams over time, which can degrade model performance if unaddressed. Prior work has catagorized drift as abrupt, gradual, incremental, or recurring, each requiring distinct detection and adaptation strategies[7]. Early detectors such as DDM focused on classification error rates[7], while EDDM targeted slower, gradual changes[14]. ADWIN monitors distributional change via an adaptive window and adjusts its size using a statistical test to avoid manual window parameters[15]. Stream-based classification with drift detectors has been well-studied; however, analogous evaluations for unsupervised anomaly detection remain scarce[8,15–19].

2.2. Unsupervised Anomaly Detection Techniques. Isolation Forest identified anomalies by recursively partitioning data; outliers require fewer partitions. It scales well to high-dimensional data and is commonly used in industrial CPS. One-Class SVM constructs a boundary around normal data by maximising the margin from the origin in feature space; its performance is sensitive to kernel choice and parameters (e.g., ν, γ). Other techniques such as local outlier factor and clustering-based methods have been applied in static scenarios, but their drift resilience is under-explored[1,5,6,20–22].

2.3. Drift-Aware Anomaly Detection. Recent efforts have embedded drift awareness into anomaly detection using three recurring ideas. First, windowed retraining or accuracy-updated ensembles aim to recover quickly by refreshing models on short post-drift windows, at the cost of instability if the window is contaminated[7,8] Second, change-detector-gated policies use signals such as ADWIN to trigger adaptation; these signals indicate when to adapt rather than what is anomalous, so naïve use as labels risks systematic errors[15,17]. Third, hybrid or physics-informed scoring uses residuals from simple physical expectations as inputs to data-driven detectors, thereby improving the identifiability of genuine faults under shift[1]. Our contribution is a minimally controlled benchmark that isolates drift and anomaly factors and systematically compares (i) static detectors, (ii) the drift-signaled use of ADWIN, and (iii) scheduled retraining under the same streaming protocol. A summary of these adaptive strategies is provided in Table 1.

2.4. Turbofan Engine Prognostics and Anomaly Detection. NASA's C-MAPSS dataset[13] is a canonical benchmark for prognostic health management (PHM) and anomaly detection in aerospace engines. Several studies have focused on remaining useful life (RUL) estimation using supervised or semi-supervised deep learning. Anomaly detection work on C-MAPSS often uses autoencoders or probabilistic approaches but typically assumes a stationary distribution. To the best of our knowledge, no prior work has robustly simulated concept drift within C-MAPSS for unsupervised anomaly detector benchmarking. Therefore, we adopted a minimal, reproducible protocol that isolates drift and anomaly factors for a fair comparison in Sections 4–5.

3. METHODOLOGY

3.1. Dataset Description. We used NASA's C-MAPSS FD001 subset[13] containing multivariate time-series sensor data from 100 simulated turbofan engines. Each engine unit produces a sequence of observation cycles until failure. Per cycle, there are 21 sensor measurements and three operational settings. For simplicity, we focused the main analysis of Unit 1 and replicated the protocol in Units 2–3 (Section 4.5). We excluded three operational settings and used only the 21 sensor features.

3.2. Anomaly Injection. To emulate rare fault events, we injected synthetic anomalies into 5% of the data points. Let N be the total number of cycles in unit 1. We sampled indices without replacement to select anomalies of approximately N × 0.05 anomalies. We distributed the anomalies evenly: one-third of the anomalies in cycles <80 (pre-drift), one-third in cycles 80–120 (drift window), and one-third in cycles >120 (post-drift). At each selected index, we added Gaussian noise with a mean of 0.5 and standard deviation of 0.1 to sensors 1–5. These rows are markedwith an is_anomaly flag (1) leaving the remainder at zero. Figures 1–3 illustrate the sensor distributions in the pre-, mid-, and post-drift phases.

3.3. Concept Drift Simulation. We simulated gradual drift in sensors 1–5 over cycles 80–120 by applying a multiplicative factor.

$$m(c) = 1.0 + 0.0125 \cdot (c - 80), \quad 80 \le c \le 120 \tag{1}$$

Table 1. *Summary of adaptive strategies for non-stationary streams (from Related Work)*

Approach	Adaptation trigger	Labels needed	Typical pros	Typical cons/risks	Example references
Windowed	Time/window schedule or accuracy decay	No (unsupervised)	Fast recovery after drift; simple to implement	Unstable if the window is contaminated; window size sensitive	[7,8–10]
Drift-gated policies (e.g., ADWIN)	Change detector fires	No (uses change flag)	Adapts only when needed; avoids constant retraining	Flags indicate when to adapt, not anomalies; misalignment risk	[14,15,17]
Hybrid / physics-informed residuals	Residual vs simple physical model	No (unsupervised)	Better fault identifiability under shift	Requires a domain model; residual drift is still possible	[1,5,6,22]

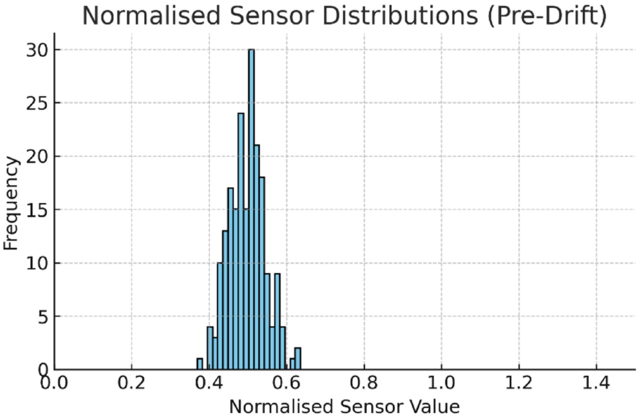


Figure 1. Normalized sensor distributions during the pre-drift phase (cycles <80); x-axis fixed to 0–1.5

Here, m(80) = 1.0, m(120) = 1.5, and m(c) = 1.0 for c <80 and m(c) = 1.5 for c >120; the multiplier is applied only to sensors 1–5. Each value $s_i(c)$ for i = 1…5 is replaced by $s_i(c)\cdot m(c)$; sensors 6–21 remain unchanged.

3.4. Data Normalization. In real-world streaming pipelines, the normalization parameters are often fixed during deployment to avoid leakage of future information. We fitted a MinMaxScaler to all 21 sensor features using only the pre-drift data (cycles <80). The scaler maps each feature to [0, 1] based on its minimum and maximum values in the pre-drift window. We then applied the same scaler to transform every cycle (including drift and anomalies), ensuring that no future statistics influence the model decisions. After normalization, data at cycles >120 (drifted sensors 1–5) often exceed the [0, 1] range in raw terms, but are represented relative to the original pre-drift bounds. Figures 1–3 show the histograms of the normalized sensor distributions across the three segments.

3.5. Model Configurations.

3.5.1. Isolation Forest (IF). We used the PyOD implementation of Isolation Forest [11] with 100 estimators and contamination = 0.05 (matching the evaluated anomaly rate). The model was trained solely on the pre-drift normalized data (cycles <80).

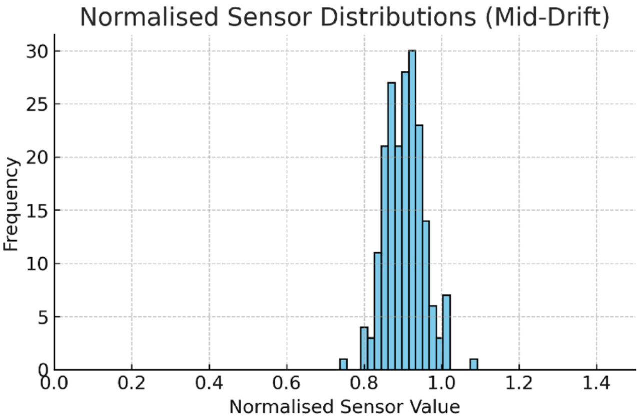


Figure 2. Normalized sensor distributions during the mid-drift phase (cycles 80–120); x-axis fixed to 0–1.5

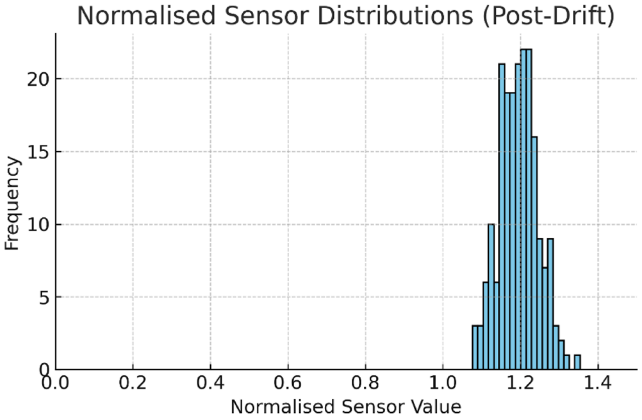


Figure 3. Normalized sensor distributions during the post-drift phase (cycles >120); x-axis fixed to 0–1.5

3.5.2. One-Class SVM (OCSVM). We employ PyOD's OCSVM [12] with an RBF kernel, $\nu = 0.05$ (matching contamination), and $\gamma =$ 'scale'. Similar to the IF, we trained the OCSVM on the pre-drift normalized data. This creates a hypersphere (or hyper-surface) enclosing normal points, where the points outside are flagged as anomalies.

3.5.3. ADWIN. ADWIN is a streaming drift detector that adaptively maintains a sliding window whose size adjusts to detect distribution changes with confidence δ. We instantiated ADWIN with $\delta = 0.002$. As we streamed each new cycle, we computed the sum of sensors 1–5 and input it into the ADWIN update() method. If ADWIN reported a change in cycle c, we flagged that single cycle as an anomaly proxy (binary 1); otherwise, we flagged it at 0. This is a deliberately naive usage: ADWIN is not designed to detect individual point anomalies but rather to drift in the aggregated statistic. We included this to evaluate whether raw change flags correlated with the anomaly events.

3.6. Streaming Experiment Protocol. We simulated the streaming by iterating through the cycles in order. The pipeline is as follows:

1. **Training (cycles <80):** IF and OCSVM were fit to pre-drift normalized data (cycles <80).
2. **Drift simulation (cycles 80–120):** m(c) was gradually appllied to sensors 1–5 and then normalized using the pre-drift scaler.
3. **Anomaly injection (all cycles):** At pre-selected cycles, Gaussian noise was added to sensors 1–5.
4. **Inference:** At each cycle c, the 21-dimensional normalized feature vector x(c) was extracted, static detector predictions $y_a(c)$ (IF or OCSVM, both trained on cycles <80). Were obtained, and $\sum(i=1)^5$ xi(c) was fed into ADWIN to obtain drift flag d(c)
5. **Evaluation segments:**
    - **Pre-drift:** c <80
    - **Mid-drift:** $80 \leq c \leq 120$
    - **Post-drift:** c >120
6. **Metrics:** For each model and segment, Precision, Recall, and F1 were computed between the predicted labels and ground truth as is_anomaly.

Table 2. *Isolation Forest performance by drift phase (retraining at cycle 120)*

Phase	Precision	Recall	F1
Pre-drift	0.857	0.857	0.857
Mid-drift	0.117	0.208	0.140
Post-drift	0.080	0.080	0.080
Post-retrain	0.080	0.080	0.080

Note:. The model trained on cycles of <80. Retraining was performed at cycle 120 using the mixed post-drift data.

7. **Retraining (cycle 120):** At $c = 120$, all data up to $c = 120$ (drifted normal + anomalies)were collected, IF was retrained on this mixture, and the post-drift ($c > 120$) was evaluated to measure recovery.
8. **Recovery latency:** For ADWIN, the earliest post-drift cycle where a 10-cycle rolling F1 $\geq 0.9 \times$ pre-drift F1 was computed; for IF with retraining, recovery is immediate at cycle 120 if F1 post-retrain $\geq 0.9 \times$ pre-drift F1.

3.7. Reproducibility and Availability.

Implementation uses scikit-learn[23], Matplotlib[24], and PyOD[25]. The codes and ready-to-run Google Colab are available (repository branch: feat/multi- unit-juri-revision). The artifacts for this revision are as sollows: results/juri_revision/ (cross_unit_summary .csv; unit_{1,2,3}_metrics.json; fig5_unit _{1,2,3}_adwin_v._sanomalies.png). In Google Colab, the data path /content/cmapss_data/CMaps/.

4. EXPERIMENTAL RESULTS

Across Units 1–3, the same pattern holds (Table 4): the Isolation Forest degrades sharply after drift, the One-Class SVM remains weak throughout, and ADWIN's change flags do not coincide with the anomaly timestamps.

4.1. Isolation Forest Performance.

Table 2 shows that Isolation Forest performs well under stationary conditions (F1 = 0.857), but fails under drift, with performance collapsing to F1 = 0.140 mid-drift and further degrading post-drift. Retraining at cycle 120 does not help, due to overlap between drifted normal and anomalous patterns.

4.2. One-Class SVM Performance.

As shown in Table 3, the OC-SVM performed worse than theIF across all phases, starting at a low (F1 = 0.364) and collapsing entirely under drift. Its sensitivity to boundary noise and overfitting from the RBF kernel prevents recovery even with retraining.

4.3. ADWIN as Anomaly Proxy.

ADWIN failed to detect anomalies in any drift phase (Table 4, F1 = 0). although it flags some distributional changes mid-drift, these do not align with anomaly timestamps. This mismatch indicates that change detection does not equate to anomaly detection.

Table 3. *One-Class SVM performance by drift phase*

Phase	Precision	Recall	F1
Pre-drift	0.444	0.308	0.364
Mid-drift	0.143	0.130	0.136
Post-drift	0.080	0.080	0.080

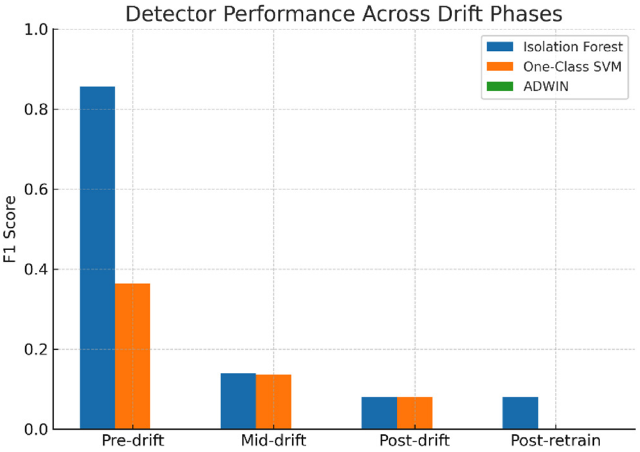


Figure 4. Detector performance across drift phases, reported as F1 score. Isolation Forest collapses after drift, One-Class SVM performs poorly throughout, and ADWIN fails to align change flags with anomalies

4.4. Recovery Latency Analysis.

The phase-wise detector performance is summarized in Figure 4 (introduced in Section 4.3); we used it to analyze the recovery latency. We define recovery latency as the earliest cycle $t \geq 120$ where a 10-cycle rolling F1 $\geq 0.9 \times$ pre-drift F1. For ADWIN, pre-drift F1 = 0, so any non-zero F1 could be counted; however, ADWIN never scored above zero, so no recovery was observed. For IF with scheduled retraining at cycle 120, post-retraining F1 = 0.080, whereas $0.9 \times$ pre-drift F1 = 0.771. IF did not meet this threshold, indicating no meaningful recovery. In effect, neither streaming drift detection nor fixed retraining restores acceptable anomaly detection performance.

To provide a consolidated view of detector performance across all phases, Figure 4 reports the F1 scores for the Isolation Forest, One-Class SVM, and ADWIN. The comparison confirms three key findings: (i) static detectors collapse once drift sets in, (ii) simple retraining fails to restore performance, and (iii) drift detectors such as ADWIN, which aresensitive to distributional changes, do not align with anomaly incidence.

4.5. Cross-Unit Robustness.

To gauge whether the observed failure was idiosyncratic to an individual engine or a universal phenomenon, we applied our protocol to other FD001 units. Table 4 presents the results for Units 2 and 3. The outcome closely resembles Unit 1: The Isolation Forest attains fair pre-drift accuracy but fails once drift begins, the One-Class SVM consistently performs below par across periods, and ADWIN generates change flags irrelevant to anomalies, resulting in zero F1 across the board. These reproducible tendencies across multiple engines verify that the failure modes we observed are not isolated artifacts but inherent weaknesses of stationary detectors and naive drift signals under non-stationary conditions.

5. DISCUSSION

5.1. Limitations of Static Detectors.

Static detectors trained only on pre-drift data degrade sharply once a moderate

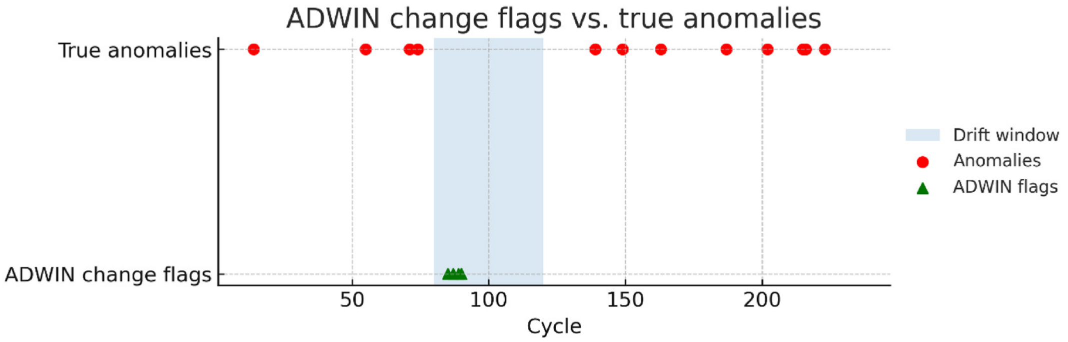


Figure 5. ADWIN change flags versus injected anomalies on FD001 (sensors 1–5); drift window shaded for cycles 80–120.Sub-subsection 1

Table 4. *Detector performance across multiple FD001 units*

Unit	Detector	Pre-drift F1	Mid-drift F1	Post-drift F1
1	Isolation Forest	0.857	0.140	0.080
1	One-Class SVM	0.364	0.136	0.080
1	ADWIN	0.000	0.000	0.000
2	Isolation Forest	0.750	0.182	0.069
2	One-Class SVM	0.167	0.178	0.069
2	ADWIN	0.000	0.000	0.000
3	Isolation Forest	0.667	0.095	0.127
3	One-Class SVM	0.400	0.093	0.127
3	ADWIN	0.000	0.000	0.000

Note: Units 1–3 of the FD001 dataset. Anomalies were injected at 5% across sensors 1–5. Drift window = 80–120 cycles. The results are consistent with those of Unit 1: The Isolation Forest collapses under drift, One-Class SVM performs poorly throughout, and ADWIN fails to align drift flags with the anomalies.

drift hits the key sensor channels. Isolation Forest performs well in static settings, but its F1 drops to 0.140 mid-drift; the One-Class SVM performs weakly before drift and collapses thereafter. Scheduled retraining at cycle 120 is ineffective because the drifted normal data and injected anomalies are confounded; therefore, retraining does not separate the evolving normal range from the actual faults.

Mechanistically, the failure modes are different. Isolation Forest's training-time partitions approximate pre-drift quantiles, and when the means shift for sensors 1–5, large regions of now-normal readings fall into sparse partitions and are flagged as outliers. One-class SVM, a margin-based method, shifts its decision function toward the new mean, and with increased variance after drift, nominal and anomalous scores overlap, driving both false positives and false negatives [2,3,11,12].

Sensor-wise effects mirror this story: mid-drift right-shifts are concentrated on sensors 1–5 by design (Figures 1–3), whereas sensors 6–21 remain comparatively stable and contribute less to errors. In other words, IF falters because its partitions no longer fit the data, while OCSVM falters because its boundary adapts poorly when the variance itself is non-stationary.

5.2. Misapplication of Drift Detectors.

As configured, ADWIN is designed to signal significant shifts in aggregated statistics and not isolate single-cycle anomalies. Its failure (F1 = 0) confirms that the raw change flags should not be used directly as anomalous flags. As Figure 5 shows, ADWIN produces change flags around the onset of drift, yet none coincide with the anomaly timestamps. This mismatch underlines the fact that distributional change signals, which are useful as adaptation triggers, do not substitute anomaly labels. ADWIN detects changes in the mean of a data stream by maintaining adaptive windows; in our setup, its change points are correctly aligned with the drift interval (cycles 80–120), but these signals reflect distributional change, not anomalous behavior. Instead, drift detectors trigger contextual adaptation steps. For example, one might pause anomaly scoring, collect a clean window of data after a drift for retraining, or resume detection. Future work may explore pipelines in which ADWIN change signals automatically initiate a retraining process on a filtered subset of non-anomalous data [15, 18, 19].

5.3. Implications for Real-World Deployment.

Predictive maintenance systems cannot rely on static unsupervised detectors in environments subject to slow, predictable degradation. The lag between the drift onset and retraining, in addition to the confounding effect of drifted normal points with false anomalies, underscores the need for real-time drift-aware mechanisms. As Figure 4 shows, no detector maintained a usable performance after the drift onset (see Figure 5 for the drift window), highlighting the operational risk of naive deployments. The same collapse repeats across Units 1–3 (Table 4). Possible approaches include the following:

- **Ensemble detectors:** combine multiple anomaly detectors trained over different time windows and weigh them based on the drift confidence.
- **Drift-gated retraining:** uses drift detector signals to temporarily suspend anomaly scoring, gather a verified normal reference window (e.g., from maintenance cycles), and retrain detectors on clean data.
- **Feature weighting:** dynamically adjusts feature importance; sensors known to drift (such as temperature readings) can be downweighted in the anomaly score once drift is detected.
- **Hybrid physics-AI models:** incorporate physics-based models or digital twins to anchor the expected sensor behavior and augment them with data-driven detectors to isolate anomalies above physical thresholds.

5.4. Limitations and Future Directions.

This study focuses on FD001 (single operating condition) with additive

Gaussian anomaly injections of a fixed magnitude across sensors 1–5 and a multiplicative drift from cycles 80 to 120. These choices stabilize the benchmark but limit generalizability: real faults often follow multi sensor context-dependent patterns. Future work should extend the protocol to FD002–FD004 (multiple operating regimes) and other datasets, incorporate richer anomaly forms (e.g., engineered fault signatures and correlated patterns), and evaluate deployment constraints at the edge by reporting per-cycle latency, memory footprint for sliding windows, and retraining cost, along with detection delay and false-alarm rate. Benchmarking under resource-limited hardware clarifies the trade-offs between faster recovery and constrained computing.

Beyond static or periodically retrained detectors, future studies should focus on systems that can adapt as objects change. This means moving toward hybrid frameworks that react to drift the moment they appear, rather than after the fact. One promising approach is drift-triggered ensembles, which update or replace their models when sensing a shift. The other is online learning methods that gently reshape decision boundaries as new data flows. There is also a growing interest in physics-aware deep networks — for instance, long short-term memory (LSTM) autoencoders paired with digital twin residuals — that can model complex behaviors without losing interpretability. Testing these adaptive ideas on multi regime datasets such as FD002–FD004 under real-time conditions would reveal how well they balance quick responsiveness with reliability and efficiency at the edge.

6. CONCLUSION

We presented a comprehensive benchmark of unsupervised anomaly detectors under concept drift in a streaming turbofan engine scenario. By injecting drift and anomalies into NASA's C-MAPSS FD001 data, we show that static detectors degrade sharply (IF: pre-drift $F1 = 0.857$ to mid-drift $F1 = 0.140$), naive retraining fails, and streaming drift detectors (ADWIN) cannot serve as direct anomaly proxies ($F1 = 0$). These findings highlight the need for drift-aware anomaly-detection strategies that separate evolving normal signals from genuine fault events. We have publicly released our code and drift simulation protocol to encourage replication, extension, and innovation.

AFFILIATIONS AND AUTHOR DETAILS

Undergradute Author and Corresponding Author

Rakan M. AlKhulaif – *Department of Applied Computer Science, King Saud University, Riyadh 11451, Saudi Arabia;* 0009-0008-4131-3325
Email: 442105575@student.ksu.edu.sa

ACKNOWLEDGEMENTS

This study did not receive any specific funding.

CONFLICTS OF INTEREST

The author declares no conflicts of interest.

REFERENCES

(1) Chandola, V., Banerjee, A. & Kumar, V. Anomaly detection: A survey. *ACM Comput. Surv.* **41**, 15 (2009).

(2) Gama, J. et al. A survey on concept drift adaptation. *ACM Comput. Surv.* **46**, 44 (2014).

(3) Lu, J. et al. Learning under concept drift: A review. *IEEE Trans. Knowl. Data Eng.* **31**, 2346–2363 (2019).

(4) Widmer, G. & Kubat, M. Learning in the presence of concept drift and hidden contexts. *Mach. Learn.* **23**, 69–101 (1996).

(5) Hundman, K. et al. Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding Proc. KDD 387–395 (2018).

(6) Malhotra, P. et al. LSTM-based encoder-decoder for multi-sensor anomaly detection. *arXiv preprint arXiv:1607.00148v2* (2016).

(7) Gama, J., Sebastião, R. & Rodrigues, P. On evaluating stream learning algorithms. *Mach. Learn.* **90**, 317–346 (2013).

(8) Brzezinski, D. & Stefanowski, J. Reacting to different types of concept drift: The accuracy updated ensemble. *IEEE Trans. Neural Netw. Learn. Syst.* **25**, 81–94 (2013).

(9) Domingos, P. & Hulten, G. Mining high-speed data streams. *Proc. KDD* **2000**, 71–80.

(10) Gama, J., Castillo, G. & Rodrigues, P. Incremental learning of decision trees. *Proc. ACM SAC* **2006**, 1105–1110.

(11) Liu, F. T., Ting, K. M. & Zhou, Z.-H. Isolation Forest. *Proc. IEEE Int. Conf. Data Mining* **2008**, 413–422.

(12) Schölkopf, B. et al. Estimating the support of a high-dimensional distribution. *Neural Comput.* **13**, 1443–1471 (2001).

(13) Bifet, A. & Gavaldà, R. Learning from time-changing data with adaptive windowing. *Proc. SIAM Int. Conf. Data Mining* **2007**, 443–454.

(14) Saxena, A., Goebel, K., Saha, B. & Eklund, N. Damage propagation modelling for aircraft engine run-to-failure simulation. *Proc. PHM 2008*, 1–9 (2008).

(15) Gama, J., Medas, P., Castillo, G. & Rodrigues, P. Learning with drift detection. *SBIA* **2004**, 286–295.

(16) Kifer, D., Ben-David, S. & Gehrke, J. Detecting change in data streams. *Proc. VLDB* **2004**, 180–191.

(17) Baena-García, M. et al. Early drift detection method. *KDD Work. Knowl. Discov. Data Streams* **2006**, 77–86.

(18) Basseville, M. & Nikiforov, I. V. *Detection of Abrupt Changes: Theory and Application*. (Prentice Hall, 1993).

(19) Page, E. S. Continuous inspection schemes. *Biometrika* **41**, 100–115 (1954).

(20) Breunig, M. M. et al. LOF: Identifying density-based local outliers. *Proc. ACM SIGMOD* **2000**, 93–104.

(21) Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. *Proc. KDD* **1996**, 226–231.

(22) Lai, G., Chang, W.-C., Yang, Y. & Liu, H. Modeling long- and short-term temporal patterns with deep neural networks for time series forecasting. *Proc. SIGIR* **2018**, 95–104.

(23) Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

(24) Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

(25) Zhao, Y., Nasrullah, Z. & Li, Z. PyOD: A Python toolbox for scalable outlier detection. *J. Mach. Learn. Res.* **20**, 1–7 (2019).

https://doi.org/10.64589/juri/214614

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article *Computer Science*

KF-2025: Deep Learning-Based Classification of Canteen Food Trays Using Custom-Collected Dataset

Saad Obada Alshami and Ali Nasir*

Cite **https://doi.org/10.64589/juri/214613**

Submitted: May 31, 2025 **Revised:** November 06, 2025 **Accepted:** November 23, 2025

ABSTRACT

University canteens serve large student populations within short time windows, necessitating efficient and automated food service solutions. This study presents KF-2025, the first real-world, tray-based food image dataset collected directly from the King Fahd University of Petroleum and Minerals (KFUPM) students' canteen to facilitate automated self-checkout and smart cafeteria applications. The dataset comprises 3,222 manually annotated images across 26 food classes, derived from over 1,000 tray recordings under real canteen conditions. Four transfer learning models, i.e., ResNet-50, ResNet-152, MobileNetV2, and EfficientNet-B0 were trained and evaluated. The best performance was achieved by ResNet-152 using the Adam optimizer, attaining 97.25% accuracy, 97.62% precision, 93.78% recall, and 94.53% F1-score. The high performance across all models confirms the robustness and practical utility of KF-2025 for real-world deployment, including on resource-constrained devices. Beyond its immediate application to canteen automation, KF-2025 provides a foundation for advancing context-aware food recognition, intelligent campus systems, and data-driven nutritional monitoring, thereby contributing to the broader development of smart and sustainable institutional environments.

Keywords: food recognition, deep learning, image classification, ResNet, efficientNet, transfer learning, custom dataset, smart canteen

1. INTRODUCTION

As Artificial Intelligence and deep learning evolve rapidly, several domains of daily life are becoming subjects of research. For example, industrial productivity[1], transportation[2], and banking[3] have already undergone transformations due to deep learning. One promising area for deep learning application is food consumption, which is an essential part of daily human activities. Accordingly, food classification powered by deep learning models has become a prominent research area to allow the detection and classification of food types, offering practical solutions to improve daily lives. For instance, food classification can help reduce waiting times in restaurants and canteens by enabling self-checkout, and can help estimate caloric intake.

However, food classification faces several challenges, including overlapping food categories[4], complex appearance of food items (where even identical ingredients may vary in shape and color)[5], and differences between the environments in which food images are captured and those in which they are recognized[6]. Such factors can significantly hinder model accuracy and classification performance.

University canteens are among the most complex and challenging environments for food recognition, where food appearance in terms of shape, size, and color is inconsistent. The same dish can contain several ingredients, and identical food classes may differ significantly from one canteen to another, limiting the utility of other food datasets. In addition, many food classes are missing from available datasets. At King Fahd University of Petroleum and Minerals (KFUPM), thousands of students visit the canteen daily, particularly during breaks, leading to long waiting times. Therefore, automating the checkout process could enhance students' experiences by saving time and improving convenience.

To address this need, this study aims to construct the first real-time tray-based food image dataset, KF-2025, collected from KFUPM's student canteen. This dataset covers most of the food classes presented daily in the cafeteria. Furthermore, this study provides experimental results when training and testing multiple AI classification models on this dataset after preprocessing and annotating, contributing to a complete food recognition pipeline that can be deployed in the future for automated self-checkout systems.

2. RELATED WORK

In the field of food recognition, advancing artificial intelligence (AI) models has garnered attention due to their potential impact on daily life. A typical pipeline for developing food recognition systems involves constructing or utilizing existing food datasets, preprocessing and annotating images, and training models on those datasets to assess the performance. Food classification is

Figure 1. Plov from CAFD (left) and Rice from UEC-FoodPix (right). Despite both being rice-based, their appearance differs greatly, illustrating the domain mismatch across cuisines

highly influenced by the quality of the dataset and the techniques used to annotate it.

Some well-known food recognition datasets include Food-101[7], which contains 101,000 images across 101 food categories; UEC-FoodPix Complete[8], which contains 10,000 images; and UNIMIB2016[9]. These datasets were collected under varying conditions. Food-101 was captured in a controlled laboratory environment, whereas UNIMIB2016 was captured in a real-world university canteen, reflecting different environmental settings. Additionally, food datasets vary in terms of food classes and regional cuisine, such as the CAFD dataset[10], which focuses on Central Asian food, and the UEC-FoodPix dataset, which focuses on Eastern Asian food. Although both datasets represent Asian food, they are substantially different; they share no common food classes and similar base classes differ significantly in appearance, as illustrated in Figure 1. For example, while both "Rice" (from UEC-FoodPix) and "Plov" (from CAFD) are rice-based dishes, their presentation, accompaniments, and overall appearance differ significantly. This highlights that using food datasets from different cuisines can result in major mismatches when applied to a local context.

Two notable image datasets have been proposed for Middle Eastern food: the MEFood and EMFID platform databases. MEFood is a large-scale dataset covering 38 categories, with images collected from search engines, such as Google, and social media platforms, such as Instagram, to increase representativeness[11]. The EMFID provides a comprehensive collection of food

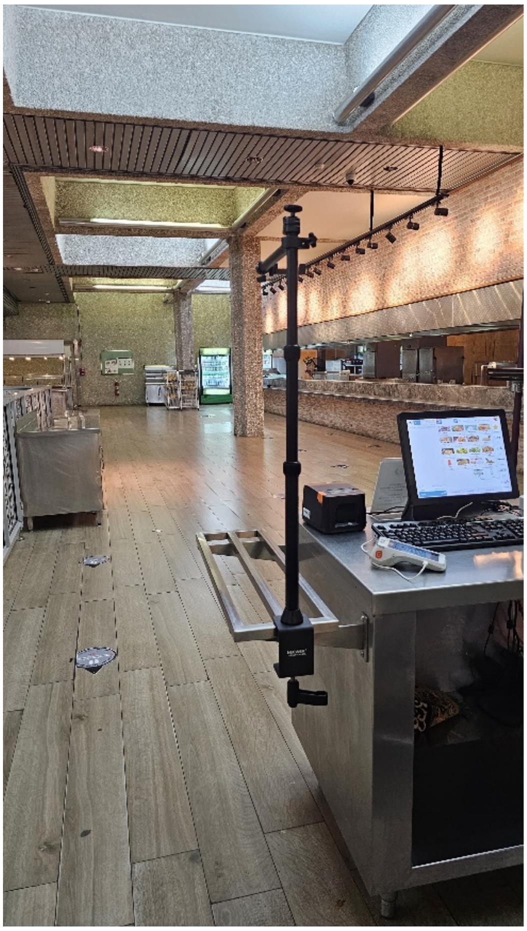
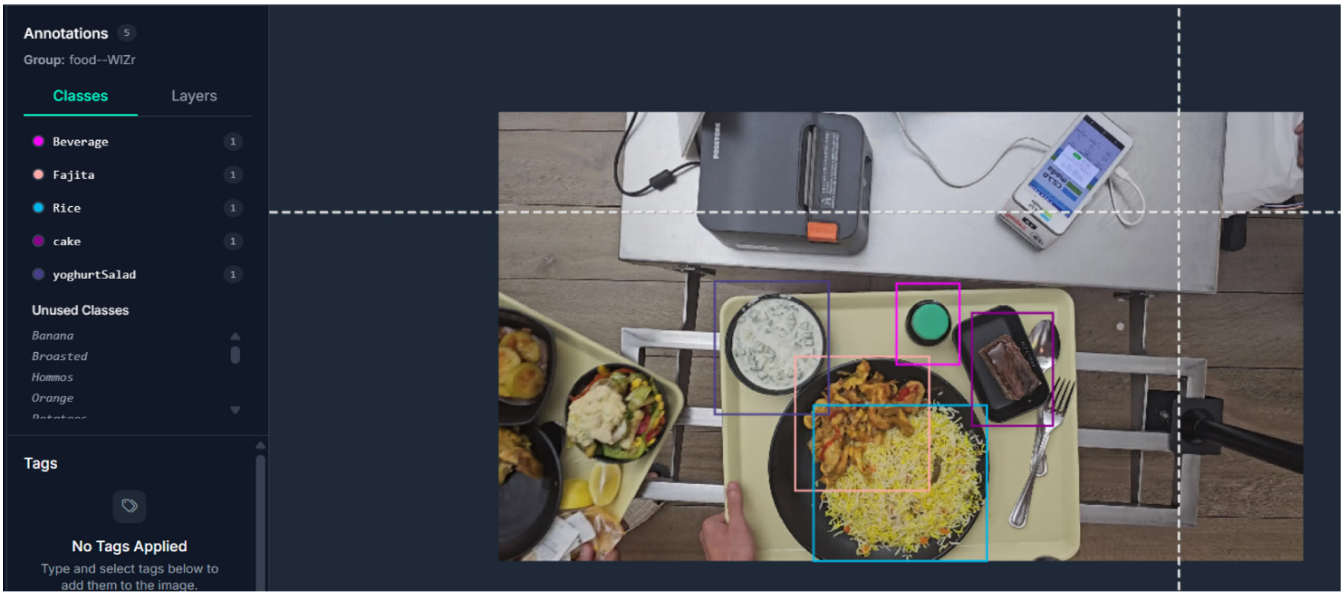


Figure 2. Camera setup used to capture tray images

Figure 3. Sample tray image with multiple food items



No Tags Applied

Type and select tags below to add them to the image.

Figure 4. Bounding boxes used via Roboflow to define the boundaries of each food class

samples from Eastern Mediterranean countries for educational and research purposes[12]. While both datasets include a variety of regional food classes, their images were not captured in real-life environments such as canteens, which reduces their suitability for deployment in practical tray-based applications.

In addition to the dataset content, the annotation strategy plays a crucial role in the performance of food recognition systems. Several techniques have been adopted to annotate food images, each offering different levels of detail and computational requirements. For instance, in Food-101, images are grouped by class, with each image placed in a folder that corresponds to its category. Although this method is simple, it limits the granularity of information in each image. In contrast, semantic segmentation has been applied in some studies[9], where the food content in an image is isolated by separating it from the background. This method allows for more detailed labeling but is computationally expensive. In addition, the use of bounding boxes to isolate individual food items in trays, where food items were marked within trays using the Roboflow platform[10], enables the recognition of multiple food items in a single image, an essential feature of tray-based food classification.

Furthermore, manual annotation has been shown to outperform automatic and AI-assisted annotations in terms of accuracy. For instance, in the UEC-FoodPix Complete dataset, the model trained on fully manual annotations achieved higher accuracy (0.668) than the model trained with automatic labeling (0.560)[8]. This demonstrates the importance of accurate and reliable annotations, especially in complex environments, such as

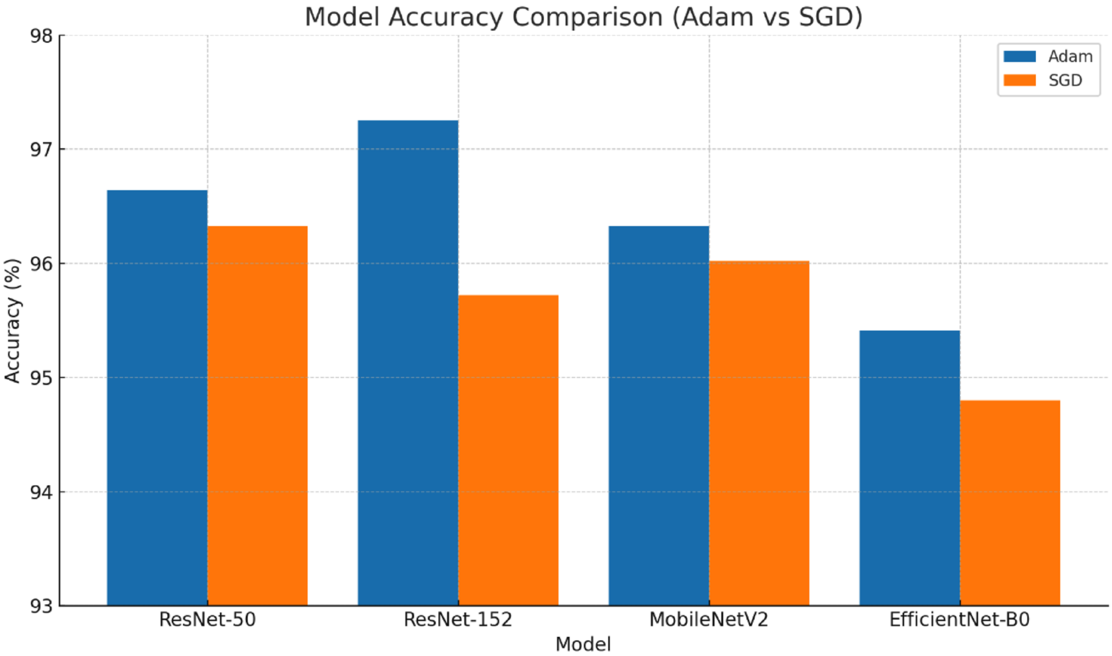


Figure 5. Model accuracies on KF-2025 using Adam and SGD

Table 1. *Evaluation metrics of the models trained on the KF-2025 dataset using SGD and Adam optimizers*

Model	Optimizer	Learning Rate	Accuracy	Precision	Recall	F1-score
ResNet-50	Adam	0.001	96.64%	95.47%	93.74%	93.33%
ResNet-50	SGD	0.001	96.33%	92.35%	91.02%	89.74%
ResNet-152	Adam	0.001	97.25%	97.62%	93.78%	94.53%
ResNet-152	SGD	0.001	95.72%	94.45%	90.05%	89.76%
MobileNetV2	Adam	0.001	96.33%	96.01%	91.20%	92.09%
MobileNetV2	SGD	0.001	96.02%	90.25%	90.16%	89.37%
EfficientNet-B0	Adam	0.001	95.41%	92.40%	91.13%	89.74%
EfficientNet-B0	SGD	0.001	94.80%	95.43%	88.34%	89.08%

university canteens, where food items often overlap and vary in appearance.

Finally, the choice of models and training techniques significantly affects the performance of food recognition systems. For example, early studies that used the Food-101 dataset achieved an accuracy of approximately 50.76%[7]. However, later research utilizing more advanced models and transfer learning techniques achieved accuracies of 80%[13] and 93.26 %, respectively[14]. More recent studies have employed pretrained architectures, such as Residual Networks (ResNet) and EfficientNet, further enhancing food recognition accuracy, reaching 97.54%[15]. These improvements highlight the continuous advancements in the field and the importance of model selection for achieving higher performance.

Earlier studies contained a greater variety of models reporting high metric results. For instance, DenseNet-246 achieved a top-1 accuracy on ImageNet[16]. ResNeXt-101 achieved an accuracy of 79.6 on the ImageNet further[17]. The vision transformer achieved 88.6% accuracy when pretrained on a JFT-300M[18]. However, these models are computationally heavy compared to other networks such as ResNet and MobileNet, making them less optimal for online applications when computational ability is constrained.

3. METHODOLOGY

3.1. Data Collection.

In the KFUPM student restaurant, there are four checkout points where students pay for meals. At one of these points, a phone (Samsung Galaxy S23 Ultra) was mounted on a stand to capture videos (4K resolution, 30 FPS) of the trays from a top-down perspective, ensuring that all items on the tray were visible and clearly detectable. The camera setup is shown in Figure 2.

Using the FFmpeg tool, the recorded videos were decomposed into individual frames. After discarding frames in which the tray contents were obscured by people or suffered from noise and motion blur, 1,039 usable images were extracted. Given that each tray typically includes multiple food components (e.g., rice, dessert, and beverages), as illustrated in Figure 3, it was essential to annotate each item individually to enable more precise downstream processing.

3.2. Annotation.

The next step was to upload the selected images to Roboflow, a platform that provides all the necessary tools for building and deploying computer vision models, including annotation and segmentation. Many KF-2025 images contain more than one type of food on a single plate, making food classification based on whole dishes impractical. Therefore, bounding boxes were used to isolate each food item. In this step, 711 images were manually annotated into 28 food classes using bounding boxes, as illustrated in Figure 4.

3.3. Preprocessing.

The annotated 711 images were randomly split into 563 training, 74 validation, and 74 test images, comprising approximately 80% for training, 10% for validation, and 10% for testing. Subsequently, the annotations were exported in YOLOv8 format. Using the bounding box coordinates, each food item was cropped based on its class labels. For example, all cropped regions labeled as “cake” were grouped under the “cake” class.

Owing to the varying availability of food items in the canteen, there was an inherent imbalance in the dataset. Commonly served items such as beverages, rice, and yoghurt had the most images, with the beverage class containing the most cropped images at 458. In contrast, rarely served items had fewer images, with burgers having the fewest at 6. Due to the extremely low number of samples in the ‘burger’ and ‘coleslaw’ classes, they were removed from the dataset. This resulted in a final dataset of 26 classes and 3,222 cropped images, including training, validation, and test sets.

3.4. Trained Models.

Three transfer-learning models were selected to represent different levels of architectural complexity: MobileNetV2 and EfficientNet-B0 (lightweight), ResNet-50 (moderate), and ResNet-152 (complex).

3.4.1. ResNet (50 and 152).

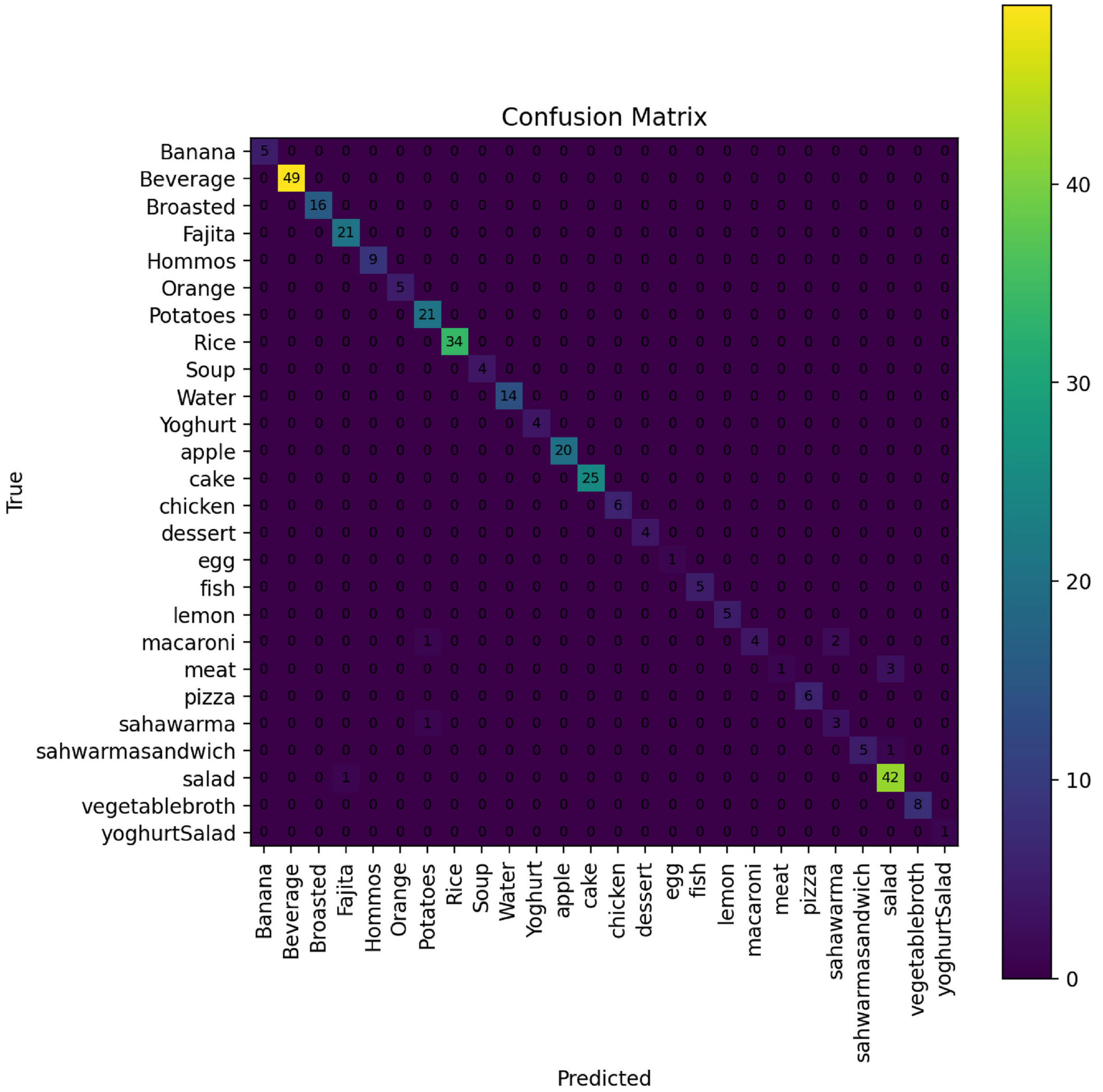
ResNets are convolutional neural networks (CNNs) pretrained on over a million images and are capable of classifying them into 1,000 object categories. ResNet-50 contains 50 layers (48 convolutional layers, one maximum pooling layer, and one average pooling layer)[19] while ResNet-152 consists of 152 layers[20].

3.4.2. MobileNetV2.

A lightweight CNN with 28 layers introduced by Google in 2017 was designed for mobile and embedded devices[21]. It has significantly fewer parameters than other CNNs while maintaining competitive accuracy[22].

3.4.3. EfficientNet-B0.

EfficientNet is a CNN developed using a multi-objective neural architecture search to optimize both accuracy and computational efficiency (FLOPS). The EfficientNet family consists of eight models (B0–B7), where the number of parameters increases with the model complexity[23].



	Banana	Beverage	Broasted	Fajita	Hommos	Orange	Potatoes	Rice	Soup	Water	Yoghurt	apple	cake	chicken	dessert	egg	fish	lemon	macaroni	meat	pizza	sahawarma	sahwarmasandwich	salad	vegetablebroth	yoghurtSalad
Banana	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beverage	0	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Broasted	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fajita	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hommos	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Orange	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Potatoes	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rice	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Soup	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Yoghurt	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
apple	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
cake	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0	0	0	0	0
chicken	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
dessert	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
egg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
fish	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0
lemon	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
macaroni	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4	0	0	2	0	0	0	0
meat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0
pizza	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0
sahawarma	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0
sahwarmasandwich	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1	0	0
salad	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	42	0	0
vegetablebroth	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0
yoghurtSalad	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Figure 6. Confusion matrix of ResNet-152 with the Adam optimizer (the best-performing model). The majority of test samples were correctly classified, with some exceptions noted in the Salad and Shawarma classes

3.5. Training Setup.

Setting the learning rate to 0.001 and batch size to 32, SGD and Adam optimizers were used to train the models on the KF-2025 dataset. Adam was selected for its fast convergence and minimal tuning needs[24], which are suitable for smaller datasets, whereas SGD was included for its strong generalization with sufficient tuning[25]. The models were trained using 15 epochs of fine-tuning, and the entire model was trained. The trained models were then evaluated using accuracy, precision, recall, and F1-score to assess both model performance and the overall quality of the dataset.

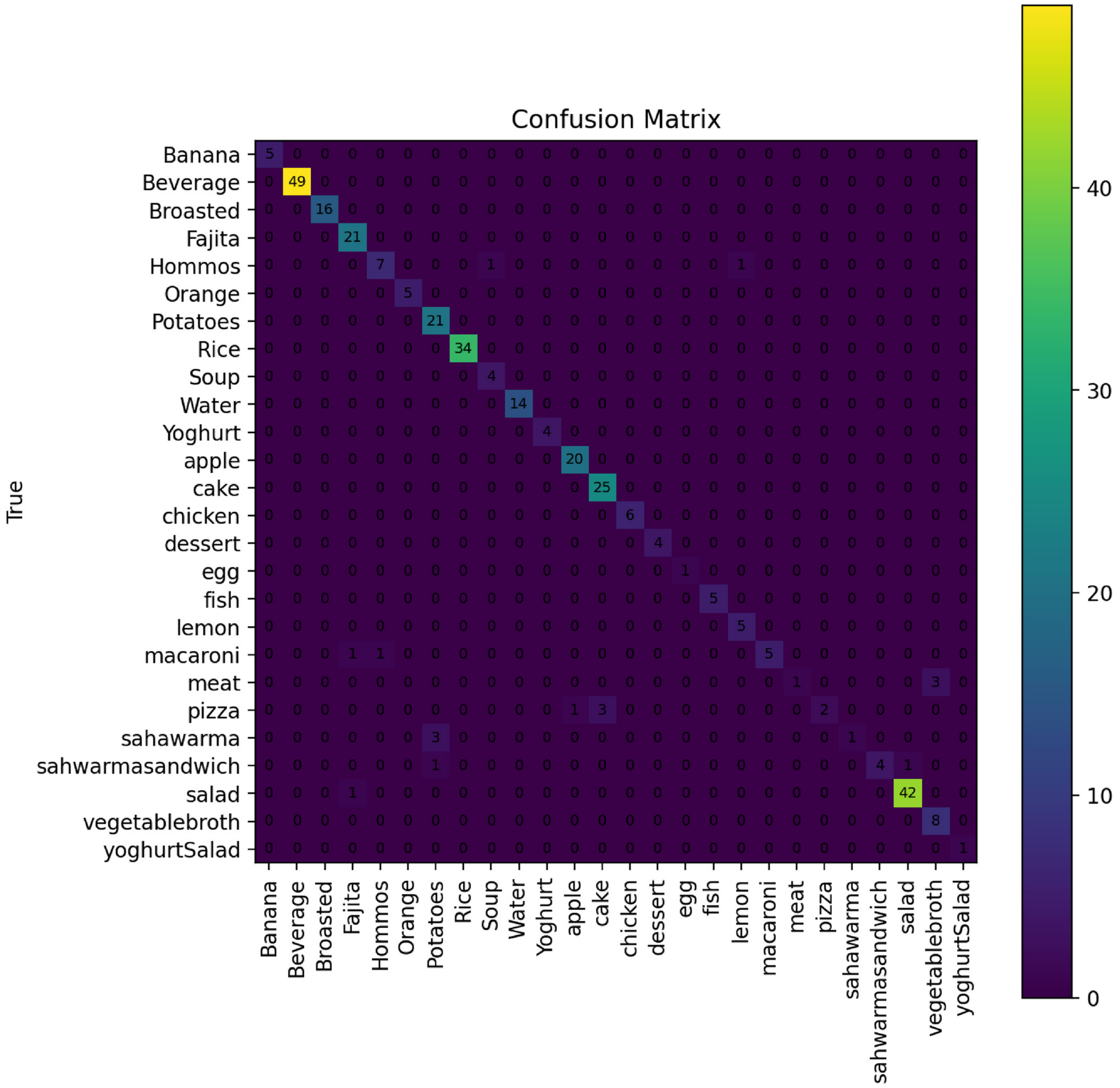
Although the accuracy metric reflects the overall accuracy of the model in recognizing food classes, it can be misleading in imbalanced datasets, where certain classes dominate. Therefore, precision and recall were also considered to ensure that the model was not biased toward the more frequent classes while neglecting the rare ones. Finally, the F1-score provides a balanced measure by combining precision and recall, thereby offering a more inclusive view of the model's performance across all classes.

4. RESULTS

The experimental results are shown in Table 1 and Figure 5.

In addition to the test metric results, it was important to examine the confusion matrices of the experiments to identify the classes that were more frequently misclassified. The confusion matrices for the three selected experiments representing the different models and optimizers are provided in Figures 6 to 9.

Given that the KF-2025 dataset was newly introduced, it is important to compare its results with established food classification benchmarks such as Food-101. To ensure fairness, the same models and optimizers were trained on both datasets for five epochs, with consistent hyperparameters. Table 2 presents the best-performing model-optimizer combinations for each dataset based on accuracy, precision, recall, and F1-score.



	Banana	Beverage	Broasted	Fajita	Hommos	Orange	Potatoes	Rice	Soup	Water	Yoghurt	apple	cake	chicken	dessert	egg	fish	lemon	macaroni	meat	pizza	sahawarma	sahwarmasandwich	salad	vegetablebroth	yoghurtSalad
Banana	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beverage	0	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Broasted	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fajita	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hommos	0	0	0	0	7	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Orange	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Potatoes	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rice	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Soup	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Yoghurt	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
apple	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
cake	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0	0	0	0	0
chicken	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
dessert	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
egg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
fish	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0
lemon	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
macaroni	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0
meat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	3	0
pizza	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	2	0	0	0	0	0
sahawarma	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
sahwarmasandwich	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	0	0
salad	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	42	0	0
vegetablebroth	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0
yoghurtSalad	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Figure 7. Confusion matrix of MobileNetV2 using the SGD optimizer. The matrix shows notable misclassification of Shawarma as Potatoes multiple times

5. DISCUSSION

As shown in Table 1, for the same optimizer type, ResNet architectures consistently outperformed the other models. This can be attributed to the use of residual blocks with skip connections, which enables deeper networks to learn more complex and abstract features effectively[26]. Among the models tested, ResNet-152 achieved the highest performance when paired with the Adam optimizer, likely owing to its increased depth and greater number of trainable parameters compared with ResNet-50. However, ResNet models are more computationally intensive than lightweight architectures such as MobileNetV2 and EfficientNetB0, which limits their deployment on devices with limited processing power. Nonetheless, the fact that the lowest accuracy recorded across all models was 94.80% suggests that even lightweight models can achieve a strong performance when efficiency is a priority.

Regarding the optimizer choice, the models trained with Adam generally achieved a higher accuracy than those trained with

Table 2. *Performance comparison between Food-101 and KF-2025 datasets using their best-performing model and optimizer configurations over five training epochs*

Dataset	Model	Optimizer	LR	Accuracy	Precision	Recall	F1-score
Food-101	ResNet-152	SGD	0.001	80.05%	80.50%	80.05%	80.08%
KF-2025	MobileNetV2	Adam	0.001	94.46%	92.03%	92.03%	92.21%

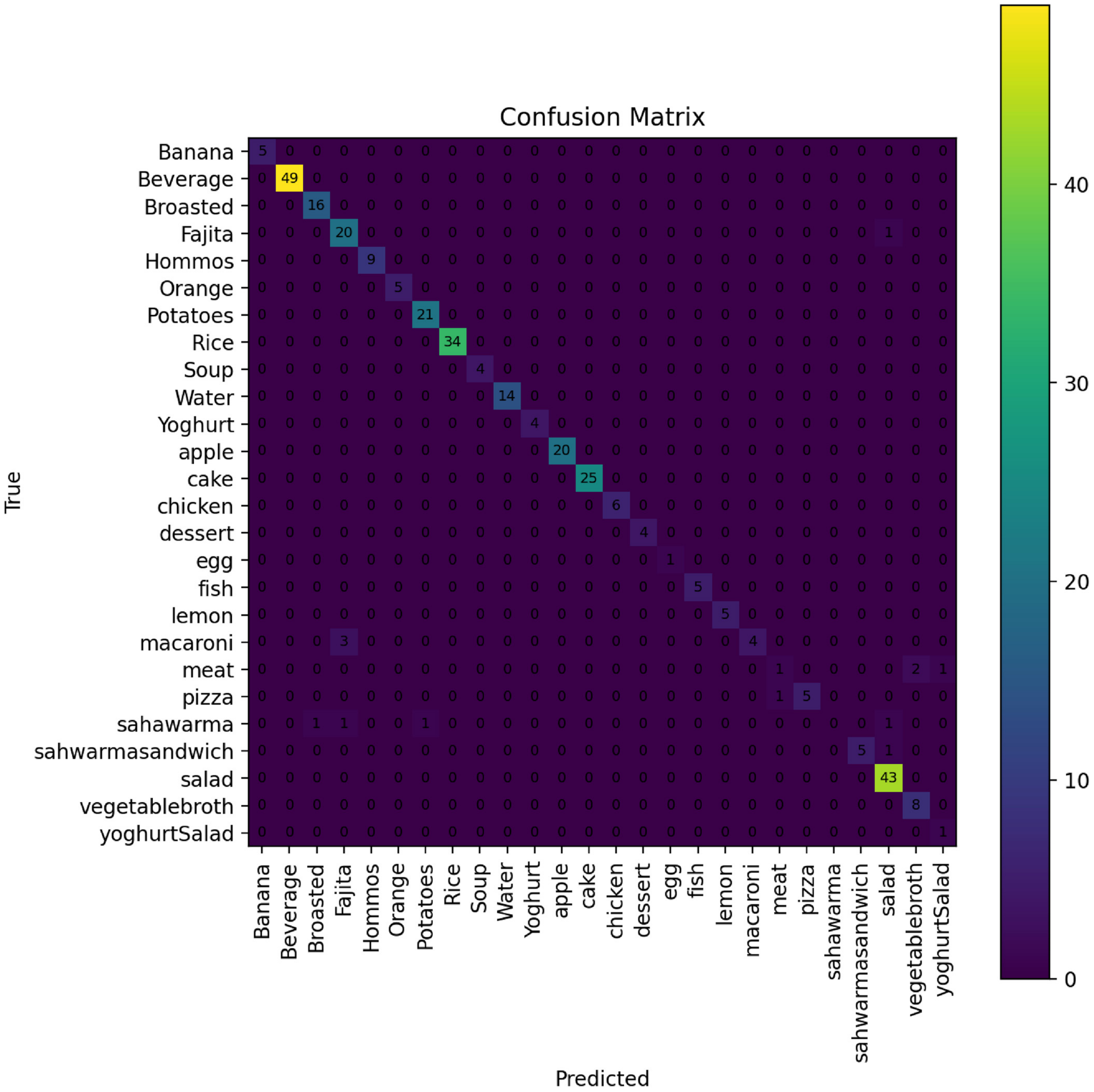


Figure 8. Confusion matrix of EfficientNet-B0 with the Adam optimizer. In addition to misclassification of the Shawarma class, the matrix shows multiple instances of Macaroni being misclassified as Fajita

SGD. This may be attributed to Adam's adaptive learning mechanism and hyperparameters, which typically require minimal tuning[24], making it well suited for scenarios involving relatively small training datasets, as in this study.

Although the precision, recall, and F1-scores were generally high, their lower values compared with the accuracy suggest that the models may favor the majority classes, which is expected given the class imbalance in the dataset. Reducing this imbalance by ensuring more uniform class distributions could help improve these metrics and bring them closer to overall accuracy.

As shown in the confusion matrices (Figures 6, 7, 8, and 9), most classes were correctly identified. However, some misclassifications occurred, likely owing to similarities in ingredients and the visual appearance of certain food items served in the canteen. For instance, in Figure 7, the model misclassified the "Shawarma" class as "Potatoes" three times. This confusion is likely due to the fact that shawarma is often served with fries in the same dish, leading to overlapping visual features. Figure 8 illustrates this point, showing how both classes appear visually similar when served together.

Additionally, as shown in the confusion matrices, particularly in Figure 9, and in the misclassified samples illustrated in Figure 11, images labeled as "Macaroni" were frequently predicted as belonging to other classes. This issue arises from the notable intra-class variability within the Macaroni category, in which plates prepared with different sauces and ingredients are grouped under the same label. Introducing more fine-grained class definitions, for example, by separating macaroni dishes according to sauce type or main ingredients, could help the model capture distinctive features more effectively and improve the classification accuracy.

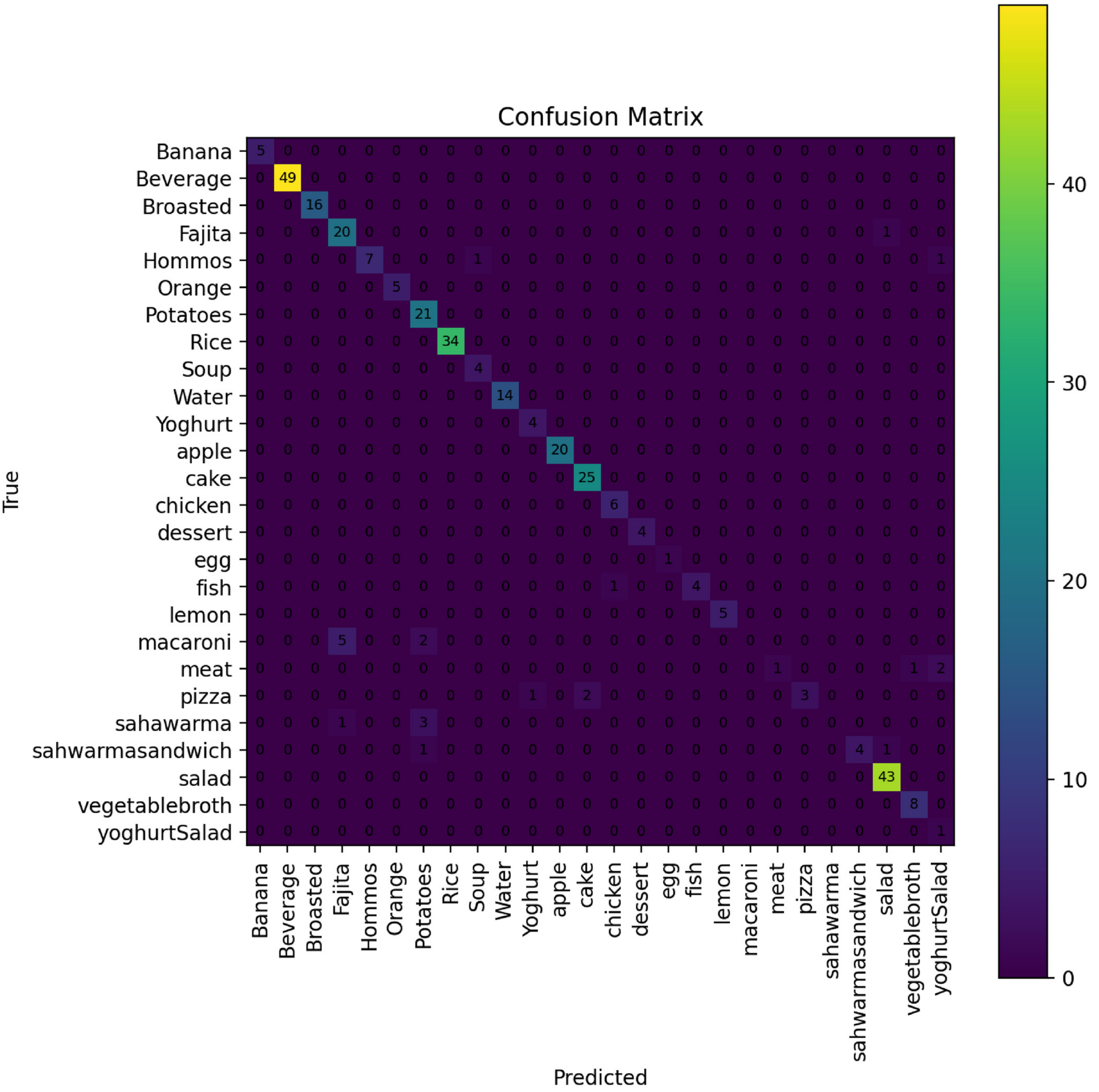


Figure 9. Confusion matrix of ResNet-50 with the SGD optimizer. The confusion matrix confirms the misclassification problem with the Macaroni and Shawarma classes, given the relatively high number of errors

Table 2 presents a performance comparison between Food-101 and KF-2025 under the same training conditions and models. The superior performance of KF-2025 can be attributed to several factors. First, KF-2025 includes fewer classes, which reduces overall classification complexity. Second, Food-101 contains images collected from diverse internet sources, introducing high intraclass variability in visual appearance. Therefore, effective feature extraction for Food-101 requires extensive training.

Furthermore, the close match between accuracy and F1-score in Food-101 suggests a relatively balanced class distribution, whereas the KF-2025 dataset suffers from a class imbalance. This imbalance likely contributes to the performance gap between the accuracy and F1-score in KF-2025. Thus, increasing the number of samples in underrepresented classes may improve model fairness and yield more consistent metrics, similar to the results observed for Food-101.

Despite its small size and imbalance, the strong results of the KF-2025 highlight its practical effectiveness for real-world canteen classification tasks, especially in settings with a limited number of food categories.

5.1. Limitations and Future Work.

Despite the promising results, the current dataset includes only 26 food classes, which do not fully represent the complete variety of meals served in student restaurants. Some food types were excluded due to the limited number of available samples, which may have affected the overall inclusiveness of the dataset. Additionally, the dataset suffers from class imbalance, where certain food categories are overrepresented compared to others. This imbalance may have influenced the model's performance, particularly in underrepresented classes. Furthermore, techniques that typically improve model generalization, such as data augmentation, were not applied in this study. To address these limitations, future work will involve

Figure 10. Examples showing visual similarity between Shawarma (left) and Potatoes (right), which may contribute to misclassification

collecting additional images across different days of the week to cover full menu variations, using polygon-based annotations for more accurate labeling, applying data augmentation techniques to enhance model robustness, and balancing class distributions during data collection.

A key focus of future work should be the deployment of trained models and associated challenges. While ResNet models can be deployed on embedded platforms, they typically require computationally advanced devices, such as Jetson Xavier, to achieve adequate performance[27]. In actual canteen settings, deploying multiple units at different checkout points can significantly increase costs, and installation may be constrained by limited space. In contrast, lightweight architectures, such as MobileNetV2 and EfficientNet-B0, can be deployed on smaller and more cost-effective devices, such as Raspberry Pi[28], although they generally report slightly lower accuracy. Given the importance of efficient and real-time recognition processes, high-accuracy heavy models with powerful hardware versus lightweight models with affordable devices should be evaluated to determine the most practical setup for canteen environments.

6. CONCLUSIONS

This study presented KF-2025, the first real-world, tray-based food image dataset collected directly from the students' canteen at the King Fahd University of Petroleum and Minerals. KF-2025 captures real, context-specific images from the KFUPM student restaurant, targeting automation use cases such as self-checkout and nutritional monitoring. The dataset was constructed by recording high-resolution videos at the checkout points, extracting thousands of frames, and manually annotating food items

Figure 11. Examples of Macaroni images misclassified due to high intra-class variability, where items with different shapes, colors, and features are grouped under the same class

using bounding boxes in RoboFlow. After preprocessing, 3,222 cropped food images across 26 classes were prepared for training, validation, and testing.

To evaluate the suitability of the dataset for real-world deployment, four pre-trained models, ResNet-50, ResNet-152, MobileNetV2, and EfficientNet-B0, were trained using both the Adam and SGD optimizers. The best performance was achieved by ResNet-152 with Adam, reaching 97.25% accuracy, 97.62% precision, 93.78% recall, and 94.53% F1-score. Across all the models, Adam consistently outperformed SGD, highlighting its robustness for relatively small and imbalanced datasets. The strong performance across multiple architectures reflects the clarity of the collected images, quality of the annotations, and overall reliability of KF-2025. Despite its limited size compared with large-scale public datasets, KF-2025 proved effective for food classification tasks in a real university environment.

This dataset provides a valuable foundation for future research and applications, including real-time food recognition, automated checkout systems, and nutrition monitoring. Further improvements, such as increasing the sample size to address class imbalance, applying data augmentation, and testing on resource-constrained devices, can enhance its practicality and impact for smart canteen system applications.

7. ETHICAL AND PRIVACY CONSIDERATIONS

Because video recording took place at a sensitive location, namely the checkout point of the university canteen, specific ethical and privacy concerns were considered. These included potential exposure to students' faces and any visible banking information. The following measures were implemented to ensure full compliance with privacy and ethical standards:

1. The camera was positioned to capture only the food trays from a top-down angle, ensuring that the students' faces were not included in any recordings.
2. All extracted frames used for the dataset were reviewed to confirm that no credit cards, phones, or personal banking details appeared in any image.
3. The data collection process was carried out with approval from the KFUPM Food Services, and the dataset remained private and confidential, shared only for academic evaluation purposes, and not publicly released.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Saad Obada Alshami – *Department of Control and Instrumentation Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;*
0009-0007-5809-1215
Email: s202182770@kfupm.edu.sa

Corresponding Author

Ali Nasir – *Research Mentor, Assistant Professor, Department of Control and Instrumentation Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;*
0000-0002-7789-428X
Email: ali.nasir@kfupm.edu.sa

ACKNOWLEDGEMENTS

This research was conducted with the support of the SDAIA-KFUPM Joint Research Center for Artificial Intelligence, which provided access to practical equipment and high-specification remote workstations for conducting the experiments. The Department of Control and Instrumentation Engineering and the KFUPM Undergraduate Research Office (URO) provided additional guidance and support.

DATA AVAILABILITY STATEMENT

The dataset will be accessible upon reasonable request from the authors.

REFERENCES

(1) Brunelli, L., Masiero, C., Tosato, D., Beghi, A. & Susto, G. A. Deep learning-based production forecasting in manufacturing: A packaging equipment case study. *Proc. Manuf.* **38**, 248–255 (2019).

(2) Abosaq, H. A. et al. Unusual driver behavior detection in videos using deep learning models. *Sensors* **23**, 311 (2022).

(3) Clements, J. M., Xu, D., Yousefi, N. & Efimov, D. Sequential deep learning for credit risk monitoring with tabular financial data. *arXiv preprint* arXiv:2012.15330 (2020).

(4) Deng, L. et al. Mixed dish recognition with contextual relation and domain alignment. *IEEE Trans. Multimed.* **24**, 2034–2045 (2021).

(5) Wu, X., Fu, X., Liu, Y., Lim, E.-P., Hoi, S. C. H. & Sun, Q. A large-scale benchmark for food image segmentation. *arXiv preprint* arXiv:2105.05409 (2021).

(6) Lu, Y. et al. A multi-task learning approach for meal assessment. *arXiv preprint* arXiv:1806.10343 (2018).

(7) Bossard, L., Guillaumin, M. & Van Gool, L. Food-101 – Mining discriminative components with random forests. *Lect. Notes Comput. Sci.* **8694**, 446–461 (2014).

(8) K. Okamoto and K. Yanai, "UEC-FoodPIX Complete: a Large-Scale Food Image Segmentation Dataset," in Lecture notes in computer science, 2021, pp. 647–659. doi: 10.1007/978-3-030-68821-9_51.

(9) Ciocca, M., Napoletano, P. & Schettini, R. Food recognition: A new dataset, experiments, and results. *IEEE J. Biomed. Health Inform.* **21**, 588–598 (2017).

(10) A. Karabay, A. Bolatov, H. Varol, and M.-Y. Chan, "A central Asian Food dataset for personalized dietary interventions," Nutrients, vol. 15, no. 7, p. 1728, Mar. 2023, doi: 10.3390/nu15071728.

(11) M. Y. Ansari and M. Qaraqe, "MEFood: a Large-Scale Representative benchmark of quotidian foods for the Middle East," IEEE Access, vol. 11, pp. 4589–4601, Jan. 2023, doi: 10.1109/access.2023.3234519.

(12) "EMFID." https://emfid.org/frontend/web/index.php?r=site/goals

(13) V, G., Vutkur, P. & V, P. Food classification using transfer learning technique. *Glob. Transit. Proc.* **3**, 225–229 (2022).

(14) Tripathi, P. Transfer learning on deep neural network: A case study on Food-101 food classifier. Int. J. Eng. Appl. Sci. Technol. 5, (2021).

(15) Philip, A. Optimizing Food-101 classification with transfer learning: A fine-tuning approach using EfficientNetB0. *Int. J. Intell. Inf. Syst.* **13**, 59–77 (2024).

(16) G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," arXiv (Cornell University), Jan. 2016, doi: 10.48550/arxiv.1608.06993.

(17) S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, “Aggregated residual transformations for deep neural networks,” arXiv (Cornell University), Jan. 2016, doi: 10.48550/arxiv.1611.05431.

(18) A. Dosovitskiy et al., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” arXiv (Cornell University), Jan. 2020, doi: 10.48550/arxiv.2010.11929.

(19) C. K. K. Reddy et al., “A fine-tuned vision transformer based enhanced multi-class brain tumor classification using MRI scan imagery,” Frontiers in Oncology, vol. 14, Jul. 2024, doi: 10.3389/fonc.2024.1400341.

(20) “Abnormality Detection in Mammography using Deep Convolutional Neural Networks,” IEEE Conference Publication — IEEE Xplore, Jun. 01, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8438639

(21) Y. Li, J. Xue, K. Wang, M. Zhang, and Z. Li, “Surface Defect Detection of Fresh-Cut Cauliflowers Based on Convolutional Neural Network with Transfer Learning,” Foods, vol. 11, no. 18, p. 2915, Sep. 2022, doi: 10.3390/foods11182915

(22) A. G. Howard et al., “MobileNets: efficient convolutional neural networks for mobile vision applications,” arXiv (Cornell University), Jan. 2017, doi: 10.48550/arxiv.1704.04861.

(23) M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” arXiv (Cornell University), Jan. 2019, doi: 10.48550/arxiv.1905.11946

(24) D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” arXiv (Cornell University), Jan. 2014, doi: 10.48550/arxiv.1412.6980.

(25) P. Zhou, J. Feng, C. Ma, C. Xiong, S. C.-H. Hoi, and W. E, “Towards theoretically understanding why SGD generalizes better than ADAM in deep learning,” arXiv (Cornell University), Jan. 2020, doi: 10.48550/arxiv.2010.05627.

(26) Singh, P. Detection of Gender in Crowds Using ResNet Model. J. Electrical Systems, 389–408 (2024).

(27) Y. Han, Z. Nan, S. Zhou, and Z. Niu, “DVFS-Aware DNN Inference on GPUs: Latency modeling and Performance analysis,” arXiv (Cornell University), Feb. 2025, doi: 10.48550/arxiv.2502.06295.

(28) R.-Y. Ju, T.-Y. Lin, J.-H. Jian, and J.-S. Chiang, “Efficient convolutional neural networks on Raspberry Pi for image classification,” arXiv (Cornell University), Jan. 2022, doi: 10.48550/arxiv.2204.00943.

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article
Computer Science/Electrical Engineering

Data-Driven Deep Learning Fault Diagnosis Model for Smart Grids Under Load Demand Uncertainty

Khalid Alfuwail, Hamza M. Anwar, Md Shafiullah*, Md Mahfuzur Rahman, and Mohamed Ali Abido

Cite https://doi.org/10.64589/juri/215357

Submitted: March 20, 2025 **Revised:** November 25, 2025 **Accepted:** December 09, 2025

ABSTRACT

Faults are common hazards in electricity distribution networks and should be diagnosed immediately to prevent harm and restore the power supply. This study proposes a hybrid approach for fault diagnosis that combines deep learning models and advanced signal-processing approaches. The proposed method creates a standard four-node test distribution grid in MATLAB/Simulink, extracts features using discrete wavelet transform (DWT) and short-time Fourier transform (STFT), and trains deep learning models to identify, categorize, and pinpoint faults. The results show that while both STFT and DWT perform well in fault detection, STFT demonstrates superior robustness in fault localization and categorization under noisy conditions owing to its uniform time–frequency resolution. The proposed framework uniquely integrates load demand uncertainty, measurement noise, and a hybrid MATLAB–Python pipeline, providing a more realistic evaluation than previous studies. The proposed models are not affected by variations in the pre-fault loading conditions, fault inception angle, or fault resistance. The proposed framework is suitable for deployment in real urban distribution grids with significant electrical interference. The main advantage is the reduced fault location time, which reduces outages and improves reliability.

Keywords: deep neural networks, fault detection, distribution grids, time frequency analysis, short-time fourier transform, wavelet transform

1. INTRODUCTION

Electricity is a necessity in the modern world, serving as the basis for power generation, transmission, and distribution. However, as energy demand increases, the grid becomes more crowded and thus more prone to malfunctions, such as short circuits, which can cause power outages[1–3]. These faults can disrupt daily routines and essential tasks. Therefore, it is essential to provide steady and seamless electrical transmission from utilities to consumers. Identifying the location of a grid breakdown quickly is crucial for minimizing customer minutes lost and accelerating electricity restoration. Effective methods are required to accurately identify, categorize, and pinpoint defects to ensure the safety and efficiency of electrical grids[4–6].

Grid fault detection techniques are categorized into knowledge-based, traveling-wave-based, and impedance-based methods[7]. The impedance-based approach estimates the fault distance using voltage and current measurements, offering simplicity and cost-effectiveness[8]. However, it suffers from inaccuracies owing to iterative calculations, dynamic loads, noise, and laterals[9]. The traveling wave method analyzes fault-generated pulses using the arrival time, wave velocity, phase, and polarity to locate faults from one or both line ends[10,11]. Although effective for long transmission lines, it is unsuitable for short distribution lines (accuracy: 100–500 m) and requires costly infrastructure[12].

Knowledge-based approaches, particularly machine learning (ML), are promising for fault diagnosis. ML techniques adapt by identifying patterns in the training data without explicit programming, thereby proving effective for fault locations in distribution grids[13]. They excel in handling fault complexities and can be integrated with signal processing to obtain hybrid solutions. For example, Mamuya[2] combined a wavelet transform (WT) for feature extraction with artificial neural networks (ANN) to enhance the fault detection accuracy. The versatility of ML and data-driven learning make it a robust alternative to traditional methods.

Wavelet-based approaches: Rizeakos et al.[19] combined continuous wavelets transform (CWT) with CNNs for fault location, whereas Naidu et al.[21] integrated discrete wavelets transform (DWT) with an ANN optimized via particle swarm optimization for high-impedance fault (HIF) location. Iliyaeifar and Hadaeghi[22] paired DWT with an extreme learning machine (ELM) for HVDC systems, and Rezaee Ravesh et al.[23] used wavelet packets transform with an ANN for hybrid transmission lines. While these approaches capture multi-resolution signal features effectively, most have been validated only on HVDC or hybrid systems with limited testing under noisy AC distribution conditions.

Fourier-based approaches: Grcić et al.[24] applied short-time Fourier transform (STFT) with ML classifiers in DC microgrids, but lacked ANN-based classification, whereas Shafiullah et al.[7] used S-transform for optimized fault location in distribution grids. These studies indicate the potential of STFT for fault diagnosis but leave open the question of its comparative performance against wavelet methods under noise and load variations.

Other ML-based approaches: Lan et al.[14] used empirical mode decompositions (EMD) with a 1D-CNN for HVDC gearbox fault detection, although only in simulations. Yu et al.[15] introduced a directed tree-based traveling wave method and Babayomi and Oluseyi[16] applied an adaptive neuro-fuzzy inference system, with both studies having limitations in repeated fault handling. Advanced models, such as histogram-based gradient boosting (HGB)[17] and graph convolutional networks (GCNs)[18], have shown promise in specific scenarios, but have not been benchmarked against hybrid STFT/DWT approaches.

Research gap: Across these studies, there has been limited validation of realistic and noisy AC distribution systems, insufficient analysis of repeated faults, and no systematic comparison between STFT and DWT integrated with deep learning for simultaneous detection, classification, and location tasks under-load demand uncertainty. This study addresses these limitations.

Most knowledge-based methods rely on MATLAB toolboxes, such as Statistics and Machine Learning[25] and Deep Learning[26]. While effective, exploring open-source alternatives, such as Python, could improve the accessibility and comparison of model efficacy[27]. Python offers powerful ML libraries (scikit-learn, TensorFlow, and Keras)[28–32] and signal processing tools with simpler syntax than MATLAB. Combining MATLAB signal processing with Python ML capabilities may optimize fault detection in power grids. Future work should leverage Python open-source libraries to enhance fault identification while maintaining performance.

Current fault diagnosis methods often overlook key factors, such as dynamic load variations, measurement noise, and practical deployment constraints. This study addresses these limitations by proposing a hybrid fault diagnosis approach for distribution networks that explicitly incorporates the load demand uncertainty, multiple noise levels, and a new set of statistical features. This combination, along with MATLAB–Python integration, enables a more comprehensive and realistic evaluation than existing studies. This study employs Python-based deep learning models and makes the following contributions:

- The uncertainty around the fault information and load demand is considered when modeling the test distribution feeders.
- From the collected data, features are extracted using DWT and STFT.
- Deep neural networks with several hidden layers are used to create smart defect diagnostic models (detection, classification, and localization) using Python and its modules.
- For fault location, the train-test split for both the complete dataset and individual fault types must be varied to determine the split that yields the best results.
- The sensitivity of the resulting models is evaluated while considering the network-related uncertainty and measurement noise.

The remainder of this paper is structured as follows: Section 2 details the deep learning and signal processing methods for fault detection. Section 3 describes the distribution feeder model. Section 4 explains the feature extraction and proposed fault detection approach. Section 5 compares the model results. Finally, Section 6 summarizes the study and provides the conclusions.

2. METHODS

2.1. Signal Processing Techniques.

Signal processing transforms signals to reveal key components[33]. For grid fault diagnosis, mathematical transformations (e.g., STFT and DWT) extract features from the current signals, which are then fed to the ML algorithms. This study utilizes both STFT and DWT as feature extraction tools. The following section explains the theoretical foundations of these transformers.

2.1.1. Short-time Fourier Transform.

The Fourier transform decomposes signals into frequency components but loses temporal information[34]. STFT overcomes this problem by analyzing equal-length signal segments and providing time-localized frequency data[34]. It multiplies the signal $x(t)$ by a short-duration window function $w(t)$ (e.g., Blackman or Hann). The STFT of $x(t)$ is expressed as follows:

$$X(\tau, f) = \int_{-\infty}^{+\infty} x(t)\, w(t-\tau)\, e^{-i2\pi ft} dt \tag{1}$$

where t is time, f is frequency, and $w(t-\tau)$ is the window function centered at τ. This enables joint time-frequency analysis for fault detection.

Most computer applications, including MATLAB, use a fast Fourier transform (FFT) to compute STFT features[35]. The STFT of a discrete signal $x(n)$ is determined as follows:

$$X_m(f) = \sum_{n=-\infty}^{+\infty} x(n)\, w(n-mR)\, e^{-i2\pi fn} \tag{2}$$

The STFT process uses a frame index n, window $w(t)$ centered at mR, and R sample intervals[36]. As $w(n)$ slides along the signal length M, the discrete Fourier transform (DFT) computes each windowed segment. Overlap $L = M - R$ occurs as adjacent segments hop R samples, creating overlapping analysis windows for a continuous time-frequency representation[37].

2.1.2. Wavelet transform.

WT overcomes the Fourier transform limitations by providing a multi-resolution time-frequency analysis[38]. Using variable-width wavelets derived from a mother wavelet $\psi(t)$ (e.g., Morlet or Daubechies), it captures signal components at different scales. Unlike the fixed windows in STFT, wavelets adapt their widths to be narrow for high frequencies and wide for low frequencies. CWT slides these scalable wavelets along the signal, multiplying and analyzing them from high to low frequencies. Mathematically, the CWT of signal $f(t)$ is represented as $F(\tau, s)$, where τ is translation and s is scale. This approach enables precise localization of both transient and sustained features in fault signals:

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t)\, \psi^{*}\left(\frac{t-\tau}{s}\right) dt \qquad (3)$$

CWT uses $\psi^{*}\left(\frac{t-\tau}{s}\right)$ *as the daughter wavelet, where* ψ is the complex conjugate of the mother wavelet, s controls the scale/frequency, and τ determines the position[39]. For computational efficiency, DWT discretizes this approach. The DWT $D[a,b]$ of signal $f[t_m]$ analyzes discrete translations (b) and dyadic scales ($a = 2^j$), enabling efficient multi-resolution analysis, as follows:

$$D[a,b] = \frac{1}{\sqrt{b}} \sum_{n=0}^{p-1} f[t_m]\, \psi\left[\frac{t_m - a}{b}\right] \qquad (4)$$

where a and b are both integers, the scale parameter is $s = b$, and the translation parameter is $\tau = a$[39]. To determine the approximation and detail coefficients (low- and high-frequency components, respectively) of the input signal, DWT uses a wavelet-based technique known as multi-level decomposition, in which the signal is processed through many low- and high-pass filters [4]. An example of this process is shown in Figure 1.

2.2. Deep Neural Network.

ML enables systems to learn patterns from training data and make predictions. Inspired by biological neurons, ANNs are fundamental ML models composed of interconnected (input, hidden, and output) layers. Feed-forward neural networks (FNNs), which are the simplest ANNs, process data through weighted connections and biases between layers (Figure 2)[40]. During training, the input data propagates forward by adjusting the weights to minimize errors. Deep neural networks (DNNs) extend this approach to multiple hidden layers for complex pattern recognition. Typically, datasets are split into training and testing subsets to evaluate their performance. These architectures form the basis of advanced fault diagnosis systems.

Neural networks apply activation functions (ReLU, sigmoid, tanh, and softmax) to weighted inputs, thus introducing nonlinearity into complex pattern learning[41]. During training, the output layer errors (predicted versus desired) drive backpropagation, which is an optimization process that adjusts weights and biases to minimize loss[4]. This iterative refinement enables the DNNs to model intricate relationships. After training, the models are evaluated using unseen test data by comparing predictions against known outputs and quantifying performance through error metrics. The combination of forward propagation (feature transformation) and backward error correction (gradient-based optimization) allows deep networks to progressively improve their predictive accuracy through exposure to labeled training samples.

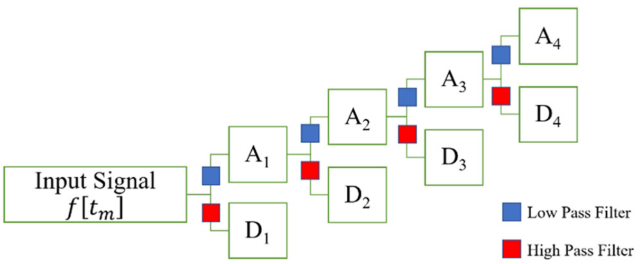


Figure 1. An example of multi-level decomposition up to four levels. Before each level, the signal is passed through a high-pass filter, from which the detailed coefficients (D_i) are obtained, and a low-pass filter, from which the detailed coefficients (A_i) are obtained

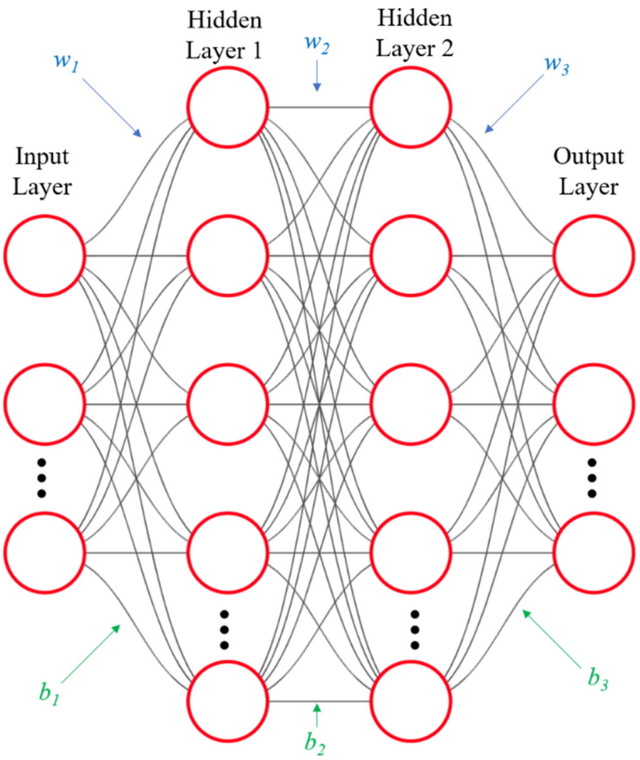


Figure 2. DNN model used for fault detection (binary), fault classification (seven fault types), and fault location (continuous value between 5–15 km). The input vector contains either 108 STFT-based features or 144 DWT-based features per sample

3. DISTRIBUTION NETWORK MODEL UNDER LOAD DEMAND UNCERTAINTY

Figure 3 shows the 60 Hz test feeder with transformers, a lumped load, a transmission line, and four nodes. The distribution grid parameters are listed in Table 1. All data in this study were generated using MATLAB/Simulink "Simscape Electrical" with discrete power systems blockset (PSB)[42]. Fault events were labeled by type, location, resistance, and inception angle directly from the simulation parameters, ensuring an accurate ground truth for training and testing. A 20 kHz sampling rate (333 samples/cycle) and two-block transmission lines enabled fault simulations at multiple locations. The dataset contained 14,000 samples per signal processing method (STFT and DWT), evenly distributed across seven fault types and non-fault cases, with four noise levels (0, 20, 30, and 40 dB). This resulted in a balanced dataset of 64,000 samples (4 noise levels × 8 conditions × 2,000 samples).

4. PROPOSED FAULT DIAGNOSIS METHOD

This section details the steps involved in developing the proposed fault diagnostic approach. It covers the application of flaws; the parameters of signal processing techniques, including the

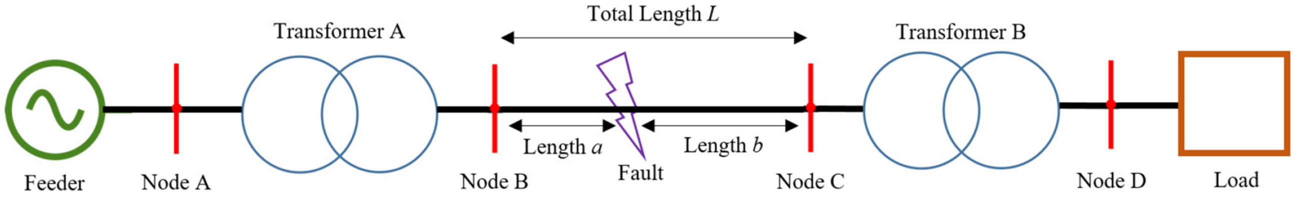


Figure 3. Modeled four-node distribution grid used for simulation. The system consists of transformers, a lumped load, and transmission lines. Faults are applied at different locations along the distribution line between Nodes B and C

features extracted from them; the preparation of data; and the construction of deep learning models for fault diagnosis.

4.1. Fault Application.

Three-phase fault currents were recorded from Node B (Figure 3) for various fault types (AG, BG, CG, ABG, BCG, ACG, and ABCG) applied between nodes B and C, capturing one pre-fault and one post-fault cycle. The fault parameters (type, distance, resistance, and inception angle) and load demands were varied systematically. The non-fault data were generated by varying only the load. The MATLAB/Simulink "awgn" function created signals with 0, 20, 30, and 40 dB noise levels[4]. Figure 4 shows a sample fault current waveform. This comprehensive dataset enables a robust analysis of fault characteristics under different system conditions and noise environments.

4.2. Feature Extraction.

This study generated 64,000 samples (16,000 per noise level) using STFT and DWT in MATLAB, covering seven fault types and normal conditions. Each noise level (0, 20, 30, and 40 dB) contained 2,000 samples per fault scenario ($8\ \mathit{conditions} \times 2,000 = 16,000$). The subsequent sections detail the feature extraction parameters for both signal processing methods.

4.2.1. STFT parameters.

The MATLAB "stft" function processed three-phase current signals using a 128-sample Hann window with 96-sample overlap and 128-point FFT. This produced a time-frequency matrix for each phase. From each matrix, the following statistical features were extracted for each phase:

- The maximum values and energies within each of the 17 frequency columns, resulting in 34 features (17 maxima and 17 energy values).

Table 1. *Test four-node distribution grid parameters*[4]

Component	**Parameter**	**Value**
Distribution Line	Length	20 km
	Positive Sequence Resistance	0.1153 Ω/km
	Positive Sequence Inductance	1.05 mH/km
	Positive Sequence Capacitance	11.33 nF/km
	Zero Sequence Resistance	0.413 Ω/km
	Zero Sequence Inductance	3.32 mH/km
Lumped Load	Active Power	6 MW
	Reactive Power	3 MVar (Inductive)
System Variation	Load Demand Uncertainty	±5%
Fault Parameters	Impedance (R~f~)	0.1 – 15 Ω
	Inception Angle	0–360°
	Location	5–15 km
Transformer A	Voltage Ratio	33/11 kV
Transformer B	Voltage Ratio	11/4.16 kV
Transformers A and B	Power Rating	12 MVA
	Winding 1 (R, L)	0.0434 Ω, 3.454 mH
	Winding 2 (R, L)	68.88 µΩ, 5.481 µH
	Magnetization Branch (R~m~, L~m~)	26042 Ω, ∞ H
	Connection	Y~g~-Δ

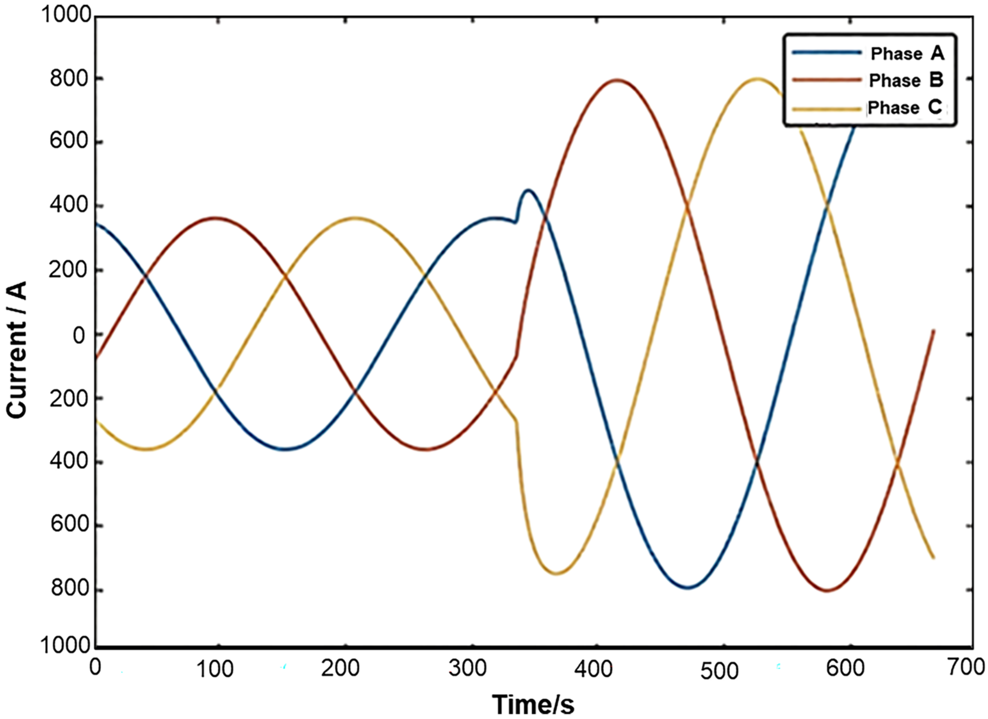


Figure 4. Simulated three-phase current waveform for an ABCG (three-phase-to-ground) fault

- The standard deviation of the maximum values across the time frames (rows) and frequency bins (columns), yielding two features.
- The overall mean magnitude of the time-frequency matrix, adding one final feature.

This resulted in 36 features per phase (17+17+1+1). Using three-phase signals, each sample produced 108 features (36×3). This feature set effectively captured the time-frequency characteristics and statistical properties of fault transients. The frequency range was centered during STFT computation. This feature extraction method captured both time-frequency characteristics (via STFT matrices) and statistical properties, enabling comprehensive fault representation. This approach systematically transformed raw current signals into a discriminative feature set suitable for ML classification while maintaining physical interpretability through energy and deviation metrics.

4.2.2. DWT parameters. MATLAB *wavedec*, *appcoef*, and *detcoef* functions performed seven-level DWT decomposition using 'db4' wavelet, extracting one approximate and seven detailed coefficients. Six statistical measures (std, mean, entropy, energy, kurtosis, and skewness) were computed for each coefficient, yielding 144 features. The implementation details are provided in Shafiullah et al.[4].

4.3. Input/Output Features and Data Preparation. The extracted features were stored as .mat files and imported into Python using the SciPy *loadmat* function[45], and converted to Pandas DataFrames[31]. Each dataset included (1) binary fault indicators (1/0), (2) fault type labels $(1 - 7)$, and (3) fault distances from Node B. For each noise level $(0, 20, 30,\ and 40\ dB)$, separate datasets were created by combining all fault scenarios $(seventypes + normal)$, resulting in eight datasets per signal processing method (STFT/DWT). Data were shuffled to prevent pattern memorization. Feature normalization (0–1 range) was performed using the scikit-learn *MinMaxScaler*[30], which is particularly critical for regression tasks, such as fault distance prediction. This preprocessing pipeline ensured compatibility with Python deep learning frameworks while maintaining the physical meaning of the fault characteristics. The structured datasets enabled both classification (fault detection/typing) and regression (location estimation) tasks, with normalized features improving the model convergence during training. The input to the deep learning models consisted of statistical features extracted from STFT or DWT, whereas the outputs varied by task: binary labels (1 = fault, 0 = no fault) for detection, multiclass labels (1–7) for fault types, and continuous values (5–15 km) for fault location. No manual feature engineering was performed and the features were fully derived from the raw current signals.

4.4. Deep Learning Models. The Keras library[29] was used to build three ML models per signal processing method: fault detection, classification, and localization. The classification/localization models used DNNs with two optimized hidden layers (to balance complexity and performance). The hyperparameters were systematically tuned via trial-and-error to achieve optimal results. Deep learning architectures were selected because of their enhanced pattern-recognition capabilities in complex fault-diagnosis tasks.

4.5. Performance Evaluation Metrics. Model performance was evaluated using task-specific metrics: accuracy for fault detection/classification (correct predictions/total samples) and regression metrics (R^2, mean absolute error (MAE), and root mean square error (RMSE)) for fault location[4]. Location predictions were denormalized to the 5–15 km range before metric

Table 2. *Detailed feature extraction pipeline from three-phase current signals*

Step	Description	STFT-Based Features	DWT-Based Features
1. Input	Raw three-phase current signals (Ia, Ib, and Ic), sampled for one pre-fault and one post-fault cycle.	Identical Input	Identical Input
2. Signal Transform	Apply the signal processing technique.	**STFT** ●Creates a time-frequency matrix for each phase. ●Parameters: Hann window (128 samples), 96-sample overlap, 128-point FFT.	**DWT** ●Decomposes the signal of each phase into coefficients. ●Parameters: 'db4' wavelet, seven levels of decomposition.
3. Raw Output	The direct result of the transform.	A complex-valued matrix X(f, t) for each phase, showing the signal strength at different frequencies (f) and times (t).	A set of eight coefficient vectors for each phase: ●A7: Approximation coefficients (low-frequency trend). ●D1-D7: Detailed coefficients (high-frequency details, D1 is the highest frequency).
4. Feature Calculation (the key difference)	Conversion of the raw output into a set of numbers (features) for the artificial intelligence (AI) model.	For the time-frequency matrix of each phase: 1. Maxima: Find the maximum value in each of the 17 frequency columns → 17 features. 2. Energy: Calculate the energy in each of the 17 frequency columns → 17 features. 3. Standard Deviations: - std(max_values_across_time) → 1 feature - std(max_values_across_frequency) → One feature → Subtotal per phase: 17 + 17 + 1 + 1 = 36 features.	For the set of coefficients of each phase: ●Take all eight coefficient vectors (A7, D1, D2, D3, D4, D5, D6, and D7). ●For each vector, calculate six statistical measures: Standard Deviation, Mean, Entropy, Energy, Kurtosis, and Skewness. → Subtotal per phase: 8 vectors ×6 statistics = 48 features.
5. Total Features per Sample	Combine features from all three phases.	3 phases ×36 features = 108 features.	3 phases ×48 features = 144 features.
6. Final Step	Prepare the final dataset for the AI model.	All features from all samples are normalized to a [0, 1] range.	All features from all samples are normalized to a [0, 1] range.

calculation. To ensure reproducibility despite Keras stochastic weight initialization, a fixed random seed (42) was used for all the experiments. The optimal models were selected after multiple trials to confirm their stable performance, and the reported results were obtained from these finalized models. The optimal models were selected after multiple trials when the performance stabilized, thereby ensuring consistent metrics. This approach accounted for the inherent variability of neural networks, while identifying the most reliable configurations. The evaluation framework provided a comprehensive assessment of both classification precision (accuracy) and regression fidelity (R^2, MAE, and RMSE) for a complete fault diagnosis.

4.6. Deep Learning Models Development. The models were optimized using systematic hyperparameter tuning. The neuron counts in the hidden layers followed heuristic rules[46], with STFT-based models using 108 neurons (matching input features) in layer 1 and 55 neurons (half input count) in layer 2. The DWT models employed different input fractions (96, 144, and 72 - two-thirds ratios). The fault location models used similar proportional schemes. Table 2 lists the tested neuron combinations. The selection process evaluated the performance across these configurations, adhering to established neural network design principles[46]. Optimal architectures balanced complexity and generalization, with neuron counts systematically derived from the input dimensionality rather than from arbitrary choices.

For fault classification, optimal neuron configurations were determined through performance comparisons. The STFT model achieved the peak accuracy with 108 (layer 1) and 60 (layer 2) neurons, showing minimal loss (Figure 5). Comparable results were obtained for the 140/60, 140/70, and 120/55 combinations. The DWT model performed best with 130 and 80 neurons (Figure 6). The regression models followed the same optimization process but used R^2 and RMSE metrics instead of accuracy/loss. Table 3 lists the neuron counts selected for the STFT/DWT regression models. This systematic approach balanced model complexity and predictive performance while avoiding overfitting, with architectural decisions guided by empirical results rather than arbitrary selection.

Table 3. *Iterations of hidden layer neuron counts for the fault location and classification models*

Signal Processing Technique		Fault Classification (Multi-class Classification)		Fault Location (Regression)	
	STFT	DWT	STFT	DWT	
Number of	Layer 1	90, 100, 111, 120, 130, 140	120, 130, 144, 150, 160	90, 100, 111, 120, 130, 140	100, 110, 120, 130, 144, 150
hidden neurons	Layer 2	30, 40, 55, 60, 70	50, 60, 72, 80, 96	30, 40, 55, 60, 74, 80	50, 60, 72, 80, 96

This study systematically evaluated the activation functions and found that ReLU and tanh were the most effective for hidden layers across models. The output layers used the following task-specific functions: sigmoid (binary detection), softmax (multi-class classification), and linear (regression for fault location). The Adam optimizer was employed universally, except for the DWT classification model, in which RMSprop yielded better results.

For fault location, identified as the most critical challenge, additional optimization was performed on the following: (1) training sample size and (2) train-test splits. The complete dataset contained 14,000 samples (2,000 per fault type), of which 9,800 samples (70%) were initially allocated for training. Experimental increments (10% steps) revealed optimal training sizes: 7,840 samples (80%) for the STFT models (noting a performance dip at 90%) versus the full 9,800 samples for the DWT models.

Figure 7 shows how model performance (measured by R^2) varied with the quantity of training data. The 70/30% train-test split provided sufficient validation data while maximizing learning. This rigorous optimization process, particularly for fault location tasks, ensured a robust model performance by balancing data utilization against overfitting risks. The results demonstrate that the optimal training set sizes can vary significantly between different signal processing approaches in fault diagnosis systems.

This study evaluated train-test splits in 10% increments to optimize model performance. For STFT models, a 70/30% train-test split proved optimal, whereas the DWT models performed best with an 80/20% split. This analysis was conducted for each fault type using the previously determined ideal training sample sizes. Figure 8 shows the performance of the DWT model for one fault type, revealing the 80/20% split yielded the lowest RMSE, with the 70/30% and 90/10% splits showing comparable results. A systematic evaluation confirmed that these splits maximized generalization across both signal processing methods while preventing overfitting. Similar patterns emerged for the other fault types, although their plots were excluded for brevity.

A binary classification problem, “fault equivalent to 1” or “no-fault equivalent to 0,” was the outcome of fault detection. The fault type, denoted by a number between one and seven, was the outcome of the multi-class classification problem. It is

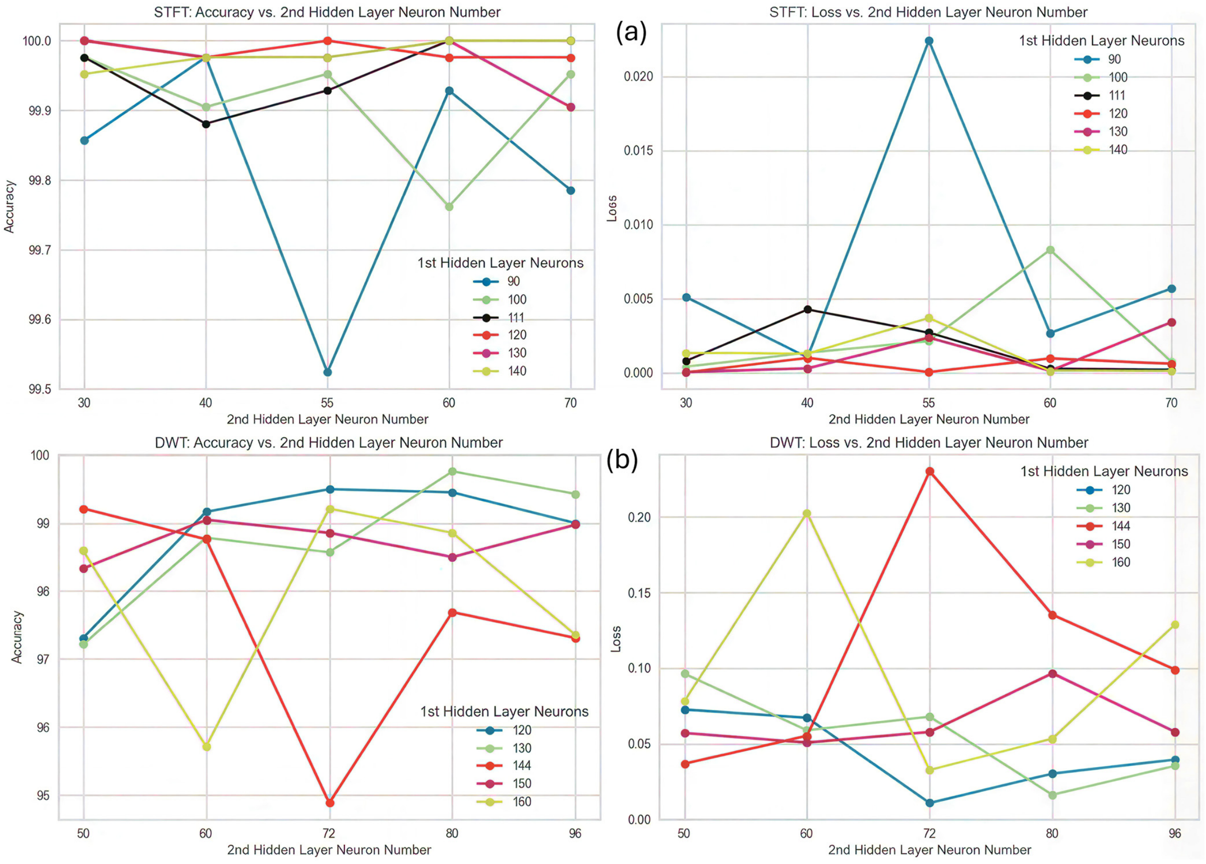


Figure 5. Model accuracy and loss for various neuron counts in the (a) STFT-based and (b) DWT-based fault classification models

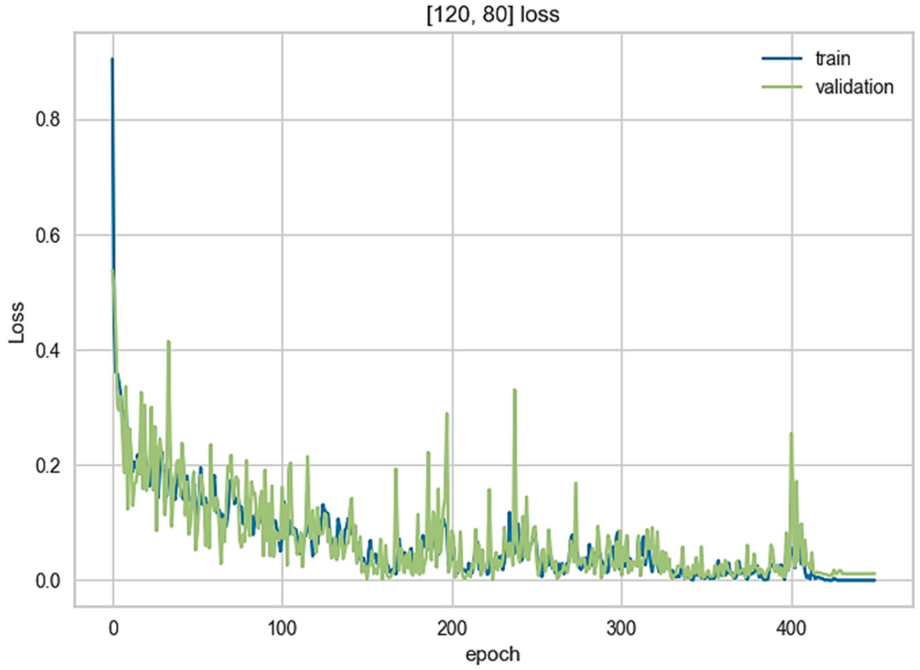


Figure 6. Training curve of the optimum DWT classification model

worth noting that non-faulty samples from the dataset were eliminated prior to fault classification. A decimal value that should fall between 5–15 (the range in kilometers where the faults are placed) was produced by the regression issue of fault placement. This number may exceed this range depending on the accuracy of the model. The selected parameters of the deep learning models are presented in the Table 4.

The fault detection and classification models used relatively few epochs because additional iterations provided diminishing returns in accuracy while increasing the computation time. The optimal epoch count was determined when the training curve plateaued. For these models, the training data were further split (70/30%), with 33% allocated for validation, to plot the learning curves. The 70/30% train-test split followed the conventional practice. Figure 9 shows a flowchart of the complete fault diagnosis methodology, summarizing the systematic approach from data processing to model evaluation. This balanced design ensured efficient training without overfitting while maintaining robust performance validation using dedicated validation datasets.

Table 4. *Fault detection model parameters*

Task	Model	Hidden Layers & Neurons	Activation Functions	Loss Function	Optimizer	Learning Rate	Batch Size	Epochs
Fault Detection	STFT-based	10	Hidden: ReLU Output: Sigmoid	Binary Cross-Entropy	Adam	0.001	32	20
	DWT-based	10	Hidden: ReLU Output: Sigmoid	Binary Cross-Entropy	Adam	0.001	32	20
Fault Classification	STFT-based	[111, 60]	Hidden 1: Tanh Hidden 2: ReLU Output: Softmax	Categorical Cross-Entropy	Adam	0.001	64	50
	DWT-based	[130, 80]	Hidden 1: Tanh Hidden 2: ReLU Output: Softmax	Categorical Cross-Entropy	RMSprop	0.0005	64	450
Fault Location	STFT-based	[120, 55]	Hidden 1: Tanh Hidden 2: ReLU Output: Linear	Mean Squared Error	Adam	0.001	128	1000
	DWT-based	[110, 60]	Hidden 1: Tanh Hidden 2: ReLU Output: Linear	Mean Squared Error	Adam	0.001	128	1000

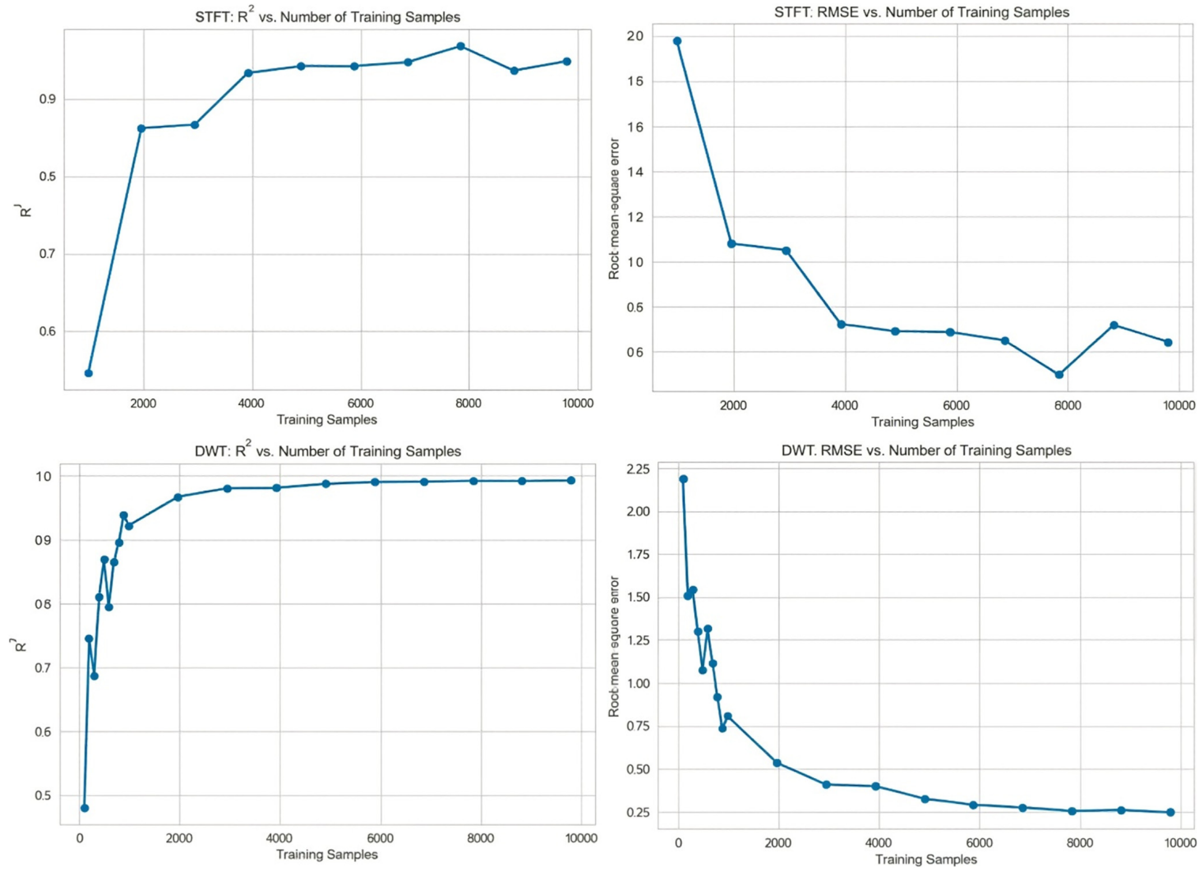


Figure 7. R^2 values and RMSEs of both regression models for various numbers of training samples

5. RESULTS AND DISCUSSION

This section presents the outcomes of the developed deep learning method combined with advanced signal processing techniques for fault diagnosis in the distribution grid. First, Table 5 shows that both the STFT and DWT models accurately detected faults under both noisy and noise-free measurements. In all cases, they were able to distinguish between faulty and non-faulty cases. Given the ease of discovering faults, this was anticipated. The human eye can easily distinguish between defective and non-faulty current signals. As shown in Figure 4, the current surges sharply at a fault, resulting in a significant increase in the amplitude in the current plot versus time.

Both the STFT and DWT models achieved a high accuracy (99.66% and 98.11%, respectively) across noise levels (0–40 dB). STFT delivered flawless classification (100%) under noise-free conditions and near-perfect accuracy (>99.6%) with noise, excelling for multi-phase faults. This advantage arises from the uniform time–frequency resolution of STFT, which better preserves the transient signatures of faults under noise, making it more resilient for localization tasks than the

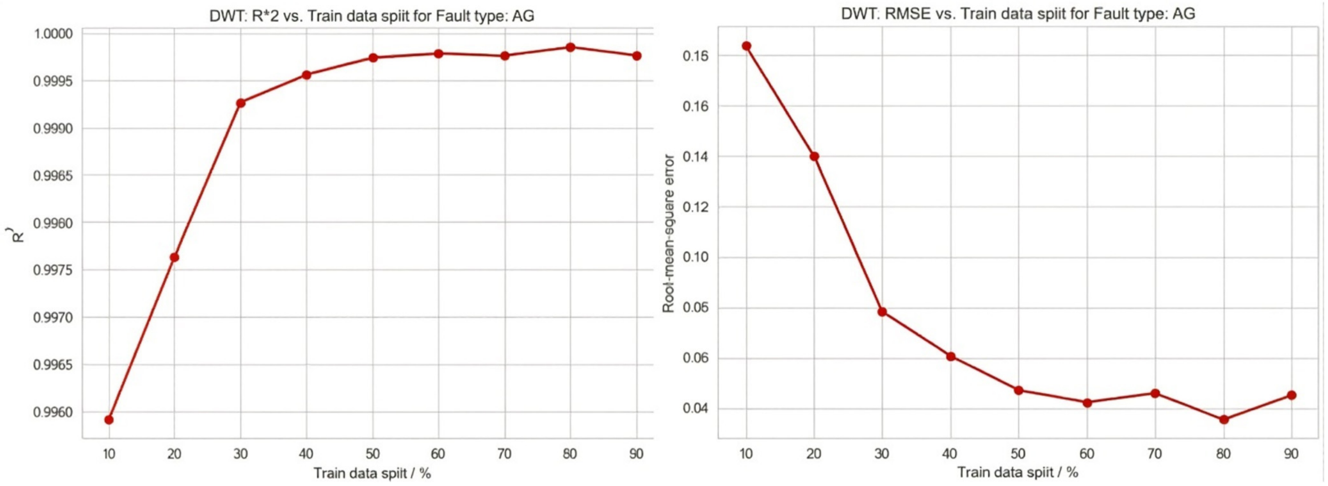


Figure 8. R^2 values and RMSEs for various training data splits of the DWT regression model for the fault type “AG”

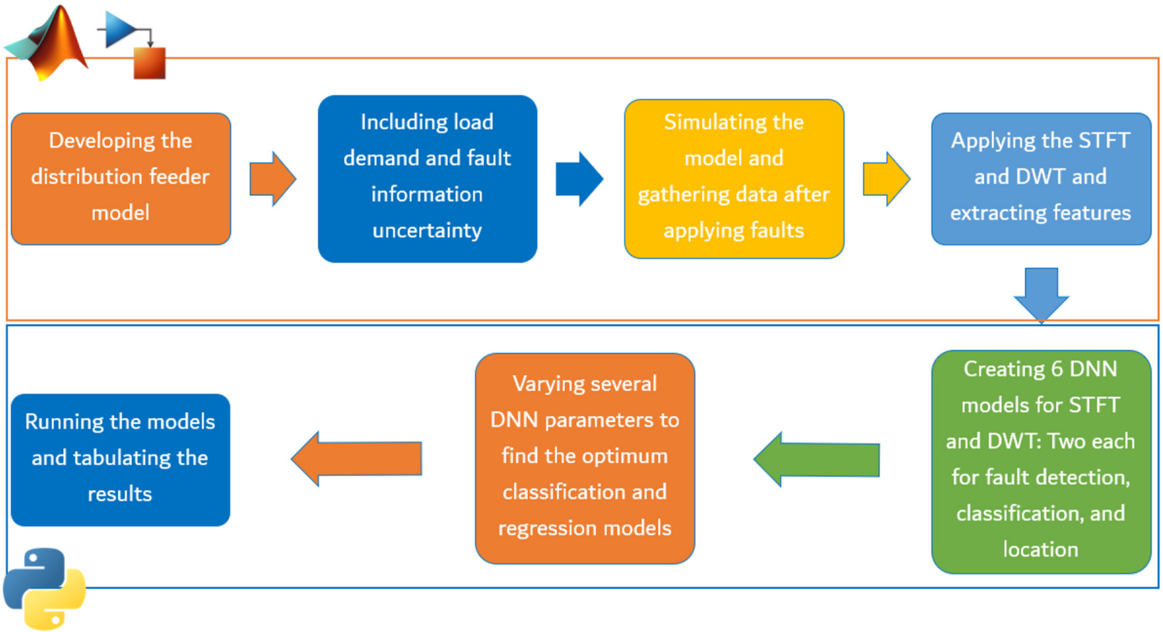


Figure 9. Complete fault diagnosis method

scale-dependent resolution of DWT. Minor dips occurred in the BG (20/40 dB) and ABCG (20 dB) faults. DWT also performed robustly, with perfect results in noise-free and 40 dB cases but slight inaccuracies in the CG, ACG, and BCG classifications. Both models maintained consistent accuracy (STFT: 99–100%, DWT: 98–100%), demonstrating strong noise immunity. The superior performance of STFT suggests that its time-frequency analysis better captures fault features, although both methods are reliable (<2% error) for real-world applications.

In addition to the accuracy, the precision, recall, and F1-score were computed for each fault type to evaluate the classification performance more rigorously. The STFT model achieved a macro-average precision of 0.997, recall of 0.996, and F1-score of 0.996 at a noise level of 30 dB. The confusion matrix (Figure 10) confirmed the high separability between classes, particularly for

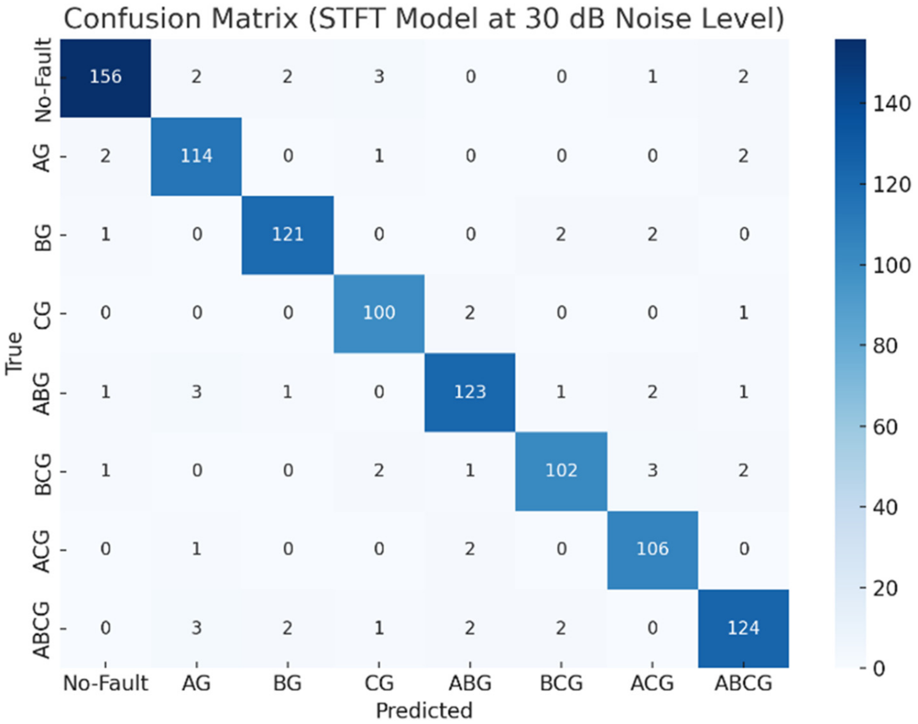


Figure 10. Confusion matrix for fault classification using the STFT model under 30 dB noise. The model achieved strong separability across all fault types, with minimal misclassification

Table 5. *Precision of the models for fault classification and detection*

Fault type	**0 dB**	**20 dB**	**30 dB**	**40 dB**				
	STFT	DWT	STFT	DWT	STFT	DWT	STFT	DWT
No-Fault	100%	100%	100%	100%	100%	100%	100%	100%
AG	100%	100%	100%	100%	100%	100%	100%	100%
BG	100%	100%	99.66%	100%	100%	100%	99.82%	100%
CG	100%	99.84%	100%	99.67%	100%	100%	100%	100%
ABG	100%	100%	100%	100%	100%	100%	100%	100%
BCG	100%	100%	100%	98.78%	99.83%	100%	100%	100%
ACG	100%	100%	100%	99.67%	100%	98.11%	100%	100%
ABCG	100%	100%	99.82%	100%	100%	100%	100%	100%

the three-phase and ground faults. The high classification accuracy, along with a low R^2 value for fault location, reflects the difference in task complexity; classification is a discrete labeling task that is less sensitive to noise, whereas regression-based location estimation is highly sensitive to signal distortion. In particular, DWT loses location accuracy in noisy environments owing to its scale–frequency trade-off, whereas STFT maintains its performance by retaining more uniform time–frequency details.

The fault location analysis revealed distinct performance differences between the STFT- and DWT-based models. Although both achieved excellent results under noise-free conditions (DWT: $R^2 = 0.995$, STFT: $R^2 = 0.973$), their noise robustness varied significantly. The STFT model maintained a consistently high accuracy across all noise levels ($R^2 > 0.89$, $RMSE < 0.93$), demonstrating superior stability. In contrast, the DWT performance degraded sharply with noise, with R^2 dropping to 0.203 and RMSE >2.5 at higher noise levels.

The R^2 metric (range: $0 - 1$) indicated that STFT predictions closely matched actual values ($\geq$89% variance explained), while DWT became unreliable in noisy conditions. RMSE measurements (lower=better) showed that the average location error of STFT remained below 0.93 regardless of noise, whereas DWT errors tripled in noisy scenarios. This suggests that the time-frequency localization of STFT better preserves the fault distance information under signal degradation.

Notably, DWT outperformed STFT under ideal (noise-free) conditions, achieving a near-perfect $R^2(0.995)$. However, this advantage disappeared even with moderate noise. The robust performance of the STFT model across all test conditions makes it suitable for practical implementations where noise is inevitable. These results highlight the critical importance of testing fault location algorithms under realistic and noisy conditions rather than ideal scenarios.

The severe performance degradation of the DWT-based model at 20 dB noise ($R^2 = 0.2029$) can be attributed to the fundamental properties of the wavelet transform. DWT provides multi-resolution analysis, offering a high time resolution but poor frequency resolution at high frequencies, and vice-versa. Under significant noise, the high-frequency noise components can corrupt the detail coefficients (D1–D4) that are crucial for pinpointing the precise transients caused by a fault. Unlike STFT, which maintains a uniform time-frequency resolution across the spectrum, the adaptive windowing of DWT may amplify the influence of noise in specific sub-bands, effectively drowning out the salient features needed for accurate distance regression. This inherent sensitivity makes the DWT model vulnerable to noisy conditions, whereas the consistent windowing approach of STFT is far more robust for the fault location task.

Both models exhibited unexpected performance variations across noise levels. Surprisingly, 40 dB noise yielded better results than 30 dB; STFT achieved higher R^2 (0.97 versus 0.96) and lower RMSE (0.52 versus 0.60), with similar trends for DWT. However, increasing the noise from 0–40 dB degraded the overall performance. The R^2 value of the STFT model decreased from 0.973 (noise-free) to 0.968 (40 dB), whereas its RMSE remained stable (<0.93). The DWT model suffered more severe degradation: R^2 decreased sharply from 0.995 (0 dB) to 0.903 (40 dB), and the RMSE increased from 0.202 to 0.902.

Table 6 shows that higher noise (20 dB) caused greater performance loss than moderate levels (30/40 dB). STFT demonstrated resilience (R^2>0.89 at 40 dB), while DWT became unstable (R^2 as low as 0.203), suggesting the time-frequency analysis of STFT better preserves the fault location accuracy under distortion. The non-linear noise impact (inverse trend at 30/40 dB) warrants further study. The consistency of STFT favors noisy environments, though real-world validation is needed. Computational costs from STFT-deep learning integration may challenge

Table 6. *Fault location model results*

	STFT				**DWT**			
Performance Metrics	0 dB	20 dB	30 dB	40 dB	0 dB	20 dB	30 dB	40 dB
R^2	0.9733	0.8966	0.9561	0.9675	0.9951	0.2029	0.7069	0.9032
RMSE	0.4701	0.9213	0.6070	0.5221	0.2015	2.5703	1.5842	0.9018
MAE	0.3351	0.6611	0.4259	0.3719	0.1296	1.7758	1.0648	0.6145

real-time use; solutions such as model pruning or edge optimization could be useful. Future work should enhance interpretability via tools such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model Agnostic Explanation (LIME) to build trust in critical infrastructure applications.

Practical implications: For deployment in real distribution grids, the proposed models can be integrated with phasor-measurement units (PMUs) or feeder-mounted sensors to provide near-real-time fault detection, classification, and localization. The strong noise resilience of the system, particularly with the STFT-based features, suggests its viability in urban environments with significant electrical interference. The main operational benefit is the reduced fault location time, which directly shortens outage durations and improves service reliability. The challenges include ensuring computational efficiency for edge devices and validating the performance using field data, which will be the focus of future work.

6. CONCLUSIONS

In this study, an intelligent fault diagnosis system that combines STFT/DWT signal processing with deep learning was developed for distribution grids. The MATLAB/Simulink test feeder simulated diverse faults under varying loads and noise levels (0–40 dB). STFT extracted 108 time-frequency features, whereas DWT derived 144 wavelet-based features per sample. These features trained deep learning models for fault detection, classification, and location. The evaluation showed that both techniques achieved high accuracy; fault detection was near-perfect, whereas classification favored STFT (99.66% versus 98.11% for DWT). For location, the STFT demonstrated superior noise robustness (R^2 >0.89 across all noise levels) compared to DWT, whose performance degraded sharply (R^2 : 0.995–0.203) under noise. The consistent RMSE of STFT (<0.93) confirmed its reliability in practical noisy environments. Future work should explore (1) validation using larger real-world grids with distributed renewable generation, (2) optimization of STFT parameters (window type/length) for enhanced feature extraction, and (3) alternative transforms such as Stockwell for comparative analysis. The adaptability of the framework to evolving grid complexities, including the high penetration of inverter-based resources, can further strengthen its diagnostic utility.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Authors

Khalid Alfuwail – *Department of Aerospace Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0009-0002-4442-3454
Email: g202045220@kfupm.edu.sa

Hamza M. Anwar – *Department of Control & Instrumentation Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia; Interdisciplinary Research Center for Sustainable Energy Systems, King Fahd University of Petroleum & Minerals, Dhahran 31216, Saudi Arabia;*
Email: g201928690@kfupm.edu.sa

Corresponding Author

Md Shafiullah – *Research Mentor, Assistant Professor, Department of Control & Instrumentation Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0000-0003-2282-5663
Email: shafiullah@kfupm.edu.sa

Authors

Md Mahfuzur Rahman – *Assistant Professor, Department of Information and Computer Science, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia; Interdisciplinary Research Center for Intelligent Secure Systems, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia* 0000-0002-2871-9119
Email: mdmahfuzur.rahman@kfupm.edu.sa

Mohamed Ali Abido – *Professor, Department of Electrical Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia; Interdisciplinary Research Center for Sustainable Energy Systems, King Fahd University of Petroleum & Minerals, Dhahran 31216, Saudi Arabia;* 0000-0001-5292-6938
Email: mabido@kfupm.edu.sa

DECLARATIONS OF COMPETING INTEREST

The authors affirm that they have no known competing financial interests or interpersonal conflicts that could have affected the research presented in this study.

ACKNOWLEDGEMENTS

The authors acknowledge research support received from the IRC for Sustainable Energy Systems under Grant No. INSE2512, at King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia.

REFERENCES

(1) Taft, J. D. *Fault Intelligence: Distribution Grid Fault Detection and Classification.* (Richland, Sep. 2017).

(2) Mamuya, Y. D., Lee, Y. D., Shen, J. W., Shafiullah, M. & Kuo, C. C. Application of machine learning for fault classification and location in a radial distribution grid. *Appl. Sci.* **10**, (2020).

(3) Zhang, C. L. *et al.* Fault diagnosis and removal for hybrid power generation systems based on an ensemble deep learning diagnostic method with self-healing strategies. *Int. J. Hydrogen Energy* **109**, 1297–1313 (2025).

(4) Shafiullah, M., AlShumayri, K. A. & Alam, M. S. Machine learning tools for active distribution grid fault diagnosis. *Adv. Eng. Softw.* **173**, (2022).

(5) Shafiullah, M., Abido, M. A. & Al-Mohammed, A. H. Intelligent fault diagnosis for distribution grid considering renewable energy intermittency. *Neural Comput. Appl.* **34**, 16473–16492 (2022).

(6) Wu, H. & Wang, T. A novel fault diagnosis method of smart grids based on artificial immune spiking neural P systems considering false data injection attacks. *Electr. Power Syst. Res.* **244**, 111572 (2025).

(7) Shafiullah, M., Abido, M. A. & Abdel-Fattah, T. Distribution grids fault location employing ST based optimized machine learning approach. *Energies* **11**, (2018).

(8) Stefanidou-Voziki, P., Sapountzoglou, N., Raison, B. & Dominguez-Garcia, J. L. A review of fault location and classification methods in distribution grids. *Electr. Power Syst. Res.* **209**, 108031 (2022).

(9) Shafiullah, M. & Abido, M. A. A Review on Distribution Grid Fault Location Techniques. *Electr. Power Compon. Syst.* **45**, 807–824 (2017).

(10) Jia, H. An Improved Traveling-Wave-Based Fault Location Method with Compensating the Dispersion Effect of Traveling Wave in Wavelet Domain. *Math. Probl. Eng.* **2017**, 1–11 (2016).

(11) Baseer, M. A. Travelling Waves for Finding the Fault Location in Transmission Lines. *J. Electr. Electron. Eng.* **1**, 1–19 (2013).

(12) Dutta, R. & Samantaray, S. R. Assessment of impedance based fault locator for AC micro-grid. *Renew. Energy Focus* **26**, 1–10 (2018).

(13) Labrador Rivas, A. E. & Abrão, T. Faults in smart grid systems: Monitoring, detection and classification. *Electr. Power Syst. Res.* **189**, 106602 (2020).

(14) Lan, S., Chen, M. J. & Chen, D. Y. A Novel HVDC Double-Terminal Non-Synchronous Fault Location Method Based on Convolutional Neural Network. *IEEE Trans. Power Deliv.* **34**, 848–857 (2019).

(15) Yu, K. *et al.* A Novel Traveling Wave Fault Location Method for Transmission Network Based on Directed Tree Model and Linear Fitting. *IEEE Access* **8**, 122610–122625 (2020).

(16) Babayomi, O. O. & Oluseyi, P. O. Intelligent Fault Diagnosis in a Power Distribution Network. *Adv. Electr. Eng.* **2016**, 1–10 (2016).

(17) Sahu, S. K. *et al.* Machine learning based adaptive fault diagnosis considering hosting capacity amendment in active distribution network. *Electr. Power Syst. Res.* **216**, 109025 (2023).

(18) Chen, K., Hu, J., Zhang, Y., Yu, Z. & He, J. Fault Location in Power Distribution Systems via Deep Graph Convolutional Networks. *IEEE J. Sel. Areas Commun.* **38**, 119–131 (2020).

(19) Rizeakos, V. *et al.* Deep learning-based application for fault location identification and type classification in active distribution grids. *Appl. Energy* **338**, 120932 (2023).

(20) Hong, Y. Y. & Cabatac, M. T. A. M. Fault Detection, Classification, and Location by Static Switch in Microgrids Using Wavelet Transform and Taguchi-Based Artificial Neural Network. *IEEE Syst. J.* **14**, 2725–2735 (2020).

(21) Naidu, K. *et al.* Optimized artificial neural network to improve the accuracy of estimated fault impedances and distances for underground distribution system. *PLoS One* **15**, e0227494 (2020).

(22) Iliyaeifar, M. M. & Hadaeghi, A. Extreme learning machine-based fault location approach for terminal-hybrid LCC-VSC-HVDC transmission lines. *Electr. Power Syst. Res.* **221**, 109487 (2023).

(23) Rezaee Ravesh, N., Ramezani, N., Ahmadi, I. & Nouri, H. A hybrid artificial neural network and wavelet packet transform approach for fault location in hybrid transmission lines. *Electr. Power Syst. Res.* **204**, 107721 (2022).

(24) Grcić, I., Pandžić, H. & Novosel, D. Fault detection in dc microgrids using short-time fourier transform. *Energies* **14**, 277 (2021).

(25) The MathWorks Inc. *Statistics and Machine Learning Toolbox*. (mathworks.com).

(26) The MathWorks Inc. *Deep Learning Toolbox*. (mathworks.com).

(27) Van Rossum, G. & Drake, F. L. *Python 3 Reference Manual*. (CreateSpace, Scotts Valley, CA, 2009).

(28) Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

(29) Chollet, F. *Keras*. (keras.io).

(30) Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

(31) McKinney, W. Data Structures for Statistical Computing in Python. in *Proceedings of the 9th Python in Science Conference* 56–61 (2010).

(32) Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

(33) Oppenheim, A. V. & Schafer, R. W. *Discrete Time Signal Processing*. (Prentice Hall, 1998).

(34) Kehtarnavaz, N. *Digital Signal Processing System Design: LabVIEW-Based Hybrid Programming*. (Elsevier, 2008).

(35) Heideman, M. T., Johnson, D. H. & Burrus, C. S. Gauss and the History of the Fast Fourier Transform. *IEEE ASSP Mag.* **1**, 14–21 (1984).

(36) Smith, J. O. *Spectral Audio Signal Processing*. (W3K Publishing, 2011).

(37) The MathWorks Inc. *stft*. (mathworks.com).

(38) Nicoll, A. *The Wavelet Transform for Beginners*. (YouTube).

(39) Srivastava, M. *Improving Signal Resolution and Reducing Experiment Time in Electron Spin Resonance Spectroscopy via Data Processing Methods*. (Cornell University, 2018).

(40) LeNail, A. NN-SVG: Publication-Ready Neural Network Architecture Schematics. *J. Open Source Softw.* **4**, 747 (2019).

(41) Brownlee, J. *Using Activation Functions in Neural Networks*. (Machine Learning Mastery).

(42) The MathWorks Inc. *Simscape Electrical Toolbox*. (mathworks.com).

(43) Kachhepati, B. *Application of Short Time Fourier Transform (STFT) in Power Quality Monitoring and Event Classification*. (New Mexico State University, 2016).

(44) Mishra, S., Bhende, C. N. & Panigrahi, B. K. Detection and Classification of Power Quality Disturbances Using S-Transform and Probabilistic Neural Network. *IEEE Trans. Power Deliv.* **23**, 280–287 (2008).

(45) Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

(46) Heaton, J. in *Introduction to Neural Networks with Java, 2nd Edition* (ed. Smith, K.) 159–159 (Heaton Research, Inc, 2008).

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article *Computer Science*

Static Analysis of Code Quality in Open-Source Python Projects

Bashir Adam Ahmed Ali

Cite https://doi.org/10.64589/juri/215034

Submitted: September 15, 2025 **Revised:** November 09, 2025 **Accepted:** December 02, 2025

ABSTRACT

Ensuring high code quality is essential for the reliability and sustainability of widely used open-source software, yet empirical, multi-metric assessments of popular Python libraries remain scarce. This study presents a systematic static analysis of code quality in two prominent open-source Python projects, Requests and BeautifulSoup4. We employed a suite of static analysis tools—Radon, Pylint, and Flake8—to measure metrics including lines of code (LOC), cyclomatic complexity, maintainability index, Pylint scores, and PEP8 violations across all project modules. Automated Python scripts were developed to ensure a reproducible and comprehensive data collection process. Our analysis reveals a nuanced picture of code quality: while both projects exhibit good overall maintainability, we identified specific modules with high cyclomatic complexity and a significant number of style violations. For instance, dammit.py in BeautifulSoup4 recorded 318 PEP8 violations and a maintainability index of 40.08. The results underscore the practical value of static analysis for pinpointing refactoring candidates and enforcing coding standards.

We recommend that open-source projects integrate these tools into continuous integration pipelines to proactively manage technical debt and enhance long-term maintainability. All code and data for replicating this study are publicly available. This work provides a novel, reproducible, and multi-tool assessment that delivers actionable, module-level insights for maintainers, addressing a gap in empirical Python code quality research.

Keywords: code quality, static analysis, open-source, Python, Python Enhancement Proposal-8, maintainability

1. INTRODUCTION

Software quality is a critical determinant of a software system's reliability, maintainability, and long-term sustainability. High-quality code reduces maintenance costs, minimizes defects, and improves developer productivity. Python, as one of the most popular programming languages, powers a vast ecosystem in web development, automation, data science, and machine learning. Consequently, the code quality of foundational open-source Python projects is of paramount practical importance, as they underpin countless academic and commercial applications [1].

Despite the existence of coding standards like PEP8, code quality in open-source projects can vary significantly due to factors such as diverse contributor expertise and evolving project requirements. Poor quality often manifests as high cyclomatic complexity, low maintainability, and style inconsistencies, which collectively increase maintenance effort and risk[2]. Static analysis tools offer an automated, non-execution-based method to evaluate code quality. Tools like Radon, Pylint, and Flake8 can measure key metrics related to complexity, maintainability, and style compliance [3–5].

While previous research has extensively evaluated static analysis for languages like Java and C++ [6,7], and some studies have focused on the capabilities of individual Python tools [8], there is a scarcity of empirical, multi-metric studies that systematically assess code quality across popular, real-world Python libraries. This gap limits our understanding of the actual state of code quality in the Python open-source ecosystem. Recent surveys highlight the growing need for such empirical validations in modern software engineering contexts [9].

This study aims to address this gap by conducting a reproducible static analysis of two widely used Python libraries: BeautifulSoup4 and Requests. To guide our investigation, we formulate the following research questions:

- **RQ1:** How do code quality metrics, such as cyclomatic complexity and maintainability index, vary across different modules within BeautifulSoup4 and Requests?
- **RQ2:** What are the differences in overall code quality and adherence to PEP8 standards between these two projects?
- **RQ3:** Which specific modules are the most critical candidates for refactoring based on the aggregated metrics?

The novelty and contribution of this work are threefold. First, it provides a detailed, empirical snapshot of code quality in two critical Python libraries, moving beyond tool capability studies to deliver actionable, module-level insights for maintainers. Second, it employs a transparent and fully reproducible methodology, with all scripts and data publicly available, addressing a key shortcoming in many existing studies. Third, it synthesizes findings from multiple complementary tools (Radon, Pylint, Flake8) to present a holistic view of quality through a fully reproducible methodology—addressing a key gap in replicable Python research where prior work often focuses on a single tool or metric.

2. RELATED WORK

The critical role of software quality and the utility of static analysis are well-established in software engineering literature. Recent surveys, such as that by Iftikhar et.al[9], provide a comprehensive overview of code quality metrics and their impact on software maintainability, underscoring their continued relevance.

In the domain of open-source software, many foundational studies focused on languages like Java and C++. For example, research has analyzed trends in code smells and defect prediction across multiple Java projects[6]. However, the findings from these studies are not directly transferable to Python due to its dynamic typing and different idiomatic patterns.

Research specifically targeting Python code quality is growing. Siddik and Bezemer et al.[10] explored the correlation between code style and code quality, finding that consistent style often accompanies higher quality, which justifies the inclusion of PEP8 analysis in our study. Other works have evaluated the effectiveness of specific tools. Scalabrino et al.[11] conducted a comprehensive study on automatically assessing code understandability, a key aspect of maintainability. Recent empirical studies have applied these tools in specific contexts, such as comparing code quality in Python Jupyter notebooks and scripts for data science[12] or detecting code smells in Python software[13].

Furthermore, the detection of complex issues in Python has been a focus of recent research. Studies have adapted traditional metrics and definitions for Python's unique characteristics[14,15]. Our work directly builds on this by applying established complexity and maintainability metrics (via Radon) and code quality scoring (via Pylint) to large-scale, real-world projects, providing a validation of these concepts in practice.

This study synthesizes these strands of research by conducting a systematic, multi-faceted assessment of two large, community-driven Python projects using a suite of complementary tools. Unlike studies that focus on a single aspect of quality or a specific domain, we provide a holistic view of complexity, maintainability, and style in general-purpose libraries, identifying specific, actionable refactoring targets.

3. METHODOLOGY

This study employs a systematic and reproducible methodology to evaluate the code quality of the selected open-source Python projects. The process encompasses project selection, tool selection and configuration, metric extraction, data aggregation, visualization, and analysis.

3.1. Project Selection.

We selected two high-profile Python projects based on their popularity, active maintenance, and significance within the Python ecosystem:

- **Requests**[16] — A simple, yet powerful HTTP library for Python.
- **BeautifulSoup4**[17]— A library for pulling data out of HTML and XML files.

These projects were chosen because their widespread use makes their code quality a concern for a large community, and their manageable size allows for a comprehensive module-by-module analysis.

3.2. Tool Selection and Configuration.

To ensure a comprehensive assessment, we used a suite of established static analysis tools. The exact commands and configurations are provided in the supplementary materials to ensure reproducibility.

- **Radon (v6.0.1):** Used to compute raw metrics. We used the radon cc -a command for the average cyclomatic complexity per module and radon mi -s for the maintainability index.
- **Pylint (v3.1.0):** Used for general code quality analysis and scoring. We used a default configuration.
- **Flake8 (v7.0.0):** Used to detect PEP8 style violations. We used the default rule set without modifications.

3.2.1. Justification for tool selection.

Radon, Pylint, and Flake8 were selected for their complementary strengths, widespread adoption in the Python community, and ability to measure distinct aspects of code quality. Radon provides raw complexity and maintainability metrics, Pylint offers a holistic quality score, and Flake8 is the de facto standard for PEP8 compliance. Using their default configurations ensures reproducibility and aligns with common practitioner usage, allowing our results to be directly comparable to typical developer workflows.

3.3. Metric Selection.

We focused on the following metrics to provide a multi-dimensional view of code quality:

- **Lines of Code (LOC):** Measures module size.
- **Cyclomatic Complexity (CC):** Measures the structural complexity of functions and methods.
- **Maintainability Index (MI):** A composite metric (scale of 0-100) where higher values indicate better maintainability.
- **Pylint Score:** A normalized score (typically 0-10) for overall code quality.
- **PEP8 Violations:** A count of style guide deviations.

These metrics directly address our RQs: CC and MI (RQ1), PEP8 violations (RQ2), and their aggregation (RQ3).

3.4. Data Collection and Reproduction.

The entire data collection process was automated with Python scripts to ensure accuracy and reproducibility. The steps were:

- **Repository Cloning:** The latest code was cloned from the official GitHub repositories.
- **Static Analysis Execution:** Scripts sequentially executed Radon, Pylint, and Flake8 on every Py file.
- **Metric Extraction:** Outputs were parsed to extract numerical values for each metric.
- **Data Aggregation:** Results were compiled into structured CSV files.

3.5. Visualization.

We used Matplotlib and Seaborn to generate visualizations for intuitive analysis. All figures are placed after their first textual reference, have clear axis labels and captions, and exclude the "SUMMARY" row from module-level charts.

Table 1. *Static analysis metrics for BeautifulSoup4*

Module/File	LOC	CC (avg)	Maintainability Index	Pylint Score	PEP8 Violations
dammit.py	829	3.85	40.08	7.39	318
test_tree.py	1829	3.59	6.87	7.13	69
testing.py	592	3.72	45.19	0.0	39
__*init*__.py	406	3.17	100.0	N/A	34
element.py	1611	0	0	N/A	26
SUMMARY	7755	3.40	74.78	3.98	656

Note: A Maintainability Index (MI) of 0.0, as seen in element.py, occurs when the calculated index is non-positive due to very high complexity and/or size overwhelming the effect of comments, and is reported as 0 by the Radon tool.

- **Bar charts** of average cyclomatic complexity per module.
- **Histograms** of Pylint score distribution.
- **Scatter plots** of LOC vs. Cyclomatic Complexity.

4. RESULTS

The results of the static analysis for BeautifulSoup4 and Requests are presented below.

4.1. BeautifulSoup4. The analysis of BeautifulSoup4 covered 19 modules with a total of 7,755 LOC. The project-level averages are a cyclomatic complexity of 3.4, a maintainability index of 74.78, a Pylint score of 3.98, and 656 total PEP8 violations (Table 1). Figure 1 illustrates the average cyclomatic complexity per module, highlighting dammit.py and test_tree.py as particularly complex. The distribution of Pylint scores (Figure 2) shows a bimodal pattern, indicating variability in code quality across modules. A scatter plot of LOC versus cyclomatic complexity (Figure 3) reveals a moderate positive correlation, suggesting that larger modules tend to be more complex.

4.2. Requests. The Requests project is larger, with 34 modules and 11,248 LOC. It shows better average metrics: a lower cyclomatic complexity (2.87), a similar MI (73.26), a higher Pylint score (5.96), and fewer total PEP8 violations (533) (Table 2). Figure 4 displays the average cyclomatic complexity per module, with test-related modules showing the highest values. The distribution of Pylint scores (Figure 5) is skewed toward higher values, reflecting better overall code quality. Figure 6 shows the relationship between LOC and cyclomatic complexity, with a moderate positive correlation ($r = 0.41$).

5. DISCUSSION

This study provides empirical insights into the code quality of two foundational Python libraries, directly addressing our research questions.

5.1. RQ1: Variation of Metrics within Projects. The analysis reveals significant internal variation, identifying clear refactoring candidates. In BeautifulSoup4, dammit.py is a clear outlier with high complexity (3.85), low maintainability (40.08), and a high density of PEP8 violations. Similarly, test_tree.py is very large and complex with a critically low MI of 6.87. In Requests, test_requests.py is the primary concern, being the largest and one of the most complex modules, with an MI floored

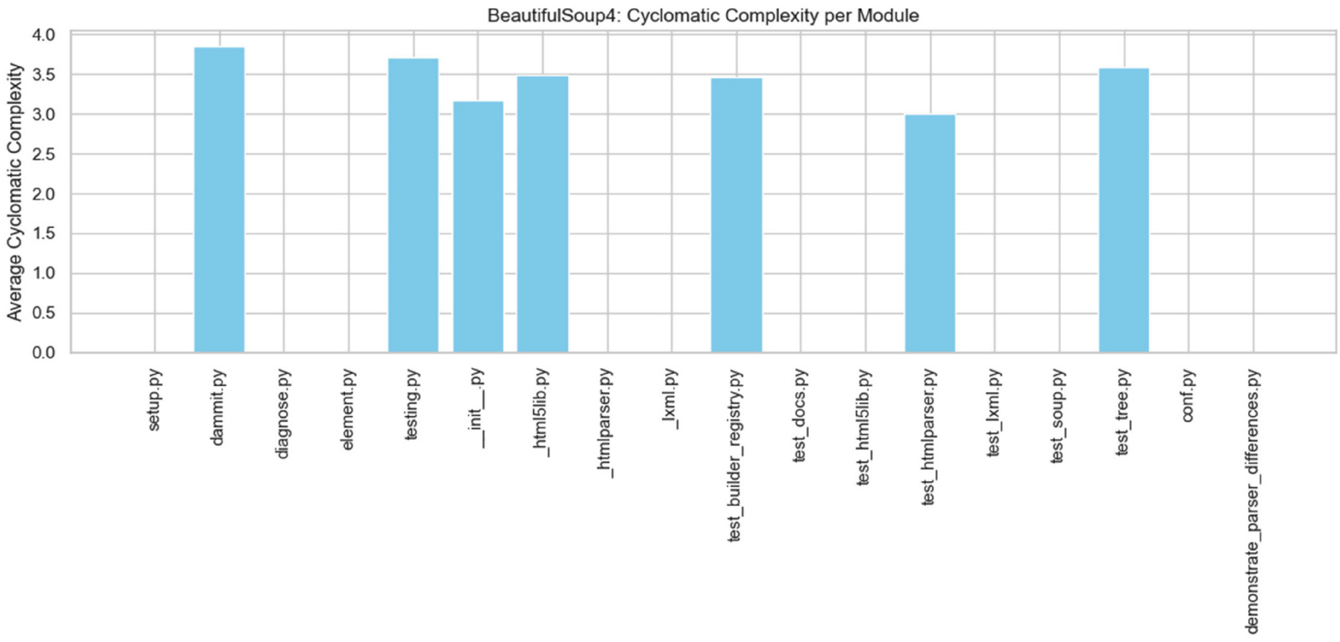


Figure 1. Average cyclomatic complexity per module in BeautifulSoup4; higher bars indicate greater structural complexity

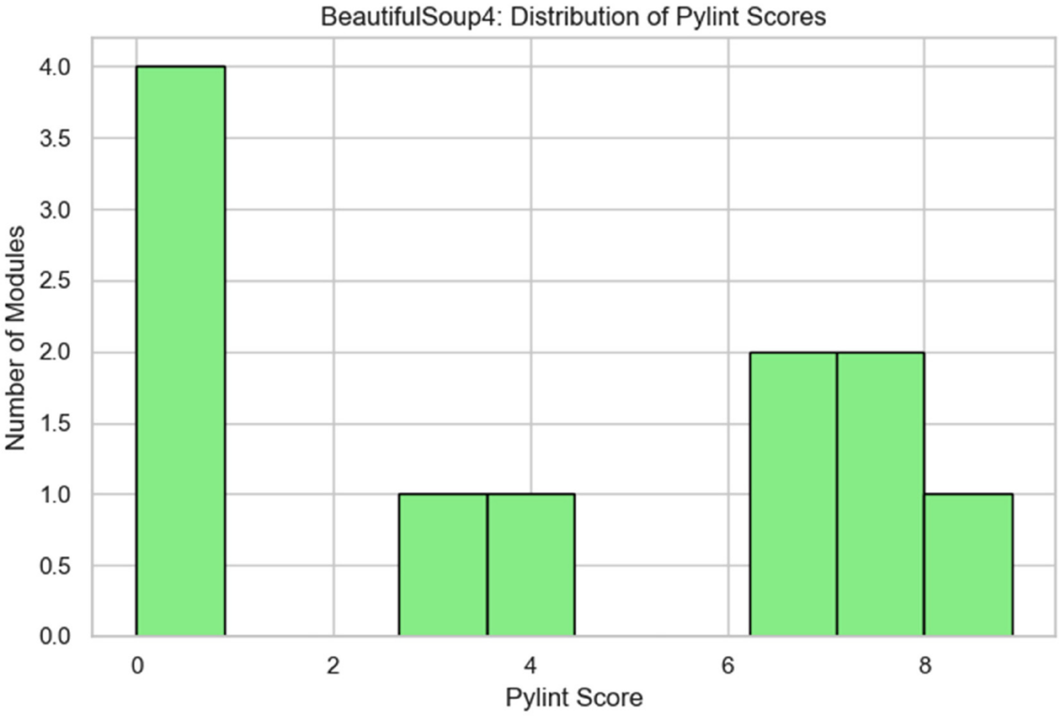


Figure 2. Distribution of Pylint scores, showing a bimodal spread across modules

at 0. This answers RQ1 by pinpointing the most complex and least maintainable modules within each project.

5.2. RQ2: Differences between Projects. Comparing the two projects, Requests demonstrates higher average code quality. It has a lower overall cyclomatic complexity (2.87 vs. 3.40) and a higher average Pylint score (5.96 vs. 3.98). While both projects have a similar number of total PEP8 violations, the distribution is different; BeautifulSoup4's violations are concentrated in a few key modules, whereas Requests' are more spread out. This suggests that Requests may have more consistent coding practices and a more mature quality assurance process, likely due to its larger and more structured contributor base.

5.3. RQ3: Refactoring Candidates. Based on the aggregated metrics, we can clearly identify the most critical refactoring candidates:

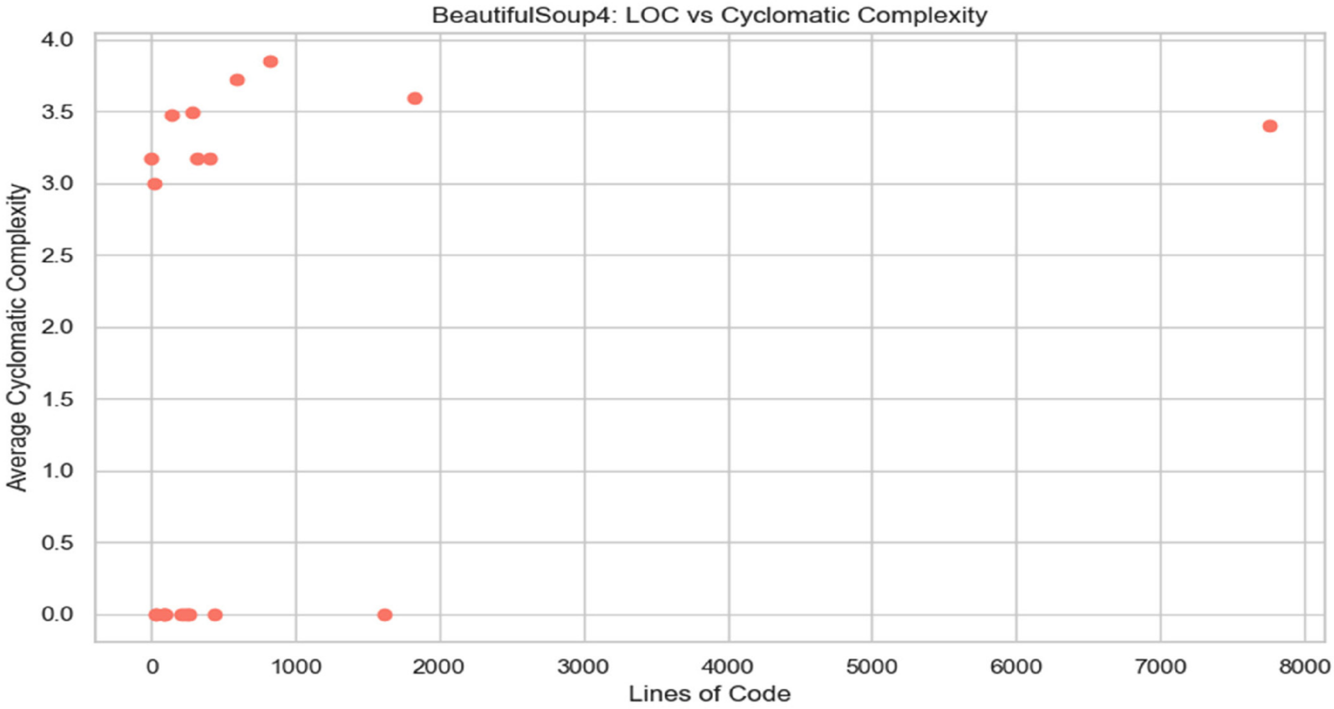


Figure 3. Relationship between lines of code (LOC) and cyclomatic complexity in BeautifulSoup4

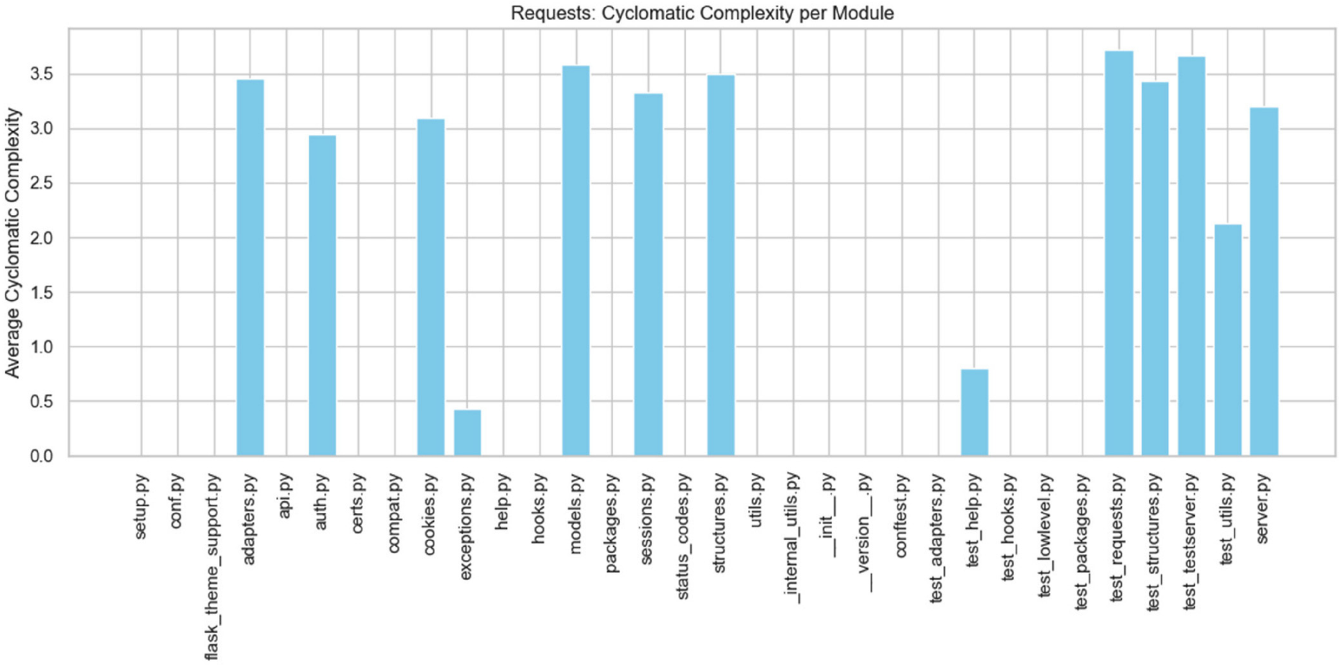


Figure 4. Average cyclomatic complexity per module in Requests; test-related modules (test_requests.py, test_testserver.py) show the highest values

- **BeautifulSoup4:** dammit.py is the highest priority, requiring immediate refactoring to simplify logic and fix style issues. test_tree.py and testing.py are also important targets.
- **Requests:** test_requests.py should be broken down into smaller, less complex test suites. models.py also shows high complexity and low maintainability, warranting investigation.This concentration of violations in complex modules aligns with literature linking style and quality[10].

5.4. Analytical Depth and Broader Implications.

A correlational analysis reveals a moderate negative relationship between cyclomatic complexity and maintainability index across all modules ($r = -0.58$, $p < 0.001$). This statistically significant correlation underscores that as code becomes more complex, its maintainability decreases, validating the interconnected nature of these metrics.

The findings have several practical implications for software sustainability:

- **Readability and Collaboration:** A high density of PEP8 violations, as seen in dammit.py, directly impairs readability. This creates a barrier to entry for new contributors and slows down collaborative development, ultimately threatening the long-term health of open-source projects.
- **Technical Debt Management:** The identified high-complexity, low-MI modules represent significant technical debt. Prioritizing their refactoring is an investment in long-term sustainability, reducing the future cost of changes and bug fixes.
- **Quality Assurance Automation:** Integrating static analysis tools into CI/CD pipelines can prevent the accumulation of technical debt by enforcing quality gates for new contributions, ensuring consistent quality across distributed teams.

6. LIMITATIONS AND ETHICAL CONSIDERATIONS

This study has several limitations. The analysis is restricted to two Python projects, which, while influential, limit the generalizability of the findings. Future work should expand to a larger and more diverse set of projects. Furthermore, static analysis provides a snapshot of code quality and does not capture runtime behavior or developer intent.

Table 2. *Static Analysis Metrics for Requests*

Module/File	LOC	CC (avg)	Maintainability Index	Pylint Score	PEP8 Violations
test_requests.py	3040	3.72	0.0	7.66	61
models.py	1039	3.58	21.69	7.83	28
adapters.py	696	3.45	47.45	6.49	36
test_testserver.py	165	3.67	49.21	8.58	5
internal_utils.py	50	0.0	100.0	10.0	2
SUMMARY	11248	2.87	73.26	5.96	533

Note: A Maintainability Index (MI) of 0.0, as seen in test_requests.py, occurs when the calculated index is non-positive due to very high complexity and size overwhelming the effect of comments, and is reported as 0 by the Radon tool.

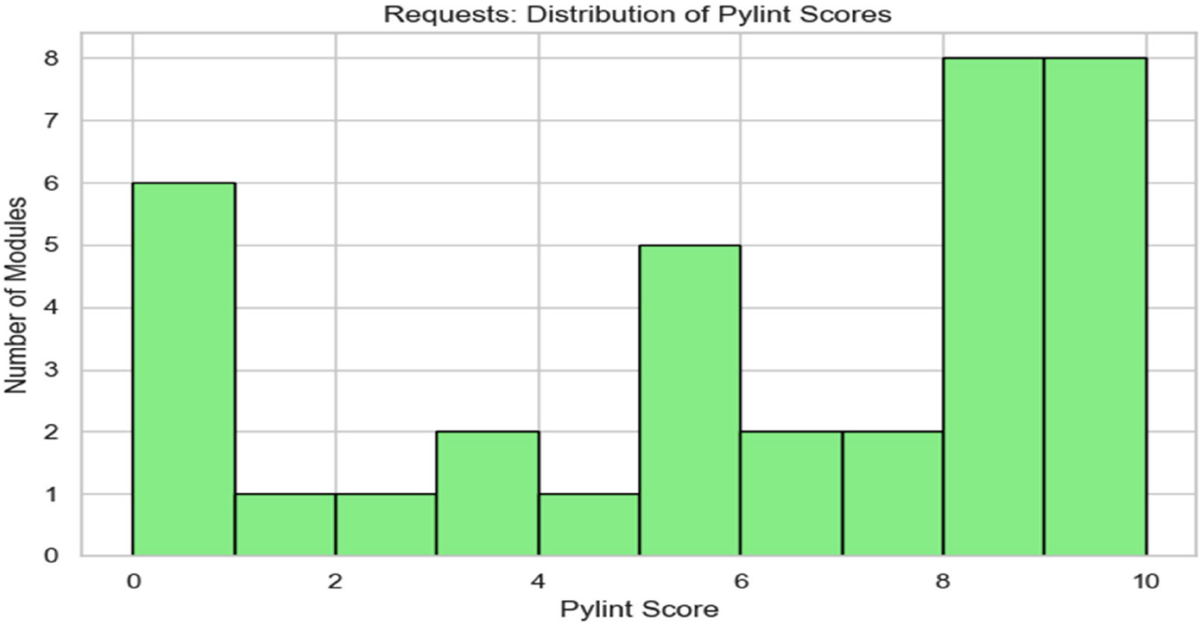


Figure 5. Distribution of Pylint scores across Requests modules; higher scores reflect better overall code quality

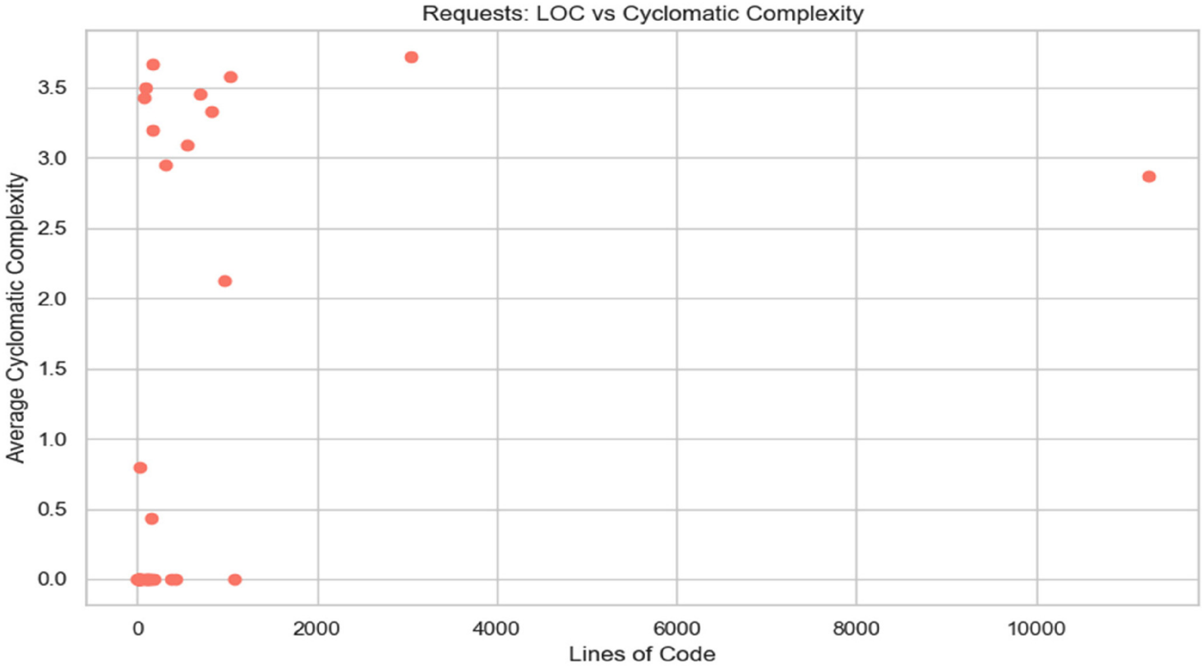


Figure 6. Relationship between lines of code (LOC) and cyclomatic complexity in Requests; moderate positive correlation ($r = 0.41$)

Regarding ethical considerations, this study exclusively used publicly available source code from open-source repositories. No private or proprietary software was analyzed. The research respects the licenses of the projects studied and aims to contribute back to the open-source community by providing insights that can help improve their code quality.

7. CONCLUSION

This study presented a reproducible static analysis of code quality in the Requests and BeautifulSoup4 projects. While generally maintainable, certain modules exhibited high complexity and numerous style violations, identifying them as prime refactoring candidates. Requests demonstrated more consistent code quality overall compared to BeautifulSoup4.

The primary contribution of this work is the detailed, empirical, and reproducible snapshot of code quality in these essential libraries, providing actionable insights for their maintainers. The significant negative correlation between complexity and maintainability reinforces the importance of monitoring these metrics. Furthermore, we contribute a transparent methodology and a full replication package for the community.

Beyond immediate maintainer benefits, this work has broader societal and educational implications. By improving code quality in foundational open-source libraries, we contribute to more reliable and sustainable software ecosystems that support education, research, and industry. The reproducible methodology and public dataset also serve as a valuable resource for students and educators teaching software engineering best practices, promoting transparency and rigor in empirical software research.

For future work, we plan to expand this analysis to a larger set of projects from different domains, including historical analysis to track quality evolution, and incorporate machine learning techniques to predict defect-prone modules based on these static metrics.

8. SUPPLEMENTARY INFORMATION

All code, raw data, and detailed instructions for reproducing the results, including the exact shell commands and environment setup, are available in our supplementary repository: https://github.com/BashirAdam/Static-Analysis-of-Code-Quality-in-Open-Source-Python- Projects. A single script (run_analysis.py) is provided to regenerate all tables and figures.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author and Corresponding Author

Bashir Adam Ahmed Ali – *Department of Software Engineering, Ostim Technical University, Ostim, 06374 Yenimahalle/Ankara, Turkiye;* 0009-0009-0267-2904
Email: 210208999@ostimteknik.edu.tr

ACKNOWLEDGEMENTS

The author would like to acknowledge the Department of Software Engineering at Ostim Technical University for their support. Special thanks also to the open-source developer communities behind BeautifulSoup4 and Requests, whose projects formed the basis of this analysis.

REFERENCES

(1) Python Software Foundation. (2024). Python. https://www.python.org.

(2) Liu, H., Jin, J., Xu, Z., Zou, Y., & Bu, Y. (2021). Deep learning-based code smell detection. IEEE Transactions on Software Engineering, 47(9), 1811–1837. https://doi.org/10.1109/TSE.2019.2936376

(3) Radon. (2024). Radon Documentation. Radon Project. https://radon.readthedocs.io.

(4) Pylint. (2024). Pylint Documentation. Python Code Quality Authority. https://pylint.org.

(5) Flake8. (2024). Flake8 Documentation. Python Code Quality Authority. https://flake8.pycqa.org.

(6) Fontana, F. A., Mantyla, M. V., Zanoni, M., & Marino, A. (2016). Comparing and experimenting machine-learning techniques for code-smell detection. Empirical Software Engineering, 21(4), 1143-1191.https://doi.org/10.1007/s10664-015-9376-6

(7) Vassallo, C., Panichella, S., Palomba, F., Proksch, S., Zaidman, A. & Gall, H. C. (2018). Context is king: The developer perspective on the usage of static analysis tools. Proceedings of SANER 2018 25th International Conference on Software Analysis, Evolution and Re-engineering, 38–49 (2018). https://doi.org/10.1109/SANER.2018.8330195

(8) McCabe, T. J. (1976). A complexity measure. IEEE Transactions on Software Engineering, SE-2(4), 308-320. https://doi.org/10.1109/TSE.1976.233837

(9) Iftikhar, U., Ali, N. B., Börstler, J., & Usman, M. (2024). A tertiary study on links between source code metrics and external quality attributes. Information and Software Technology, 165, 107348. https://doi.org/10.1016/j.infsof.2023.107348

(10) Siddik, M. S., & Bezemer, C.-P. (2023). Do code quality and style issues differ across (non-)machine learning notebooks? Yes! In 2023 IEEE 23rd International Working Conference on Source Code Analysis and Manipulation (SCAM) (pp. 72–83). IEEE. https://doi.org/10.1109/SCAM59687.2023.00018

(11) Scalabrino, S., Bavota, G., Vendome, C., Linares-Vásquez, M., Poshyvanyk, D., & Oliveto, R. (2017). Automatically assessing code understandability: How far are we? *Proceedings of the 32nd IEEE/ACM International Conference on Automated Software Engineering*, 417-427. https://doi.org/10.1109/ASE.2017.8115654

(12) Grotov, K., Tiufanov, S., Serebriakova, Y., Golubev, Y., Kovalchuk, N., & Bryksin, T. (2022). A large-scale comparison of Python code in Jupyter notebooks and scripts. In Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE ’22) (pp. 653664). ACM. https://doi.org/10.1145/3524842.3528447

(13) Chen, Z., Chen, L., Ma, W., Zhou, X., Zhou, Y., & Xu, B. (2018). Understanding metric-based detectable smells in Python software: A comparative study. Information and Software Technology, 94, 1429. https://doi.org/10.1016/j.infsof.2017.09.011

(14) AlOmar, E. A., Mkaouer, M. W., & Ouni, A. (2021). Can refactoring be self-affirmed? An exploratory study on how developers document refactoring activities. IEEE Access, 9, 21449-21467. https://doi.org/10.1109/ACCESS.2021.3055175

(15) Liu, H., Jin, J., Xu, Z., Zou, Y., & Bu, Y. (2021). Deep learning-based code smell detection. IEEE Transactions on Software Engineering, 47(9), 1811-1837. https://doi.org/10.1109/TSE.2019.2961345

(16) Reitz, K. (2024). Requests: HTTP for Humans. GitHub repository. Python Software Foundation. https://github.com/psf/requests.

(17) Richardson, L. (2024). Beautiful Soup Documentation. Crummy.com. https://www.crummy.com/software/BeautifulSoup/.

KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Article Electrical Engineering/Mechanical Engineering/Physica

Numerical Investigation of a Divided-Leg Thermoelectric Generator for Enhanced Power and Efficiency

Md. Habibur Rahman Aslam, Foyzul Karim, and Anisul Islam Suva*

Cite https://doi.org/10.64589/juri/214391

Submitted: June 2, 2025 **Revised:** July 20, 2025 **Accepted:** November 18, 2025

ABSTRACT

Given the ongoing energy crisis, researchers plus policymakers worldwide are pursuing alternative energy solutions. Thermoelectric generators (TEGs) are an attractive approach offering a sustainable and efficient means to address current energy demands. Despite the numerous TEG designs, their efficiency remains poor, posing a challenge to their widespread adoption. In this study, we introduce a leg-division method designed to enhance the efficiency of TEGs. To establish its superiority, the design was simulated using COMSOL Multiphysics 5.5, and its performance was benchmarked against a conventional rectangular leg geometry. Across the two designs, a temperature gradient of 280 K ($T_h = 573$ K, $T_c = 293$ K) was maintained, and the material volume was held constant to ensure a fair comparison. The results show that the new model achieves an open-circuit voltage of 64.3 mV, which is 1.2 times higher than that of the baseline geometry. The maximum power output reaches 49.22 mW, 2.8 times greater than the conventional model, and the maximum efficiency is 6.69%, indicating 2.1×improvement over the prior geometry. The enhanced performance of the new TEG design has the potential to redefine deployment in renewable energy generation by improving waste heat harvesting, thereby advancing sustainable power generation.

Keywords: thermoelectric generators, divided leg, geometry, energy, COMSOL

1. INTRODUCTION

In recent decades, the world has faced an escalating energy crisis driven by rising demand from urbanization, industrialization, economic growth, and rapid technological development[1]. At the same time, primary energy sources such as hydrocarbon resources are being depleted[2]. Furthermore, the combustion of fossil fuels markedly elevates CO_2 output, contributing to sea-level rise, natural disasters, and global warming[3]. In this context, renewable energy sources provide robust means for tackling these pressing challenges[4]. Thermoelectric generators (TEGs) contribute to solving the energy crisis and mitigating environmental impacts by converting waste heat into electricity and reducing the reliance on fossil fuels, thereby supporting the shift toward renewable and sustainable energy sources[5]. However, their application is limited by their bulk size, low energy conversion efficiency, material costs, and the complexity of thermal management in practical systems[6,7].

Several studies have been conducted to enhance the efficiency of TEGs, which fall into three main approaches. First, improving the thermoelectric material properties[8] to achieve a higher figure of merit (ZT) substantially boosts the performance. Second, segmentation[9] techniques offer higher output power and conversion efficiency than conventional designs. Third, designing novel geometries[10] for the device further enhances the heat transfer and overall efficiency.

Khalil et al. evaluated different geometrical shapes (five shapes: rectangular, I-leg, X-leg, trap-leg, and Y-leg) of the same volume using the finite element method. The study found that a rectangular leg model in a thermoelectric generator with a 6 mm leg length achieved the highest efficiency (5.482%) and output power, outperforming other shapes despite identical volumes[11]. Quan et al.[12] designed a TEG system with a polygonal heat exchanger and sickle-shaped fins to increase waste-heat reclamation. Transient CFD analysis under CLTC conditions showed improved performance, with power rising from 24.48 W to 34.48 W and confirmed fidelity, aiding TEG optimization for automotive use. Luo et al.[13] proposed a novel Y-type TEG configuration that increased the output power and conversion efficiency by 87.5% and 42.62% respectively, at the exhaust temperature of 450 K and velocity of 20 m/s when applying the Y-type design ($\Delta L = 0$ mm) instead of the conventional π-type (or rectangle-shaped) TEG. Ibeagwu et al.[14] studied geometries such as rect-leg, trap-leg, Y-leg, I-leg, and X-leg with varied leg forms and found that the X-leg offers the highest power density. Siddique et al. found that under a constant temperature boundary condition ($\nabla T = 10$ °C), the rectangular TEG produced 1.4 times more power than the trapezoid configuration. However, the power density of the rectangular configuration remains comparable to that of the trapezoidal one[15].

In this study, we investigated the divided-leg geometry within the device to improve its output performance. This innovative design effectively reduces the internal resistance, resulting in an elevated open-circuit voltage, enhanced power generation, and superior energy-conversion efficiency. The notable performance improvements in this structure demonstrate its substantial potential for renewable energy applications, enabling pathways to optimize TEGs in sustainable energy systems. These findings highlight the importance of geometric modifications in advancing thermoelectric technology for practical energy solutions.

Following the general background, the introduction should summarize prior work relevant to the present study[1]. This includes citing earlier results and highlighting the current state of knowledge. Authors should subsequently identify gaps, inconsistencies, or limitations in the literature that justify the need for the present research. For example, while traditional steel dowels have been widely studied, research on small-diameter FRP dowels in real-world applications remains limited[2,3]. Such gaps motivate further investigation and emphasize the originality of the current work. Alvaro et al.[4] found that the bond strength of three different GFRP rebar types—sand-coated, helically grooved, and lugged—varied based on surface enhancement, but was not significantly affected by exposure time or temperature in seawater environments.

Next, the purpose and objectives of the present study should be clearly stated. Authors should define what their research aims to investigate or demonstrate and succinctly outline methods. This creates a coherent bridge from the literature background to the specific contribution of the current paper. A well-defined research objective helps audiences understand the scope and importance of the study.

In some cases, authors may opt to add a final paragraph outlining the structure of the paper, briefly describing what each subsequent section will address. To ensure proper citation and referencing throughout the manuscript, it is strongly advised that authors use reference management software such as Mendeley, Zotero, or EndNote.[5,6]. Zotero was the reference manager employed to manage citations in this document. These tools support the Nature citation style required by the Journal of Undergraduate Research International and streamline the process of formatting in-text citations and the reference list. This section should be focused and concise, typically not exceeding 10% of the total word count of the paper.

2. DIVIDED LEG STRUCTURE

In the leg division method, each TEG leg is divided into two parts—upper and lower—with a metallic conductor incorporated between them, as illustrated in Figure 1a. These segments comprise alternating thermoelectric semiconductors that generate opposite thermoelectric potentials. Specifically, in the left leg of the proposed divided-leg geometry, the upper section contains a p-type semiconductor. Conversely, the lower section is composed of an n-type semiconductor. In p-type semiconductors, holes act as the majority charge carriers; they absorb thermal energy from the hot side and diffuse it toward the central conductor. In contrast, n-type semiconductors have electrons as the majority charge carriers, which diffuse toward the cold side. This movement results in effective hole diffusion toward the central conductor, establishing a thermoelectric potential across the leg. Thus, a positive potential develops at the central conductor of the left leg owing to the directional movement of the charge carriers. Similarly, the right leg of Figure 1a comprises an n-type semiconductor in the upper part and a p-type semiconductor in the lower part, resulting in the development of a negative potential at the central conductor. These two central conductors function as the electrical terminals of the divided-leg geometry, enabling connection to the external circuit.

Moreover, the conductor at the center serves dual roles: it acts as both an electrical junction and a thermal interface. This structure limits Joule heating and enables more efficient thermoelectric conversion by allowing localized charge separation and shorter diffusion paths for charge carriers. This design strategy is especially beneficial for microscale or densely packed TEG systems, where resistive losses can significantly affect overall performance.

The equivalent circuit diagrams for the divided-leg geometry and the rectangular leg are shown in Figure 2 a and b, respectively.

Each semiconductor (p or n), conductor, and soldering junction has its own internal resistance, denoted as R_P or R_N, R_C, and R_{Sh}. The total internal resistance of the system is obtained by

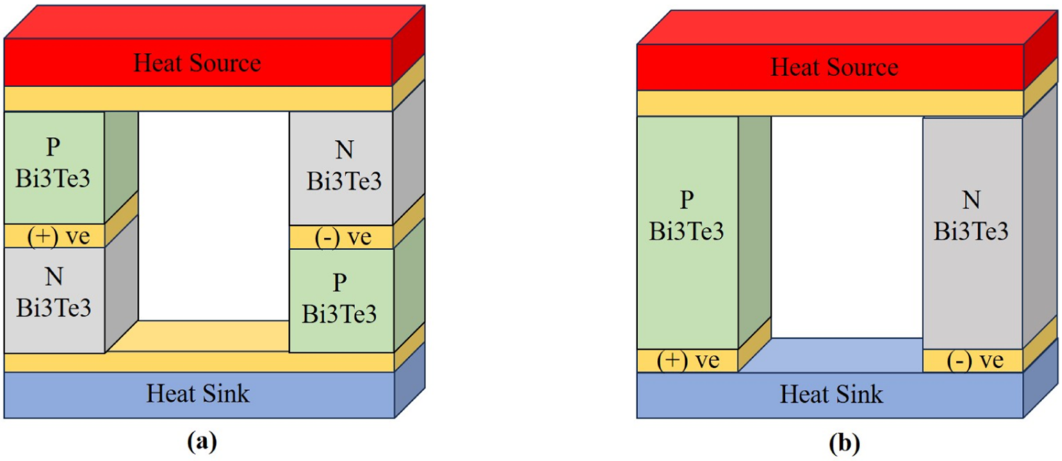


Figure 1. Schematic diagram of (a) divided-leg geometry, (b) rectangular leg geometry

summing component values. Equivalent circuit diagrams for the divided-leg and rectangular-leg geometries appear in Figure 2 (a) and (b), respectively.

The basic relationship of electrical resistance is given by the following equation:

$$\mathrm{R} = \frac{l}{\sigma A} \tag{1}$$

where R is the resistance, l is the length of the conductor, σ is the electrical conductivity, and A is the cross-sectional area. Considering that the copper conductor resistance ($\mathrm{R_C}$) and soldering resistance ($\mathrm{R_{Sh}}$) are negligible, the equivalent internal resistance of the rectangular TEG is:

$$\mathrm{R_{REC}} = \mathrm{R_P} + \mathrm{R_N} \tag{2}$$

where $\mathrm{R_P}$ and $\mathrm{R_N}$ denote the resistances of the p- and n-type legs, respectively.

In a divided-leg geometry, shortening the effective electrical path within each material segment by splitting them into equal upper and lower halves reduces the resistive losses, enhances the current flow, and improves the voltage output. Each leg forms two branches (p- and n-type) and two parallel branches, each consisting of $\mathrm{R_P}/2$ and $\mathrm{R_N}/2$ in series. Because the electrical resistance in parallel pathways combines inversely, this geometry effectively reduces the total internal resistance of the TEG to approximately one-fourth of that in a conventional series-connected rectangular leg design. Thus, a reduction in the internal resistance occurs in the leg division geometry owing to the parallel configuration.

Thus, the total internal resistance in this configuration can be written as:

$$\mathrm{R_{REC}} = \frac{\mathrm{R_P} + \mathrm{R_N}}{4} \tag{3}$$

Although the equivalent circuit diagrams in Figure 2 explicitly include $\mathrm{R_C}$ and R_{Sh}, these components are considered negligible in the analytical simplification to streamline the resistance calculations. While this is a common assumption because of their typically low values relative to the semiconductor resistance, notably, neglecting $\mathrm{R_C}$ and $\mathrm{R_{Sh}}$ may introduce minor deviations in precise modeling, especially in high-precision or miniaturized TEG systems.

This reduction in the internal resistance reduces dissipation, allowing higher current flow and improving the voltage output, thereby enhancing the overall power generation efficiency of the thermoelectric generator.

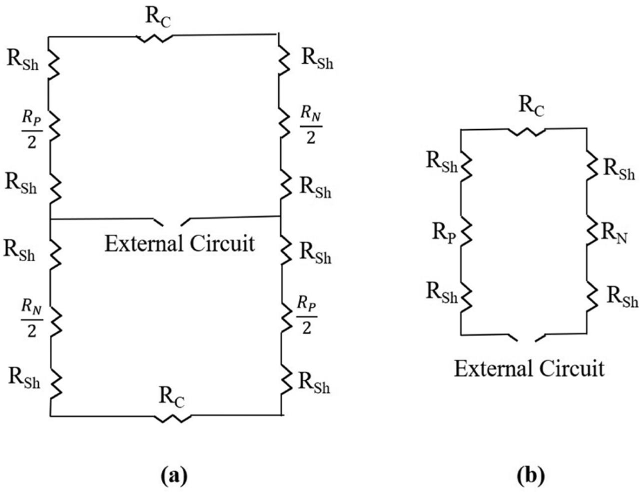


Figure 2. Circuit diagram of (a) divided-leg geometry, (b) rectangular leg geometry

3. SIMULATION

Both TEG models were simulated using COMSOL Multiphysics 5.5 by integrating the Heat Transfer and AC/DC modules to resolve their thermal and electrical characteristics. The Heat Transfer module modeled the temperature distribution and heat flow, whereas the AC/DC module solved the electrical potential and current flow. This coupled approach is critical for evaluating thermoelectric performance under realistic operating conditions. In these simulations, the heat source temperature was fixed at 573 K, and the heat sink temperature was set to 293 K.

3.1. Necessary Equations of Simulation.

Energy flows as heat from the hot side and is absorbed by thermoelectric materials, as governed by conservation of energy:

$$\rho C_{p.}u\nabla T + \nabla q = Q + Q_{ted} \tag{4}$$

where ∇q would represent the divergence of the heat flux vector, and Heat flux, $q = -k.\nabla T$

This equation describes the conductive heat transfer and includes thermoelectric heat sources, including the Joule and Peltier effects.

The electric current interface models current conservation based on Ohm's law, using the scalar electric potential as the primary variable. This enables the computation of the electric potential, current density, and electric field within the conductive materials.

$$\nabla J = Q_j \tag{5}$$

$$J = \sigma E + J_e \tag{6}$$

$$E = -\nabla V \tag{7}$$

These equations capture the response of electrical currents to electric fields and external sources within the material. This multiphysics interface combines electric currents and heat transfer in solid modules, enabling the simulation of the interrelated Peltier, Seebeck, and Thomson effects in thermoelectric devices.

$$Q = P.J \tag{8}$$

$$P = S.T \tag{9}$$

$$J_e = -\sigma S \nabla T \tag{10}$$

Nomenclature:

$\mathbf{C_p}$ – Heat capacity (J/kg.K)
J – Induced current density
ρ – Material density ($\mathrm{kg/m^3}$)
$\mathbf{J_e}$ – External current density
T – Temperature (K)
$\mathbf{Q_j}$ – Current source
Q – Volumetric heat generation from internal sources ($\mathrm{W/m^3}$)
k – Thermal conductivity

Q_{ted} – Thermoelectric heat source or sink associated with the Thomson and Peltier effects (W/m^3)
σ – Electrical conductivity
P – Peltier Coefficient
h – Convective heat transfer coefficient
α – Seebeck coefficient
ΔT – Temperature difference
E – Electric field
V – Electric potential
ZT – Dimensionless figure of merit

3.2. Materials.

To compare the performance of both models, Bismuth Telluride (Bi_2Te_3) was used as the thermoelectric material. The Seebeck coefficient, electrical conductivity, and thermal conductivity were the primary thermoelectric properties evaluated. The properties of Bi_2Te_3 appear in Table 1.

Ceramic (Al_2O_3) was used as the hot- and cold-side material because of its excellent thermal stability, high-temperature resistance, and low thermal expansion, which make it ideal for maintaining structural integrity during thermal cycling. Copper was chosen as the conductor because of its high electrical and thermal conductivities, which ensure minimal resistive losses and efficient heat transfer. For both models, the cross-sectional area was set to 1 mm^2 and the total length to 2.5 mm.

4. RESULTS AND DISCUSSION

The performance of the TEG was evaluated using key metrics: the open-circuit voltage, load terminal voltage, load current, output power, and energy conversion efficiency. The geometrical structure of the TEG strongly affects these parameters. A comparison of these parameters between the two models demonstrates the validity of the proposed design.

4.1. Thermal and Voltage Distribution.

Figure 3 a and b show the simulated thermal distributions for the proposed divided-leg geometry and conventional rectangular leg geometry, respectively. In both models, a temperature gradient of 280 K was imposed, with the hot side maintained at 573 K and the cold side maintained at 293 K. However, the divided-leg geometry ensures slightly improved thermal utilization by reducing the localized thermal resistance at the conductor junction, facilitating efficient heat transfer through the thermoelectric materials. Although the overall thermal profiles appear similar, the performance enhancement in the divided-leg design mainly arises from its lower internal resistance and optimized electrical configuration rather than from a significant difference in temperature distribution.

Table 1. *Temperature-Dependent Material Properties*[16] *of* Bi_2Te_3

Temperature (K)	Seebeck Coefficient (μ V/K)	Thermal Conductiviy (W/mK)	Electrical Conductivity (S/m)
200	168	2.4	1.4286×10^5
250	192	1.9	1.1111×10^5
300	210	1.6	0.86957×10^5
350	225	1.6	0.71429×10^5
400	237	1.75	0.58824×10^5

The potential developed at the terminal of the divided-leg geometry is 64.3 mV, as shown in Figure 3 (c), while the conventional rectangular-shaped geometry generates a terminal potential of 54 V, as shown in Figure 3 (d). The COMSOL simulation confirms that the divided-leg geometry not only improved heat distribution but also enhanced voltage development, demonstrating superior performance for thermoelectric energy conversion.

4.2. Load Current and Terminal Voltage.

When a load is connected between the two terminals of the TEG, current flows through the load. The magnitude of the current can be determined via Ohm's law.

$$I = \frac{V}{R_L + R_{int}} \tag{11}$$

Additionally, the internal resistance of the device causes a voltage drop that reduces the available terminal voltage at the load. The load-end TEG voltage is given by

$$V = \frac{R_L \times V_{OC}}{R_L + R_{int}} \tag{12}$$

Figure 4 (a) shows that the divided-leg geometry consistently yields a higher current than the rectangular geometry, particularly at lower resistances. Additionally, the divided-leg geometry again outperforms, maintaining a higher voltage across the range, with the gap most noticeable at lower resistances, as shown in Figure 4 (b). Both graphs highlight the greater electrical efficiency of the divided-leg geometry. Elevated current enhances the power output and improves the energy delivery to low-resistance loads, making the TEG more suitable for practical applications. A higher terminal voltage reduces the transmission losses, enhances interfacing to commercial electronics, and facilitates their integration into external circuits. Together, the high current and voltage optimize the power efficiency, increase the system flexibility, and expand the operational range of the TEG, making it more effective for energy harvesting and remote power generation.

4.3. Power.

This current flow generates power at the load terminals, representing the usable output power. This power is computed using the following equation:

$$P = I^2 R_L \tag{13}$$

The graph in Figure 5 shows the power developed for both the divided-leg geometry and the rectangular TEG. The divided-leg geometry achieves a maximum power output of 49.22 mW, significantly outperforming the rectangular leg shape, which yields 17.36 mW under identical conditions.

Rectangular leg geometry is commonly used owing to its structural simplicity and reliability[11,15]. Several recent studies have demonstrated that modified geometries can significantly improve the performance of TEGs. Siddique et al.[15] found that for a given restricted volume envelope, the trapezoidal-shaped TE prototype generates 1.24 times more voltage and 1.5 times more power when compared to the rectangular-shaped prototype at 30 °C hot side temperature when the cold side is exposed to the surroundings. Okirigiti et al.[17] reported that the trapezoidal-shaped legs-TEG with a top surface area of 2.5 × 2.5 mm^2 converts the

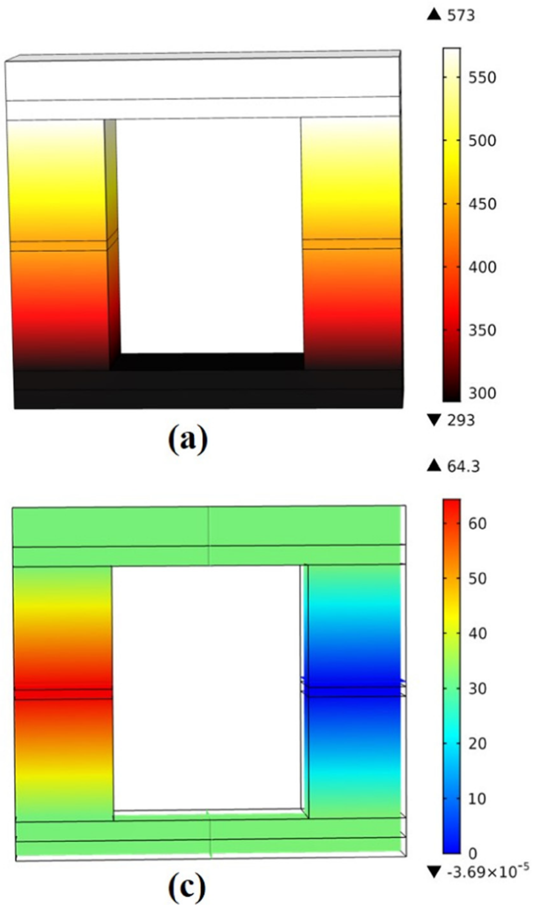


Figure 3. Thermal Distribution of divided-leg geometry, (b) Thermal Distribution of rectangular leg geometry, (c) Voltage Distribution of divided-leg geometry, (d) Voltage Distribution of rectangular leg geometry

output power of 0.60 mW from a temperature gradient (ΔT) of 200 K, which represents a two-fold increase compared to the cuboid-shaped legs (5×5 mm^2)-based TEG. Yin et al.[18], conducted a comparison-based study between the developed T-shaped and the conventional π-shaped TEGs, and reported that the specific output power was significantly increased by 90.1% under the same given thermal conditions.

Compared with existing geometric configurations, the proposed divided-leg design yields a 2.8-fold increase in the power output, outperforming all previously reported modifications. This significant enhancement underscores the superior efficacy of divided-leg geometry. Elevated power output in TEGs is critical because it directly improves the energy conversion efficiency, thereby facilitating more effective utilization of waste heat in sustainable power generation applications, including industrial processes and renewable energy systems.

4.4. Efficiency.

Electrical power is the transduction of heat energy absorbed from a heat source. The amount of absorbed heat energy is calculated using the following equation:

$$Q_h = \alpha T_h I - 0.5(I^2 R_{int}) + K(T_h - T_c) \tag{14}$$

where, $\alpha = \alpha_p - \alpha_n$, $K = \frac{k_p A_p}{L_p} + \frac{k_n A_n}{L_n}$, R_{int} = internal resistance, I = current flow.

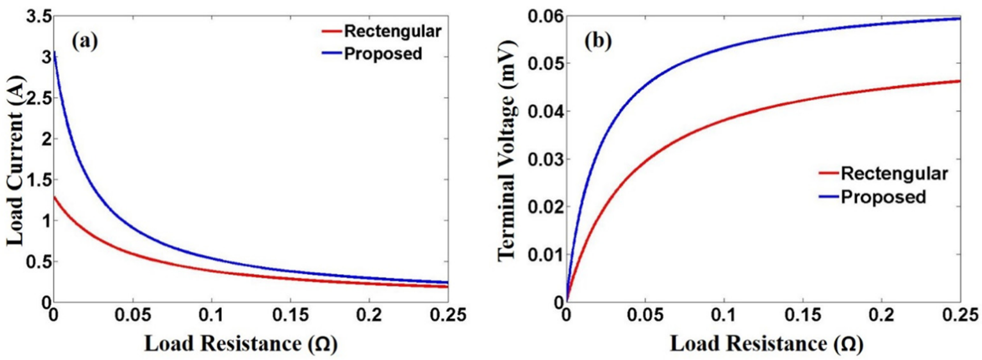


Figure 4. Graph of (a) Load Current (A) vs Load Resistance (Ω), (b) Load terminal voltage (mV) vs Load Resistance (Ω)

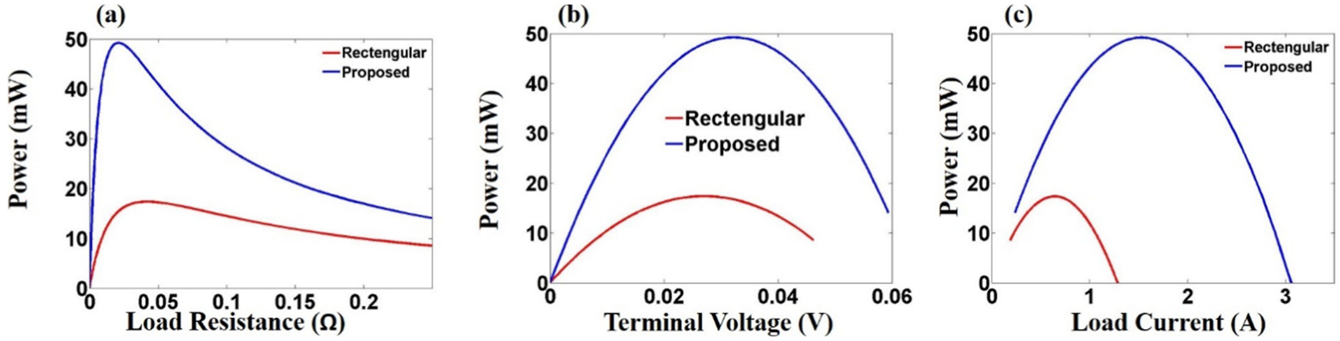


Figure 5. Graphs of (a) power vs. Load Resistance, (b) power vs. terminal voltage, and (c) power vs. load current

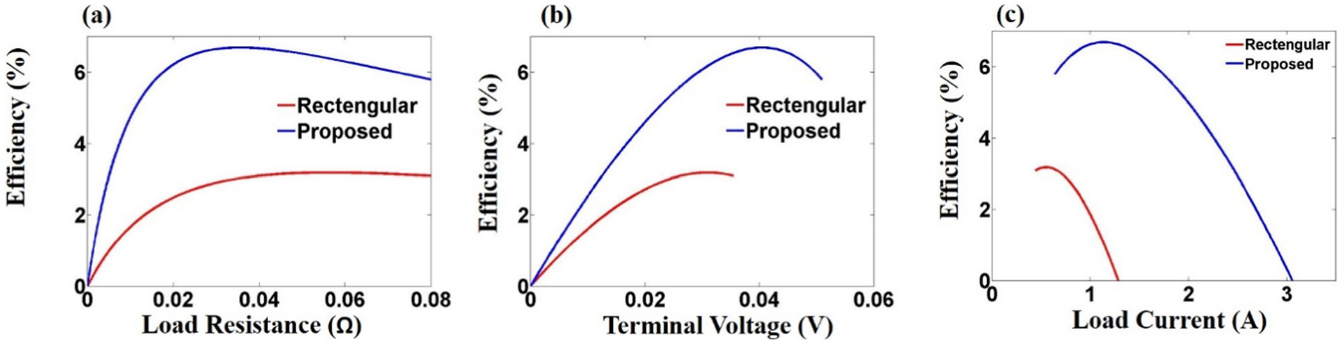


Figure 6. Graphs of (a) efficiency vs. load resistance, (b) efficiency vs. terminal voltage, and (c) efficiency vs. load current

The efficiency of this energy conversion process is given as the ratio of the output electrical power to the absorbed heat energy.

$$\eta = \frac{P}{Q_h} \tag{15}$$

High efficiency in TEGs is vital because it maximizes the energy conversion from heat to electricity, enhancing practicality, material efficiency, and overall performance. Figure 6 shows the efficiency variations of the proposed TEG unit under different operating conditions. The graph demonstrates that the rectangular leg geometry achieves a maximum efficiency of 3.18%, whereas that of the proposed divided-leg geometry reached 6.69%.

Several recent studies have highlighted the influence of leg geometry on device performance. Khalil et al.[11] found that among five geometries of equal volume, the rectangular leg attained the highest efficiency (5.482%). Chen et al.[19] introduced a cylindrical TEG incorporating hot-side shunts and an optimized heat exchanger with segmented fin-hollow channels. By adjusting the blocking plate location and fin-to-hollow height ratio, they obtained a 108.8% increase in power and a 72.8% efficiency gain with optimal performance at moderate exhaust velocities, thereby guiding future TEG designs.

However, the present work demonstrates that the proposed divided-leg geometry achieves over a two-fold improvement in maximum efficiency. This substantial improvement highlights the advantages of the divided-leg geometry, including reduced internal resistance, enhanced current flow, and improved thermal gradient utilization. Higher efficiency not only boosts the electrical output of TEGs but also improves practicality for energy harvesting applications, particularly in waste heat recovery and remote power generation.

5. CONCLUSION

The newly proposed divided-leg geometry exhibited greater thermoelectric efficacy. For a temperature gradient of $\Delta T = 280$ K ($Th = 573$ K, $Tc = 293$ K), the traditional rectangular leg structure yielded an open-circuit voltage of 54 mV, while the divided-leg geometry achieved 64.3 mV. Additionally, this divided-leg design produced a maximum power output of 49.22 mW, compared to just 17.36 mW for the rectangular geometry. Consequently, the energy conversion efficiency of the divided-leg geometry reached approximately 6.69%, significantly outperforming the 3.18% efficiency of the traditional design. These results highlight the high potential and practical viability of the divided-leg geometry, offering a promising solution for improving TEG performance. These improvements make the divided-leg geometry a promising candidate for applications in wearable electronics, vehicle exhaust heat recovery, and remote power sources.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Authors

Md. Habibur Rahman Aslam – *Department of Electrical and Electronic Engineering, Chittagong University of Engineering & Technology (CUET), Chattogram, Bangladesh;*
0009-0008-0751-9645
Email: u1902135@student.cuet.ac.bd

Foyzul Karim – *Department of Electrical and Electronic Engineering, Chittagong University of Engineering & Technology (CUET), Chattogram, Bangladesh;*
0009-0007-8318-0519
Email: u1902169@student.cuet.ac.bd

Corresponding Author

Anisul Islam Suva – *Research Mentor, Institute of Energy Technology, Chittagong University of Engineering & Technology (CUET), Chattogram, Bangladesh;* 0009-0000-7967-7548
Email: anisulislam.me@cuet.ac.bd

ACKNOWLEDGEMENTS

The authors acknowledge the support and resources from COMSOL. Extensive user forums, tutorials, and shared simulation insights enabled the modeling and validation processes in this study. The collaborative knowledge base offered by COMSOL users worldwide has materially aided in overcoming technical challenges and improving simulation accuracy.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this manuscript.

FINANCIAL DISCLOSURE

The authors received no specific funding for this work.

REFERENCES

(1) Gajdzik, Bożena, et al. "The influence of the global energy crisis on energy efficiency: A comprehensive analysis." *Energies* 17.4 (2024): 947.

(2) Höök, Mikael, and Xu Tang. "Depletion of fossil fuels and anthropogenic climate change—A review." *Energy policy* 52 (2013): 797–809.

(3) Fang, JingYun, et al. "Global warming, human-induced carbon emissions, and their uncertainties." *Science China Earth Sciences* 54 (2011): 1458–1468.

(4) Poudyal, Ramhari, et al. "Mitigating the current energy crisis in Nepal with renewable energy sources." *Renewable and Sustainable Energy Reviews* 116 (2019): 109388.

(5) Mamur, Hayati, et al. "Thermoelectric generators act as renewable energy sources." *Cleaner Materials* 2 (2021): 100030.

(6) Chen, Jinlong, et al. "Enhanced efficiency of thermoelectric generator by optimizing mechanical and electrical structures." *Energies* 10.9 (2017): 1329. (6)

(7) Khan, Salman, et al. "High power density of radiative-cooled compact thermoelectric generator based on body heat harvesting." Nano Energy 87 (2021): 106180. (7)

(8) He, Hailong, et al. "Comprehensive modeling for geometric optimization of a thermoelectric generator module." *Energy conversion and management* 183 (2019): 645–659.

(9) Yang, Wenlong, et al. "Joint optimization of leg configuration and fin structure in a two-stage segmented thermoelectric generator for enhanced 3E performance." *Applied Thermal Engineering* 271 (2025): 126354.

(10) Schilz, J., et al. "A local selection criterion for the composition of graded thermoelectric generators." *Journal of applied physics* 83.2 (1998): 1150–1152.

(11) Khalil, ALkhadher, et al. "Performance comparison of TEGs for diverse variable leg geometry with the same leg volume." *Energy* 224 (2021): 119967.

(12) Quan, Rui, et al. "Dynamic performance of a polygonal thermoelectric generator using sickle-shaped fins for automotive application." *Applied Thermal Engineering* 255 (2024): 123985.

(13) Luo, Ding, et al. "A novel Y-type thermoelectric generator configuration for the fluid waste heat recovery." *Applied Thermal Engineering* 270 (2025): 126226.

(14) Ibeagwu, Onyebuchi Isreal. "Modelling and comprehensive analysis of TEGs with diverse variable leg geometry." *Energy* 180 (2019): 90–106.

(15) Siddique, Abu Raihan Mohammad, Shohel Mahmud, and Bill Van Heyst. "Performance comparison between rectangular and trapezoidal-shaped thermoelectric legs manufactured by a dispenser printing technique." *Energy* 196 (2020): 117089.

(16) Prasad, Asutosh, and Raj CN Thiagarajan. "Multiphysics modeling and multilevel optimization of thermoelectric generator for waste heat recovery." *Proceedings of the COMSOL Conference*. 2018.

(17) Okirigiti, Momanyi Amos, et al. "Enhanced energy harvesting performance of bendable thermoelectric generator enabled by trapezoidal-shaped legs." *Journal of Power Sources* 631 (2025): 236254.

(18) Yin, Tao, et al. "Performance analysis and design optimization of a compact thermoelectric generator with T-Shaped configuration." *Energy* 229 (2021): 120652.

(19) Chen, Jie, et al. "Innovative design and numerical optimization of a cylindrical thermoelectric generator for vehicle waste heat recovery." *Energy Conversion and Management* 326 (2025): 119478.

https://doi.org/10.64589/juri/214391

KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Article *Material Science and Engineering/Physics*

Antenna Design for Aircraft Situational Awareness: Coverage of ADS-B/TCAS and UHF SATCOM Frequencies

Taqi Abdulkhaleg Albaiyat, Bazilah Binti Baharom*, and Ayman Muhammad Abdullah

Cite https://doi.org/10.64589/juri/215033

Submitted: September 14, 2025 **Revised:** November 17, 2025 **Accepted:** December 02, 2025

ABSTRACT

This paper presents the design and simulation of a broadband blade antenna for aeronautical and astronautical communication systems, including the Traffic Collision Avoidance System (TCAS), Automatic Dependent Surveillance–Broadcast (ADS-B), and UHF satellite communications (SATCOM). Using CST Studio Suite, a compact blade-type antenna was modeled to operate from 400 MHz to 1.1 GHz, covering both SATCOM (approximately 400–500 MHz) and ADS-B/TCAS (1030–1090 MHz). Performance parameters, such as reflection coefficient, surface current distribution, far-field radiation pattern, gain, and efficiency, were analyzed at multiple frequencies. The results indicate broadband impedance matching, as evidenced by an S_{11} value below -10 dBi, along with consistent gain ranging from 7.363 dBi to 8.282 dBi throughout the operational frequency band. Compared to conventional aircraft platforms with multiple narrowband antennas, a single wideband blade antenna offers distinct benefits such as reduced aerodynamic drag, lower weight and cabling, simplified certification and maintenance, and improved system integration with reduced electromagnetic interference. By combining multiple services into one compact structure, the proposed design provides an integrated, cost-effective, and aerodynamically efficient solution for modern aircraft and UAV platforms, enhancing the reliability and communication performance across critical missions. We expect the proposed preliminary design blade antenna can contribute to future aircraft technologies.

Keywords: blade antenna, automatic dependent surveillance broadcast, traffic collision avoidance system, ultra-high frequency satellite communication, aeronautical communications, airborne systems

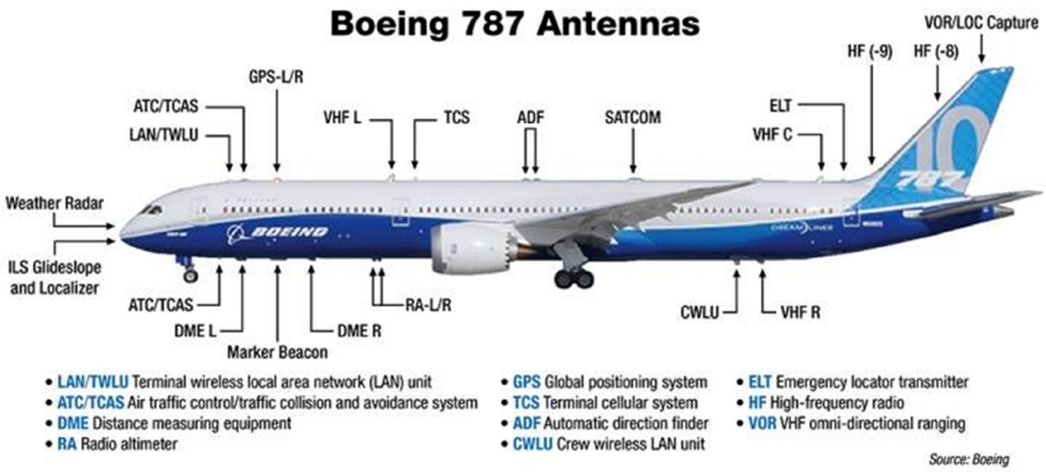
1. INTRODUCTION

The aviation and aerospace industries continue to demand high-performance antenna systems capable of supporting various communication and navigation functions within compact and aerodynamically efficient form factors. Among these, the Traffic Collision Avoidance System (TCAS), Automatic Dependent Surveillance–Broadcast (ADS-B), and UHF Satellite Communications (SATCOM) are essential for maintaining flight safety, aircraft coordination, and reliable long-range communications[1,2]. Figure 1 shows the distribution of multiple communication and navigation antennas on a Boeing 787 aircraft, where each antenna serves a specific function such as TCAS, ADS-B, and other systems, illustrating the complexity and aerodynamic impact of using multiple narrowband antennas and suggesting the advantages of one wideband antenna for reducing drag, weight, and system-integration challenges.

Conventional aircraft platforms often rely on multiple narrowband antennas to cover specific frequency ranges. While effective, this approach, increases system complexity, weight, and cost while reducing aerodynamic performance. By contrast, blade antennas, a subclass of monopole structures, balance mechanical robustness, aerodynamic integration, and reliable electrical performance across operational bands[3–5]. Recent research emphasizes their potential for integration into aircraft-fuselage conformal surfaces[6,7].

Various blade-antenna configurations have been proposed for specific communication applications. Pascawatia et al.[8] developed a blade antenna for 2.4 GHz rocket telemetry; Belen et al.[9] explored UHF-oriented blade antennas for vehicular and aeronautical systems; Sairam et al.[10] introduced broadband blade antennas for airborne platforms; and Saaquib et al.[11] incorporated air-core coils and horizontal notches to improve impedance bandwidth and matching. Silva et al.[12] embedded blade antennas in low-cost dielectric substrates to reduce costs while maintaining performance stability.

Figure 2 shows the communication architecture of a remotely piloted aircraft. The uplink commands and downlink status updates are transmitted through direct ground links or via a SATCOM relay, alongside. An ATC communication link facilitates flight coordination. This illustrates the objective of this study: to design a wideband blade antenna that operates across the range of UHF SATCOM and ADS-B/TCAS frequencies, integrating these critical links into a single-antenna structure. By consolidating multiple communication services into one



Source: Boeing

Figure 1. Communication, navigation, and surveillance antennas on the Boeing 787 aircraft

antenna, the design reduces reliance on separate narrowband antennas, improving aerodynamic efficiency and simplifying system integration for both manned/unmanned aircraft platforms.

However, designing a single compact-blade antenna that seamlessly covers 400 MHz to 1.1 GHz (crucial for UHF SATCOM and ADS-B/TCAS) remains challenging. Therefore, this study aims to design and analyze wideband blade antenna using CST Studio Suite, and emphasizes on impedance stability, gain consistency, radiation efficiency, and directional behavior across the targeted spectrum [13–15].

2. ANTENNA DESIGN

This section details the antenna design and choice of substrate material. The proposed blade-type antenna structure is mounted on a rectangular ground plane optimized to achieve wideband coverage suitable for aeronautical and astronautical communication applications. Figure 3 shows the geometries of the preliminary and proposed designs. The ground plane is 400 × 850 ×10 mm in size, providing a stable platform for the radiating element and acting as a reflector to enhance performance. The blade radiator has a total height of approximately 490 mm and thickness of 10 mm, and it is tilted 45° relative to the ground plane. This tilt improves radiation efficiency and reduces impedance mismatch, which is common challenge in wideband monopole configurations. An SMA feed pin excites the radiator, ensuring proper coupling with the ground plane to support broadband operation across the targeted frequency range of 0.4–1.1 GHz.

The blade antenna was modeled using a perfect electric conductor (PEC) material for the metallic structure, which is used in aerospace antenna design owing to its excellent electrical and mechanical characteristics. The PEC plate with a thickness of 10-mm ensures robust structural integrity under aerodynamic loading while minimizing ohmic losses. The initial shape and dimensions of the preliminary blade antenna were selected based on insights from previous studies on broadband monopole and blade antennas [8,12] , which revealed that radiator geometry and feed positioning are crucial for achieving multiband operation. Building on this foundation, the design was further refined and optimized to meet the research objectives. The selected

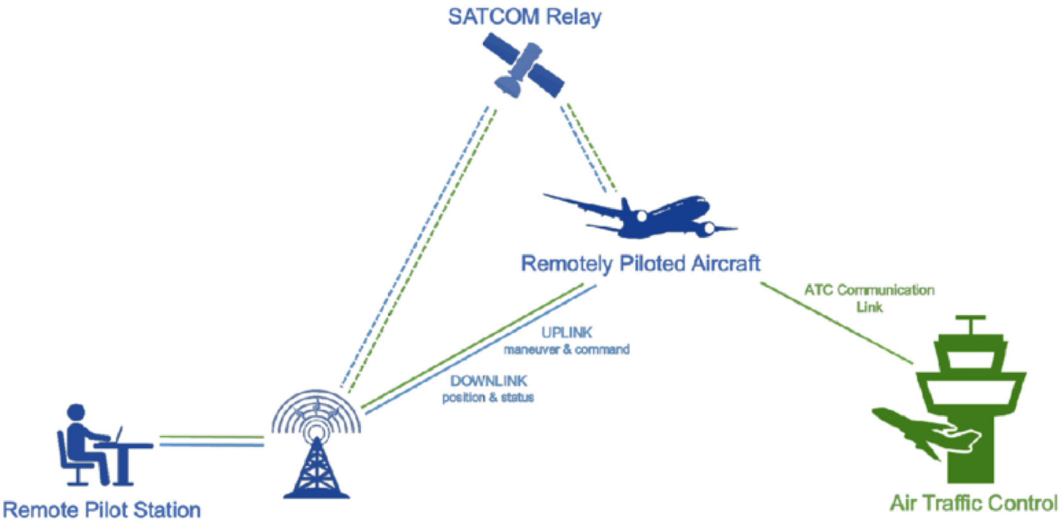
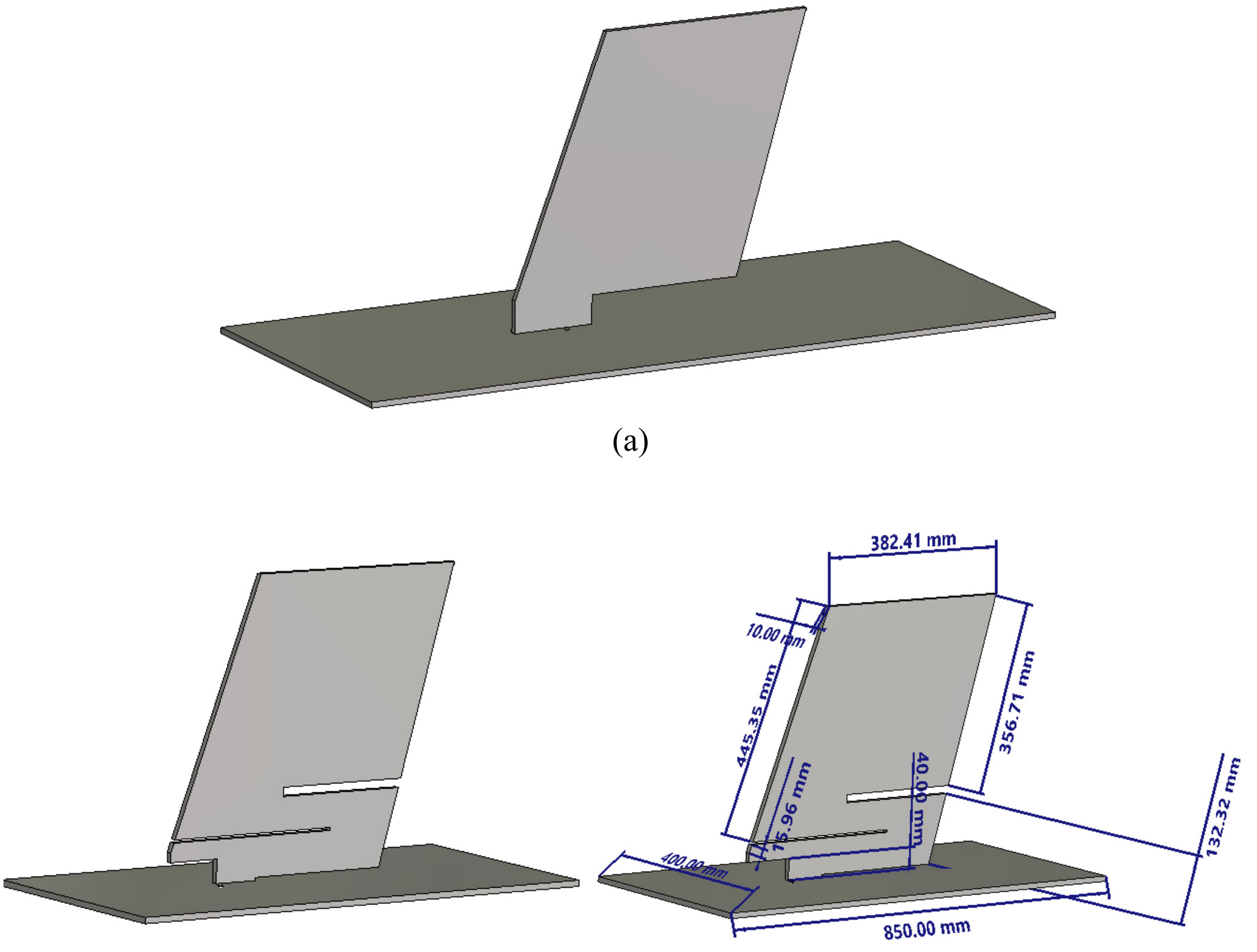


Figure 2. Communication links of a remotely piloted aircraft using a SATCOM relay, ground station, and air-traffic control



(a)

(b)

Figure 3. Configuration of blade antenna design: (a) preliminary design and (b) optimized blade antenna design: (left) clean version and (right) with dimensions

configuration enables the antenna to cover the UHF SATCOM (400–500 MHz) and ADS-B/TCAS (1030–1090 MHz) bands.

2.1. Optimization Method.

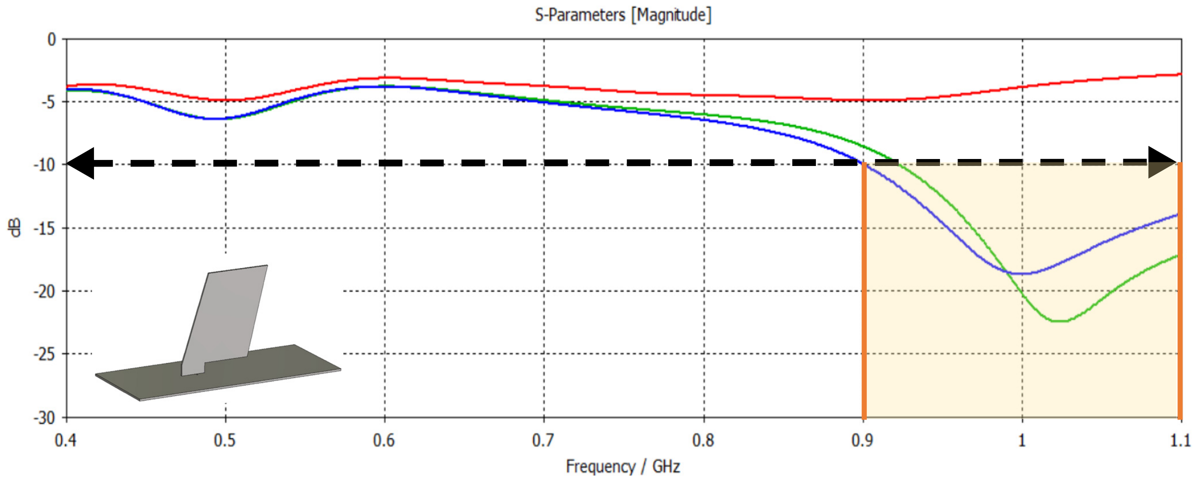
The preliminary antenna geometry, which was designed manually based on a detailed review of existing broadband blade antenna designs reported in the literature [8–12]. To further enhance the antenna's performance, a parametric sweep optimization performed using CST Studio Suite. The optimization process began with the basic geometry shown in Figure 3 (a). Successive refinements were then introduced based on the observed electromagnetic behavior. First, a horizontal slit was added near the upper part of the blade at the back to broaden the impedance bandwidth. Next, the slit position and length were adjusted toward the blade's upper back corner to improve matching across the target frequency bands. To enhance the multiband performance, a two-stage parametric sweep was executed in which the slit length was varied from 20 to 80 mm in 5 mm increments, while the slit width was adjusted between 2 and 6 mm. Performance was evaluated using the reflection coefficient (S_{11}) <–10 dB, voltage standing wave ratio (VSWR) <2, and gain stability at 0.43, 0.5, 1.03, and 1.09 GHz.

To better understand the effect of each modification on the affected performance, the surface current distributions were analyzed at every iteration. These current maps revealed regions of strong current concentration around the feed and ground-plane junctions. Based on these observations, an additional ground-plane cutout surrounding the SMA connector region was introduced to reduce local current crowding and improve the overall impedance characteristics.

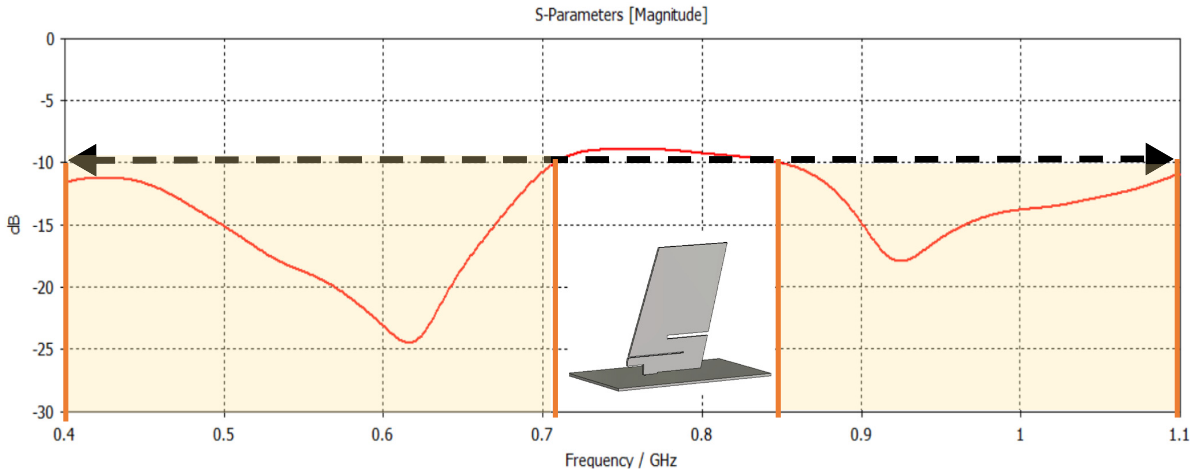
Finally, a second slit was added and optimized in the lower front section of the blade, to complete the design evolution. This final configuration Figure 3 (b) provides the most favorable combination of wideband impedance matching, stable radiation behavior, and peak gain performance. The optimization of the blade antenna geometry was through a structured parametric sweep analysis using CST Studio Suite. The optimized geometry provided the widest impedance bandwidth and maintained a peak gain exceeding 7.3 dBi across all required frequency bands.

2.2. Material, Feed Selection, and Simulation Setup Considerations.

The antenna is installed using an SMA/N-type connector and bonded to the fuselage ground plane,



(a)



(b)

Figure 4. Reflection coefficient (S_{11}) versus frequency of the proposed blade antenna: (a) preliminary blade antenna and (b) optimized blade antenna

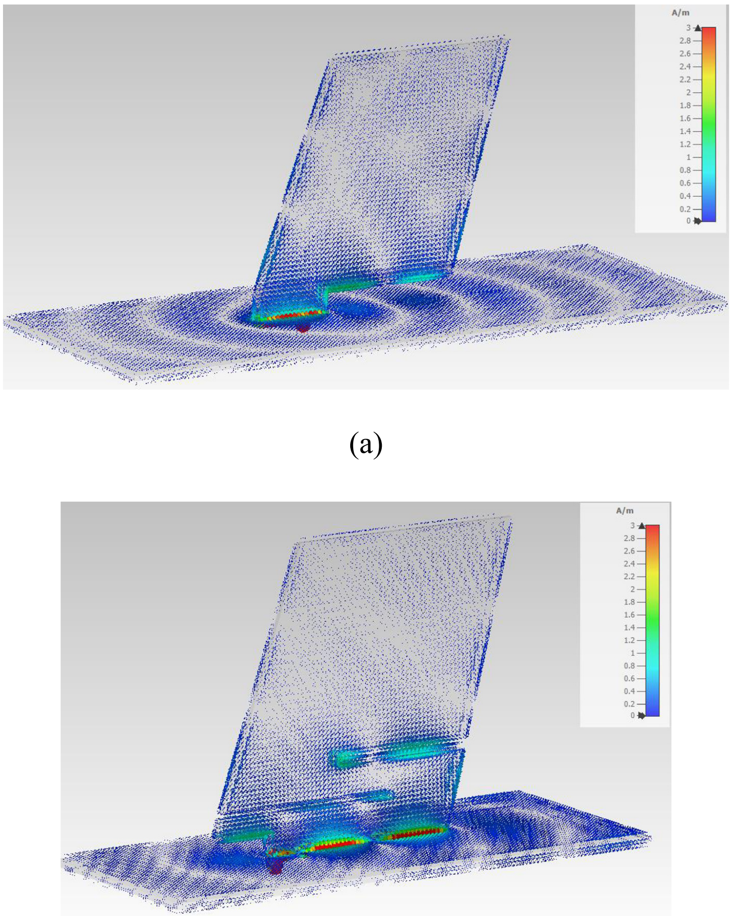
following the same concept used in[8–12]. In future simulations, a thin dielectric radome may be included to protect the antenna assembly from environmental factors such as rain, lightning, and aerodynamic forces that could impact structural integrity.

3. PERFORMANCE EVALUATION

Simulations were performed in CST Studio Suite using open boundary conditions to replicate the free-space performance. Key metrics such as S_{11}, surface current distribution, far-field radiation patterns, gain, and VSWR were evaluated. Representative frequencies at 0.43 GHz, 0.50 GHz, 1.03 GHz, and 1.09 GHz were selected for detailed analysis, corresponding to the operational ranges of SATCOM and ADS-B/TCAS. This systematic approach ensures that the antenna meets the broadband performance requirements while maintaining stable gain and radiation efficiency across multiple bands.

Figure 4 compares the reflection coefficient and frequency. Figure 4 (a) shows that the preliminary blade antenna exhibits a limited operational bandwidth, and its reflection coefficient values (S_{11}) are not consistently below the −10 dBi threshold across the desired frequency range. This indicates poor impedance matching and reduced efficiency at several frequencies. By contrast, Figure 4 (b) shows the results for the optimized blade antenna design, where the S_{11} curve demonstrates a much wider impedance bandwidth with values well below −10 dBi over a broad frequency range. This confirms that the optimization successfully improved the matching characteristics and enabled the antenna to support wideband operation. This improvement is primarily attributed to the introduction of slits on both the front and back of the blade radiator. These slits act as perturbations to the surface current paths, effectively generating additional resonances and enhancing multiband and broadband performance.

Figure 5 shows the current distributions for the preliminary and proposed blade antenna designs. Figure 5 (a) illustrates the



(a)

(b)

Figure 5. Surface current distribution at 1.03 GHz: (a) preliminary blade antenna and (b) optimized blade antenna

current distribution for the preliminary blade antenna design, where strong surface currents are concentrated near the feed region, with a relatively less effective distribution across the entire radiator. This limits the ability of the antenna to radiate efficiently at multiple frequency bands. In contrast, Figure 5 (b) demonstrates the optimized blade antenna, where the surface current is more evenly distributed across the blade structure, including regions around the introduced slits. The slits force the current to flow along multiple longer paths, creating additional resonant modes. This directly extends the bandwidth and improves the impedance matching.

Figure 6 shows the VSWR response of the optimized blade antenna over the frequency range of 0.4–1.1 GHz. The VSWR curve stays below the standard threshold of 2.0 for most of the operating band, which indicates good impedance matching between the antenna and the feed line. This ensures that minimal power is reflected toward the source, thereby improving radiation efficiency. At approximately 0.6 GHz and 0.9 GHz, the VSWR approaches its lowest points, demonstrating strong resonance and excellent matching at these frequencies. Although the curve slightly exceeds 2.0 in some regions, the majority of the band of interest remains within acceptable limits, confirming that the antenna is capable of supporting wideband operation.

The improved VSWR performance results from the optimized blade geometry, particularly the slit modifications, which introduce multiple current paths and resonant modes. This effectively broadens the bandwidth and enhances the suitability of the antenna for applications such as UHF SATCOM (400–500 MHz), TCAS/ADS-B (1030–1090 MHz), and UAV communication links.

Figure 7 illustrates the simulated 1D (E-plane cut) and 3D radiation patterns of the optimized blade antenna at the four frequency points. At the lower frequencies 0.43 and 0.50 GHz, the antenna demonstrates wide coverage with a more toroidal radiation pattern at the elevation plane, which is characteristic of monopole-like behavior, and is omnidirectional primarily in the azimuth plane. While at higher frequencies 1.03 and 1.09 GHz, the radiation becomes slightly more directive in some side directions with stronger main lobes. At these two frequencies, the peak gain is not located along the boresight ($\theta = 0°$)

Table 1. *Overall performance of the optimized blade antenna design*

Frequency point (GHz)	**0.43**	**0.50**	**1.03**	**1.09**
Peak gain (dBi)	7.753	7.363	8.282	8.197
VSWR	1.75	1.42	1.55	1.74
Reflection coefficient (S_{11})	−11.26	−15.14	−13.38	−11.41

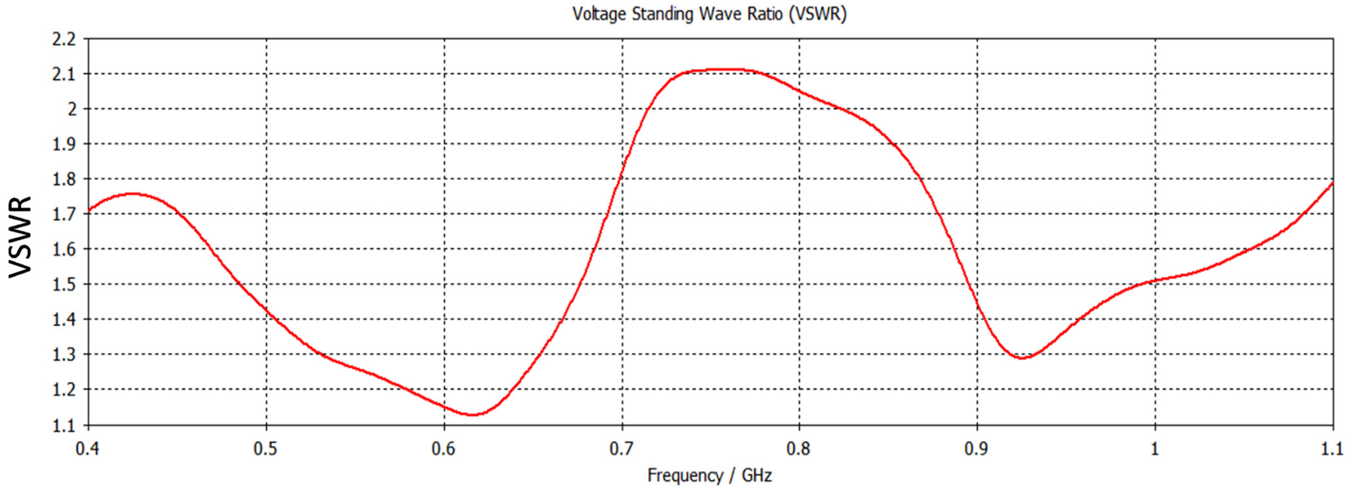


Figure 6. VSWR of the optimized blade antenna

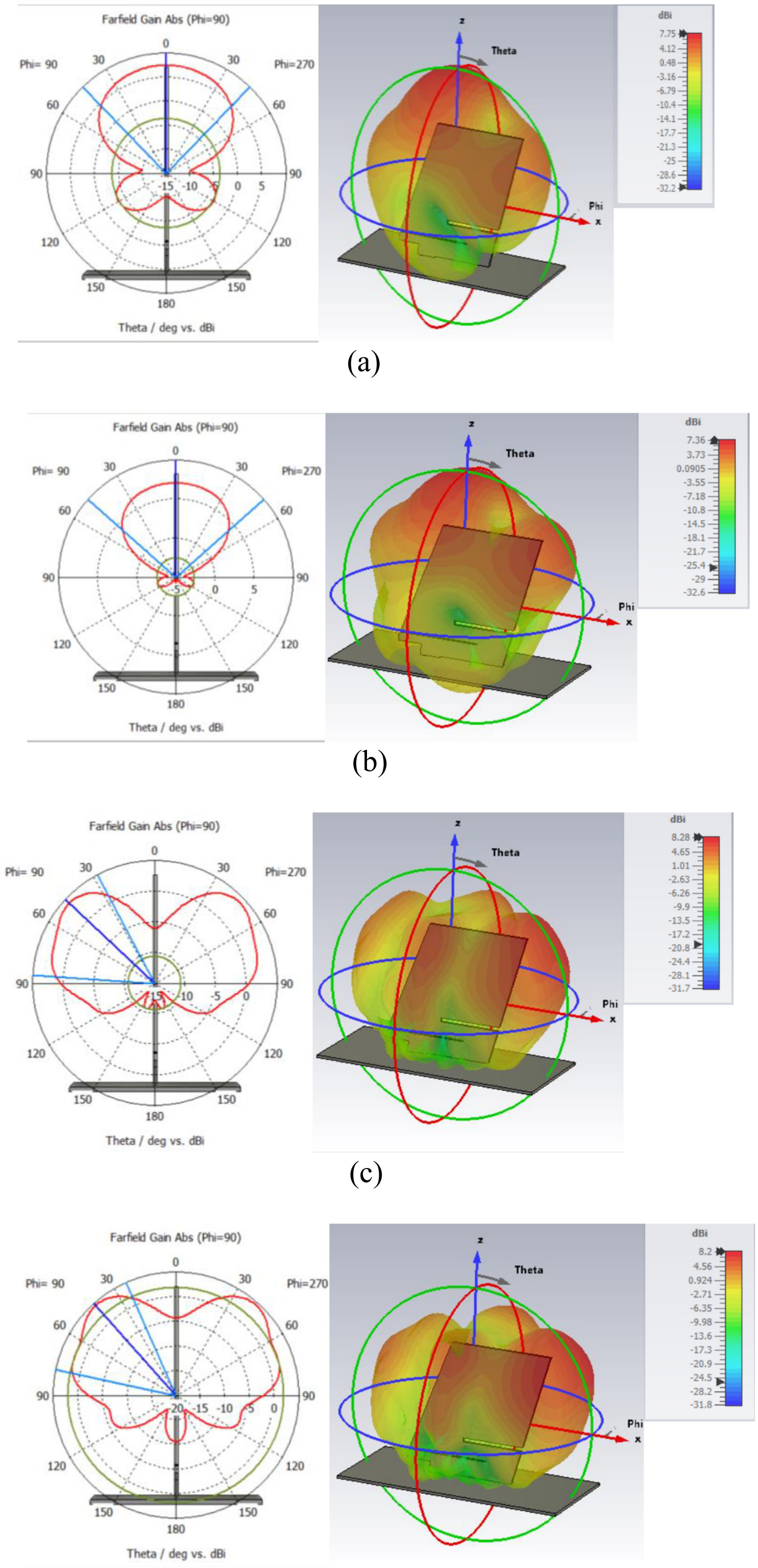
but is slightly shifted (∼25°–30° off-axis). This does not affect the ADS-B/TCAS performance, because these systems require toroidal or hemispherical coverage and do not focus on zenith gain. For the UHF SATCOM (400–500 MHz), the radiation remains closer to the boresight, ensuring reliable satellite link connectivity.

Table 1 summarizes the performance of the proposed blade antenna at the four frequency points: 0.43, 0.50, 1.03, and 1.09 GHz. The gain varies between 7.363 dBi and 8.282 dBi, with the highest values recorded at 1.03 GHz and 1.09 GHz. However, the radiation patterns in Figure 7 show that at these higher frequencies, the peak gain does not occur along the boresight, but instead shifts toward the side lobes, indicating that the antenna radiates more strongly in the off-axis directions. This is consistent with the slight asymmetry observed in the 3D radiation plots. The VSWR values remain below 2 across all frequencies, with the best impedance matching at 0.50 GHz (1.42), while the reflection coefficients are better than –10 dBi in all cases, reaching –15.14 dBi at 0.50 GHz.

Table 2 compares the proposed blade antenna with several representative designs reported in the literature. The existing designs by Belen et al.[9], Sairam et al.[10], and Saaquib et al.[11] generally prioritize either wideband behavior or specific aerospace applications, but they also exhibit notable limitations in gain, size, or functional coverage. For example, Sairam et al.[10] reported a compact fiber-reinforced polymer radome operating between 0.5–2.0 GHz, but the measured gain remains relatively low (3–5 dBi), making it more suitable for air-core EW and broadband systems than long-range UHF SATCOM. Similarly, the composite trapezoidal design by Belen et al.[9] covers 600–900 MHz with moderate gain (4.5–6.2 dBi) but lacks the operational bandwidth needed for simultaneous UHF SATCOM and ADS-B/TCAS support. The notched aluminum blade by Saaquib et al.[11] achieves good performance in the 835–962 MHz region; however, its gain (4.73–6.05 dBi) remains lower than that of the

Table 2. *Key performance metrics of comparison methods*

Reference	Frequency range	Gain (dBi)	Bandwidth	Physical size	Applications
Belen et al., 2018	600–3000 MHz	4.5–6.2 dBi (measured)	Ultra-wide (~150%)	Printed trapezoidal blade, multi-slot geometry	Very wideband UHF blade antenna
Sairam et al., 2018	0.5–2.0 GHz	3–6 dBi (measured)	Very wide (~150%)	~60 ×60 ×35 mm radiator (on 1.2 m ground plane)	Air-core coil, fiber-reinforced polymer radome, airborne electronic warfare applications
Saaquib et al., 2017	835–962 MHz	4.73 dBi (simulated) / 6.05 dBi when aircraft-mounted	~14% (928 MHz center)	0.4 mm aluminum sheet blade (SolidWorks + CST design)	Notched blade with parasitic element
This study	400–1100 MHz	7.36–8.28 dBi	>60%	400 ×850 mm ground plane; 490 mm height	Covers both UHF SATCOM (400–500 MHz) + ADS-B/TCAS (1030–1090 MHz)



(a)

(b)

(c)

(d)

Figure 7. Radiation pattern as a (left) 1D E-plane cut and (right) 3D view of the optimized blade antenna at (a) 0.43 GHz; (b) 0.50 GHz; (c) 1.03 GHz; and (d) 1.09 GHz

optimized design, and its effective bandwidth is still limited to a single communication service.

By contrast, the optimized blade antenna developed in this study demonstrates a substantially broader operational range (400–1100 MHz) while simultaneously providing higher peak gain values (7.36–8.28 dBi) across the operational band. This represents a significant improvement in both radiation efficiency and multiservice compatibility. Within a single compact structure (400 ×850 mm ground plane and 490 mm height), the antenna successfully covers two critical aviation communication bands: the UHF SATCOM band (400–500 MHz), which requires robust hemispherical coverage for satellite connectivity, and the ADS-B/TCAS band (1030–1090 MHz), which requires reliable broadcast and surveillance performance. The achieved bandwidth, which exceeds 60%, further highlights the capability of the antenna to support diverse airborne systems without requiring multiple antennas or a larger fuselage volume. Overall, the results in Table 2 show that the proposed design outperforms previous state-of-the-art antennas in terms of gain, bandwidth, and dual-band coverage, thereby offering a more efficient and integrated solution for modern aviation platforms.

4. CONCLUSIONS AND FUTURE WORK

This study demonstrated the feasibility of a compact wideband blade antenna designed to support essential aeronautical communication functions. Using CST Studio Suite 2024, the antenna was simulated to evaluate the gain, surface current, and efficiency performance over the full 400 MHz to 1.1 GHz range. The results showed that the peak gain ranges from 7.363 to 8.282 dBi across the operational band, indicating stable efficiency and strong current distribution, making this design a strong candidate for future aerospace communication systems.

As mentioned above, the simulation model assumes a PEC. However, the real-world implementation of the antenna relies on aerospace-grade materials whose electromagnetic behavior differs from that of an ideal PEC. Real metallic structures, such as aluminum alloys or copper-plated composites, exhibit finite conductivity and surface resistance, which introduce small but measurable ohmic losses. Ramos et al.[16] provide a detailed electromagnetic characterization of aerospace materials, showing that factors such as surface finish, conductivity, composite lay-up, resin type, and bonding interfaces influence the overall electromagnetic performance of real structures. Their results demonstrated that even highly conductive aerospace materials can experience variations in shielding effectiveness, surface currents, and EM depending on the material composition and assembly conditions. These findings highlight the importance of considering real material conductivity rather than relying solely on PEC assumptions.

Therefore, the PEC-based simulations reported in this work represent an upper bound on antenna performance. When practical aerospace-grade conductors are used, the realized gain is expected to decrease by only an approximate of 0.2–0.4 dBi, which remains well within typical modelling and fabrication tolerances. This small deviation confirms that the PEC-based simulation results are accurate and reliable for the design assessment presented in this study. By addressing the need for a unified antenna across multiple aviation bands, this study contributes a practical solution for reducing onboard component load and improving aerodynamic efficiency. The simplicity of the design, combined with its robust broadband characteristics, paves the way for real-world implementation on both civil and military aircraft platforms.

This antenna type is suitable for a wide range of airborne applications. For TCAS/ADS-B (1030–1090 MHz), its high gain and directional performance support long-range aircraft identification and avoidance systems. For UHF SATCOM (~400–500 MHz), the antenna offers sufficient bandwidth and toroidal radiation patterns, making it a viable choice for satellite uplinks and emergency beacons. Furthermore, for unmanned aerial vehicles and remotely piloted aircraft, the compact profile allows seamless integration into unmanned platforms without compromising stealth or aerodynamic efficiency. Further research can explore dual-port or tunable versions for dynamic frequency allocation, as well as integration with conformal fuselage structures for enhanced stealth in defense applications.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Taqi Abdulkhaleg Albaiyat – *Aerospace Engineering Department, King Fahd University of Petroleum & Minerals, Saudi Arabia;* 0009-0006-7859-8710
Email: s202038980@kfupm.edu.sa

Corresponding Author

Bazilah Binti Baharom – *Research Mentor, Interdisciplinary Research Centre for Aviation and Space Exploration, King Fahd University of Petroleum & Minerals, Saudi Arabia;* 0000-0003-3193-6197
Email: bazilah.baharom@kfupm.edu.sa

Author

Ayman Muhammad Abdullah – *Interdisciplinary Research Centre for Aviation and Space Exploration, King Fahd University of Petroleum & Minerals, Saudi Arabia;* 0000-0002-9365-7364
Email: aymanma@kfupm.edu.sa

ACKNOWLEDGEMENTS

The authors acknowledge the support provided by the Interdisciplinary Research Center for Aviation and Space Exploration at the King Fahd University of Petroleum and Minerals. The corresponding author also extends his appreciation to his advisor for guidance throughout the RES 200 Independent Research course, particularly for assistance in understanding antenna operations in aviation and aircraft applications, conducting the literature review, and supporting the antenna design process.

REFERENCES

(1) D. De and P. K. Sahu, “A novel approach towards the designing of an antenna for aircraft collision avoidance system,” *Int. J. Electron. Commun. (AEÜ)*, vol. 71, pp. 53–71. (2017).

(2) S. D. Ilcev, “Airborne satellite navigation and other integrated antenna systems,” in *Proc. Int. Conf. Electrical, Electronics, and Optimization Techniques (ICEEOT)*. (2016).

(3) A. V. SE, S. J and T. Kavitha, “Conformal Antenna for Aircraft Applications,” *2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, Bangalore, India, pp. 1–7. (2023).

(4) N. V. Laxminarayen, G. Lokeshwaran, K.S. Murugan and Dr. P. Jothilakshmi, “Design of conformal antenna for aircraft applications,” in *Proc. Int. Conf. Smart Structures and Systems (ICSSS)*. (2017).

(5) M. Pallavi, P. Kumar, T. Ali, S. B. Shenoy and L. Sharma, “Design and Analysis of Patch Antenna with T-Shape DGS for Aircraft Surveillance Applications,” *2022 6th International Conference on Green Technology and Sustainable Development (GTSD)*, Nha Trang City, Vietnam, pp. 1157–1163. (2022).

(6) Z. Ma and X. Huang, “Design of a circularly polarized microstrip antenna for aircraft tracking based on BeiDou III compatible with multi-navigation system,” *Micromachines*, vol. 14, no. 11, p. 2083. (2023).

(7) Ö. Dündar and M. F. Ateş, “Use of microstrip antennas in aerospace applications: L5 band satellite communication example,” *Aerospace Research Letters*, vol. 3, no. 1, pp. 69–78. (2024).

(8) A. Pascawati et al., “Blade Antenna 2.4 GHz for Rocket Applications," *2021 International Conference on Computer System, Information Technology, and Electrical Engineering (COSITE)*, Banda Aceh, Indonesia, pp. 204–208. (2021).

(9) A. Belen, H. P. Partal, S. Ördek and M. A. Belen, “Blade antenna design for UHF applications,” *2018 26th Signal Processing and Communications Applications Conference (SIU)*, Izmir, Turkey, pp. 1–4. (2018).

(10) C. Sairam, T. Khumanthem, S. Ahirwar and S. Singh, “Broadband blade antenna for airborne applications,” *2011 Annual IEEE India Conference*, Hyderabad, India, pp. 1–4. (2011).

(11) N. Saaquib, T. Sarker, N. Rahman, L. E. Khan and P. K. Saha, “Design and development of a wide band monopole blade antenna for aircraft navigation and communication,” *2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT)*, Himeji, Japan, pp. 1–5. (2017).

(12) Silva, W.F., Bianchi, I., & Santos, T.P, “Design of a blade antenna embedded in low-cost dielectric substrate,” *J. Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 16, pp. 180193. (2017).

(13) M. Pallavi et al., “A review on gain enhancement techniques for vertically polarized mid-air collision avoidance antenna for airborne applications,” *IEEE Access*, vol. 9, pp. 30761–30774. (2021).

(14) D. De and N. Chattoraj, “A review: Theoretical analysis of TCAS antenna,” in *Proc. IEEE Int. Conf. Green Computing, Communication and Electrical Engineering (ICGCCEE)*. (2014).

(15) H. Li, W. Kinsner, Y. Wang, B. Palma and A. Tay, “Airborne Radar Based Collision Detection and Avoidance System for Unmanned Aircraft Systems in a Varying Environment," *2021 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Cleveland, OH, USA, pp. 43–48. (2021).

(16) Ramos. D, Cidrás. J, Plaza. B, Moravec. C, de la Torre. A, Frövel. M.R.K, Poyatos. D, “ Novel Electromagnetic Characterization Methods for New Materials and Structures in Aerospace Platforms”, *Materials*, vol. *15*, 5128. (2022).

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article *Mechanical Engineering*

Techno-economic Assessment of a Coupled Solar-photovoltaic Water Heating System for a Residential Complex in Saudi Arabia

Ahmad Alzughaibi and Awad B.S. Alquaity*

Cite https://doi.org/10.64589/juri/215229

Submitted: June 17, 2025 **Revised:** October 16, 2025 **Accepted:** December 05, 2025

ABSTRACT

Solar energy offers a sustainable pathway to meet residential demands for hot water and electricity. However, comparative community-scale assessments of solar thermal and photovoltaic (PV) systems remain limited. This study presents a techno–economic comparison of solar water heaters and PV systems to meet the combined hot water and electricity needs of a 2000-home residential community in Dhahran, Saudi Arabia. Using Renewable Energy and Energy Efficiency Technology Screening (RETScreen) software, the annual hot water load for a five-occupant household was modeled and a detailed energy audit was conducted to estimate the annual electricity demand. Community-scale requirements were then derived by scaling the results for an individual home by 2000. Two solar water heaters (ThuSolar and SunEarth) and three PV modules (LONGi Solar, Jinko Solar, and REC Solar) were compared in fixed-tilt and one-axis tracking configurations across five metrics: annual energy output, capital expenditure, CO_2 reduction, payback period, and land use. SunEarth water heaters achieved the highest thermal yield of 6.19 million kWh, offsetting 3.44 million kg of CO_2, with a land footprint of 7400 m^2 of land and payback period of 14.6 years. Jinko Solar with one-axis tracking generated the highest electricity output of 63.37 million kWh, offsetting 35.22 million kg of CO_2, with a land footprint of 142568 m^2 and payback period of 12.3 years. These findings highlight that optimized solar thermal and PV technologies can play complementary roles in delivering cost-effective and sustainable energy solutions for large residential communities.

Keywords: solar photovoltaics, solar water heater, net zero homes, payback period, CO_2 reduction

1. INTRODUCTION

Global energy consumption has increased significantly in recent decades, accompanied by a marked increase in fossil fuel consumption. In 2023, the global energy requirements reached approximately 180,000 TWh, an almost six-fold increase from that in 1950 (approximately 30,000 TWh) in Figure 1[1], with fossil fuels remaining the predominant energy source. This heavy reliance on fossil fuels has detrimental consequences, including oil spills, water and air pollution, resource depletion, and, most critically, climate change. Thus, an urgent global impetus has emerged to transition toward a more sustainable energy paradigm. These challenges are directly linked to the United Nations' Sustainable Development Goals (SDGs), particularly SDG 7 (Affordable and Clean Energy), SDG 12 (Responsible Consumption and Production), and SDG 13 (Climate Action).

In alignment with these global imperatives, Saudi Arabia has boldly committed to achieving net-zero greenhouse gas emissions by 2060[2]. This strategic vision includes the deployment of carbon capture technologies in both existing and forthcoming gas turbines and the extensive use of combined cycle power plants, which will enhance energy efficiency to curtail fuel consumption, as well as large-scale investments in renewable energy infrastructure. One focal area of this national strategy is the decarbonization of the residential sector. Residential appliances, particularly air conditioners and water heaters, contribute significantly to household energy consumption as shown in Figure 2. To address this demand sustainably, Saudi Arabia is pursuing a multifaceted approach that incorporates solar photovoltaic (PV) systems, solar thermal water heating technologies, wind energy, and improved building insulation along with the optimization of existing thermal processes and energy cycles. Collectively, these measures will contribute to the Kingdom's long-term objective of an energy-resilient and low-carbon future.

A pivotal strategy for decarbonizing residential communities involves replacing conventional electric water heaters with solar thermal water heating systems. These systems are broadly categorized into two primary types: active and passive solar heating. Active solar heating systems, while more complex and costlier, offer superior efficiency owing to the incorporation of pumps and electronic controls that actively circulate the heat-transfer fluids. In contrast, passive systems rely on natural convection and gravity-driven flow, which make them simpler and less expensive as well as less efficient[4]. The residential market offers three primary configurations of solar water heaters: flat-plate collectors, integrated collector storage (ICS) systems, and evacuated-tube solar collectors. These configurations vary considerably in terms of efficiency, cost, and suitability depending on the climate and

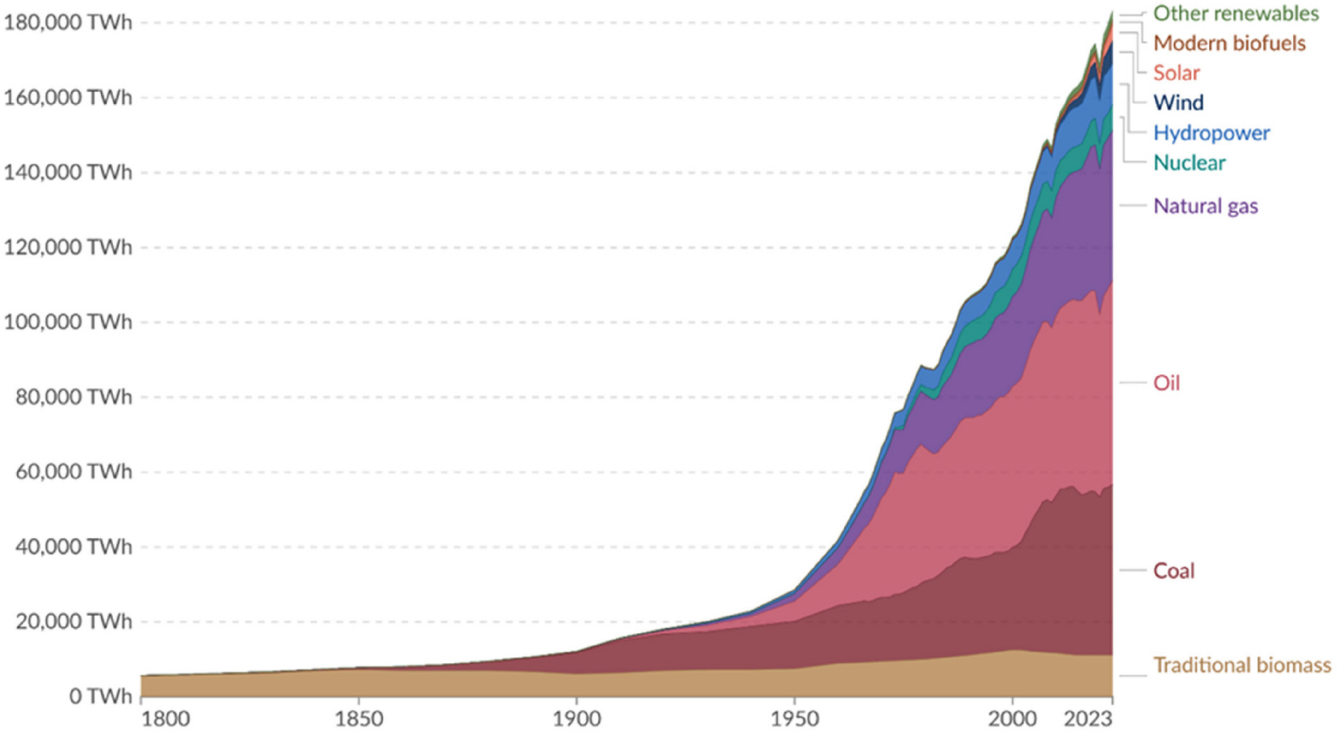


Figure 1. Global primary energy consumption by source

operational conditions[5]. Flat plate collectors are among the most commonly available solar water heaters and were modeled in this study.

Similarly, solar PV technology has emerged as a cornerstone of global energy transition and continues to experience exponential growth driven by decreasing costs and improved efficiency. Solar PV systems are composed of interconnected semiconductor cells, each capable of generating approximately 1–2 watts of electrical power. Solar PV technologies include monocrystalline silicon, polycrystalline silicon, and thin-film modules, each of which differs in terms of efficiency, cost, and application. Monocrystalline silicon PV modules are among the most efficient and thus were modeled in this study[6].

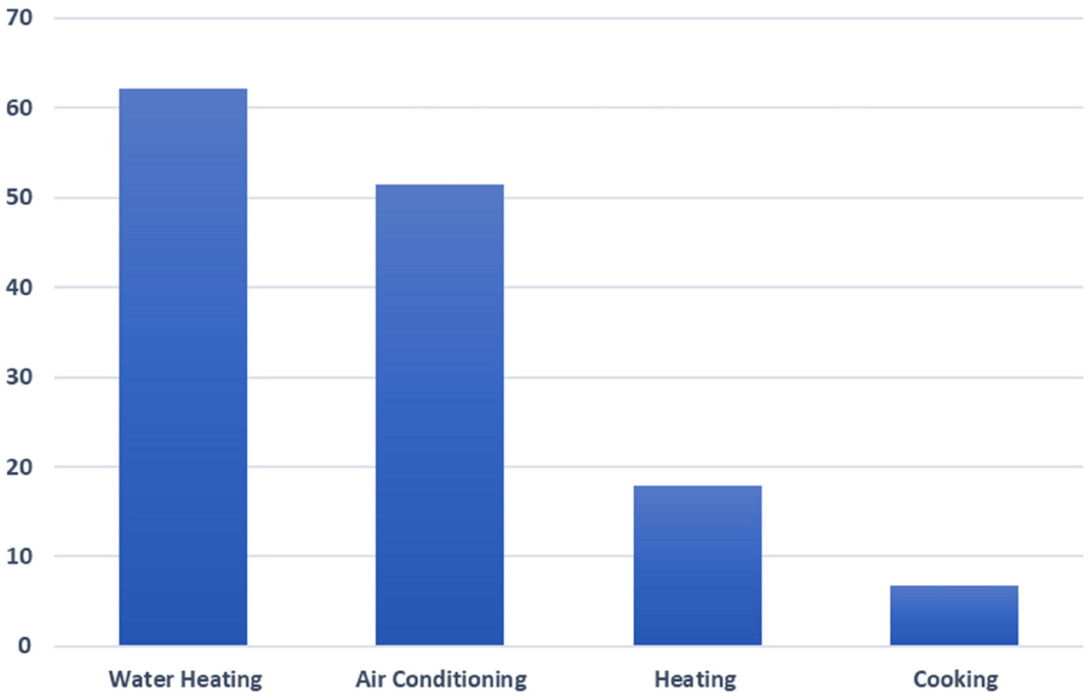


Figure 2. Average weekly operating hours of electrical appliances in the kingdom in 2023[3]

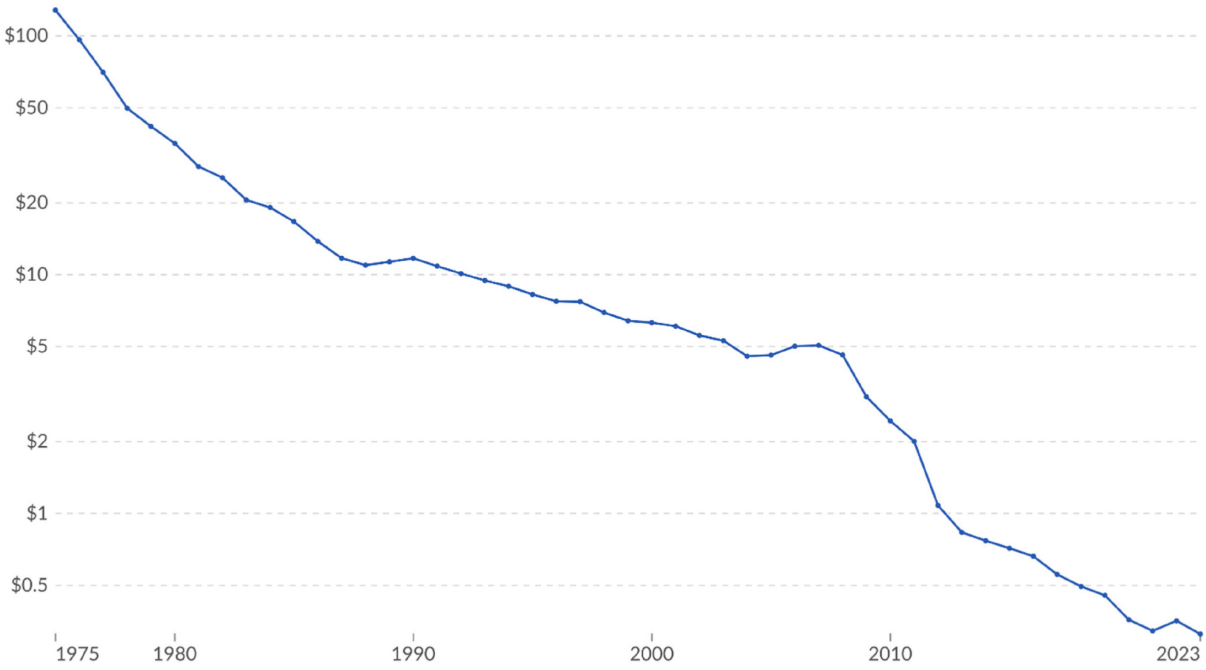
The cost of solar PV technology has drastically declined in recent decades, while its efficiency and scalability have continuously advanced. As illustrated in Figure 3, the average price of solar PV panels in 2023 dropped to approximately \$0.31/watt, which is only 4.8% of the cost in 2000—which stood at \$6.41/watt. This substantial cost reduction underscores the growing economic viability of solar PV systems and strengthens the case for accelerating the transition toward solar energy. However, declining cost alone is not sufficient for nationwide deployment. Comprehensive feasibility studies guided by rigorous data and statistical analyses are essential for governments to assess the reliability and long-term cost-effectiveness of solar PV infrastructure.

A considerable body of research has addressed the technical and economic performance of solar PV and solar thermal technologies. Previous studies have provided critical insights into the optimal deployment of solar technologies in Saudi Arabia. Abd-ur-Rehman and Al-Sulaiman[8], conducted a comprehensive techno-economic evaluation of evacuated tube and glazed flat-plate solar water heater systems across 10 Saudi cities and found significant regional variations in solar viability and economic feasibility. Almohammadia and Allouhi[9] proposed optimal grid-connected PV installations for residential buildings in 19 Saudi cities and highlighted the key trade-offs between renewable fractions and net present costs. Their findings emphasized the need to balance economic viability and renewable energy utilization to achieve sustainable energy transitions. Alanazi[10] explored hybrid renewable energy systems consisting of wind turbines, PV panels, batteries, and diesel generators in rural Saudi Arabia and concluded that despite considerable solar and wind potential, stand-alone renewable systems require supplementary storage or backup generators to ensure a reliable electricity supply. Recent advances in hybrid systems include developing solar PV-powered thermoelectric solar air conditioning and modeling their performance using multiple artificial intelligence models[11]. Other innovative studies have coupled solar PV/thermal collector systems with electrolytic hydrogen production units, and they reported hydrogen production yields as high as 4.41 kg/day[12,13]. Furthermore, the use of solar dish-Stirling systems has also been explored in detail, with reported efficiencies approaching 32% in arid climates; thus, these systems may play a key role in distributed energy systems[14].

While these studies have highlighted the performance potential of solar thermal, PV, and hybrid solar approaches, most of them focused on component-level demonstrations. Thus, community-scale comparative techno-economic assessments of solar thermal and PV systems for meeting heating and electrical loads in regions with high solar insolation such as Saudi Arabia are still limited in scope.

Therefore, this study provides a detailed techno-economic assessment of solar water heaters and PV systems specifically tailored for a 2000-home residential community in Dhahran, Saudi Arabia. The objectives of this study were as follows:

- Comprehensively compare the techno-economic performances of two commercial solar water heaters (ThuSolar and SunEarth) and three widely used PV modules (LONGi Solar, Jinko Solar, and REC Solar).
- Evaluate the influence of fixed-tilt and one-axis tracking configurations across key metrics, including annual energy output, capital investment, CO_2 emission reductions, payback time, and land requirements.
- Identify optimal solar energy solutions capable of sustainably and economically meeting both hot water and electrical energy demands, thus supporting Saudi Arabia's long-term



Data source: IRENA (2025); Nemet (2009); Farmer and Lafond (2016) OurWorldinData.org/energy | CC BY

Note: Data is expressed in constant 2024 US\$ per watt.

Figure 3. Solar PV panel prices in USD per watt over the years[7]

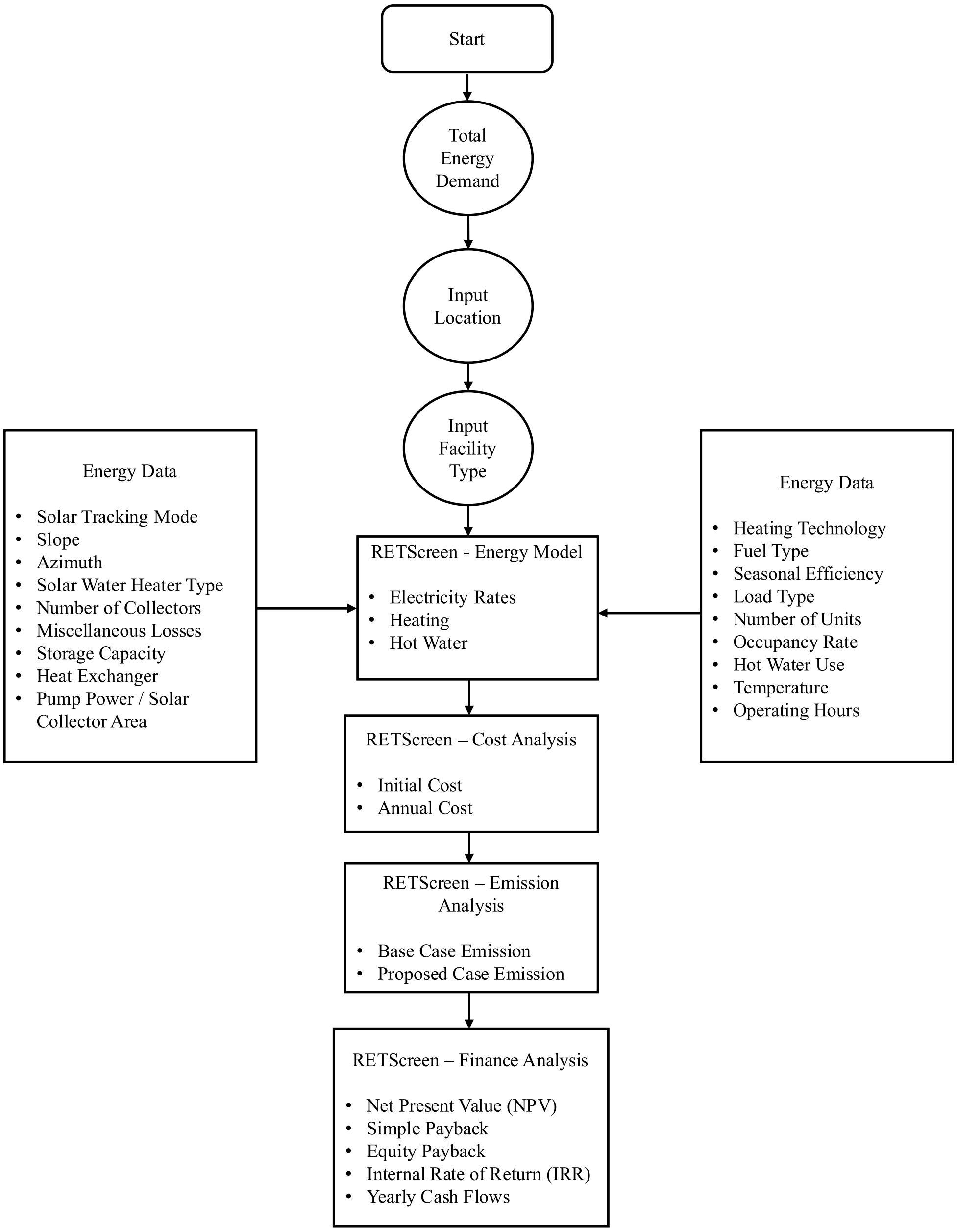


Figure 4. Solar water heater flowchart

decarbonization goals and contributing toward achieving the relevant SDGs (7 and 13).

The findings highlight how optimized solar thermal and PV technologies can play complementary roles in delivering cost-effective and sustainable energy solutions for large residential communities

2. RESEARCH METHODOLOGY

The research methodology used in this study consists of two primary stages. In the first stage, detailed energy demand data were collected from a representative household sample to quantify the hot water and electricity requirements that could be supplied by solar water heater and PV systems. In the second stage, Renewable Energy and Energy Efficiency Technology Screening (RETScreen) software was utilized to identify suitable technologies and conduct comprehensive cost, emission, and financial analyses for the selected location in Dhahran, Saudi Arabia. The overall methodological framework is illustrated in detail in Figures 4 and 5.

Figure 6 shows a simplified 2D schematic of the system. The diagram illustrates the major components, beginning with the solar PV farm located at distance from the community. The PV farm is connected to a direct current (DC) combiner box and inverter, which conditions and converts the generated electricity into alternating current (AC) power. Transmission lines carry electricity to the community, where service disconnects are installed to ensure safety during maintenance and emergency operations.

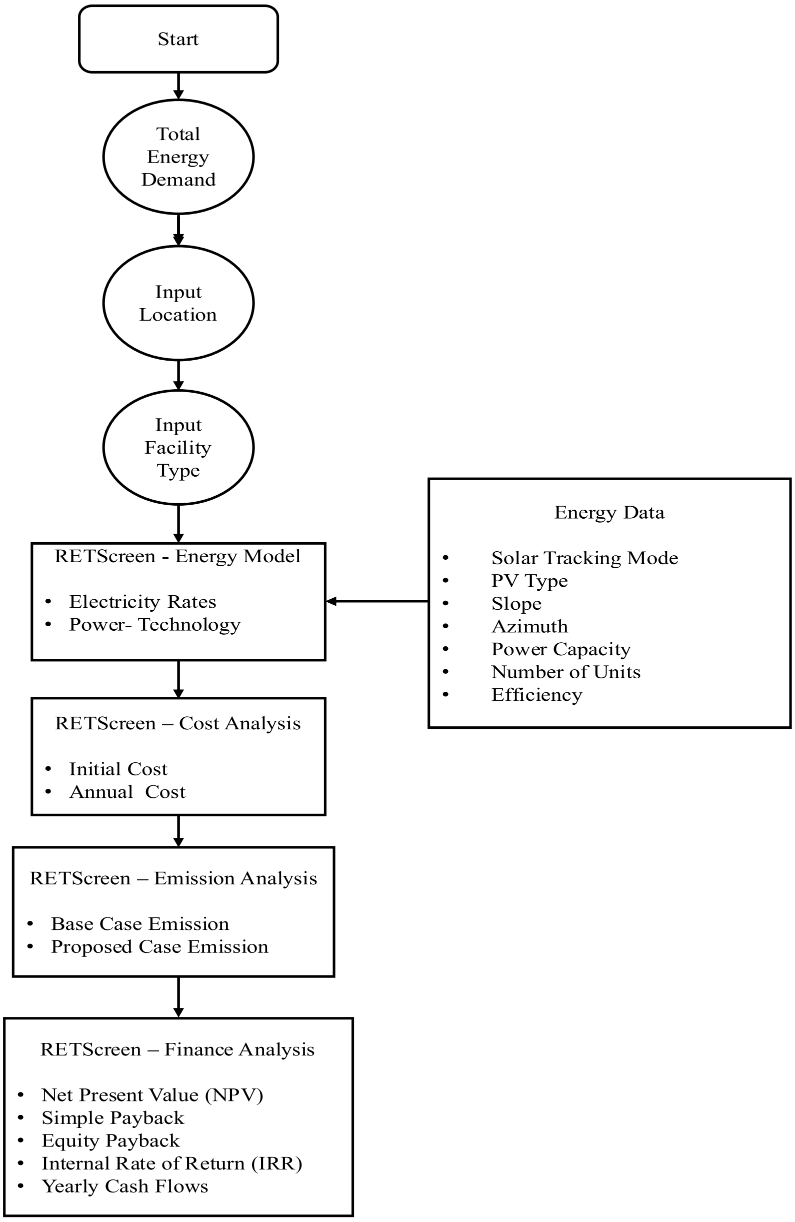


Figure 5. Solar PV flowchart

The schematic highlights the solar water heating system within the community. This includes a solar water heater collector, circulation pump, water storage tank, valves, and piping, all of which are integrated to provide hot water to residents.

2.1. Total Energy Demand.

A representative household of five individuals was selected as the basis for estimating the annual energy consumption. The most frequently used appliances were identified, along with their power ratings (in W) and average daily operating hours in both winter and summer. To simplify the analysis, a single representative day from each season was used to represent the overall seasonal demand. The daily energy consumption for each summer and winter day was then calculated using Eq. (1):

$$\text{Consumption per day (kWh)} = \text{Usage Hours} \times \text{Quantity} \times \text{Rating} \div 1000 \tag{1}$$

Annual energy demand was determined by averaging the calculated daily consumption for summer and winter days and multiplying it by 365.

Table 1 presents detailed energy consumption data calculated for a single representative household. These data served as the basis for determining the required heating and electricity generation from the solar water heating and PV systems. Based on the above data, the total annual energy consumption of the representative households was 45,008 kWh. Assuming that all households consume the same amount of energy, the total annual energy consumption for 2,000 households was calculated as 9, 00, 16, 000 kWh.

As solar water heaters are assigned to meet hot water demand, the energy consumption of water heaters, estimated to be 20,805 kWh per household, was subtracted from the total household load to determine the net energy demand to be met by the solar PV system. The PV system was then sized to generate electricity

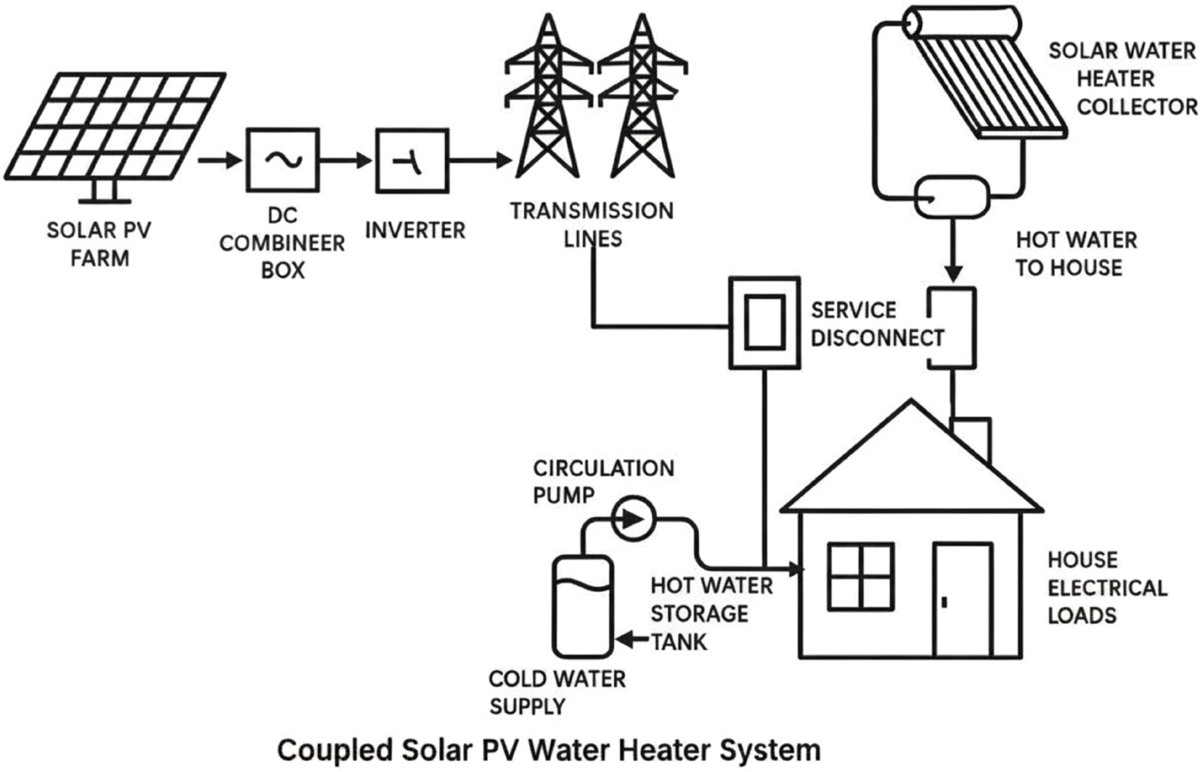


Figure 6. 2D schematic of coupled solar PV water heating system

for this reduced demand of 48, 406, 000 kWh, as calculated using Eq. (2).

$$\begin{aligned}&\textit{Reduced Energy Demand}\\&= \textit{Annual Energy Consumption for } 2,000 \textit{ Houses}\\&- \textit{Annual Consumption by Solar Heater} \times 2,000 \qquad (2)\end{aligned}$$

2.2. RETScreen Software.

In this study, the RETScreen software was employed to model, analyze, and design both solar water heating systems and solar PV systems. RETScreen is a comprehensive software platform designed to evaluate the technical performance, economic feasibility, and greenhouse gas (GHG) emissions of a wide range of clean energy projects. It has been extensively validated as a reliable tool for predicting the energy yields of solar PV and solar water heating systems. González-Peña et al.[15], evaluated RETScreen, SAM, PVGIS, and other PV software using 12 years of operational data from three PV plants in Castile and León, Spain, and reported that RETScreen's predictions for fixed-, single-axis-, and dual-axis systems showed good agreement with the measured data, with discrepancies remaining below 10%. Similarly, the solar water heater model was validated against both theoretical predictions and experimental data from ten solar water heater systems in Ontario, Canada. For medium to heavy loads approaching 380 L/day, RETScreen overestimated the collector performance by less than 15%, which remained within the acceptable range for feasibility studies[16]. In this study, RETScreen was applied to design and assess the performance, environmental benefits, and financial payback periods of solar water heating and PV systems under the climatic and economic conditions of Dhahran, Saudi Arabia. This approach ensures that the proposed solutions are technically viable and environmentally sustainable, and it offers practical insights into their long-term economic returns.

2.3. Solar Heater System.

2.3.1. Solar water heater theory.

In this study, a flat-plate collector in Figure 7 was employed to satisfy the water heating requirements. In this type of system, potable water flows through a network of pipes embedded within a dark-colored absorber plate enclosed in insulated and glazed housing. The absorber plate captures both direct and diffuse solar radiation and converts it into thermal energy, which is transferred to the water circulating through the embedded tubes. The heated water is then stored in a dedicated storage tank for domestic use. Flat-plate collectors are designed to operate efficiently across a wide range of solar conditions; however, their thermal performance is influenced by ambient factors, such as air temperature, wind speed, and cloud cover. This type of solar water heater was selected for this study owing to its wide commercial availability, proven reliability, and compatibility with standard residential water heating systems.

2.3.2. Solar water heater components.

A typical solar water heating system consists of several integral components that work together to ensure efficient thermal capture and utilization. The solar collector absorbs solar energy and transfers it to the water, thereby increasing the water temperature. The heated water is stored in a storage tank until it is required for domestic use. A pump circulates the water through the system, maintaining a consistent flow to maximize heat absorption. Controls and sensors are used to activate, deactivate, and monitor the system based on the water temperature and demand. Usually, a heat exchanger is included to transfer heat to the domestic water via a secondary fluid, thereby preventing freezing in colder climates.

Table 1. *Electricity consumption baseline data*

Appliances	Summer Usage (h)	Winter Usage (hours)	Quantity	Rating (W)	Summer Consumption/ day (kWh)	Winter Consumption/ day (kWh)	Annual Energy Consumption (kWh)
Refrigerator	6	6	1	425	2.55	2.55	930.8
Dishwasher	2	2	1	NA	-	-	265.0
Air conditioner	12	2	5	NA	0.00	0.00	11705.0
Electric water pump	2	2	1	1100	2.20	2.20	803.0
Electric heater	0	4	2	2000	0.00	16.00	2920.0
TV	4	3	1	210	0.84	0.63	268.3
Laptop	6	6	2	200	2.40	2.40	876.0
Clothes washer	0.571	0.571	1	NA	0.00	0.00	442.0
Dryer	0.2857	0.2857	1	NA	0.00	0.00	560.0
LED lights	6	8	133	6	4.79	6.38	2038.9
Microwave	1	1	1	900	0.90	0.90	328.5
Electric kettle	1	1	1	2200	2.20	2.20	803.0
Mist fan	12	0	1	100	1.20	0.00	219.0
Oven	2	2	1	2800	5.60	5.60	2044.0
Water heater	1	18	5	1200	6.0	108	20805
						Total	**45,008**

However, the system used in this study did not use antifreeze fluids or heat exchangers, making it well-suited for regions such as Dhahran, where freezing temperatures are rare. Pipes and fittings facilitate water flow and complete the closed-loop system, thereby enabling continuous operation[18].

2.4. Solar Water Heater Methodology.

2.4.1. Location. The city selected for this project was Dhahran, which is located in the Eastern Province of Saudi Arabia, at 26.3° N a latitude and 50.2° E longitude. Monthly weather data for Dhahran are shown in Figure 8.

The facility type selected in RETScreen software is "individual measure," and the type is "solar water heater." All subsequent inputs provided by the software are presented in detail below.

2.4.2. Energy model. The next step involved entering the data for the conventional electric water heater, which served as the baseline system. As this system operates solely on electricity, no fuel cost is involved, and only electricity expenses are considered. In Saudi Arabia, the electricity tariff is 0.18 SAR/kWh, which is 0.048 $/kWh after converting to US dollars using a conversion rate of 3.75. Therefore, the base technology for water heating was an electric heater using electricity as the fuel type, with a tariff of 0.048 $/kWh, and a seasonal efficiency of 100%.

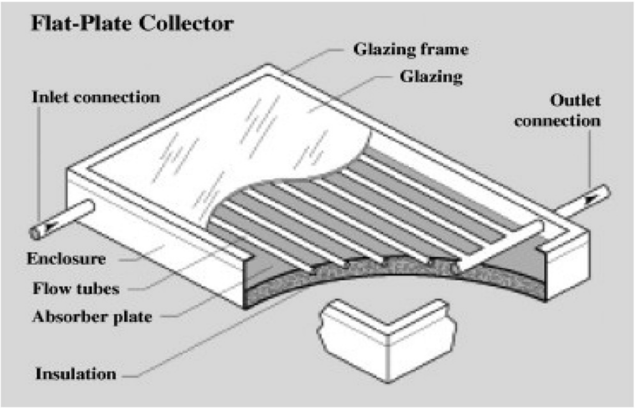


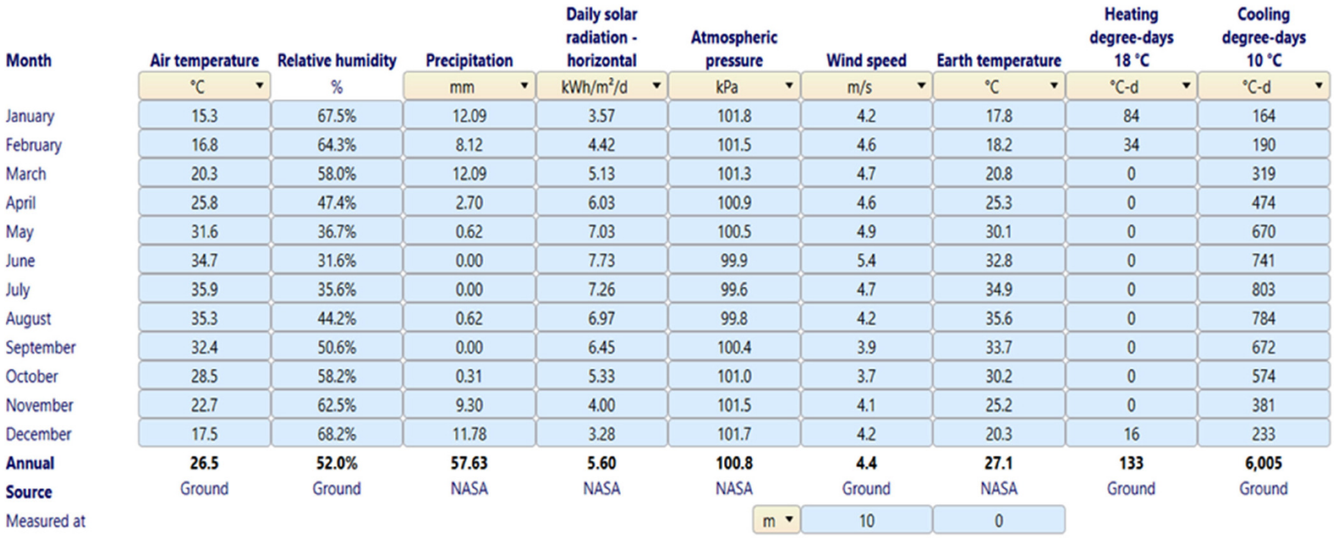
Figure 7. Flat-plate collector[17]

Next, the hot water demand was defined to represent a typical household with five occupants, an occupancy rate of 90%, and a water temperature of 70 °C. These inputs were used to estimate the thermal load that the solar water heater would need to meet.

For the solar water heater configuration, a fixed solar tracking system was selected, with a collector slope of 22° (angle between the tracking axis and the horizontal axis) and an azimuth of 0° (horizontal angle at which the collector is facing, measured clockwise from true north). Glazed flat-plate solar collectors manufactured by ThuSolar and SunEarth were selected and the specifications are shown in Figures 9–11. Miscellaneous losses of 2% were assumed, and the storage capacity was designed to meet a daily demand of 270 L. No heat exchanger or pump power was included because these components are not required under Dhahran's climatic conditions. The thermal efficiency of the solar water heating system was calculated using Eq. (3).

$$\eta_c = F_R \times K_{\tau\alpha} \times \tau \times \alpha - F_R \times U \times \frac{T_{w,in} - T_{air}}{G} \tag{3}$$

where η_c is the collector efficiency, F_R is the heat removal factor, $K_{\tau\alpha}$ is the incident angle modifier, τ is the transmittance of the cover, α is the absorptance of the absorber plate, U is the overall heat transfer coefficient, $T_{w,in}$ is the inlet temperature of the water, T_{air} is the ambient air temperature, and G is the solar irradiance.

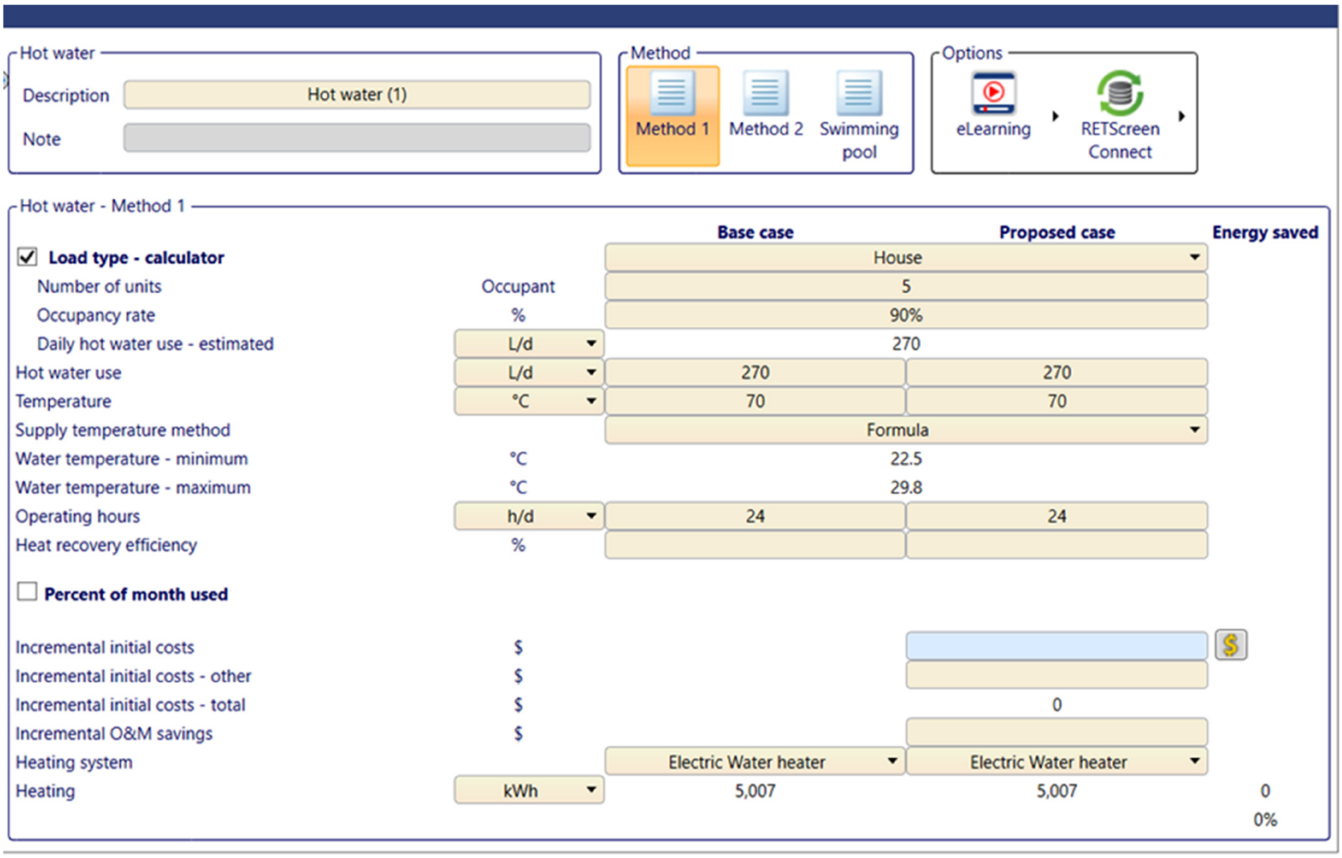


Month	Air temperature	Relative humidity	Precipitation	Daily solar radiation - horizontal	Atmospheric pressure	Wind speed	Earth temperature	Heating degree-days 18 °C	Cooling degree-days 10 °C
	°C	%	mm	kWh/m²/d	kPa	m/s	°C	°C-d	°C-d
January	15.3	67.5%	12.09	3.57	101.8	4.2	17.8	84	164
February	16.8	64.3%	8.12	4.42	101.5	4.6	18.2	34	190
March	20.3	58.0%	12.09	5.13	101.3	4.7	20.8	0	319
April	25.8	47.4%	2.70	6.03	100.9	4.6	25.3	0	474
May	31.6	36.7%	0.62	7.03	100.5	4.9	30.1	0	670
June	34.7	31.6%	0.00	7.73	99.9	5.4	32.8	0	741
July	35.9	35.6%	0.00	7.26	99.6	4.7	34.9	0	803
August	35.3	44.2%	0.62	6.97	99.8	4.2	35.6	0	784
September	32.4	50.6%	0.00	6.45	100.4	3.9	33.7	0	672
October	28.5	58.2%	0.31	5.33	101.0	3.7	30.2	0	574
November	22.7	62.5%	9.30	4.00	101.5	4.1	25.2	0	381
December	17.5	68.2%	11.78	3.28	101.7	4.2	20.3	16	233
Annual	**26.5**	**52.0%**	**57.63**	**5.60**	**100.8**	**4.4**	**27.1**	**133**	**6,005**
Source	Ground	Ground	NASA	NASA	NASA	Ground	NASA	Ground	Ground
Measured at					m	10	0		

Figure 8. Dhahran monthly weather report

2.5. Solar Photovoltaic. The analysis of the solar PV system followed a workflow similar to that of the solar water heating system; however, the facility type was defined as a power plant with a photovoltaic configuration. The photovoltaic specifications were selected to maximize efficiency, and PV modules from three different manufacturers–LONGi Solar, REC Solar, and Jinko Solar–were used for comparison, as presented in Section 3.2.

2.5.1. Solar photovoltaic theory. This study employed monocrystalline silicon PV modules, which were chosen for their high efficiency and suitability in applications where space is limited. The cells absorb photons from sunlight with energy equal to or greater than the silicon bandgap, whereas photons with lower energies are either transmitted or reflected. The monocrystalline structure was formed from a single, highly purified silicon crystal doped to produce an n-type layer (phosphorus-doped)



Hot water

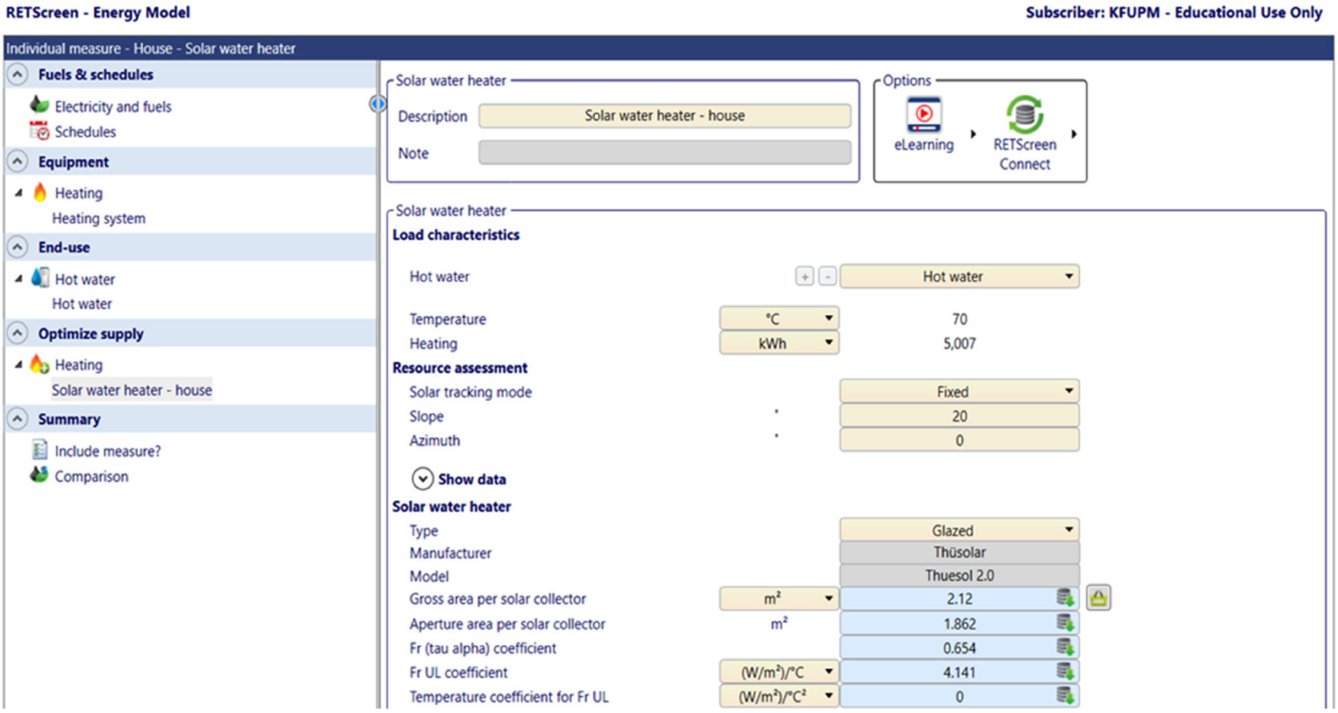
Description: Hot water (1)

Note:

Hot water - Method 1

		Base case	Proposed case	Energy saved
☑ **Load type - calculator**		House		
Number of units	Occupant	5		
Occupancy rate	%	90%		
Daily hot water use - estimated	L/d	270		
Hot water use	L/d	270	270	
Temperature	°C	70	70	
Supply temperature method		Formula		
Water temperature - minimum	°C	22.5		
Water temperature - maximum	°C	29.8		
Operating hours	h/d	24	24	
Heat recovery efficiency	%			
☐ **Percent of month used**				
Incremental initial costs	$			
Incremental initial costs - other	$			
Incremental initial costs - total	$		0	
Incremental O&M savings	$			
Heating system		Electric Water heater	Electric Water heater	
Heating	kWh	5,007	5,007	0
				0%

Figure 9. Hot water data



Solar water heater

Description	Solar water heater - house
Note	

Solar water heater

Load characteristics

Hot water		Hot water
Temperature	°C	70
Heating	kWh	5,007

Resource assessment

Solar tracking mode		Fixed
Slope	°	20
Azimuth	°	0

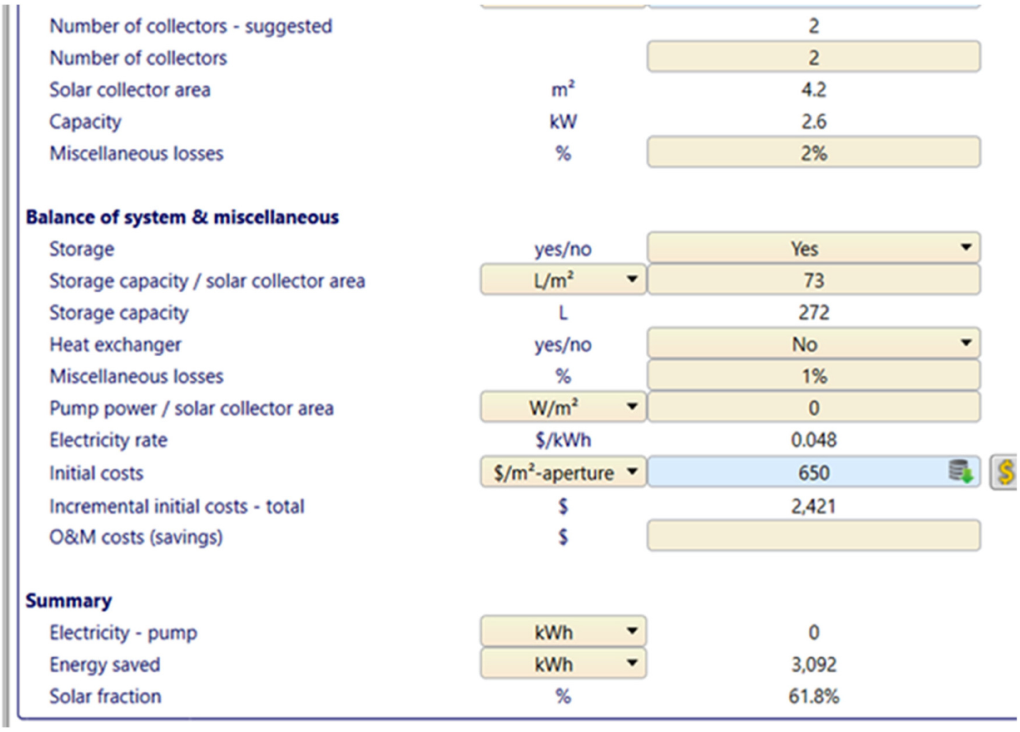
Solar water heater

Type		Glazed
Manufacturer		Thūsolar
Model		Thuesol 2.0
Gross area per solar collector	m²	2.12
Aperture area per solar collector	m²	1.862
Fr (tau alpha) coefficient		0.654
Fr UL coefficient	(W/m²)/°C	4.141
Temperature coefficient for Fr UL	(W/m²)/°C²	0

Figure 10. ThuSolar water heater data part 1

and a p-type layer (boron-doped), thereby creating a p–n junction. When photons strike the cell, the electrons in the n-type layer are excited and migrate toward the p-type layer, establishing an electric field across the junction. This process generates DC power, which is then converted into AC power through an inverter for household use or grid integration. With typical efficiencies ranging from 20% to 22%, monocrystalline PV modules deliver excellent performance; however, they also have higher manufacturing costs[19].

*2.5.2. **Solar photovoltaic components.*** A standard solar PV system consists of several key components that work together to enable efficient solar power generation and utilization. At the core of the system are solar PV cells, which capture solar energy



Number of collectors - suggested		2
Number of collectors		2
Solar collector area	m²	4.2
Capacity	kW	2.6
Miscellaneous losses	%	2%

Balance of system & miscellaneous

Storage	yes/no	Yes
Storage capacity / solar collector area	L/m²	73
Storage capacity	L	272
Heat exchanger	yes/no	No
Miscellaneous losses	%	1%
Pump power / solar collector area	W/m²	0
Electricity rate	$/kWh	0.048
Initial costs	$/m²-aperture	650
Incremental initial costs - total	$	2,421
O&M costs (savings)	$	

Summary

Electricity - pump	kWh	0
Energy saved	kWh	3,092
Solar fraction	%	61.8%

Figure 11. ThuSolar water heater data part 2

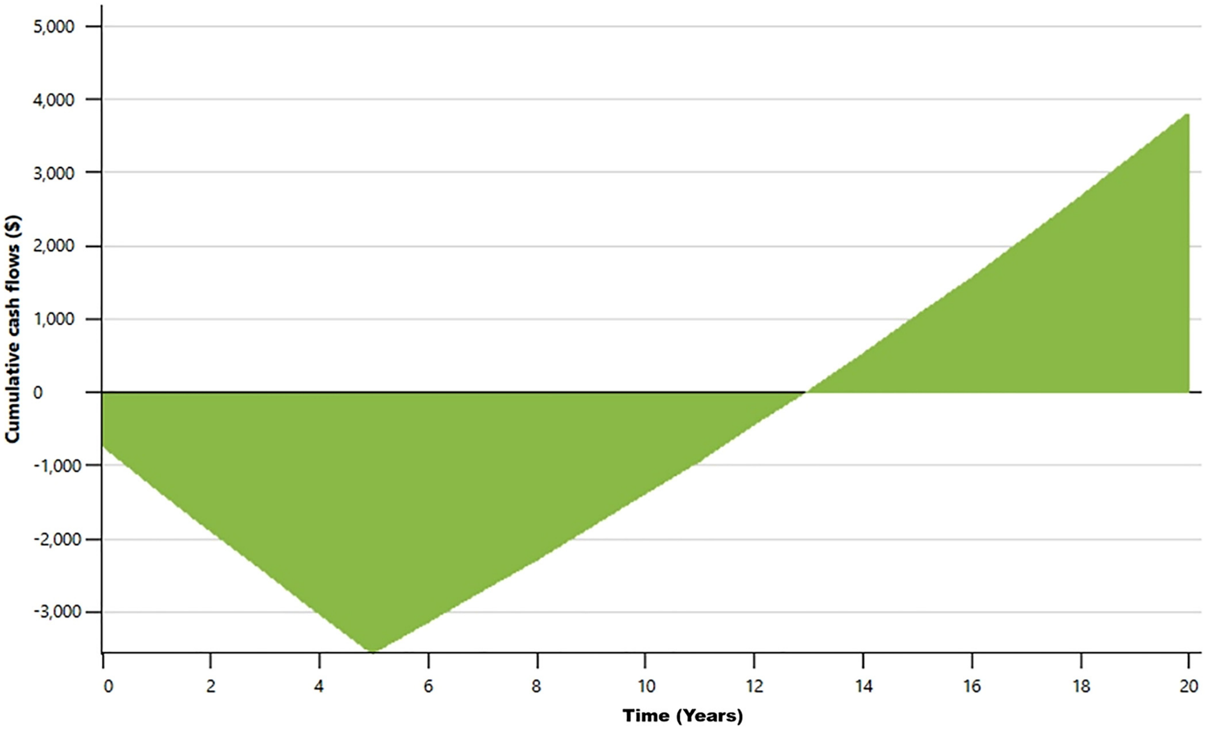


Figure 12. ThuSolar cumulative cash flows graph

and convert it to DC electricity. The cells are secured within a mounting structure to ensure stability and optimal orientation. The generated DC is then converted into AC power by an inverter, making it suitable for residential and commercial use.

A small battery is also typically included to store excess energy for use during periods of low solar power generation. Miscellaneous components, such as cables and connectors, ensure electrical connectivity among all the parts of the system. DC and AC disconnects are incorporated to allow for the safe isolation of the inverter during maintenance. A solar charge controller regulates the current and voltage to prevent battery overcharging, whereas a combiner box aggregates the electrical output from multiple solar panels connected in parallel, simplifying the wiring and improving the overall system efficiency[20].

2.5.3. Solar photovoltaic methodology. A key factor in the solar PV system analysis is the efficiency equation, which relates the power output to the maximum available solar radiation, as expressed in Eq. (4).

$$\eta_{cell} = \frac{P}{AG} \tag{4}$$

where P is the power capacity of the cell, A is the area of the solar collector, and G is the incident solar irradiation.

The methodology for the solar PV system followed steps similar to those outlined in Section 2.4, including the selection of the location, electricity prices, and number of PV panels to meet the required electrical load.

3. RESULTS & DISCUSSION

3.1. Solar Water Heater. Figure 9 presents the hot water demand calculations for a single household of five occupants with an occupancy rate of 90%, which translates to a daily requirement of 270 L of hot water. The water temperature was set to 70 °C, and the water heater was assumed to operate for 24 h per day during winter. With these inputs, the total power required for heating was calculated to be 5,007 kWh.

The device specifications were selected as shown in Figures 10 and 11. The solar tracking type was set to *fixed* because it is the most cost-effective option. The tilt angle (slope) of the collector was determined based on the latitude of the selected location. For Dhahran, which has a latitude of 26.3°, the acceptable range for optimal efficiency is between 20° and 27°. An initial slope of 22° was used and subsequently optimized based on the resulting solar fraction and heating output. The azimuth was set to 0°to ensure an optimal orientation.

The type of solar water heater was *glazed* because it offers higher efficiency than unglazed types and is less expensive than evacuated-tube solar heaters. The selected model was the ThuSolar 2.0, with the specifications shown in Figures 10 and 11. The air temperature and solar irradiation were assumed to have average values of 25.8 °C and 6.03 kWh/m^2 /day, respectively, as obtained from Figure 8. These values resulted in a thermal efficiency of 35% for the ThuSolar 2.0 solar water heater using Eq. (3). Based on the system parameters, two collectors were suggested, which provided a total collector area of 4.2 m^2. For efficient operation, the miscellaneous losses were assumed to be 2% and the storage capacity was assumed to be 73 L/m^2, resulting in a total of 272 L/day, which closely matched the calculated requirement of 270 L/day.

No heat exchangers were included because Dhahran's climate does not present freezing risks. Similarly, no pump power was required. The initial cost was set at \$650/m^2 (aperture area) for a solar water heater with storage, resulting in a total incremental initial cost of $2,421. Under these conditions, the system achieved an output of 3,088 kWh with a solar fraction of 61.7%. The final step involved selecting the optimal slope to maximize the solar fraction and energy savings. After several trials, the optimal slope was determined to be 25°.

The system saved 3,092 kWh out of the 5,007 kWh required, with an initial cost of $2,421. Based on the energy savings of 3,092 kWh for a single home, the total power savings for 2000 homes was $6,184,000$ kWh. The financial parameters considered for both this module and the SunEarth module included a fuel escalation rate of 3% and an inflation rate of 2.5%. A discount rate of 12% and reinforcement rate of 9% were applied to the analysis. The project life was assumed to be 20 years, with a debt ratio of 85% and a debt term of 5 years. This results in a simple payback period of 13.6 years shown in Figure 12, and an equity payback period of 12.9 years.

The use of the ThuSolar water heating system resulted in a reduction in carbon emissions of 1719.15 kg CO_2 per household, based on a grid emission factor of 0.556 kg CO_2/kWh in Saudi Arabia, and energy savings of 3,092 kWh, as calculated using Eq. (5):

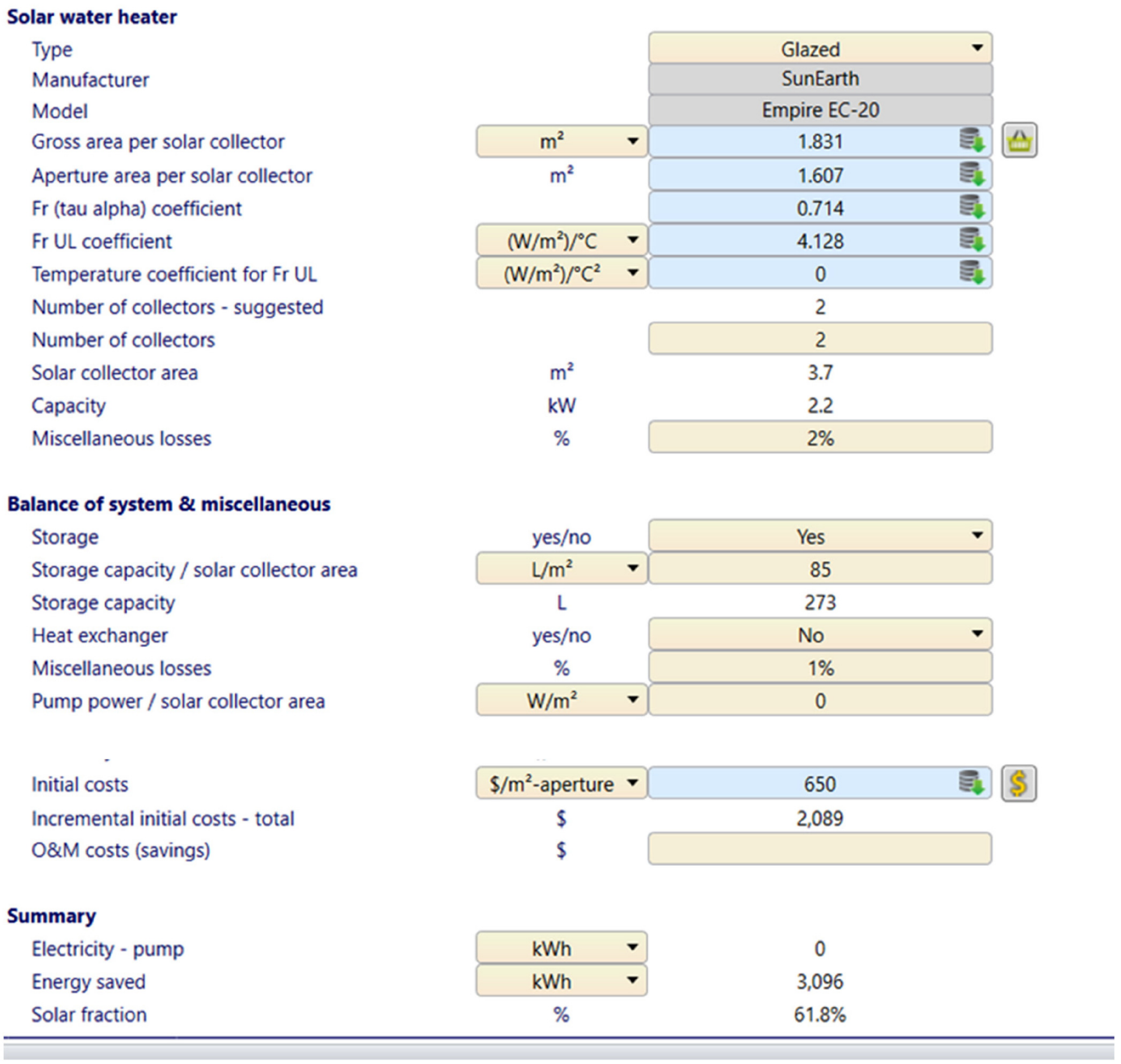
$$Reduction\ in\ CO_2 = Emissions\ factor \times Heating\ Load\ Reduction \tag{5}$$

For 2,000 houses, the total reduction in CO_2 emissions was calculated as $3,438,304\ kg\ CO_2$.

An additional performance assessment was conducted using a SunEarth solar water heater module featuring different design properties, which achieved a solar fraction of 61.8%. Using the appropriate parameters for the SunEarth module in Eq. (3), the thermal efficiency was calculated as 41%.

To satisfy the hot water demand, the storage capacity per solar collector area was adjusted to 85 L/m^2 as in Figure 13.

The SunEarth module was marginally superior to the ThuSolar module based on its higher thermal efficiency, heating value of 3,096 kWh, initial cost of $2,089, and simple payback period of 12.7 years as shown in Figure 14. The equity payback period was 12.2 years. Following a similar methodology, the CO_2 emissions



Solar water heater		
Type		Glazed
Manufacturer		SunEarth
Model		Empire EC-20
Gross area per solar collector	m²	1.831
Aperture area per solar collector	m²	1.607
Fr (tau alpha) coefficient		0.714
Fr UL coefficient	(W/m²)/°C	4.128
Temperature coefficient for Fr UL	(W/m²)/°C²	0
Number of collectors - suggested		2
Number of collectors		2
Solar collector area	m²	3.7
Capacity	kW	2.2
Miscellaneous losses	%	2%
Balance of system & miscellaneous		
Storage	yes/no	Yes
Storage capacity / solar collector area	L/m²	85
Storage capacity	L	273
Heat exchanger	yes/no	No
Miscellaneous losses	%	1%
Pump power / solar collector area	W/m²	0
Initial costs	$/m²-aperture	650
Incremental initial costs - total	$	2,089
O&M costs (savings)	$	
Summary		
Electricity - pump	kWh	0
Energy saved	kWh	3,096
Solar fraction	%	61.8%

Figure 13. Data for the solar water heater by SunEarth

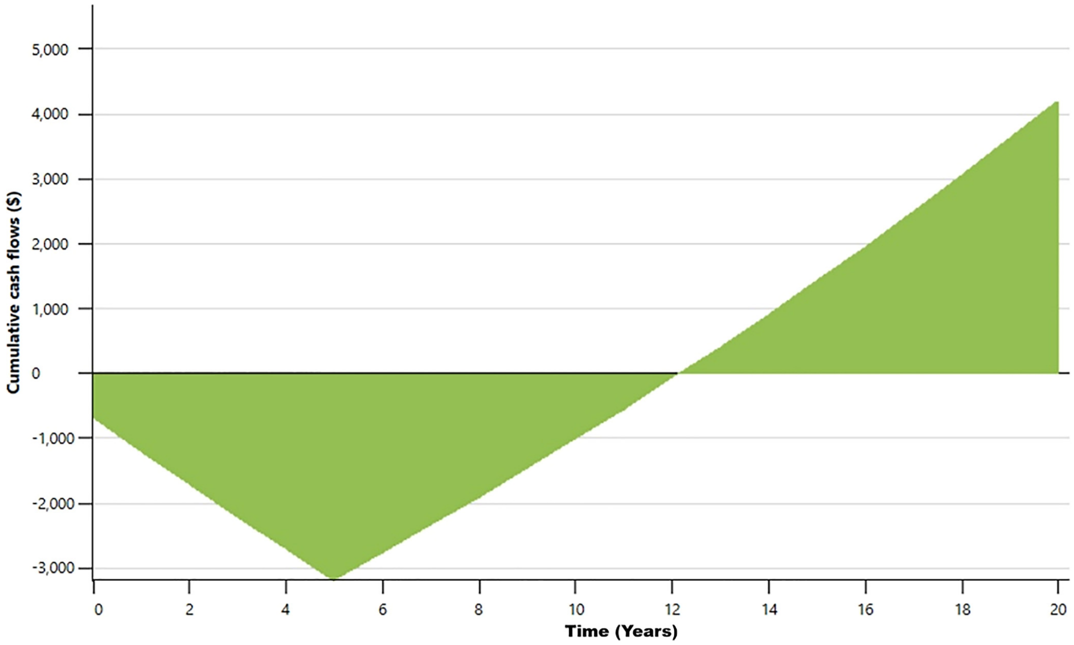


Figure 14. SunEarth cumulative cash flow graph

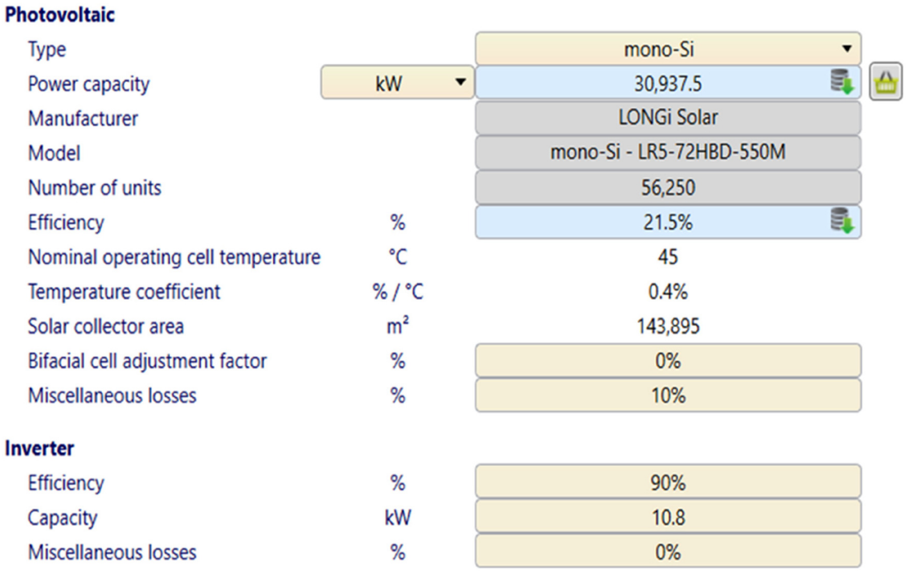
reduced by this module were calculated as $1,721.376\ kg\ CO_2$ per household. For a residential community of 2,000 houses, the use of the SunEarth water heating system resulted in an estimated total heating energy savings of 6,192,000 kWh. Therefore, the total reduction in CO$_2$ emissions was estimated as 3,442,752 kg CO$_2$, based on the emission factor and energy savings determined for a single household.

Based on the above comparative analysis, the SunEarth module is more suitable for this community because it offers higher power generation, lower costs, and a shorter payback period.

3.2. Solar Photovoltaic Power Plant. This section presents a comparative performance analysis of the three different solar panels. Figure 15 shows the details of the first solar panel used in this study. A fixed tracking mode was initially selected for the panels, with a slope of 22°and an azimuth of 0°, which provided maximum power generation. The first module was manufactured by LONGi Solar (mono-Si-LR5-72HBD-550M), with a rated power output of 550 W and an efficiency of 21.5%.

The number of units required to generate 484,060,000 kWh was 56,250, which had a total capacity of 30,937.5 kW and a capacity factor of 17.9%. The total area occupied by the panels was 143,895 m^2. The nominal operating cell temperature was 45 °C, which is an appropriate value for a city with hot summer weather such as Dhahran.

The initial cost of a solar PV power plant with a capacity between 10,000 kWh and 100,000 kWh was $1,100. With these parameters, the total power exported to the grid was calculated as 48,643,447 kWh, which generates revenue of $ 2,334,885 as Figure 16 shows. This power plant produces an additional



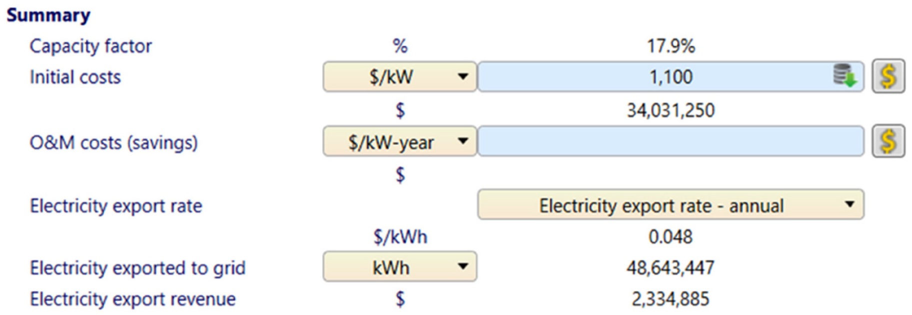
Photovoltaic

Type		mono-Si
Power capacity	kW	30,937.5
Manufacturer		LONGi Solar
Model		mono-Si - LR5-72HBD-550M
Number of units		56,250
Efficiency	%	21.5%
Nominal operating cell temperature	°C	45
Temperature coefficient	% / °C	0.4%
Solar collector area	m^2	143,895
Bifacial cell adjustment factor	%	0%
Miscellaneous losses	%	10%

Inverter

Efficiency	%	90%
Capacity	kW	10.8
Miscellaneous losses	%	0%

Figure 15. LONGi solar panels specification's part 1



Summary

Capacity factor	%	17.9%
Initial costs	$/kW	1,100
	$	34,031,250
O&M costs (savings)	$/kW-year	
	$	
Electricity export rate		Electricity export rate - annual
	$/kWh	0.048
Electricity exported to grid	kWh	48,643,447
Electricity export revenue	$	2,334,885

Figure 16. LONGi solar panels specification's part 2

237,447 kWh to compensate for the decrease in power output under poor weather conditions and efficiency degradation over time. The financial parameters for all panels used here included a fuel escalation rate of 2%, an inflation rate of 2%, and a reinvestment ratio of 9%. The project life was assumed to be 25 years, resulting in a net annual cash flow of $2,334,885. Based on the selected financial parameters, the simple payback period for this system was 14.6 years, Figure 17, whereas the equity payback period was 12.7 years. The reduction in CO_2 emissions was calculated using the same methodology as used previously for solar water heaters and based on energy savings of 48,643,447 kWh, which produced a value of $27,045,756.53$ *kg* CO_2. If one-axis tracking is used for the same number of panels, then the generated electricity would increase to 63,122,098 kWh.

Two additional panels were analyzed in comparison with the panel from LONGi. The second panel was manufactured by REC Solar (mono-Si-REC350NP2 Black 350W), which has a rated power output of 350 W and an efficiency of 19.1%, which were lower than those of the LONGi solar panels. The number of units required to generate the required electrical demand of 484,060,000 kWh was 88,500, with a total capacity of 30,937.5 kW and a capacity factor of 17.9%. The total area occupied by the solar panels was estimated at 162,173 m^2 Figure 18. The total electricity exported to the grid was 48,702,409 kWh, generating a revenue of $2,337,718. This system also produced an additional 296,409 kWh of power.

Based on the same financial parameters applied for the LONGi solar panels, the payback period for the REC PV panel was calculated as 14.6 years as in Figure 19, with an equity payback period of 12.7 years. This was expected because both the cost and revenue were lower than those of LONGi Solar. The net annual cash flow was $2,337,716, and the total reduction in CO_2 emissions for these panels for 48,702,409 kWh of electrical energy was calculated to be $27,078,539.4$ *kg* CO_2. If one-axis tracking is used for the same number of panels, the generated electricity would increase to 63,198,609 kWh.

The last module analyzed in this study was manufactured by Jinko Solar (mono-Si-JKM615N-78HL4-V 615W), which has a rated power output of 615 W and an efficiency of 22%, making it the most efficient panel among the three considered here. Owing to the larger power output per unit, only 50,500 units were needed to generate the required electrical demand of 484,060,000 kWh with a total capacity of 31,057.5 kW and a capacity factor of 17.9%. These panels also occupied the smallest total area of 141,170 m^2 Figure 20. They

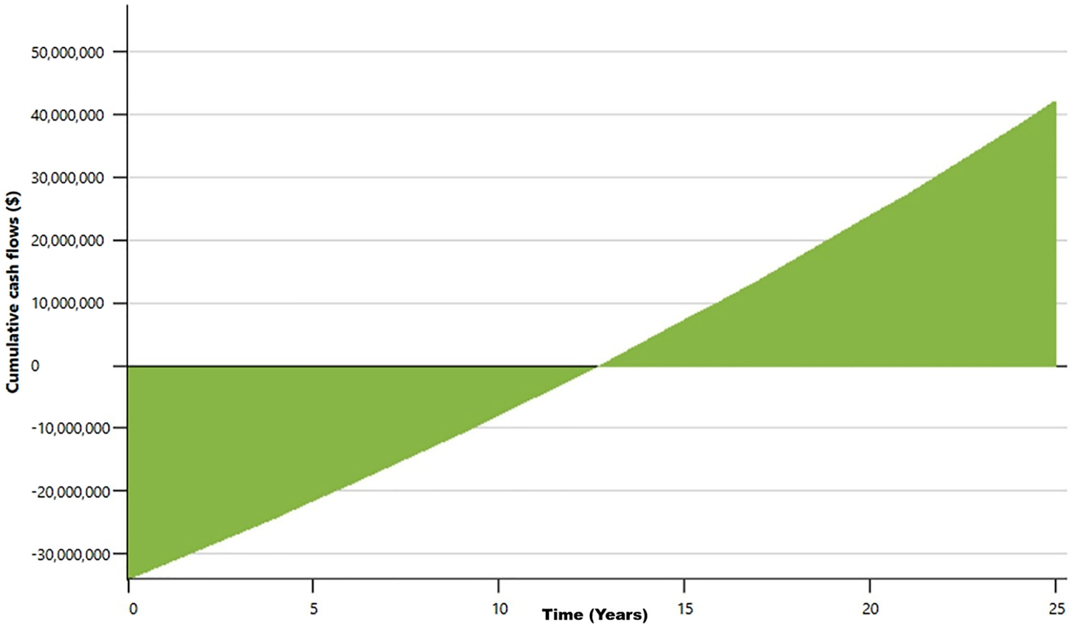
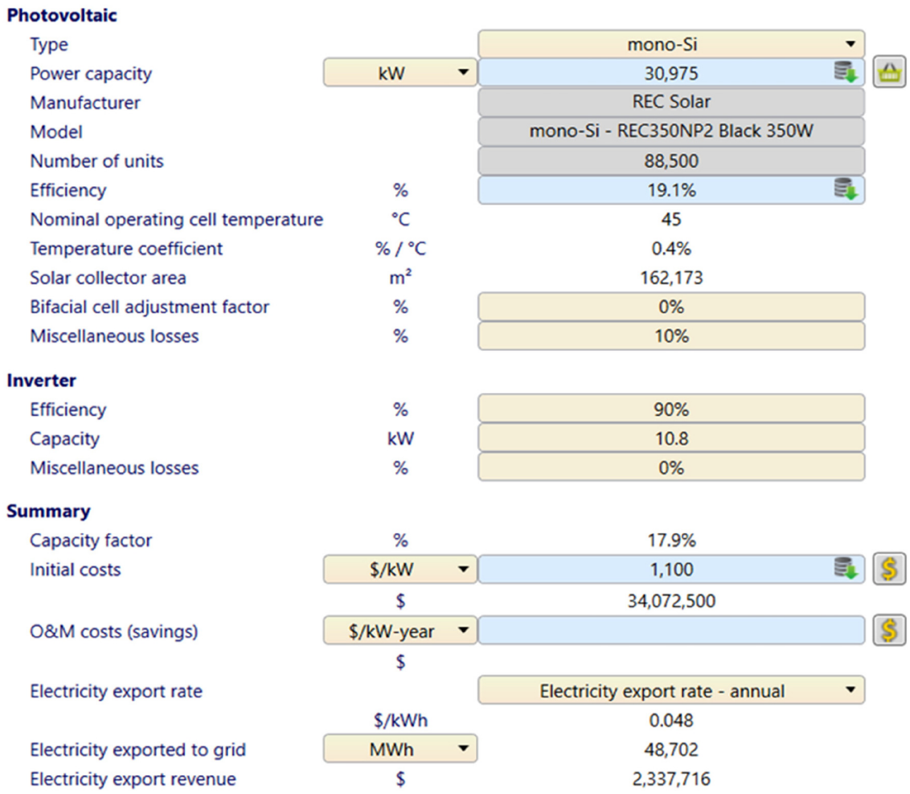


Figure 17. LONGi solar financial analysis



Photovoltaic		
Type		mono-Si
Power capacity	kW	30,975
Manufacturer		REC Solar
Model		mono-Si - REC350NP2 Black 350W
Number of units		88,500
Efficiency	%	19.1%
Nominal operating cell temperature	°C	45
Temperature coefficient	% / °C	0.4%
Solar collector area	m²	162,173
Bifacial cell adjustment factor	%	0%
Miscellaneous losses	%	10%
Inverter		
Efficiency	%	90%
Capacity	kW	10.8
Miscellaneous losses	%	0%
Summary		
Capacity factor	%	17.9%
Initial costs	$/kW	1,100
	$	34,072,500
O&M costs (savings)	$/kW-year	
	$	
Electricity export rate		Electricity export rate - annual
	$/kWh	0.048
Electricity exported to grid	MWh	48,702
Electricity export revenue	$	2,337,716

Figure 18. REC solar panels specifications

generated 48,832,125 kWh of electrical energy, with a total cost of $34,163,250 and exported electricity revenue of $2,343,942. The panels also produced an additional electrical energy of 426,125 kWh.

Using the same financial parameters as in the previous panels, the simple payback period for the Jinko solar system was calculated as 14.6 years as in Figure 21, with an equity payback period of 12.7 years and simple payback of 14.6 years, which are consistent with the previous models. The net yearly cash flow was calculated as $2,367,149. The reduced CO_2 emissions for these panels based on their electricity generation of 48,832,125 kWh equaled 27,150,661.5 *kg* CO_2. If one-axis tracking is used for the same number of panels, the generated electricity would increase to 63,366,935 kWh.

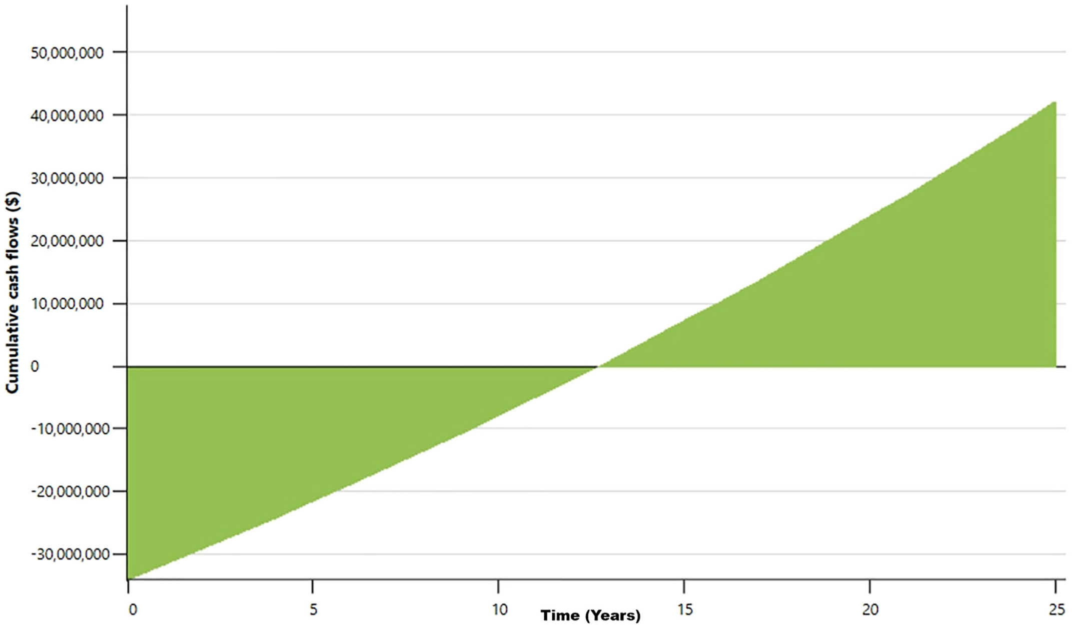
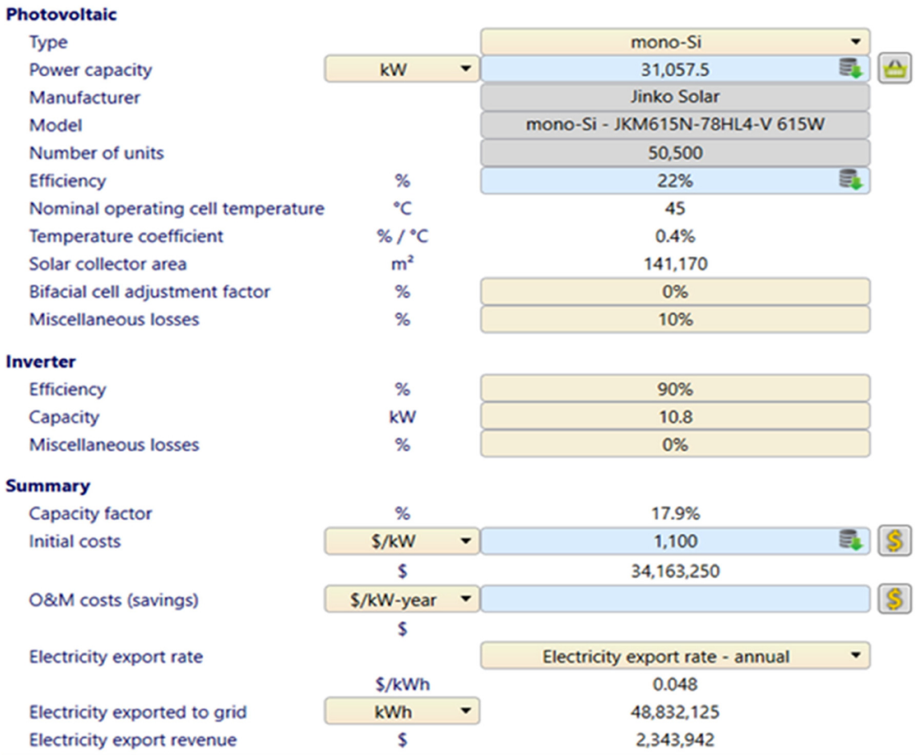


Figure 19. REC solar cumulative cash flows graph



Photovoltaic		
Type		mono-Si
Power capacity	kW	31,057.5
Manufacturer		Jinko Solar
Model		mono-Si - JKM615N-78HL4-V 615W
Number of units		50,500
Efficiency	%	22%
Nominal operating cell temperature	°C	45
Temperature coefficient	% / °C	0.4%
Solar collector area	m²	141,170
Bifacial cell adjustment factor	%	0%
Miscellaneous losses	%	10%
Inverter		
Efficiency	%	90%
Capacity	kW	10.8
Miscellaneous losses	%	0%
Summary		
Capacity factor	%	17.9%
Initial costs	$/kW	1,100
	$	34,163,250
O&M costs (savings)	$/kW-year	
	$	
Electricity export rate		Electricity export rate - annual
	$/kWh	0.048
Electricity exported to grid	kWh	48,832,125
Electricity export revenue	$	2,343,942

Figure 20. Jinko solar panels specifications

Finally, the performance of all three PV panels was evaluated under both fixed-axis and one-axis tracking configurations, and the results are summarized in Table 2. Although one-axis tracking increases the initial investment, it also enhances energy generation across all modules. The additional output more than offset the higher capital cost, thereby reducing the simple payback period from 14.6 years (fixed-axis) to 12.3 years. Moreover, the higher electrical output from the same number of units resulted in approximately a 30% greater reduction in GHG emissions and 30% increase in electricity export for all panels, with Jinko Solar demonstrating the best performance in both categories.

Based on the above results, Jinko solar panels emerged as the most suitable option for decarbonizing a 2000-home residential community because they offer the highest efficiency and greatest power generation and require the lowest number of panels (51,000 units). Its smaller land footprint and reduced panel requirements also minimize end-of-life waste

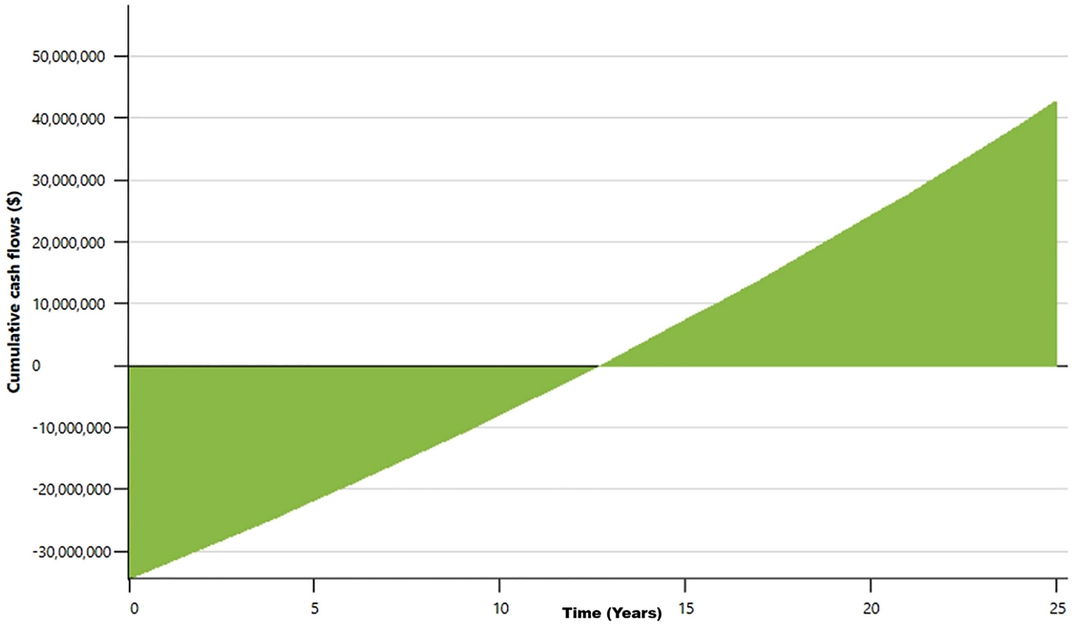


Figure 21. Jinko solar financial analysis

Table 2. *Module tracking mode comparison*

Module	Payback period (years)	Electricity exported (kWh)	Land area needed (m^2)	GHG reduction (kg CO_2)
LONGi Solar (Fixed)	14.6	48,643,447	143,895	27,045,756.5
LONGi Solar (One-axis)	12.3	63,122,098	143,895	35,095,886.5
REC Solar (Fixed)	14.6	48,702,409	162,173	27,078,539.4
REC Solar (One-axis)	12.3	63,198,609	162,173	35,138.427.2
Jinko Solar (Fixed)	14.6	48,832,125	142,568	27, 150, 661.5
Jinko Solar (One-axis)	12.3	63,366,935	142,568	35,215,336

streams, which align strongly with long-term sustainability goals.

Although the detailed comparative analysis conducted using RETSCREEN provides valuable insights, it is important to acknowledge several limitations inherent to the modelling approach. RETSCREEN relies on monthly average climate data and, while it predicts annual energy generation reliably, it may not fully capture hourly solar variability or temperature-dependent performance effects specific to Dhahran. Additionally, community-scale results were derived by scaling a single household model to 2,000 homes, which inherently overlooks variations in household behavior, load diversity, and installation conditions. Financial metrics such as payback period and CO_2 reduction are also based on simplified economic assumptions that do not reflect potential market fluctuations or future policy changes. These limitations suggest that while the present findings are robust at a screening level, more detailed simulations and field-based validation would enhance confidence in the long-term performance projections.

4. CONCLUSIONS

This study presents a comprehensive techno-economic evaluation of solar thermal water heating and solar photovoltaic (PV) systems designed to meet the hot water and electrical energy needs of a 2,000-home residential community in Dhahran, Saudi Arabia. Detailed simulations using RETScreen showed that the SunEarth solar water heater achieved an annual thermal yield of 6.19 million kWh with a land footprint of only 7,400 m^2, corresponding to 836.5 kWh/m^2/year. In contrast, the Jinko Solar single-axis tracking PV system delivered 63.37 million kWh of electrical energy annually over a land footprint of 142,568 m^2, corresponding to 444.5 kWh/m^2 /year. While 100% conversion of PV electricity to heat would produce a comparable thermal energy supply, the solar thermal water heaters demonstrated approximately 88% higher land-use efficiency for water heating applications, which is consistent with their higher conversion efficiencies compared to solar panels. This highlights the complementary roles of these technologies in meeting heating and electrical energy needs and emphasizes the value of dedicated solar water heaters in communities with limited land availability. An optimized hybrid strategy in which land is allocated to each system according to the specific heating or electrical demands could maximize total renewable energy yields and CO_2 mitigation per unit area.

From a carbon abatement perspective, SunEarth's system reduced CO_2 emissions by 3.44 million kg annually, whereas Jinko Solar's PV modules reduced CO_2 emissions by 35.22 million kg annually due to their larger absolute energy output. The payback periods for the SunEarth and Jinko systems were 14.6 years and 12.3 years, respectively, indicating that PV systems retain a marginally faster return on investment. The use of high-performance solar PV panels reduced the total panel count, thereby lowering the end-of-life waste streams. Specifically, Jinko Solar achieved an electrical energy yield of approximately 63 million kWh with 5,250 fewer panels than LONGi Solar and 37,500 fewer panels than REC Solar, reducing both waste and recycling costs at the end of their lifetimes. This aligns well with the United Nations' SDG 12 as well as the broader national goals of minimizing resource consumption and waste generation.

Future work can expand the present analysis across multiple climatic zones in the Kingdom to evaluate performance variations under different solar resource profiles, ambient temperatures, and seasonal demand patterns. Additionally, integrating thermal storage and implementing hybrid PV–thermal technologies could further enhance the practical applicability of these systems. Finally, detailed life cycle assessments, including embodied carbon and end-of-life recycling, should be conducted to strengthen sustainability metrics.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Ahmad Alzughaibi – *Mechanical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0009-0001-8720-7527
Email: ahmad.zuhair.alzughaibi@gmail.com

Corresponding Author

Awad B.S. Alquaity – *Research Mentor, Assistant Professor, Mechanical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0000-0002-9306-0154
Email: awad.alquaity@kfupm.edu.sa

ACKNOWLEDGEMENTS

The authors acknowledge the support provided by the Mechanical Engineering Department at King Fahd University of Petroleum and Minerals.

REFERENCES

(1) Ritchie, H. *How have the world's energy sources changed over the last two centuries?* Our World in Data (2021). https://ourworldindata.org/global-energy-200-years.

(2) Climate Action Tracker. *Saudi Arabia |Climate Action Tracker.* Climate Action Tracker (2025). https://climateactiontracker.org/countries/saudi-arabia.

(3) General Authority for Statistics. *Household Energy Statistics 2023* (2023). https://www.stats.gov.sa/documents/20117/2435281/Household+Energy+Statistics+2023+En

(4) Crank, R. *What is Solar Water Heating and How Does it Work?* Countryside (2019). https://www.iamcountryside.com/self-reliance/what-is-solar-water-heating-and-how-does-it-work.

(5) U.S. Department of Energy. *Solar Water Heaters.* Energy.gov (2022). https://www.energy.gov/energysaver/solar-water-heaters.

(6) textit. U.S. Energy Information Administration. *Photovoltaics and Electricity.* EIA (2024). https://www.eia.gov/energyexplained/solar/photovoltaics-and-electricity.php

(7) textit. Ritchie, H., Rosado, P. & Roser, M. *Data Page: Solar photovoltaic module price.* Our World in Data (2025). https://archive.ourworldindata.org/20250909-093708/grapher/solar-pv-prices.html

(8) Abd-ur-Rehman, H. M. & Al-Sulaiman, F. A. Optimum selection of solar water heating systems based on their comparative techno-economic feasibility study for the domestic sector of Saudi Arabia. *Renew. Sustain. Energy Rev.* **62**, 336–349 (2016).

(9) Almohammadi, K. M. & Allouhi, A. Techno-economic assessment and optimization of grid-connected solar PV systems in Saudi Arabia's building sector. *Utilities Policy* **93**, 101885 (2025).

(10) Alanazi, M. Optimal sizing of stand-alone hybrid energy system for development of rural and remote areas in Saudi Arabia. *Case Stud. Chem. Environ. Eng.* **12**, 101257 (2025).

(11) Almodfer, R., Zayed, M. E., Abd Elaziz, M., Aboelmaaref, M. M., Mudhsh, M. & Elsheikh, A. H. Modeling of a solar-powered thermoelectric air-conditioning system using a random vector functional link network integrated with jellyfish search algorithm. *Case Stud. Therm. Eng.* **31**, 101797 (2022).

(12) Abd Elaziz, M., Senthilraja, S., Zayed, M. E., Elsheikh, A. H., Mostafa, R. R. & Lu, S. A new random vector functional link integrated with mayfly optimization algorithm for performance prediction of solar photovoltaic thermal collector combined with electrolytic hydrogen production system. *Appl. Therm. Eng.* **193**, 117055 (2021).

(13) textit. El-Hadary, M. I., Senthilraja, S. & Zayed, M. E. A hybrid system coupling spiral type solar photovoltaic thermal collector and electrocatalytic hydrogen production cell: Experimental investigation and numerical modeling. *Process Saf. Environ. Prot.* **170**, 1101–1120 (2023).

(14) Zayed, M. E., Zhao, J., Elsheikh, A. H., Li, W., Sadek, S. & Aboelmaaref, M. M. A comprehensive review on Dish/Stirling concentrated solar power systems. *J. Clean. Prod.* **283**, 124664 (2021).

(15) González-Peña, D., Garcıia-Ruiz, I., Dıiez-Mediavilla, M., Dieste-Velasco, M. I. & Alonso-Tristán, C. Photovoltaic Prediction Software: Evaluation with real data from Northern Spain. *Appl. Sci.* **11**, 5025 (2021).

(16) RETScreen International. *Clean Energy Project Analysis.* Natural Resources Canada (2004).

(17) Brownson, J. R. S. *Flat-plate collector – an overview. ScienceDirect Topics* (2014). https://www.sciencedirect.com/topics/engineering/flat-plate-collector

(18) White, J. *Components and principals of a solar hot water system. JR Gas and Water* (2025). https://www.jrgasandwater.com.au/blogs/news/components-and-principals-of-a-solar-hot-water-system.

(19) Centre for Sustainable Energy. *Solar PV. CSE* (2025). https://www.cse.org.uk/advice/a-complete-guide-to-solar-pv

(20) AltE Store. *Components for Your Solar Panel (Photovoltaic) System.* (2025). https://www.altestore.com/pages/components-for-your-solar-panel-photovoltaic-system

KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Article *Computer Science*

Cutting Parameter, Cutting Force, and Tool Vibration Correlations for Predicting Surface Roughness During 42CrMo4+QT Turning

Farah Mohammed Fallata, Muhammad Fadhlan Afif*, and Ahmed Aly Diaa Sarhan*

Cite https://doi.org/10.64589/juri/214967

Submitted: October 17, 2025 **Revised:** October 28, 2025 **Accepted:** December 1, 2025

ABSTRACT

This study processed and analyzed data describing tool holder force and acceleration during the high-accuracy turning of high-strength alloy steel (42CrMo4+QT) workpieces under different cutting conditions to identify the features most strongly correlated with surface quality, which was defined as the machined surface roughness *Ra* in this study. The time-domain force signals were filtered to remove noise before extracting their average, maximum, minimum, and standard deviation features, and the frequency-domain acceleration signals were filtered and analyzed to obtain spectral features reflecting dynamic vibrations during cutting. The correlations between the extracted features and *Ra* were evaluated using the Pearson and Spearman rank coefficients; the results indicated that the tool corner radius, feed rate, and depth of cut exerted critical influences on *Ra*. Furthermore, the combined average cutting force features in all directions exhibited better correlation with *Ra* the acceleration features. Indeed, the corner radius, feed rate, depth of cut, and combined average multidirectional cutting force predicted *Ra* with adequate accuracy and significance. It can serve as reliable indicators for predicting workpiece surface quality as a result. This conclusion can inform the development of efficient methods for real-time monitoring and quality control during the turning of high-strength alloy steel-workpieces.

Keywords: machining, turning, cutting force, vibration, surface roughness

1. INTRODUCTION

Surface quality is a critical consideration during workpiece turning operations, particularly when machining high-strength alloy steel parts commonly used in the aerospace, automotive, energy, and defense industries that must provide high dimensional accuracy, functional reliability, and fatigue resistance. Notably, the surface roughness of a part has been found to influence its wear, fatigue life, and tribological function[1]. This has motivated investigations into which manufacturing parameters exert the most influence on surface roughness.

Although the performance of high-strength steel materials has long been considered sufficient, the process used to machine parts from such materials must be improved to ensure that the full benefits of the material properties are realized. This is especially true when conducting high accuracy turning operations using smaller work engagement than applied in general turning operations. Critically, the optimal cutting parameters must be employed during the turning process to ensure efficient manufacturing and satisfy quality requirements. However, the nonlinear mechanical behaviors, work hardening, and elevated cutting parameters associated with machining high-strength alloys can affect the wear and surface finish of the resulting parts. These conditions can be characterized by process parameters such as cutting force, vibration, and heat[2]. Therefore, determination of which parameters should be used to characterize turning operations is a critical subject of research, particularly for processes employing small cutting depths and high-strength alloys with relatively low surface roughness ranges.

This study accordingly analyzed the raw signal dataset provided by Diaz et al.[3] to investigate the effects of cutting parameters, forces, and vibrations on surface quality because the authors of this dataset have not considered vibration data in their published research[4,5]. The Pearson and Spearman rank correlation analysis were used to identify the critical features in this study as they have been widely applied to similar numerical datasets[6,7]. The results of this study are expected to provide a basis for the selection of critical features when developing machine learning models to predict surface quality. Indeed, appropriate feature selection can reduce the complexity of a prediction model without sacrificing its performance while simplifying data management and storage requirements.

This remainder of this paper is organized as follows: Section 2 describes the methods employed by Diaz et al.[3] to collect the data used in this study; Section 3 details the methodology used for signal processing and feature extraction; Section 4 presents the applied analysis methods and resulting correlations between the surface roughness and cutting parameter, force, and vibration features; finally, Section 5 summarizes the conclusions of this study and their contributions to future research.

Table 1. *Chemical composition of 42CrMo4 steel*[8]

C	Mn	Si	P	S	Cu	Cr	Ni	Al	Mo	V	Ti	Sn
0.42	0.64	0.21	0.013	0.009	0.02	1.04	0.06	0.026	0.185	0.006	0.0013	0.003

2. DATA COLLECTION

This section describes the methods employed by Diaz et al.[3] to collect the data used in this study. Although this information does not represent the present work, it provides critical details describing how the data were obtained and informs the signal processing methods applied in this study.

2.1. Material Properties.

The workpieces were made of 42CrMo4+QT steel, which possesses high impact and fatigue resistance[3,8]. The chemical composition and mechanical properties of this material are listed in Table 1[8] and Table 2[4], respectively.

2.2. Experimental Scheme.

The turning process (cutting) parameters defined in Figure 1 comprise the insert radius (r), cutting depth (a), feed (f), and cutting speed (v), which was determined based on the spindle rotational speed (n) as follows:

$$v = n\pi D \tag{1}$$

where D is the workpiece diameter.

The forces and vibrations acting on the tool holder were measured using sensors to obtain the cutting forces in all three directions (Fx, Fy, and Fz) and tool vibrations in the x- and y-directions (as acceleration ax and ay), as shown in Figure 1.

2.3. Machine and Equipment.

The specifications of the computer numerical control (CNC) lathe used by Diaz et al.[3] for the turning operations are described in Table 3, and the specifications of the cutting inserts used to perform the turning operations are listed in Table 4[3,9,10]. A confocal microscope was used with the settings detailed in Table 5[3] to measure the surface profile after each workpiece was machined and thereby determine its surface roughness Ra, as depicted in Figure 1. The turning process parameters (cutting force and vibration signals) were recorded using the data acquisition system specified in Table 6[3]. Finally, the LabView software was used to collect the measured cutting parameter data from the data acquisition system.

2.4. Design of Experiment.

Diaz et al.[3] varied the cutting parameters applied during the turning process using the combinations shown in Table 7. The final dataset contained 68 records comprising 34 combinations of f, v, and a values with 2 different $r$values.

3. DATASET SIGNAL PROCESSING

3.1. Cutting Force Signals.

Savitzky–Golay filtering[11,12] was performed to cancel out the noise present in the raw signals. A typical result of applying this filtering process is shown in Figure 2, indicating that the third filtering pass adequately eliminated the influence of noise.

The raw signals in each record comprised several components corresponding to different steps in the turning process, as shown in Figure 3. The correlation analysis conducted in this study focused on the data collected during the finishing operations by applying time-interval windows to isolate the desired raw signal from each operation. This segmentation strategy excluded the first- and last-contact signals of each operation from the analysis because sudden peaks that may produce outlier data are often observed in these regions.

The mean (*Avg*), maximum (*Max*), minimum (*Min*), and standard deviation (*Std*) metrics of the Fx, Fy, and Fz data in each window were extracted, and the average value of each metric from all four windows was considered to represent each feature that record. In addition, the combined average forces in all three directions, defined as $AvgFc$, was determined follows:

$$AvgFc = \sqrt{AvgFx^2 + AvgFy^2 + AvgFy^2} \tag{2}$$

3.2. Vibration Signals.

For the same time interval (window), the acceleration signal is segmented and passed to the Fast-Fourier Transform (FFT) and the subsequent vibration analysis[13,14]. These signals were decomposed into multiple components with frequencies ranging from zero to half of the sampling frequency. The FFT output consists of complex values; the absolute value (magnitude) of these complex numbers represents the amplitude of the signal at each specific frequency. This allows for the identification of dominant frequencies within the tool holder's vibration signal.

Figure 4 shows that the dominant frequency was identified at 3124 and 9375 Hz, representing the natural frequencies of the cutting tool which are critical factors in exciting chatter and thereby increase surface roughness. The frequencies with the highest amplitudes in the x- and y-directions—9375 Hz for both in this case—were defined as ωnx and ωny and passed to the dataset for correlation analysis.

The values of the different features extracted by the signal processing methods detailed in this section are listed in Table 8; the Ra values were obtained directly from the dataset[3].

Table 2. *Mechanical properties of 42CrMo4 steel*[4]

Mechanical properties	Value
Yield stress (MPa)	Stainless-steel
Tensile stress (MPa)	1097
Hardness (HV10)	310–370
Impact toughness (KV)	83.5 at 20 °C

Table 3. *Specifications of the CNC lathe used in the experiments*[3]

Model	JATOR TAJ-42
Machine software	Fagor 8055 T
Main speed engine power	11/15 kW
Maximum spindle speed	3000 rpm
Spindle nose	DIN 55,026 (A5)

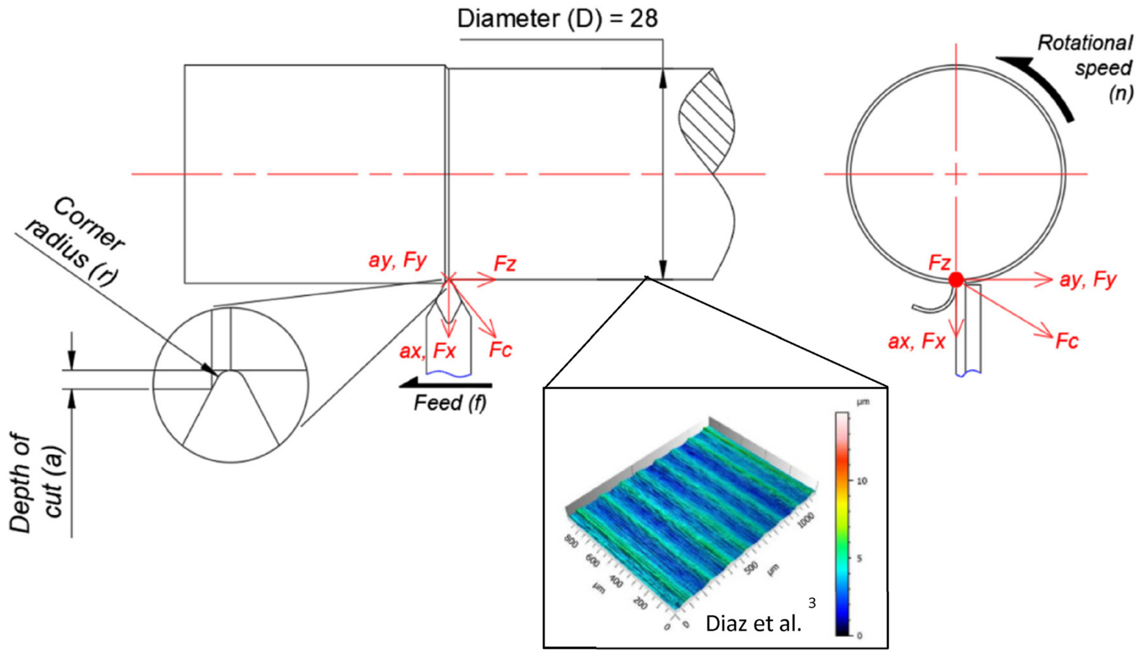


Figure 1. Cutting parameters, cutting forces, and surface roughness profile associated with the applied turning process

Table 4. *Specifications of the cutting inserts used in the experiments*[3,9,10]

Insert manufacturer	**Sandvik Coromant**	**Sandvik Coromant**
Model	**DCMX 11 T3 04-WF 4325**	**DCMX 11 T3 08-WF 4425**
Tip radius (mm)	0.4	0.8
Main cutting-edge angle (°)	93	93
Recommended *a* (mm)	1	1.5
Recommended *v* (m/min)	345–475	305–420
Recommended *f* (mm/rev)	0.07–0.30	0.12–0.4

4. FEATURE CORRELATION ANALYSIS

A correlation analysis was performed to investigate the influences of the obtained features on *Ra*. First, the Pearson correlation coefficient (Rp_{xy})[6,7] for each pair of features was obtained as follows:

$$Rp_{xy} = \frac{\sum_{i=1}^{m} (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{m} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{m} (y_i - \bar{y})^2}} \tag{3}$$

where x_i is the value of i-th record for feature x, $\bar{x}$ is the average of all feature x values, y_i is the value of i-th record for feature y, $\bar{y}$ is the average of all feature y values, and m is the total number of features.

The Pearson correlation coefficient ranges from −1 to 1; a positive coefficient indicates that an increase in feature x results

Table 5. *Leica DCM3 confocal microscope settings*[3]

Applied standard	UNE-EN ISO 21920-3
Objective magnification	10x
Measured parameter	Ra
Selected cut-off (λ_c) (mm)	0.80
Evaluation length (mm)	5.60
Evaluation width (mm)	0.95
Evaluation area (mm^2)	5.32

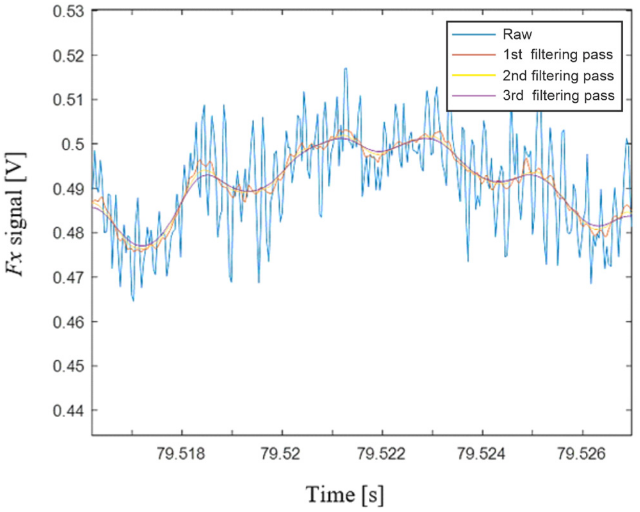


Figure 2. Results of Savitzky–Golay filtering on an example cutting force signal

Table 6. *Data acquisition system details*[3]

Procurement module		**NI 9269/NI 9234**	
Acquisition chassis		**NI cDAQ-9174**	
Parameter	**Register units**	**SI units**	**Conversion**
Acceleration X, Y	Gravity (g)	m/s^2	$9.8\ (m/s^2)/g$
Force X	V	N	100 N/V
Force Y	V	N	100 N/V
Force Z	V	N	100 N/V

in an increase in feature *y*, whereas a negative coefficient indicates that an increase in feature *x* results in a decrease in feature *y*, and the magnitude of the coefficient reflects the strength of the correlation. Note that the Pearson correlation assumes that the relationship between the two considered features is linear and reflects only the closeness, not the slope of that relationship.

Therefore, the Spearman rank correlation coefficient (Rs_{xy})[15] was also evaluated to determine whether a nonlinear (monotonic) relationship existed between features. Notably, Spearman rank correlation can compare abnormally distributed or outlier features. First, the numerical feature data were ranked (sorted) and the possibility of ties between records was considered and addressed[15], then the Spearman rank correlation was determined as follows:

$$Rs_{xy} = \frac{\sum_{i=1}^{m}\left(Rx_i - \overline{Rx}\right)\left(Ry_i - \overline{Ry}\right)}{\sqrt{\sum_{i=1}^{m}\left(Rx_i - \overline{Rx}\right)^2}\sqrt{\sum_{i=1}^{m}\left(Ry_i - \overline{Ry}\right)^2}} \tag{4}$$

where Rx_i and Ry_i are the values of *i*-th ranked features *x* and *y*, respectively, and $\overline{Rx}$ and $\overline{Ry}$ are the mean ranks for all records of features *x* and *y*, respectively.

The Spearman rank correlation coefficient varies within the same range as the Pearson correlation coefficient, and its sign indicates the same relationships. Evaluating the correlations between features using both the Pearson and Spearman rank coefficients can provide insight into the relationships between features, as shown in the example in Figure 5. The features that have low Pearson and Spearman rank coefficients (third quadrant) should be flagged as features that might cause noise on the prediction models. However, if at least one coefficient gives high value, the relationship between the two features can be estimated as shown.

The Pearson and Spearman rank correlation coefficients between all features were calculated using Equations (3) and (4), respectively, and are shown in the heat maps in Figure 6 and 7, respectively. The diagonal values in both figures are all equal to 1, confirming perfect correlation between identical features. Note that these heat maps show not only the correlations between the features and *Ra*, but also those among the features themselves.

Therefore, the heat map values in the row and column corresponding to correlations with *Ra* were used to generate the

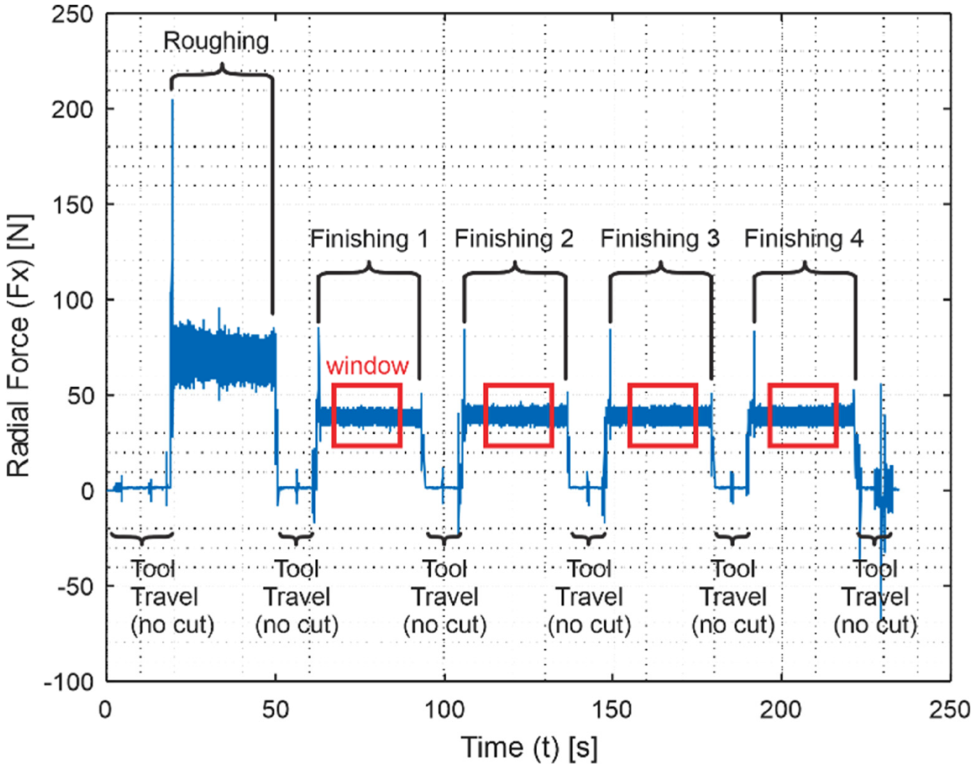


Figure 3. Signal segmentation strategy

Table 7. *Cutting parameters applied during the turning experiments*[3]

Parameter		Units	Values				
Feed rate	f	mm/rev	0.05	0.1	0.15	0.2	0.25
Cutting speed	v	m/min	76	107	138	169	200
Depth of cut	a	mm	0.05	0.1	0.15	0.2	0.25
Insert radius	r	mm	0.4	0.8	-	-	-

Run	r	a	v	f	Run	r	a	v	f
1	0.4	0.05	138	0.15	**35**	0.8	0.05	138	0.15
2	0.4	0.05	138	0.15	**36**	0.8	0.05	138	0.15
3	0.4	0.10	107	0.10	**37**	0.8	0.10	107	0.10
4	0.4	0.10	107	0.10	**38**	0.8	0.10	107	0.10
5	0.4	0.10	107	0.20	**39**	0.8	0.10	107	0.20
6	0.4	0.10	107	0.20	**40**	0.8	0.10	107	0.20
7	0.4	0.10	169	0.20	**41**	0.8	0.10	169	0.20
8	0.4	0.10	169	0.20	**42**	0.8	0.10	169	0.20
9	0.4	0.10	169	0.10	**43**	0.8	0.10	169	0.10
10	0.4	0.10	169	0.10	**44**	0.8	0.10	169	0.10
11	0.4	0.15	76	0.15	**45**	0.8	0.15	76	0.15
12	0.4	0.15	76	0.15	**46**	0.8	0.15	76	0.15
13	0.4	0.15	138	0.05	**47**	0.8	0.15	138	0.05
14	0.4	0.15	138	0.05	**48**	0.8	0.15	138	0.05
15	0.4	0.15	138	0.15	**49**	0.8	0.15	138	0.15
16	0.4	0.15	138	0.15	**50**	0.8	0.15	138	0.15
17	0.4	0.15	138	0.15	**51**	0.8	0.15	138	0.15
18	0.4	0.15	138	0.15	**52**	0.8	0.15	138	0.15
19	0.4	0.15	138	0.15	**53**	0.8	0.15	138	0.15
20	0.4	0.15	138	0.15	**54**	0.8	0.15	138	0.15
21	0.4	0.15	138	0.25	**55**	0.8	0.15	138	0.25
22	0.4	0.15	138	0.25	**56**	0.8	0.15	138	0.25
23	0.4	0.15	200	0.15	**57**	0.8	0.15	200	0.15
24	0.4	0.15	200	0.15	**58**	0.8	0.15	200	0.15
25	0.4	0.20	107	0.10	**59**	0.8	0.20	107	0.10
26	0.4	0.20	107	0.10	**60**	0.8	0.20	107	0.10
27	0.4	0.20	107	0.20	**61**	0.8	0.20	107	0.20
28	0.4	0.20	107	0.20	**62**	0.8	0.20	107	0.20
29	0.4	0.20	169	0.20	**63**	0.8	0.20	169	0.20
30	0.4	0.20	169	0.20	**64**	0.8	0.20	169	0.20
31	0.4	0.20	169	0.10	**65**	0.8	0.20	169	0.10
32	0.4	0.20	169	0.10	**66**	0.8	0.20	169	0.10
33	0.4	0.25	138	0.15	**67**	0.8	0.25	138	0.15
34	0.4	0.25	138	0.15	**68**	0.8	0.25	138	0.15

bar plot shown in Figure 8(a). The cutting force features clearly exhibited greater correlations with *Ra* than the vibration features. As a result, *AvgFc* can be used as the sole feature to differentiate and characterize the effects of various process parameters on *Ra* in the considered dataset. A bar plot showing the correlations between all features and *AvgFc* is accordingly provided in Figure 8(b). This result agrees with the conclusions of many previous studies that cutting forces can be used to characterize surface roughness[17–20].

In terms of the correlations between the key features and *Ra*, f clearly exerted the most influence on *Ra*, with Pearson and Spearman rank correlation coefficients of 0.71 and 0.68, respectively, followed by r, with coefficients of -0.41 and -0.44, respectively. Furthermore, most of the Pearson and Spearman rank correlation coefficients for the features were nearly identical, except for those corresponding to *StdFx, StdFy,* and *StdFz*. Together, these results suggest that the relationship between f and *Ra* is likely linear, whereas the relationships between the other features and *Ra* are nonlinear, complex, or reflect no correlation at all.

Interestingly, nearly all features, including *AvgFc*, were positively correlated with *Ra*, whereas $r, a,$and$v$ were negatively (although weakly) correlated with *Ra*. The reasons for these correlations can be observed in the empirical equations for estimating *Ra*, given by

$$Ra = \frac{0.0032f^2}{r} \tag{5}$$

and *Fc*, given by

$$Fc = k_c A \tag{6}$$

Table 8. *Features extracted by signal processing and analysis*

Run	*r*	*a*	*v*	*f*	*AvgFx*	*MaxFx*	*MinFx*	*StdFx*	*AvgFy*	*MaxFy*	*MinFy*	*StdFy*	*AvgFz*	*MaxFz*	*MinFz*	*StdFz*	*AvgFc*	*ωnx*	*ωny*	*Ra*
1	0.4	0.05	138	0.15	41.09	48.37	33.80	1.37	62.13	71.28	54.60	3.16	-9.29	-0.27	-16.63	3.93	75.06	8759	9375	0.858
2	0.4	0.05	138	0.15	40.84	46.79	34.20	1.19	54.10	60.18	49.97	0.94	-0.37	3.34	-4.52	0.90	67.79	8820	3125	1.078
3	0.4	0.1	107	0.1	39.05	46.50	32.82	1.34	43.55	47.77	39.23	0.91	-5.96	-1.52	-10.11	1.26	58.80	9375	9375	0.409
4	0.4	0.1	107	0.1	44.91	52.77	38.39	1.57	63.82	60.22	59.28	0.91	2.23	6.25	-1.85	1.22	78.07	8758	9375	0.558
5	0.4	0.1	107	0.2	51.70	108.37	-3.18	23.49	76.29	148.84	11.33	29.23	-11.48	20.10	-43.67	9.77	92.87	8809	9375	1.133
6	0.4	0.1	107	0.2	63.63	70.58	57.05	1.51	92.77	97.24	88.63	1.00	13.27	18.51	8.23	1.50	113.28	8809	9375	1.185
7	0.4	0.1	169	0.2	67.22	73.12	60.94	1.37	87.50	92.16	83.28	1.13	13.66	17.89	9.39	1.07	111.18	8720	3125	1.156
8	0.4	0.1	169	0.2	66.44	73.78	60.50	1.44	92.51	98.46	86.94	1.56	3.14	8.17	-2.25	1.81	113.94	8770	3125	1.140
9	0.4	0.1	169	0.1	49.49	59.04	43.14	1.68	59.81	66.25	54.16	1.62	3.19	7.74	-1.01	1.41	77.69	8723	9375	0.640
10	0.4	0.1	169	0.1	49.34	55.33	43.85	1.37	62.26	66.95	57.66	1.32	0.07	4.93	-4.25	1.56	79.44	9375	9375	0.561
11	0.4	0.15	76	0.15	61.69	70.96	52.39	1.93	97.96	103.86	92.30	1.46	-29.28	-18.98	-38.67	3.49	119.41	3125	3125	0.759
12	0.4	0.15	76	0.15	65.17	94.17	29.72	3.60	105.29	128.53	90.90	4.02	-31.50	-18.11	-43.99	2.41	127.77	3125	3125	0.806
13	0.4	0.15	138	0.05	37.21	54.37	22.95	2.12	38.43	49.83	31.77	2.07	35.68	51.72	16.35	7.72	64.29	3125	9375	0.306
14	0.4	0.15	138	0.05	63.88	83.64	44.55	3.34	171.72	188.87	158.84	3.45	59.91	78.82	46.75	3.91	192.76	9375	9375	0.657
15	0.4	0.15	138	0.15	58.71	65.46	52.32	1.43	91.16	96.03	85.98	1.13	-32.37	-25.75	-39.21	2.51	113.16	3125	9375	0.796
16	0.4	0.15	138	0.15	51.18	147.87	-7.28	31.68	120.23	346.49	-9.80	117.57	3.42	122.24	-62.01	59.57	130.72	9375	9375	0.849
17	0.4	0.15	138	0.15	65.63	73.24	59.50	1.44	95.03	99.79	90.90	1.01	-3.68	3.12	-10.90	2.79	115.55	9375	9375	0.675
18	0.4	0.15	138	0.15	63.68	70.57	57.57	1.42	90.90	96.58	86.12	1.17	-33.96	-26.22	-40.67	2.76	116.07	3125	9375	0.690
19	0.4	0.15	138	0.15	63.71	74.91	48.27	1.65	94.61	103.15	89.15	1.33	-34.99	-25.20	-44.66	4.20	119.31	9375	9375	0.888
20	0.4	0.15	138	0.15	56.98	63.96	50.27	1.52	74.21	79.09	69.54	1.06	13.31	19.92	6.75	2.46	94.51	9375	9375	0.870
21	0.4	0.15	138	0.25	83.31	95.09	72.27	2.12	129.99	142.76	123.11	1.97	-9.09	-1.26	-16.29	2.48	154.66	3125	9375	1.623
22	0.4	0.15	138	0.25	80.92	90.75	70.95	1.98	127.29	133.22	121.72	1.51	-9.01	-1.59	-16.26	2.50	151.10	3125	3125	1.682
23	0.4	0.15	200	0.15	73.31	81.41	65.47	1.85	94.08	100.61	86.99	1.66	-2.26	6.98	-11.00	3.91	119.29	3125	3125	0.587
24	0.4	0.15	200	0.15	73.31	82.14	63.39	1.84	96.82	104.82	89.30	1.80	-3.19	7.17	-13.11	4.23	121.49	9375	9375	0.596
25	0.4	0.2	107	0.1	53.74	62.21	46.11	1.75	90.43	97.05	83.74	1.53	-22.54	-9.76	-35.70	6.01	107.58	9375	9375	0.478
26	0.4	0.2	107	0.1	52.74	61.30	43.32	1.80	89.21	96.15	83.68	1.49	-21.21	-10.44	-31.07	4.45	105.78	3125	3125	0.574
27	0.4	0.2	107	0.2	86.94	98.85	71.76	2.53	120.73	127.90	113.22	1.85	-45.96	-31.12	-60.84	6.34	155.71	9375	9375	1.002
28	0.4	0.2	107	0.2	76.86	92.47	61.94	2.58	112.03	119.04	105.06	2.10	-28.67	-19.26	-38.84	3.79	138.85	9375	9375	1.074
29	0.4	0.2	169	0.2	92.08	104.52	78.65	2.91	125.64	133.40	117.50	2.18	18.94	30.00	7.23	3.89	156.92	3125	3125	1.111
30	0.4	0.2	169	0.2	87.29	97.80	75.56	2.42	118.32	125.01	111.60	1.65	10.89	22.89	-1.41	4.73	147.44	3125	3125	1.091
31	0.4	0.2	169	0.1	57.59	67.44	50.39	1.88	88.38	95.99	81.93	1.87	-19.57	-9.53	-28.89	3.85	107.29	3125	3125	0.432
32	0.4	0.2	169	0.1	58.44	67.98	50.69	1.93	89.01	96.43	82.82	1.78	16.21	28.04	4.79	5.22	107.71	9375	9375	0.460
33	0.4	0.25	138	0.15	69.60	79.22	60.12	1.98	109.35	114.92	102.75	1.34	-28.30	-16.64	-39.89	4.86	132.67	3125	3125	0.735
34	0.4	0.25	138	0.15	74.61	84.13	64.62	2.02	128.73	134.83	121.81	1.55	-33.15	-16.58	-49.74	7.56	152.43	8770	9375	0.783
35	0.8	0.05	138	0.15	37.93	48.45	30.25	1.45	60.89	65.90	56.58	1.03	-27.03	-23.48	-30.44	1.07	76.66	3125	3125	0.800
36	0.8	0.05	138	0.15	36.83	36.83	27.90	2.05	57.01	57.01	57.01	1.59	-35.04	-28.89	-38.63	0.86	76.39	3125	3125	1.069
37	0.8	0.1	107	0.1	32.50	44.55	22.37	1.73	44.49	51.23	31.06	1.58	-42.25	-1.92	-50.33	4.00	69.43	9375	9375	0.477
38	0.8	0.1	107	0.1	37.66	49.11	27.67	2.17	65.14	72.42	59.87	1.58	-42.77	-35.41	-49.87	3.00	86.55	9375	9375	0.551
39	0.8	0.1	107	0.2	52.20	61.96	43.69	1.74	73.44	77.71	69.23	0.98	-31.16	-25.21	-36.92	2.03	95.34	9375	9375	0.735
40	0.8	0.1	107	0.2	59.03	68.67	48.58	2.03	93.41	100.11	87.86	1.27	-27.99	18.34	-18.34	4.04	113.98	9375	9375	0.777
41	0.8	0.1	169	0.2	62.14	70.49	54.20	1.76	83.46	90.25	78.15	1.62	-30.00	-26.34	-33.56	0.96	108.29	9375	9375	0.730

(Continued)

Table 8. *Continued*

Run	*r*	*a*	*v*	*f*	*AvgFx*	*MaxFx*	*MinFx*	*StdFx*	*AvgFy*	*MaxFy*	*MinFy*	*StdFy*	*AvgFz*	*MaxFz*	*MinFz*	*StdFz*	*AvgFc*	ωnx	ωny	*Ra*
42	0.8	0.1	169	0.2	62.58	69.53	56.20	1.66	86.43	92.21	81.37	1.60	4.70	10.81	-0.99	2.02	106.81	3125	3125	0.727
43	0.8	0.1	169	0.1	43.02	51.41	37.59	1.47	65.52	72.28	60.84	1.25	18.72	-12.24	-25.39	2.64	80.58	3125	3125	0.544
44	0.8	0.1	169	0.1	42.98	53.44	36.99	1.68	64.04	71.05	58.90	1.38	-16.48	-10.11	-22.65	2.54	78.86	3125	3125	0.468
45	0.8	0.15	76	0.15	50.89	63.73	41.21	2.03	77.92	86.18	70.68	2.22	-47.64	-35.96	-58.82	5.06	104.55	9375	9375	3.235
46	0.8	0.15	76	0.15	57.91	69.28	44.37	2.18	100.64	108.43	92.94	1.91	-46.69	-33.65	-58.50	5.33	125.14	3125	3125	0.624
47	0.8	0.15	138	0.05	39.77	54.47	31.70	2.62	62.10	71.59	56.15	2.02	-37.94	-28.54	-46.71	3.93	82.93	3125	3125	0.307
48	0.8	0.15	138	0.05	44.13	67.65	24.80	3.49	65.77	82.31	58.15	2.22	-37.38	-23.87	-48.52	4.36	87.59	3125	3125	0.275
49	0.8	0.15	138	0.15	62.25	74.70	52.81	1.82	95.63	107.55	89.61	1.81	-42.18	-32.16	-53.09	4.02	121.65	9375	9375	0.488
50	0.8	0.15	138	0.15	64.79	89.57	37.63	3.10	102.95	115.56	91.21	2.79	-12.47	1.08	-26.37	4.95	122.28	3125	3125	0.555
51	0.8	0.15	138	0.15	63.14	70.31	55.73	1.61	97.41	103.36	93.10	1.12	-14.27	-2.53	-25.31	5.46	116.96	3125	3125	0.406
52	0.8	0.15	138	0.15	61.60	70.53	53.82	1.60	96.76	102.02	91.54	1.24	-13.84	-1.89	-25.71	5.67	115.53	3125	3125	0.410
53	0.8	0.15	138	0.15	60.23	67.28	53.60	1.48	94.19	98.73	89.64	1.06	-12.23	-1.49	-23.16	4.83	112.47	9375	9375	0.475
54	0.8	0.15	138	0.15	62.93	129.83	-15.83	4.28	112.73	164.98	83.24	10.70	-34.60	40.88	-56.83	17.33	133.66	3125	3125	0.565
55	0.8	0.15	138	0.25	81.74	90.22	73.09	1.89	133.78	139.31	128.13	1.37	-16.39	-5.02	-28.12	5.08	157.63	9375	9375	0.854
56	0.8	0.15	138	0.25	80.76	90.10	72.32	1.88	131.60	136.53	126.34	1.28	-16.23	-4.55	-26.89	4.83	155.25	9375	9375	0.902
57	0.8	0.15	200	0.15	65.17	73.94	57.81	1.88	97.11	104.31	89.67	1.96	-11.55	-2.27	-20.85	4.27	117.52	3125	3125	0.711
58	0.8	0.15	200	0.15	62.87	71.32	55.62	1.76	92.69	100.49	85.99	1.94	-17.47	-10.84	-24.15	2.73	113.35	9375	9375	0.600
59	0.8	0.2	107	0.1	57.01	118.66	-35.10	4.27	99.76	174.81	73.95	5.41	-22.23	33.87	-46.99	14.22	117.03	9375	9375	0.385
60	0.8	0.2	107	0.1	58.00	72.90	44.56	2.35	92.47	104.29	85.43	2.02	-54.09	-36.53	-71.44	8.44	121.82	3125	3125	0.461
61	0.8	0.2	107	0.2	80.93	94.11	94.11	2.50	135.92	141.84	128.63	1.44	17.35	42.85	-6.75	12.44	159.14	9375	9375	1.259
62	0.8	0.2	107	0.2	81.82	94.04	71.52	2.47	141.14	148.46	133.72	1.67	25.70	45.48	6.56	9.28	165.15	9375	9375	0.693
63	0.8	0.2	169	0.2	86.32	96.63	75.34	2.20	139.36	147.34	131.26	1.95	14.27	26.30	3.09	4.83	164.55	3125	3125	0.445
64	0.8	0.2	169	0.2	82.17	94.78	94.78	2.94	136.27	136.27	125.52	2.88	-26.09	-13.74	-38.48	4.81	161.25	3125	3125	0.466
65	0.8	0.2	169	0.1	67.49	120.63	13.26	3.64	90.99	159.87	70.17	4.60	5.34	39.61	-12.22	8.21	113.41	3125	3125	0.540
66	0.8	0.2	169	0.1	67.55	81.71	58.68	2.45	89.63	98.76	83.08	2.01	12.98	26.83	0.60	5.85	112.98	3125	3125	0.372
67	0.8	0.25	138	0.15	78.03	137.61	19.05	4.21	136.51	186.05	118.29	4.50	-42.20	-15.38	-62.96	6.13	162.80	3125	9375	0.367
68	0.8	0.25	138	0.15	70.23	79.17	61.15	1.99	115.06	121.22	108.68	1.48	21.92	40.01	3.47	8.72	136.57	3125	3125	0.535

where k_c is the specific cutting constant (in N/mm^2) and A is the cutting cross-section. They can also be observed in the theoretical surface roughness model shown in Figure 9, in which Z is the peak-to-valley distance, and a_r is to the cutting depth recommended by the tool provider (shown for comparison). Note that Equations (5) and (6), as well as the models in Figure 9, have all been widely used in previous work[4,21,22].

Both the models in Figure 9 and results given by Equations (5) and (6) agree with the results of the correlation analysis conducted in this study: a and f exhibited strong correlations with *AvgFc* because high values for these parameters result in a higher A, increasing *Fc*. Furthermore, f and$r$respectively exhibited positive and negative correlations with *Ra*. The theoretical models illustrated in Figure 9 demonstrate the effects of these two parameters on Z, which significantly influences *Ra*. Note that the complex interactions between the tool and workpiece, as well as chip breakage, material defects, and other uncertainties will inevitably result in minor deviations from the models shown in the figure[21,23]. Finally, although the applied v directly affects the scale of the horizontal surface profile, this change only slightly influences *Ra*, explaining the low correlation between these two features.

Notably, although the analysis results indicated that a smaller a yielded a higher *Ra*, this has not typically been observed during general turning processes. This unconventional trend may be caused by the low range of $a$considered in the dataset (0.05–0.25 mm) compared with the a values applied in typical turning operations or even the a_r values recommended by tool provider (1 and 1.5 mm in Table 4)[3,9,10]. This can be observed in Figure 9, which shows that small cutting depths cause the tool to only "scratch" the workpiece surface without fully engaging it, such that little cutting actually occurs; these size effects have been reported to correlate with reduced cutting stability[4,21]. Note that these relatively shallow a values are only used to achieve close tolerance and tight *Ra* requirements during high-accuracy turning operations, while avoiding workpiece undersizing that can lead

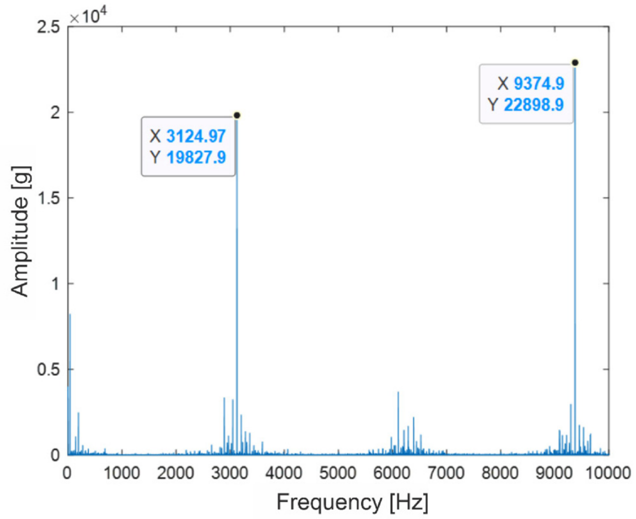


Figure 4. Frequency domain of the segmented acceleration signal

to workpiece rejection [23]. It is also known that v scales the surface profile horizontally such that it might give only slight changes to the *Ra* value. That is the reason for the low correlation value between v and Ra.

The results of the Pearson and Spearman rank correlation analyses were complemented by evaluations of the R^2 and p—values for the relationships between features, as shown in Figures 10 and 11, respectively. The R^2 values reflect the variance in the dependent variable (i.e., *Ra* or *AvgFc*) when predicted using the independent variable (i.e., each of the remaining features). Note that the R^2 analysis was based only on the Pearson correlation because R^2 is only applicable to linear relationships. The p—values reflect the statistical significance of the observed correlations, with a lower p—value indicating that the correlation is unlikely to occur by random chance.

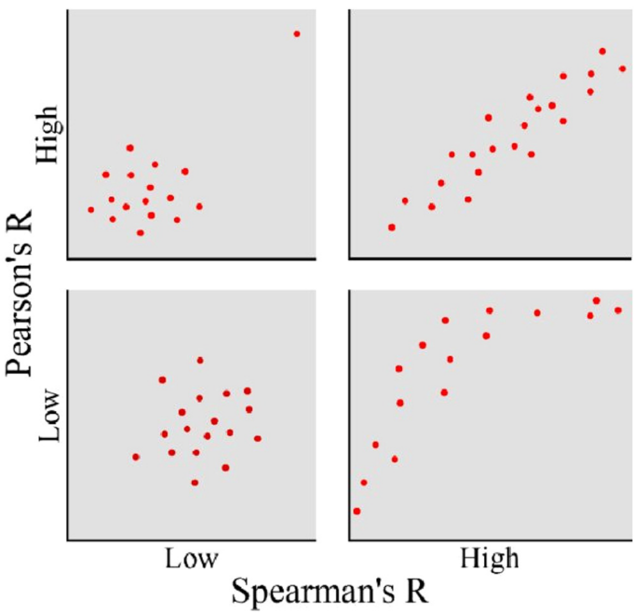


Figure 5. Relationships between two features based on Pearson and Spearman rank correlation [16]

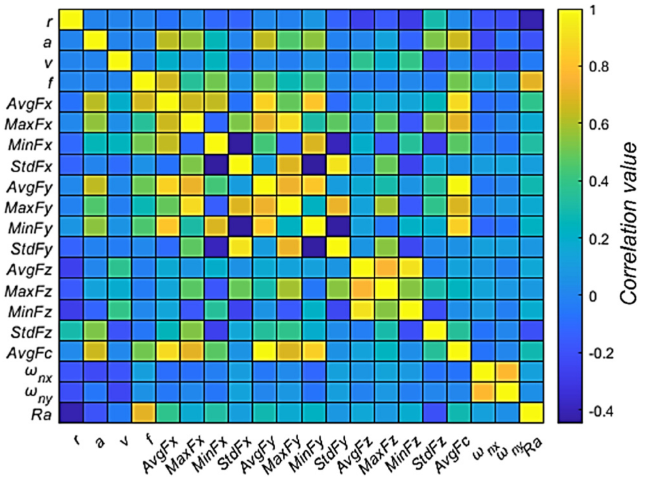


Figure 6. Pearson correlations among all features and Ra

Figure 10 indicates that the correlation between f and *Ra* exhibited the highest R^2 value of only approximately 0.5. However, relying solely on f to predict *Ra* will be insufficient, and the addition of other features, such as r and *AvgFc,* to *Ra* prediction models is recommended accordingly. Furthermore, the correlations between the various features and *AvgFc,* which is calculated by combining the average cutting forces in all directions, indicate that a and f exhibited relatively higher R^2 values than the other features. Therefore, f, r, a, and *AvgFc* were determined to exert the most influence on *Ra*. Indeed, Figure 11 shows that these four features are within the ideal p-value limit of 0.05, whereas the remaining features exhibit p-values above this limit for correlations with either *Ra* or *AvgFc.* Furthermore, note that features such as the *Max*, *Min*, and *Std* values of the forces in each direction were not reliable because the analysis indicated that they likely exhibited randomness during the turning process, affecting the integrity of their correlations with *Ra* and *AvgFc.*

These insights into the features exerting the greatest influence on *Ra* can inform efforts to increase the efficiency of turning process monitoring and data storage, which are critical concerns in the pursuit of intelligent machining [24]. Indeed, predictive models constructed to consider only a few highly correlated features can achieve superior performance and efficiency over models that

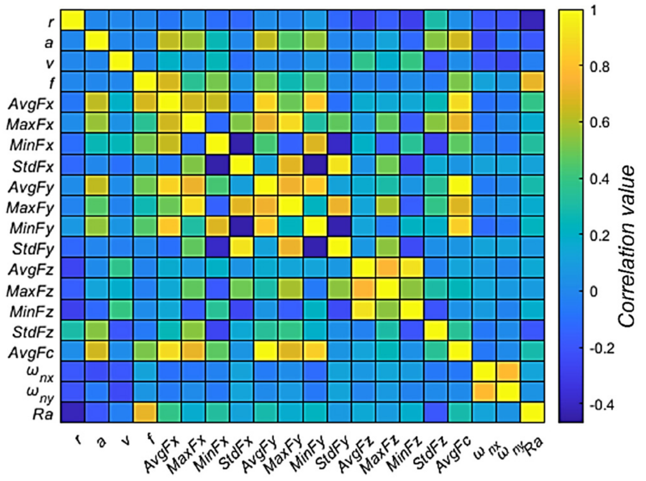
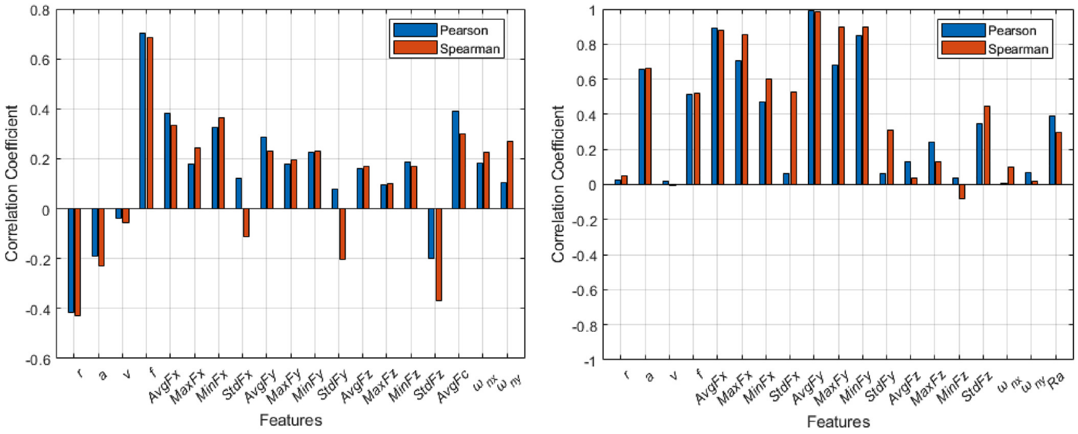


Figure 7. Spearman correlations among all features and Ra



(a) *Ra* (b) *AvgFc*

Figure 8. Pearson and Spearman rank correlations between all features and (a) Ra and (b) AvgFc

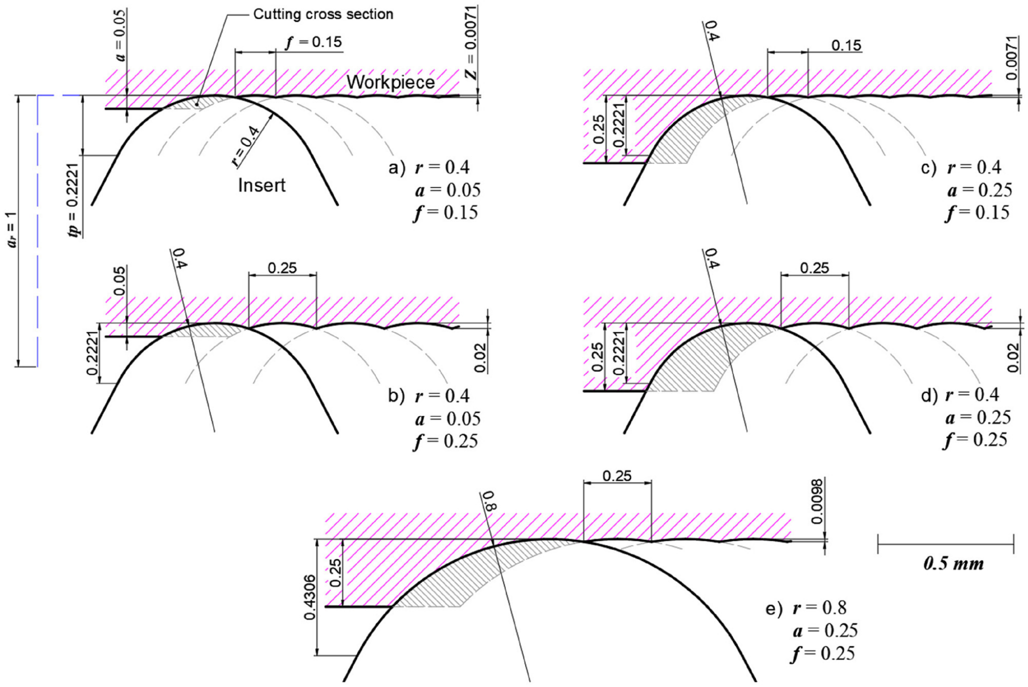
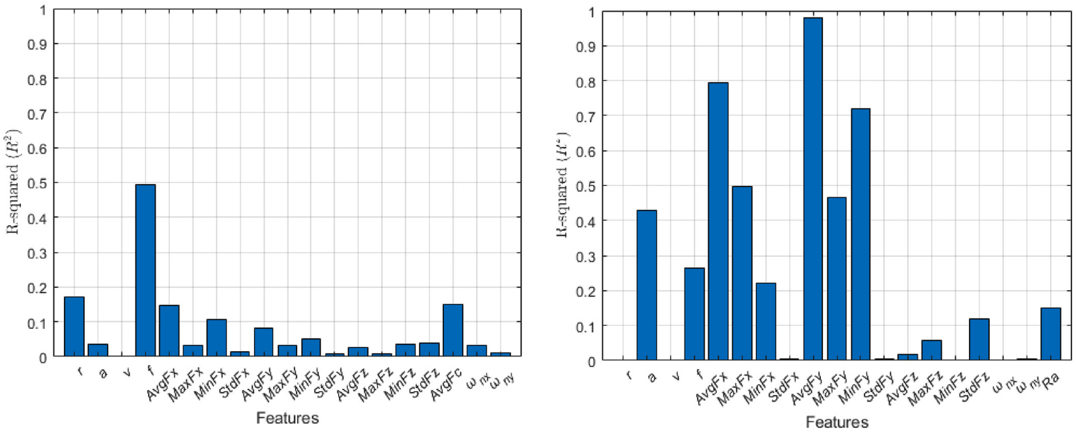
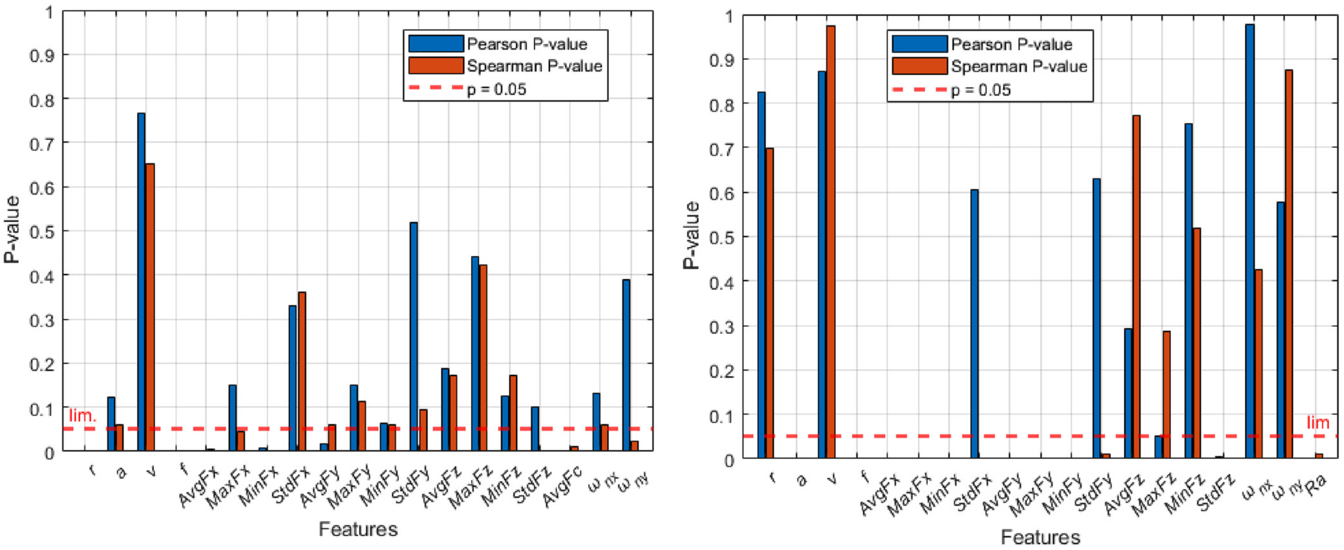


Figure 9. *Theoretical surface roughness models for various combinations of cutting parameters (units in mm)*



(a) *Ra* (b) *AvgFc*

Figure 10. R^2 values for the correlations between each feature and (a) Ra and (b) AvgFc



(a) *Ra* (b) *AvgFc*

Figure 11. p-values for the correlations between each feature and (a) Ra and (b) AvgFc

attempt to consider all measured data. Furthermore, the size of the dataset used for prediction can be significantly reduced by storing only relevant features.

5. CONCLUSIONS

This study proposed a method for processing signals and extracting relevant features from turning process data to inform a subsequent correlation analysis. First, Savitzky–Golay filtering was performed on the raw signals, then signal segmentation was undertaken to eliminate irrelevant information corresponding to roughing operations and noise from non-cutting movements while excluding outlier peaks originating from the initiation and breaking of tool contact. The mean, maximum, minimum, and standard deviation values of the segmented cutting force signals were subsequently obtained as features. The same process was applied to segment the corresponding acceleration signals are also segmented. The FFT was applied to obtain the signals in the frequency domain and determine the natural frequencies corresponding to the largest amplitude vibrations in the x- and y-directions. Finally, Pearson and Spearman rank correlation analyses were performed to determine the process parameter, cutting force, and vibration features most influencing the surface roughness *Ra*. Several features were determined to exert critical influences on *Ra*: the feed (f), corner radius (r), depth of cut (a), and average multidirectional cutting force (*AvgFc*). Among these features, only *AvgFc* directly characterized the current state of the turning operation, although all of these features give relatively high R^2 values for correlation with either *Ra* or *AvgFc*, and the p-values for the correlations between these features and either *Ra* or *AvgFc* were within 0.05. These results indicate that only f, r, a, and *AvgFc* need to be considered to predict the value of *Ra* with relatively high accuracy. Therefore, this study provides valuable insights for researchers and engineers constructing models for predicting *Ra* during turning operations, especially when conducting high-accuracy turning of high-strength 42CrMo4+QT workpieces. Table 8 presented in this study provides a convenient training or testing dataset when considering the recommended features as model inputs. Finally, the insights obtained by this study provide a useful basis for determining efficient data storage and management strategies for intelligent machining applications.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Farah Mohammed Fallata – *Mechanical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0009-0001-0611-6624
Email: s202253940@kfupm.edu.sa

Corresponding Authors

Muhammad Fadhlan Afif – *Graduate Author (PhD student), Mechanical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0009-0001-9807-546X
Email: g202304230@kfupm.edu.sa

Ahmed Aly Diaa Sarhan – *Research Mentor, Mechanical Engineering Department and Interdisciplinary Research Centre for Intelligent Manufacturing & Robotics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia;* 0000-0001-8474-8347
Email: ahsarhan@kfupm.edu.sa;

ACKNOWLEDGEMENTS

The authors thank the King Fahd University of Petroleum and Minerals and the Interdisciplinary Research Center for Intelligent Manufacturing and Robotics for providing financial and technical support. This work was funded by the KFUPM-IM & R-Funded Grant Project Code: INMR2407.

REFERENCES

(1) Pramanik, A. *et al.* Fatigue life of machined components. *Adv Manuf* **5**, 59–76 (2017).

(2) Cui, T. *Precision Machining of Hard-to-Cut Materials: Current Status and Future Directions. IJACSA) International Journal of Advanced Computer Science and Applications* vol. 15 www.ijacsa.thesai.org (2024).

(3) Díaz-Salamanca, D. *et al.* Turning of 42CrMo4+QT under different scenarios: Dataset of machining, roughness and residual stress. *Data Brief* **56**, (2024).

(4) Díaz-Salamanca, D. *et al.* Influence of turning parameters on residual stresses and roughness of 42CrMo4 + QT. *International Journal of Advanced Manufacturing Technology* **134**, 2897–2919 (2024).

(5) Diego, D.-S. *et al.* Assessing the Influence of Surface Roughness and Residual Stresses on the Fatigue Behaviour of 42CRMO4+QT Steel. in *Book of Abstracts of the 3rd International Symposium on Risk Analysis and Safety of Complex Structures* (2025).

(6) VII. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London* **58**, 240–242 (1895).

(7) N F H, M. Pearson Correlation Analysis between Radius and Surface Roughness (Ra) Measurements during Turning Processes. *International Journal of Advanced Trends in Computer Science and Engineering* **9**, 237–241 (2020).

(8) Swiss Steel Group. *Technical Datasheet 42CrMo(S)4 +QT +SH.* www.swissgreensteel.com (2023).

(9) SANDVIK Coromant. DCMX 11 T3 04-WF 4325 CoroTurn®107, insert for turning. *Product details* (2025).

(10) SANDVIK Coromant. DCMX 11 T3 08-WF 4425 CoroTurn®107, insert for turning. *Product Details* (2025).

(11) Walter, G. & Jaroslav, B. *Solving Problems in Scientific Computing Using Maple and Matlab.* (Springer Verlag, 2004).

(12) Schafer, R. What Is a Savitzky-Golay Filter? *IEEE Signal Process Mag* **28**, 111–117 (2011).

(13) Frigo, M. & Johnson, S. G. FFTW: An Adaptive Software Architecture for the FFT. in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing* 1381–1384 (1998).

(14) MATLAB. fft Documentation. *MATLAB Help Center* (2025).

(15) Bistacchi, A., Mittempergher, S., Martinelli, M. & Storti, F. On a new robust workflow for the statistical and spatial analysis of fracture data collected with scanlines (or the importance of stationarity). Preprint at https://doi.org/10.5194/se-2020-83 (2020).

(16) Zar, J. H. Spearman Rank Correlation. in *Encyclopedia of Biostatistics* (Wiley, 2005). doi:10.1002/0470011815.b2a15150.

(17) Feng, C., Li, M., Guo, H., Qiu, B. & Zhang, J. Energy efficiency identification and surface roughness prediction using cutting force signal for computer numerical controlled machine systems. *Sci Rep* **14**, 19004 (2024).

(18) Trinh, V.-L. A Review of the Surface Roughness Prediction Methods in Finishing Machining. *Engineering, Technology & Applied Science Research* **14**, 15297–15304 (2024).

(19) Khlifi, H., Abdellaoui, L. & Bouzid Saï, W. Prediction of Cutting Force and Surface Roughness in Turning Using Machine Learning. in 213–222 (2024). doi:10.1007/978-3-031-42659-9_24.

(20) Pasic, M. *et al.* Prediction and Optimization of Surface Roughness and Cutting Forces in Turning Process Using ANN, SHAP Analysis, and Hybrid MCDM Method. *Applied Sciences* **14**, 11386 (2024).

(21) Boothroyd, G. & Knight, W. A. *Fundamentals of Machining and Machine Tools Third Edition.*

(22) Horváth, R. & Lukács, J. Application of a force model adapted for the precise turning of various metallic materials. *Strojniski Vestnik/Journal of Mechanical Engineering* **63**, 489–500 (2017).

(23) Abellán-Nebot, J. V., Vila Pastor, C. & Siller, H. R. A Review of the Factors Influencing Surface Roughness in Machining and Their Impact on Sustainability. *Sustainability* **16**, 1917 (2024).

(24) Schaffer George H. ARTIFICIAL INTELLIGENCE: A TOOL FOR SMART MANUFACTURING. *American machinist & automated manufacturing* **130**, 83–94 (1986).

KFUPM
Journal of Undergraduate
Research International

Diamond Open Access

Article

Mechanical Engineering/Physics

Comparative Study of Bio-Inspired and Ingenuity Airfoils at Martian Low Reynolds Numbers

Amro Alalfi, Abdulaziz Farooq, Mohamed Takeyeldein*, and Salman Alfifi

Cite https://doi.org/10.64589/juri/214393

Submitted: September 20, 2025 **Revised:** October 8, 2025 **Accepted:** November 18, 2025

ABSTRACT

Aerial flight on Mars introduces major aerodynamic challenges due to its thin atmosphere, resulting in low-Reynolds-number (Re) conditions where conventional airfoils undergo early flow separation and increased drag. This study evaluates and compares the performance of three airfoils: bio-inspired OWL, NASA's CLF5605, and AS6097, at Re = 23000, representative of Martian rotorcraft operations. Two turbulence models, k–ω SST and SST Transition, were tested using steady RANS CFD in ANSYS Fluent with dynamic similarity applied for terrestrial conditions. The SST Transition model showed better agreement with experimental results, particularly in capturing laminar separation bubbles. Among the airfoils, CLF5605 achieved the best aerodynamic efficiency, the OWL airfoil exhibited smooth aerodynamic trends and close validation agreement, and AS6097 produced the highest lift but inconsistent drag behavior.

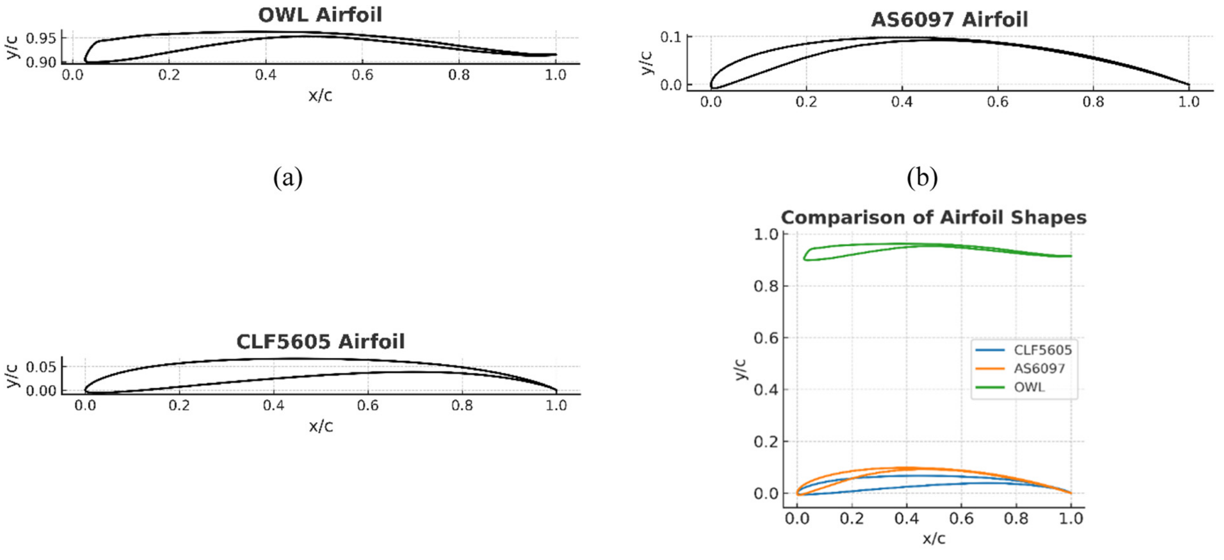
Keywords: martian unmanned aerial vehicle, low reynolds number aerodynamics, bio-inspired airfoils, ingenuity helicopter, computational fluid dynamics simulation

1. INTRODUCTION

In recent years, engineers and scientists have progressively employed advanced technologies to push the boundaries of planetary exploration techniques, enabling missions that are more efficient, autonomous, and robust in extraterrestrial environments. Historically, these missions have been conducted using telescopes, rovers, and satellite technologies[1]. However, using rovers to explore Mars has presented difficulties owing to the rugged and rocky terrain, as well as limited mobility, prompting researchers to explore aerial exploration as a means to improve maneuverability and range of examination. Aerial flight under Martian conditions presents a significant aerodynamic challenge due to the planet's atmospheric composition, which is primarily carbon dioxide and has densities as low as 1% of those on Earth. These conditions result in a markedly reduced speed of sound, confining the airfoils and rotors to operate at low Reynolds number regimes, typically ranging between 10,000 and 25,000[1,2]. In this work, 'low Reynolds number' refers to the Re = 10,000–100,000 range relevant to small-scale rotorcraft and Martian flight conditions.

The low density of Mars reduces the lift generated by the rotor, necessitating higher RPMs to account for the decreased lift per blade area. Low Reynolds numbers increase the drag coefficient because of early flow separation and turbulent reattachment, creating a laminar separation bubble (LSB). Computational studies have shown that conventional airfoils have higher drag values due to early flow separation without reattachment, causing up to a 46% decrease in the maximum lift coefficient at low Reynolds numbers[3], whereas thin, highly cambered, and unconventional airfoils exhibit superior aerodynamic performance compared to thick symmetric airfoils[4]. In 2021, NASA's Ingenuity Mars Helicopter marked the first controlled extraterrestrial flight. This achievement was facilitated by its special airfoil design, which used thin, highly cambered airfoils to improve the lift at low Reynolds numbers and rounded leading edges to delay the separation of bubbles and enhance performance in laminar regimes[5].

Efforts to develop Martian aerial vehicles began with NASA's Mini-Sniffer program in 1977, followed by more agile UAV concepts such as the Y4TR in the 2010s[6]. These developments highlight the growing need for enhanced maneuverability and set the stage for adopting bio-inspired rotor designs, as birds and insects operate under low-Reynolds-number conditions in the Earth's atmosphere, allowing for higher lift-generating capabilities on Mars. Early experimental work[7] compared the aerodynamic performance of corrugated bio-inspired airfoils to conventional airfoils at Re = 34,000 using low-speed wind tunnels and high-resolution particle image velocimetry (PIV). The results indicated that bio-inspired airfoils exhibited no apparent flow separation up to high angles of attack compared to conventional airfoils and therefore achieved better aerodynamic coefficient values. The PIV analysis revealed that the improvement stemmed from delayed laminar separation bubble (LSB) formation and sustained attached flow at higher angles of attack, attributed to the thinner, highly cambered profile of the owl airfoil, which reduced the severity of adverse pressure gradients.



(a) (b) (c) (d)

Figure 1. Airfoil geometries investigated at Martian low Reynolds numbers: (a) OWL airfoil; (b) AS6097; (c) CLF5605; and (d) comparison of the three profiles

Furthermore, multiple experiments[8,9] tested owl-inspired airfoils in a low-speed wind tunnel using force balance measurements across Reynolds numbers from 23,000 to 160,000, reporting up to a 6.7% improvement in the maximum lift compared to the NACA0012 airfoil. Several computational investigations[10–12] using both two-dimensional laminar computational fluid dynamics (CFD) and three-dimensional large-eddy simulation (LES) have consistently shown that owl-like airfoil geometries improve aerodynamic performance at low-Reynolds-number relative to canonical and other low-Re airfoils. Together, these studies demonstrate a common phenomenon: owl geometry produces strong suction peaks on the suction surface, controls the location and size of laminar separation bubbles (LSBs), and promotes pressure recovery, which increases the lift and reduces the drag over a broad range of angles of attack. When referenced collectively, these simulations provide robust evidence that tailoring camber, leading-edge radius, and upper-surface shape can deliberately manipulate LSB behavior to delay stall and enhance lift-to-drag ratio (L/D) for $\mathrm{Re} \approx 10^4$.

Given the demonstrated superiority of bio-inspired airfoils at low Reynolds number flight regimes and the lack of research on their application in Martian conditions, this study aims to computationally evaluate the effectiveness of bio-inspired airfoils for Martian flights and benchmark their performance against NASA's Ingenuity airfoil. This research seeks to answer the following primary question: Can the flow control mechanisms inherent in bio-inspired airfoils, such as delayed separation, translate effectively to the specific compressibility and density constraints of the Martian atmosphere to outperform current state-of-the-art designs?

2. METHODOLOGY

2.1. Airfoil Selection. Three airfoils were used in this study, each selected for a specific purpose. NASA's CLF5605 airfoil was included as a benchmark to enable a comparison with the other test airfoils; its geometry was obtained from NASA documentation[5]. The Owl airfoil was chosen for its proven high-lift performance and delayed stall characteristics under low-Reynolds-number conditions, using the geometry defined by Liu et al.[13]. Finally, the AS6097 airfoil was selected as a bird-inspired profile featuring a favorable camber and leading-edge geometry to investigate the computational performance trends. Its geometry was sourced from Amanda and Selig[14].

2.2. Dynamic Similarity Setup. To replicate Martian low-Reynolds-number aerodynamics without introducing numerical instability, we employed the principle of dynamic similarity. The Martian operating Reynolds number ($\mathrm{Re_M}$) was taken from the available literature: $\mathrm{Re_M} = 23,000$, corresponding to Ingenuity's characteristic chord $c = 10$ cm[5]. Direct simulation using Martian atmospheric density ($\rho_\mathrm{M} \approx 0.020\ \mathrm{kg/m^3}$) and viscosity ($\mu_\mathrm{M} \approx 1.3 \times 10^{-5}\ \mathrm{kg/(m{\cdot}s)}$) would result in prohibitively low freestream velocities and solver stiffness. Instead, all the CFD tests were performed under terrestrial air conditions, and the freestream velocity (V_E) was calculated as follows:

$$Re_E = \frac{\rho_E V_E c}{\mu_E} = Re_M \tag{1}$$

Because airfoil aerodynamics at low Mach numbers are primarily governed by the Reynolds number, this approach preserves the relevant dimensionless flow physics while avoiding the numerical challenge of simulating the Martian atmospheric conditions. To achieve a Reynolds number of 23,000 on a 10 cm chord airfoil in terrestrial conditions, assuming $\rho_\mathrm{E} = 1.225\ \mathrm{kg/m}^3$ and $\mu_\mathrm{E} = 1.81 \times 10^{-5}\ \mathrm{kg/(m{\cdot}s)}$, using Equation 1, we obtain $V_\mathrm{E} \approx 3.4$ m/s. Since chord length is preserved and the Mach number remains well below 0.3 (M≈0.01M), compressibility effects are negligible, and Reynolds number governs the aerodynamic similarity. Consequently, the Earth-based CFD simulations accurately reproduce the Martian low-Reynolds-number flow physics. Importantly, this approach also enables practical wind tunnel replication under the same velocity and chord length, allowing direct experimental validation with smoke visualization or PIV.

2.3. CFD Solver & Domain.

To model the aerodynamics of the airfoils, a steady-state CFD solver was employed in ANSYS^{TM} Fluent software version 2022 R2; all numerical simulations were performed on a high-performance workstation equipped with an Intel Core i7 processor and 32 GB of RAM to ensure efficient convergence. A second-order upwind discretization scheme was used with SIMPLE pressure–velocity coupling, which is appropriate for steady incompressible flow, to enhance solver stability. Convergence was assessed by monitoring residuals (target 10^{-5}) and integrated force coefficients. Additionally, the lift and drag coefficients (C_L,C_D, respectively) were continuously monitored during each run to ensure stable convergence.

Turbulence modeling was necessary because flows at Re=23,000 lie within the laminar–transitional regime. Two models were tested: the k–ω model, selected for its robustness in predicting near-wall behavior and low-Reynolds-number flows, and the SST (Shear Stress Transport) four-equation model, adopted for its improved ability to resolve separation bubbles and transitional phenomena, which are critical at low Reynolds numbers.

The governing equations solved were the steady, incompressible Navier–Stokes equations with Reynolds-averaged turbulence closure:

$$\nabla \cdot \vec{u} = 0 \tag{2}$$

$$\rho(\vec{u} \cdot \nabla)\vec{u} = -\nabla p + \mu\nabla^2\vec{u} + \overrightarrow{F_{turb}} \tag{3}$$

where $\mathbf{F_{turb}}$ is the modeled Reynolds stress contribution from the turbulence models.

A C-type computational domain was adopted, with Farfield boundaries sufficiently distant from the airfoil to avoid blockage. The angle of attack (AoA) was varied from -2° to 12° in increments of 2°. The inlet conditions were set at V_E =3.4 m/s (from the dynamic similarity analysis), whereas the outlet was prescribed with zero static gauge pressure. The airfoil surface was modelled as a no-slip wall.

An unstructured mesh (Figure 2) with approximately 150,000 cells was generated, with strong refinement near the airfoil wall to resolve laminar separation bubbles, incorporating at least 30 inflation layers and targeting a near-wall resolution of $y^+ \approx 10^{-5}$. Mesh quality was maintained with skewness <0.25 and orthogonal quality >0.95, ensuring solver robustness. Table 1 summarizes the CFD setup and settings used in this study. Since this study relies entirely on computational fluid dynamics simulations of inanimate airfoil geometries and involves no human or animal subjects, specific ethical approval or informed consent was not required.

3. RESULTS AND DISCUSSION

3.1. Results.

3.1.1. Aerodynamic coefficients.

Figure 3 compares the aerodynamic performance of the AS6097, OWL, and CLF5605 airfoils in terms of lift coefficient (C_L), drag coefficient (C_D), and lift-to-drag ratio (C_L/C_D) as predicted by the k–ω SST and SST Transition turbulence models over a range of angles of attack (AOA).

For the AS6097 airfoil, both turbulence models showed significant overlap in the prediction of the lift coefficient up to an AOA of 6°, after which some differences were observed past that point. The lift coefficient peaked around an AOA of 10° before exhibiting stall behavior. The drag calculation coefficient exhibited significant fluctuations in both models, with a sharp decrease at an AOA of 6°. Specifically, the k-ω SST model predicted a

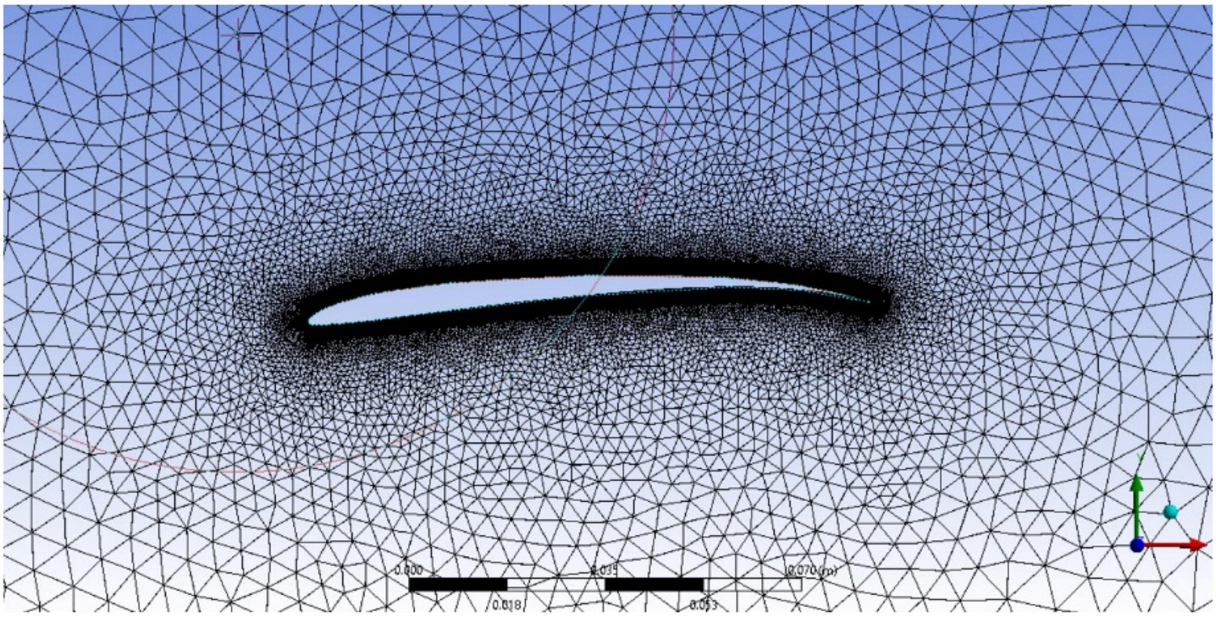


Figure 2. Unstructured triangular mesh with inflation layers around the airfoil surface

Table 1. *CFD setup*

Parameter	Specification
Solver	ANSYS Fluent, pressure-based, steady-state
Governing equations	Incompressible URANS (continuity + momentum)
Discretization	Second-order upwind for momentum & turbulence equations
Pressure–velocity coupling	SIMPLE algorithm
Convergence criteria	Residuals $<10^{-5}$; stable C_L, C_D, C_L/C_D monitoring
Turbulence models	1. Standard k–ω (robust near-wall, adverse pressure gradients) 2. Transition SST (for laminar–transitional flows)
Boundary conditions	Inlet velocity V_E =3.4 m/s; outlet static pressure = 0 Pa; airfoil = no-slip wall
Angle of Attack range	−2° to 12°, step = 2°
Meshing strategy	Unstructured grid with refinement near airfoil; ≥ 30 inflation layers; $y^+ \approx 10^{-5}$
Grid independence	3 progressively refined meshes tested; outputs compared for independence
Validation	CLF5605 and OWL vs. published data; AS6097 vs. XFOIL at the same Re

C_D of approximately 0.08, while the SST Transition model predicted a significantly lower C_D of approximately 0.06 at this angle of attack. This sharp decrease in drag, followed by a subsequent increase, led to a spike in C_L/C_D at the same AOA. While AS6097 achieved the highest maximum lift coefficient (CLmax) its drag performance was inconsistent, an effect attributed to the formation of laminar separation bubbles.

For the OWL airfoil, both models exhibited similar aerodynamic coefficients throughout the AOA range. The lift coefficient consistently increased up to the highest measured AOA of 12°, when it began to show signs of stall. A small difference in C_L/C_D was noted between AOA of 6° and 10°. A maximum C_L of approximately 1.2 is achieved at an AOA of 12°. Overall, the OWL airfoil showed the smoothest and most predictable trends, making it suitable for stable applications.

The CLF5605 airfoil also exhibited consistent aerodynamic coefficients in both models within the range of AOAs, with only minor differences. Its maximum lift coefficient was slightly lower than that of the other two airfoils, peaking at approximately 1.1 near AOA of 12°. Both turbulence models showed consistent flow behavior and aerodynamic coefficient predictions with some differences in magnitude, most notably in the calculation of the drag force on AS6097. AS6097 had the highest lift force, peaking at a C_L of approximately 1.4, whereas CLF5605 had the best aerodynamic efficiency, reaching a peak C_L/C_D of >20.

Overall, the SST Transition model provides a more conservative drag prediction and slightly different stall behavior, while the k–ω SST model predicted higher aerodynamic efficiency. Among the airfoils tested, CLF5605 consistently exhibited a superior lift-to-drag performance, highlighting its aerodynamic advantages under low-Reynolds-number conditions.

3.1.2. Grid independence study.

To validate these results, a grid convergence study was conducted to ensure that the solver results were not biased at higher mesh densities. This was done by selecting a representative case, running the solver at different inflation layers, and then comparing the resulting aerodynamic coefficients with different pressure-velocity coupling schemes. The CLF5606 airfoil at an angle of attack of 6° was selected, and the simulations were performed on inflation layers ranging from 30 to 70. The results are shown in Table 2.

The SIMPLE solver was employed for all simulations to improve solver stability. The grid-independence study demonstrated excellent convergence with this solver, as shown in Table 2. For completeness, the coupled solver was also tested and showed similar convergence behavior, with the final values differing by less than 0.4% from the SIMPLE solver results at 70 inflation layers, validating the overall approach.

3.1.3. Validation.

The comparison of the lift, drag, and lift-to-drag ratio for the OWL airfoil at Re=23000 in Figure 4 demonstrates that both turbulence models captured the expected aerodynamic behavior and showed good agreement with the reference study by Anyoji (2018). The k–ω SST model slightly overpredicted lift at low angles of attack due to its fully turbulent assumption, while the SST Transition model provides a more realistic pre-stall slope by accounting for laminar–turbulent transition. Both models, however, underpredicted the experimentally observed maximum lift. In drag prediction, both approaches underestimate the values across the range of angles; however, the SST Transition model yields slightly higher drag values that more closely align with the reference data, particularly in the pre-stall regime, where laminar separation bubbles dominate. Consequently, the computed aerodynamic efficiency (C_L/C_D) was somewhat higher than that of the reference data; however, the overall variation with the AOA was well reproduced. It is worth noting that part of the discrepancy may stem from slight geometric differences between the Liu et al. owl-like airfoil configuration used in this study and the test model used by Anyoji et al., as highlighted in their study. Nevertheless, the results demonstrate very good overall agreement, with the SST Transition model showing improved fidelity at this low Reynolds number, while the standard k–ω SST provides a sharper stall prediction.

Figure 5 presents the validation of the CFD predictions against the experimental data for the CLF5605 airfoil at Re = 23,000 under two-dimensional flow conditions. The experimental datasets were sourced from wind tunnel campaigns conducted in Japan and the United States [15]. The computational results were

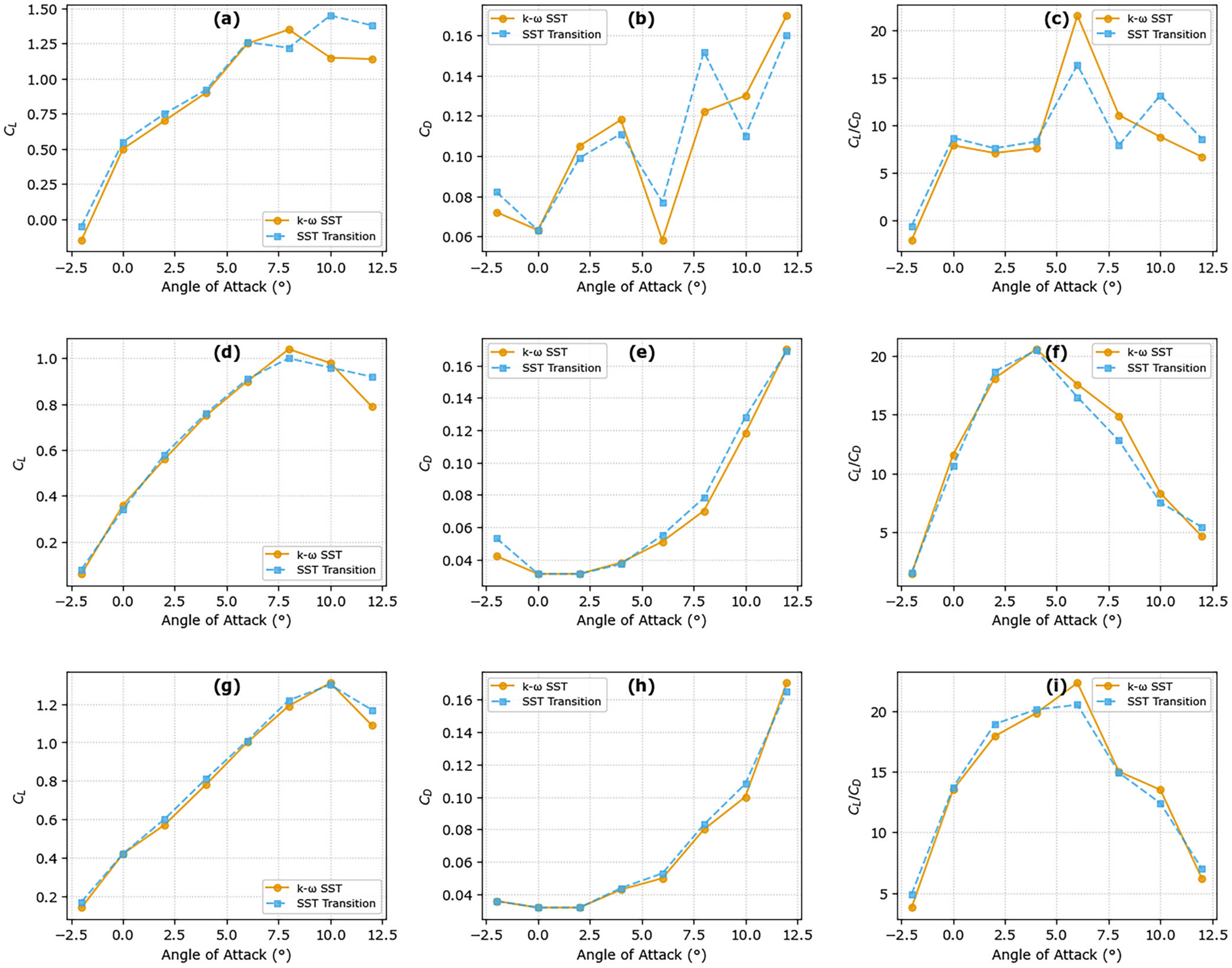


Figure 3. Comparative aerodynamic performance of AS6097, OWL, and CLF5605 airfoils under k–ω SST and SST Transition models. (a–c) AS6097: lift, drag, and lift-to-drag ratio vs. angle of attack.(d–f) OWL: lift, drag, and lift-to-drag ratio vs. angle of attack.(g–i) CLF5605: lift, drag, and lift-to-drag ratio vs. angle of attack

obtained using steady RANS simulations with both the k–ω SST model and the SST four-equation transition model.

The lift curves showed good agreement across all datasets, with the CFD models reproducing both the slope and stall onset. The maximum lift coefficient predicted by CFD ($C_{Lmax} \approx 1.32$–1.34) fell within the range of the experimental data, though slightly underpredicting the highest values observed in the Japanese dataset[15]. The minimum drag coefficient from the simulations is $C_{Dmin} \approx 0.031$–0.032, somewhat lower than the experimental drag polars, leading to an optimistic prediction of aerodynamic efficiency. The maximum lift-to-drag ratios are $C_L/C_{Dmax} \approx 22.0$ (SST) and 20.4 (SST transition), compared with lower values in the experiments, which reflects the RANS tendency to underpredict viscous drag at low Reynolds numbers.

Discrepancies in drag are expected because of the inability of RANS closures to accurately capture laminar separation bubbles, transition onsets, and small-scale unsteadiness, which dominate low-Reynolds airfoil aerodynamics. This effect is especially

Table 2. *Grid Independence Study*

	Coupled			SIMPLE		
Inflation layers	C_L	C_D	C_L/C_D	C_L	C_D	C_L/C_D
30	1.1000	0.05400	20.3704	1.0705	0.0475	22.5368
40	1.1120	0.05410	20.5545	1.1117	0.0502	22.1665
55	1.1266	0.05390	20.9017	1.1257	0.0540	20.8463
60	1.1230	0.05400	20.7963	1.1299	0.0538	21.0173
70	1.1322	0.05320	21.2820	1.3200	0.0532	21.1984

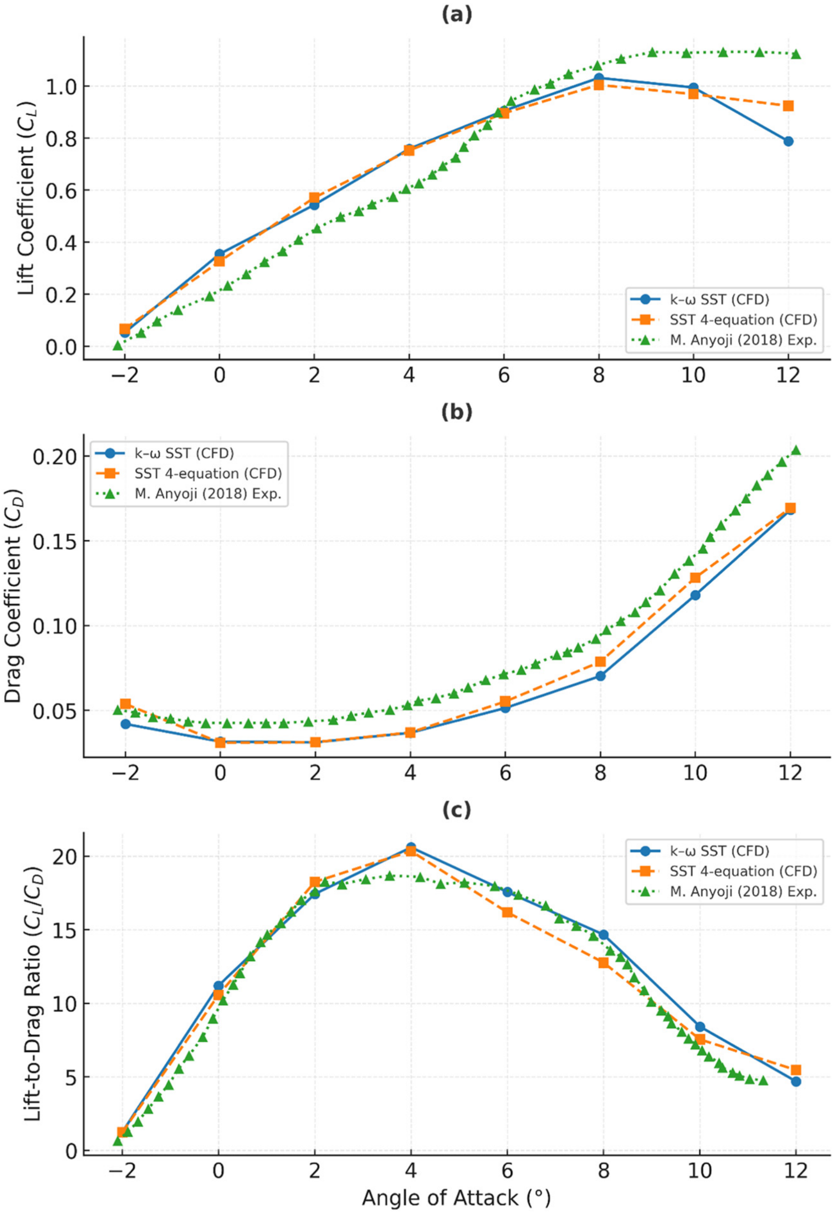


Figure 4. Aerodynamic coefficients for the OWL airfoil using $k - \omega$ SST (CFD), SST 4-equation (CFD), and M. Anyoji (2018) experimental data at Re = 23000. (a) Lift coefficient (C) vs. angle of attack.(b) Drag coefficient (C_d) vs. angle of attack. (c) Lift-to-drag ratio (C_l/C_d) vs. angle of attack

pronounced in the drag, as even minor errors in modeling the transition location can significantly alter the boundary-layer development. Nevertheless, the CFD results reproduced the main aerodynamic trends and provided credible predictions of the lift, stall behavior, and overall performance envelope, validating their use for further analysis. The primary objectives of this study—to computationally benchmark bio-inspired airfoils (OWL, AS6097) against the Ingenuity airfoil (CLF5605) and validate the CFD models against experimental data—were successfully met. The results provide a clear comparison of aerodynamic performance and modeling fidelity under the specified low-Reynolds-number conditions.

3.2. Discussion. This study drew multiple comparisons between different airfoil designs, turbulence models, and validation datasets to investigate the aerodynamic behavior under Martian Reynolds number conditions. Having presented the detailed results, it is now appropriate to take an overall view to draw broader conclusions and consolidate our understanding of the problem.

Two turbulence models were primarily employed to evaluate the predictive capability. The k–ω SST model performed well in predicting the lift coefficient (C_L) in the pre-stall regime but consistently overpredicted the drag coefficient (C_D) and deviated from complementary computational reference results at higher angles of attack. In contrast, the SST Transition model produced results more consistent with the validation datasets and demonstrated better capability in predicting turbulent reattachment in the pre-stall region. At low angles of attack, both models produced highly similar results, but as the angle of attack increased, the k–ω SST predictions began to deviate, reinforcing the SST Transition model as the more reliable choice at these Reynolds number conditions.

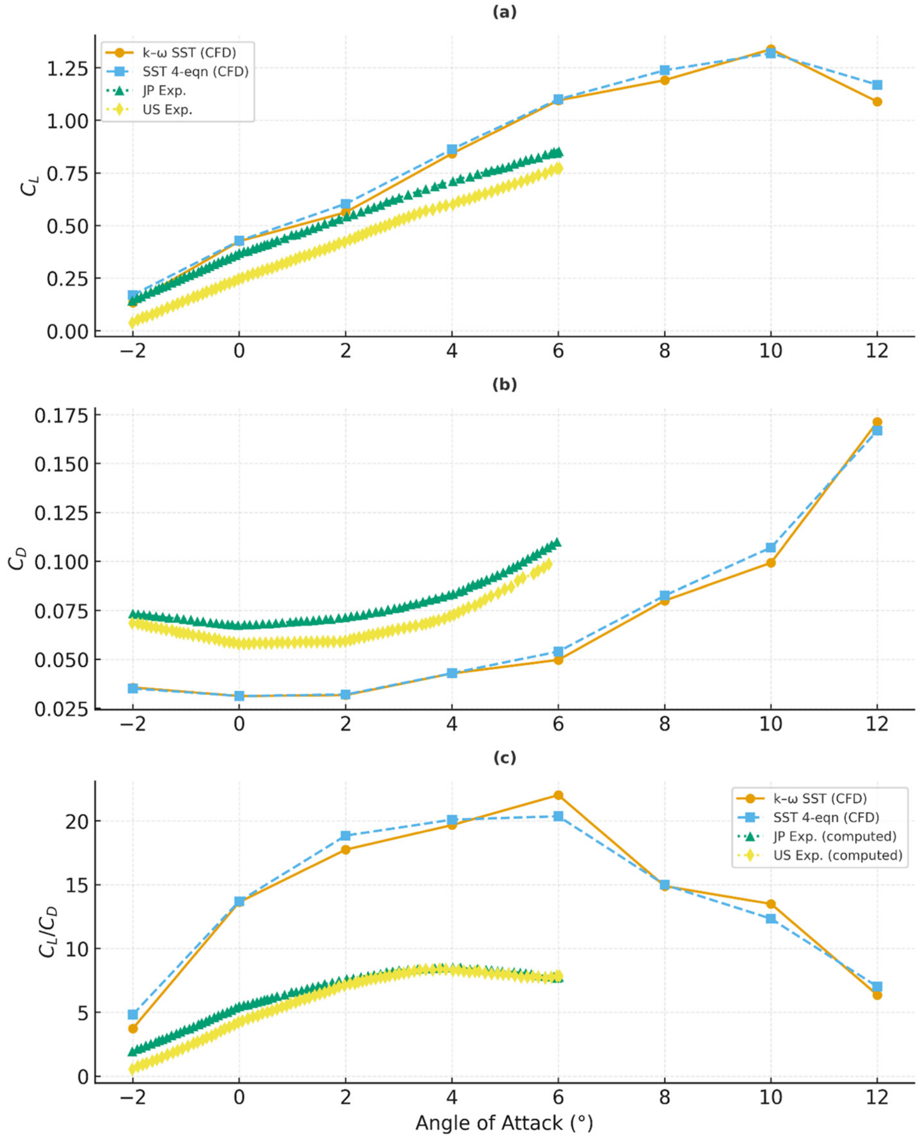


Figure 5. Aerodynamic coefficients using k–ω SST (CFD), SST 4-equation (CFD), JP Exp., and US Exp. Datasets at Re = 23000. (a) Lift coefficient (C_L) vs. angle of attack. (b) Drag coefficient (C_D) vs. angle of attack. (c) Lift-to-drag ratio (C_L/C_D) vs. angle of attack

Three airfoil configurations were tested: an OWL-like airfoil, NASA CLF5605, and AS6097. Among them, the CLF5605 exhibited the most aerodynamic efficiency, confirming its suitability and explaining its selection for NASA's Ingenuity Mars helicopter. The OWL airfoil also showed a smooth and predictable aerodynamic response and closely matched the validation datasets, suggesting its potential for stable and efficient low-Reynolds-number applications. AS6097 achieved the highest maximum lift coefficient (C_{Lmax}) among the three but produced inconsistent drag values, likely due to the formation of laminar separation bubbles. This behavior suggests the formation and reattachment of a laminar separation bubble, a phenomenon that the SST Transition model is specifically designed to capture. This airfoil could not be validated against literature data, as no comparable CFD or experimental studies were found, indicating that its performance requires further investigation.

The grid independence study confirmed that the selected mesh was sufficiently refined to resolve the aerodynamic features without introducing grid-related artifacts, ensuring that the observed differences stemmed from physical modeling rather than numerical errors.

For validation, the CLF5605 case provides a detailed comparison with wind tunnel experiments conducted in Japan and the United States. Both turbulence models captured the lift behavior and stall onset accurately. The simulations predicted $C_{Lmax} \approx 1.34$ (k–ω SST) and 1.32 (SST Transition), values within the experimental range. The minimum drag coefficients were predicted as $C_{Dmin} \approx 0.031–0.032$, slightly lower than measured experimentally, leading to optimistic lift-to-drag ratios ($C_L/C_{Dmax} \approx 22.0$ for SST and 20.4 for SST Transition). The Japanese and U.S. experimental datasets exhibited minor differences, reflecting the sensitivity of low-Reynolds-number airfoil aerodynamics to wind tunnel conditions such as turbulence intensity, surface roughness, and wall effects. The fact that CFD predictions fell within this experimental range adds confidence to the overall results.

Collectively, these findings highlight several important points. At Martian Reynolds numbers, small discrepancies in drag prediction lead to disproportionately large variations in aerodynamic efficiency (CL/CD), which is critical for vehicle performance. Although steady RANS models provide reliable predictions of the lift and stall onset, their drag prediction ability is limited by the simplified treatment of laminar separation bubbles, transition onset, and unsteady vortex dynamics. The SST Transition model reduces some of these errors by accounting for transition effects; however, higher-fidelity methods may be required for more accurate efficiency predictions.

As noted in the results, the numerical behavior of the airfoils at Re ≈ 23,000 suggests the formation and reattachment of a laminar separation bubble (LSB). This phenomenon is a critical point of uncertainty for the present study. The quantitative CD and CL values derived from the 2D RANS models must be interpreted with a clear understanding of the known limitations of these models in the low–Reynolds-number, transitional regime. The k–omega SST model, in particular, is known to struggle with this flow physics.

A recent high-fidelity benchmark study[16] directly compared 2D RANS k–omega SST simulations against experimental Particle Image Velocimetry (PIV) data for an airfoil with complex flow separation. The findings were stark: the RANS model failed to capture the true extent of the separation, resulting in a 58% error in the predicted drag coefficient (CD) and a 78% error in the lift coefficient (CL). The study diagnosed this failure as the RANS model's tendency to overestimate turbulent mixing near the wall, which artificially energized the boundary layer and prevented the physical separation that the PIV data clearly measured.

This finding is directly relevant to our Martian airfoil investigation. It implies that our RANS-predicted CD values are likely under-predicted, and the true size and behavior of the LSBs are not fully resolved.

For design applications, the results suggest that CLF5605 and OWL airfoils are the most promising candidates for Martian low-Reynolds-number flights. CLF5605 stands out for its proven efficiency and prior successful application, while the OWL configuration shows potential as a stable and predictable alternative. AS6097, although capable of producing a high lift, requires further investigation due to its inconsistent drag behavior.

Finally, this study emphasizes the importance of turbulence modeling accuracy in low-Reynolds-number aerodynamic. Future work could extend this research by employing large-eddy simulations (LES) or hybrid RANS/LES methods to better capture the laminar separation and transition phenomena. Additionally, quasi-2D or 3D wind tunnel tests with transition location measurements would provide valuable benchmarks for further validation. The study also assumed a rigid airfoil and did not model aeroelastic effects, such as the blade flapping or twisting that could occur under aerodynamic loads on a flexible Martian rotor. Together, these efforts would strengthen the predictive capability and support the development of efficient Martian aerial platforms.

4. CONCLUSIONS

This study explored the aerodynamic performance of the OWL, CLF5605, and AS6097 airfoils under Martian conditions using CFD to compare the performance of the k–ω SST and SST Transition turbulence models. Several key conclusions were drawn.

1. Modeling Fidelity and Limitations: This work confirms that selecting a transition-sensitive turbulence model (SST Transition) is critical for capturing low-Re flow physics. However, the study also reinforces that even advanced RANS models underpredict drag. This discrepancy, attributed to RANS limitations in resolving laminar separation bubbles, confirms that RANS is a reliable tool for predicting lift and stall but not for final drag and performance-efficiency calculations.Future work: High-fidelity CFD methods (LES and hybrid RANS–LES) and quasi-2D or 3D wind tunnel studies with transition measurements are recommended to refine the predictive accuracy and support rotorcraft design optimization for Martian environments.
2. Airfoil Performance Insights: While the benchmark CLF5605 was validated as the most aerodynamically efficient, the bio-inspired OWL airfoil proved to be a highly stable and reliable alternative. The AS6097, conversely, demonstrated that high lift can be negated by LSB-induced drag inconsistencies, proving that bio-inspired geometry alone is not a guarantee of superior performance.

Overall, this study's primary contribution is twofold. From a practical standpoint, it provides aerospace engineers with a direct performance comparison, validating the CLF5605's efficiency while qualifying the OWL airfoil as a viable alternative for future Martian rotorcraft. From a theoretical standpoint, the findings underscore that at Martian Reynolds numbers ($Re \approx 2.3 \times 10^4$), the mitigation of laminar separation bubbles becomes the dominant design driver, often proving more critical for aerodynamic efficiency than high-lift profiling alone.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Authors

Amro Alalfi – *Aerospace Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia;* 0009-0005-9299-5504
Email: s202254140@kfupm.edu.sa

Abdulaziz Farooq – *Aerospace Engineering, King Fahd University of Petroleum and Minerals Dhahran 31261, Saudi Arabia;* 0009-0005-5519-2833
Email: s202268420@kfupm.edu.sa

Corresponding Author

Mohamed Takeyeldein – *Research Mentor, Interdisciplinary Research Center for Aviation and Space Exploration, King Fahd University of Petroleum and Minerals Dhahran 31261, Saudi Arabia;* 0000-0002-6888-0783
Email: mohamed.elsayed@kfupm.edu.sa

Author

Salman Alfifi – *Aerospace Engineering, King Fahd University of Petroleum and Minerals Dhahran 31261, Saudi Arabia* 0009-0001-7238-0669
Email: fifis@kfupm.edu.sa

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the King Fahd University of Petroleum and Minerals (KFUPM) and the Department of Aerospace Engineering for their support.

REFERENCES

(1) Withrow, S. *et al.* Mars Science Helicopter Conceptual Design. in *ASCEND 2020* (American Institute of Aeronautics and Astronautics, Reston, Virginia, 2020). doi:10.2514/6.2020-4029.

(2) Giacomini, E. & Westerberg, L.-G. Rotorcraft Airfoil Performance in Martian Environment. *Aerospace* **11**, (2024).

(3) Winslow, J., Otsuka, H., Govindarajan, B. & Chopra, I. Basic Understanding of Airfoil Characteristics at Low Reynolds Numbers (104–105). *J. Aircr.* **55**, 1050–1061 (2018).

(4) Menon, K. & Mittal, R. Aerodynamic Characteristics of Canonical Airfoils at Low Reynolds Numbers. *AIAA J.* **58**, 977–980 (2020).

(5) Koning, W. J. F. & Dominguez, M. *Mars Helicopter Ingenuity Rotor Geometry*. http://www.sti.nasa.gov (2024).

(6) Hassanalian, M., Rice, D. & Abdelkefi, A. Evolution of space drones for planetary exploration: A review. *Prog. Aerosp. Sci.* **97**, 61–105 (2018).

(7) Hu, H. & Tamai, M. Bioinspired Corrugated Airfoil at Low Reynolds Numbers. *J. Aircr.* **45**, 2068–2077 (2008).

(8) Anyoji, M., Wakui, S., Hamada, D. & Aono, H. Experimental Study of Owl-Like Airfoil Aerodynamics at Low Reynolds Numbers. *J. Flow Control. Meas. & Vis.* **06**, 185–197 (2018).

(9) Kazemi, M. & Mani, M. Owl airfoil aerodynamic noise sources and performance compared to hawk and NACA0012 airfoils for low Reynolds applications. *Sci. Rep.* **15**, 23261 (2025).

(10) KONDO, K. *et al.* Analysis of Owl-like Airfoil Aerodynamics at Low Reynolds Number Flow. *Trans. JAPAN Soc. Aeronaut. Sp. Sci. Aerosp. Technol. JAPAN* **12**, Tk_35-Tk_40 (2014).

(11) AONO, H. *et al.* Aerodynamics of Owl-like Wing Model at Low Reynolds Numbers. *Trans. Jpn. Soc. Aeronaut. Space Sci.* **63**, 8–17 (2020).

(12) Kondo, K. *et al.* Large-Eddy Simulations of Owl-Like Wing Under Low Reynolds Number Conditions. in (American Society of Mechanical Engineers, 2013). doi:10.1115/FEDSM2013-16377.

(13) Liu, T., Kuykendoll, K., Rhew, R. & Jones, S. Avian Wing Geometry and Kinematics. *AIAA J.* **44**, 954–963 (2006).

(14) Ananda, G. K. & Selig, M. S. Design of Bird-Like Airfoils. in *2018 AIAA Aerospace Sciences Meeting* (American Institute of Aeronautics and Astronautics, Reston, Virginia, 2018). doi:10.2514/6.2018-0310.

(15) Koning, W. J. F. *et al. Experimental Results for Mars Rotorcraft Airfoils (Roamx-0201 and Clf5605) at Low Reynolds Number and Compressible Flow in a Mars Wind Tunnel.* https://ntrs.nasa.gov/citations/20240004230 (2024).

(16) Gutierrez, R., Zamponi, R., Ragni, D., Llorente, E. & Aranguren, P. On the extension of kωSST corrections to predict flow separation on thick airfoils with leading-edge roughness. *Wind Energy* **26**, 650–667 (2023).

KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Impact of Reservoir Conditions on CO_2 Miscibility in Crude Oils: Implications for CO_2 Sequestration and Enhanced Oil Recovery Projects

Noor Alsadig and Mohamed Gamal Rezk*

Cite https://doi.org/10.64589/juri/214392

Submitted: May 29, 2025 **Revised:** October 4, 2025 **Accepted:** November 18, 2025

ABSTRACT

CO_2 injections has emerged as a pivotal technique in the petroleum industry and represent a cost-effective option for enhancing oil recovery and improving environmental sustainability. The mass transfer of CO_2 into reservoir fluids is the primary mechanism responsible for CO_2 solubility trapping and oil recovery enhancement. The minimum miscibility pressure between CO_2 and crude oil determines the performance of CO_2-enhanced oil recovery (EOR) and storage projects. This study investigated the influence of reservoir conditions on the CO_2-oil miscibility during CO_2 injection into oil reservoirs. Pendant drop tests were conducted, and the CO_2-oil interfacial tension (IFT) was measured at different pressures and two temperatures. The minimum miscibility pressure (MMP) and first contact miscibility pressure (FCM) between CO_2 and crude oil were estimated from the measured IFT values using the vanishing interfacial tension technique. The measured MMP increased with the reservoir temperature from 1175 psi at T = 27 °C to 2007 psi at T = 70 °C. Similarly, the FCM pressure increased significantly with the reservoir temperature. This study provides valuable guidance for screening the suitability of oil reservoirs for CO_2 EOR and storage.

Keywords: interfacial tension, minimum miscibility pressure, carbon dioxide, pendant drop, fluid-fluid system, miscibility

1. INTRODUCTION

Carbon capture, utilization, and storage (CCUS) is emerging as a crucial tool for reducing carbon emissions while maintaining to support global energy production. CCUS offers a pathway for decarbonizing the oil and industrial sectors. Although hydrocarbon extraction represents a main source of energy, it is associated with high CO_2 emissions[2]. Hence, carbon storage is essential for reducing CO_2 emissions while meeting energy demand. CO_2 enhanced oil recovery (EOR) is considered a vital solution because it aligns with both environmental and economic priorities by reducing emissions and boosting oil recovery from mature fields[3].

Oil fields typically undergo three main production phases: primary, secondary, and tertiary recovery, each of which is designed to maximize the extraction of hydrocarbon as the reservoir conditions change[4]. In the primary recovery phase, oil naturally flows into the production well under the influence of initial reservoir pressure. During hydrocarbon production, the reservoir pressure declines, which reduces recovery. Thereafter, secondary and tertiary recovery techniques are employed, which usually involve the injection of fluids such as water or natural gas to repressurize or maintain the reservoir pressure. These methods can recover an additional 5–20% of the original oil in place (OOIP)[5]. CO_2-EOR is considered a promising tertiary recovery method owing to its ability to improve oil recovery[6]. A schematic of the CO_2 injection process is shown in Fig. 1. During CO_2 injection into oil reservoirs, CO_2 dissolves in crude oil, which reduces oil viscosity, decreases CO_2–oil interfacial tension, and increases oil volume due to swelling. Supercritical CO_2 vaporizes light and intermediate oil components at high pressures.[7]. This technique can aid in accessing trapped oil in pore spaces that the conventional method cannot reach, and further extend the field productivity life span[8].

Residual oil saturation is the fraction of oil that remains trapped in a reservoir and cannot be extracted using conventional recovery methods; however, it can be extracted by applying methods such as CO_2 gas injection[8]. This oil is retained in the reservoir because of the capillary force of the porous formation. During CO_2 injection, residual oil can be mobilized as CO_2 reduces the interfacial tension (IFT) and overcomes the capillary force, thereby improving the miscibility with crude oil and enhancing oil recovery[10]. The minimum miscibility pressure (MMP) is the minimum pressure at which the reservoir and injected fluids reach miscibility[11]. It is considered a critical parameter in EOR operations. At pressures above the MMP, oil and CO_2 gas mix and form a homogeneous phase by reducing the IFT between the two-phase boundaries[11]. Interfacial tension can be described as the force at the boundary (interface) between two immiscible fluids (e.g., oil and water or CO_2 and oil)[12]. The IFT is directly related to the capillary number, which is related to

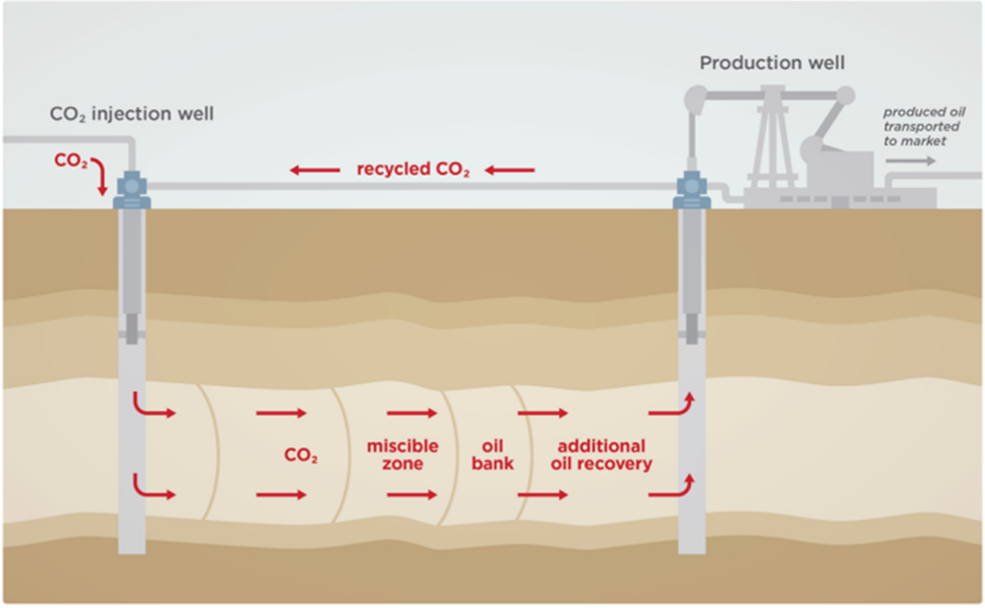


Figure 1. Schematic diagram of the CO_2-EOR injection process[9]

the viscous and capillary forces (Eq. 1)()[13]. The capillary number is inversely proportional to the IFT (σ) and directly proportional to the velocity (υ) and viscosity (μ). This indicates that the higher the IFT between the two fluids, the higher the residual oil saturation. In contrast, when the IFT between fluids is low, it allows for better miscibility, resulting in efficient hydrocarbon mobilization[7].

$$\text{Capillary No.} = \frac{\upsilon . \mu}{\sigma} \quad (1)$$

Cao and Gu[14] managed to investigate the effect of temperature on the phase behavior of a CO_2–light oil system through a series of PVT experiments. They determined the vaporization pressure of CO_2, the onset pressure for asphaltene precipitation caused by CO_2–oil interactions, and the MMP. Their results showed that the MMP was more significantly influenced by temperature than by the vaporization pressure of CO_2. Rezk et al.[7] estimated the MMP and FCM pressure for a CO_2-light oil system by measuring the IFT at various pressures. Their results demonstrated the effect of system pressure on the properties of crude oil. Li et al.[15] and Huang et al.[16] estimated the MMP of CO_2-oil systems using molecular dynamics and machine learning techniques, respectively. Their models yielded satisfactory results when validated using experimental data and empirical models. Although several studies have measured and estimated the MMP between CO_2 and crude oil, additional experimental data are required, particularly for CO_2 interactions with medium- and heavy-viscosity crude oils. Furthermore, the contribution of reservoir temperature to CO_2 miscibility in crude oil systems has not been sufficiently explored.

To address these gaps, this study aimed to experimentally determine the MMP and first-contact miscibility (FCM) pressure of CO_2 in medium-viscosity crude oil at a variety of reservoir temperatures. These findings will contribute to improved screening and evaluation of oil reservoirs for CO_2 flooding applications.

2. METHODOLOGY

2.1. Materials. This study utilized a crude oil sample collected from the Uthmaniyah oil field, which is part of the expansive Ghawar Field in Eastern Saudi Arabia (SA). It is a medium dead oil sample, and its density, viscosity, and API at atmospheric pressure and room temperature ($T = 25$ °C) were 0.876 g/cm^3, 50 cp, and 30°API, respectively. CO_2 gas with high purity (99.99%), supplied by Air Liquide Company, was also used in this study.

2.2. Experimental Apparatus. The CO_2-oil IFT was measured using axisymmetric drop shape analysis (ADSA) via the pendant drop method. The main component of the ADSA is a temperature-controlled IFT cell, as shown in Figure 2. A magnifying camera equipped with a light source was used to integrate high-resolution images that were later analyzed via Young-Laplace fitting software. The ADSA had a 5 mm diameter needle that dispensed an oil droplet inside a cell filled with CO_2. The ADSA was attached to a positive displacement pump, which was connected to a high-pressure accumulator saturated with oil. CO_2 was pressurized above the cylinder pressure using a gas booster, as shown in Figure 2.

The experimental setup included an Anton Parr 4500M density meter (as shown in Figure 3), which was utilized to appraise the crude oil density with high accuracy and precision. The meter was integrated with a temperature control system and automatic viscosity correction. The automated viscosity correction feature adjusts the density readings according to the estimated viscosity to enhance the accuracy of the density measurements. Viscosity correction was performed by combining internal algorithms, calibration data, and preinstalled estimations based on typical fluid behavior. In addition, the meter was connected to a manual pump, which allowed pressurization of the oil specimen to a range of pressures and enabled the precise measurement of density.

2.3. Density Measurement. To ensure accurate density readings, the density meter and flow lines were cleaned with

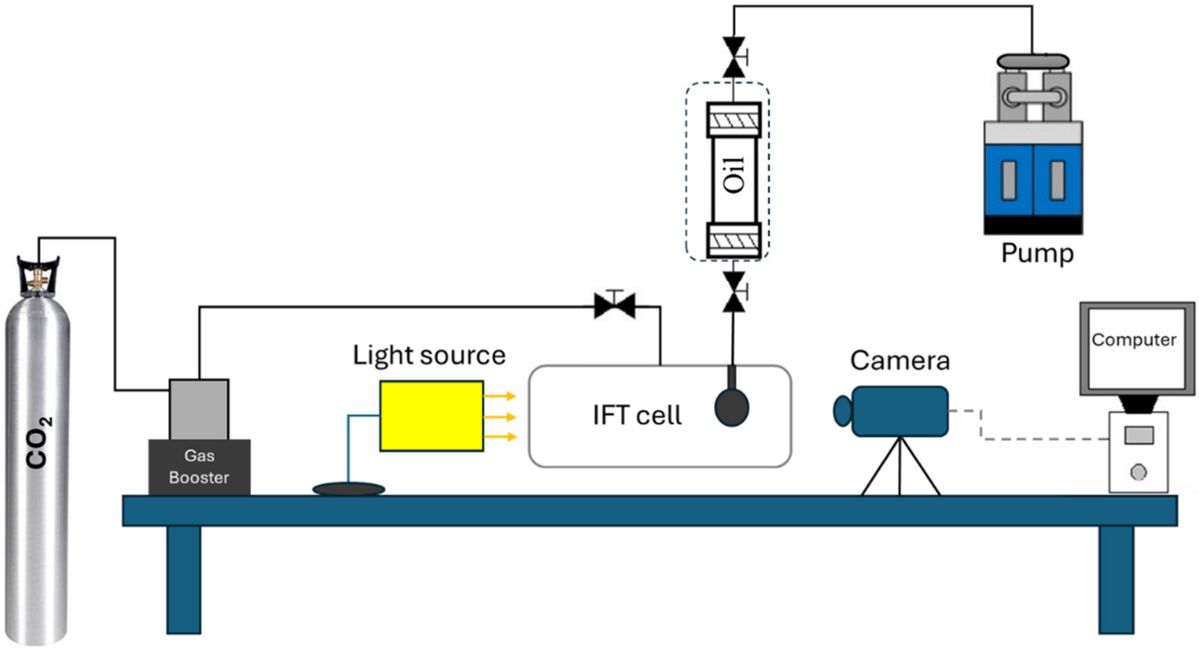


Figure 2. Schematic diagram of the IFT measurement setup (adapted from Rezk & Foroozesh[7])

toluene to dissolve and remove any previous fluid contamination. The lines were then cleaned with deionized water (DI). The cleaning process was repeated five times until a clear output of DI water was obtained and no traces of oil or toluene were observed. After the cleaning process, the DI water density was measured at temperature $T = 70$ °C and pressure varying from 400 psi to 900 psi. This was performed for calibration and to ensure accurate measurements. Finally, UTHMN crude oil density was measured at temperatures $T = 27$ °C and $T = 70$ °C and pressure ranges from 150 psi to 2800 psi. The density meter used was a vibrating U-tube filled with oil. The vibration frequency of the tube was dependent on the oil density (mass). The CO$_2$ density was obtained from the Peace software[17] and checked using the NIST database[18]. The readings were later used as input data for the IFT measurements.

2.4. IFT Measurements.

Before filling the IFT chamber with CO$_2$, the gas vessel was treated with acetone to eliminate any contamination or dust traces and then vacuumed for an hour. CO$_2$ was later transferred through the flow lines from the CO$_2$ cylinder to the ADSA to pressurize the system. The temperature progressively elevated until reaching the intended value and was maintained for 8 h to ensure thermal equilibrium stability within the system. This step was followed by loading a volume of oil sample into the fluid holder with the utmost capacity of 200 ml. The CO$_2$ cell vessel pressure was gradually increased at each step of measuring the IFT up to 750 psi, which was the maximum pressure of the cylinder. A gas booster connected to the CO$_2$ cylinder was used to perform the experiment at higher pressures.

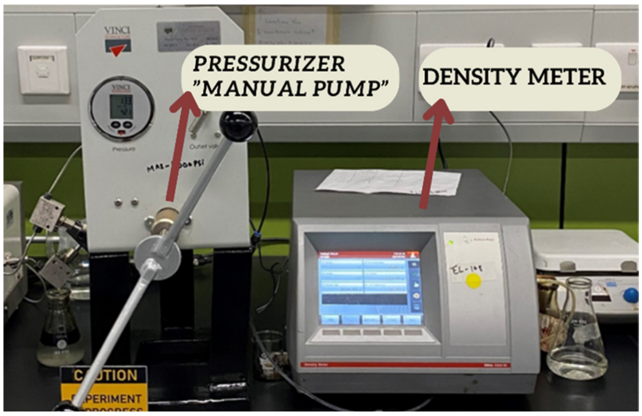


Figure 3. Density meter connected to a manual pump

After setting the desired temperature and pressure, a pendant drop of oil was created using a needle hanging in the cell. A valve was used to control the volume of the oil drops. The first 1–3 generated drops of oil were repeated to ensure the generation of asymmetric drops in the proper form for the IFT measurements. Once a stable droplet was formed, the camera began capturing pictures of the oil drops, and software was used to calculate the IFT between the oil drop and the surrounding CO$_2$.

In this study, the IFT was measured at various pressures within the range 150 psi to 2800 psi and at 28 °C and 70 °C. The CO$_2$-oil MMP and first contact miscibility (FCM) pressure were determined from the IFT vs pressure plot. The FCM is estimated by extrapolating the final linear trend in IFT, where the IFT is approaching zero, indicating that the CO$_2$ and oil phases are becoming miscible. The MMP is the lowest pressure at which a comprehensive miscibility is achieved between CO$_2$ and oil after multiple contacts.

3. RESULTS AND DISCUSSION

3.1. Fluid Density Measurements.

CO$_2$ density was obtained using the online correlation software Peace, which estimates CO$_2$ density as a function of temperature and pressure[17]. This standardized correlation is widely accepted and provides an accurate and reasonable density approximation for high-pressure systems.

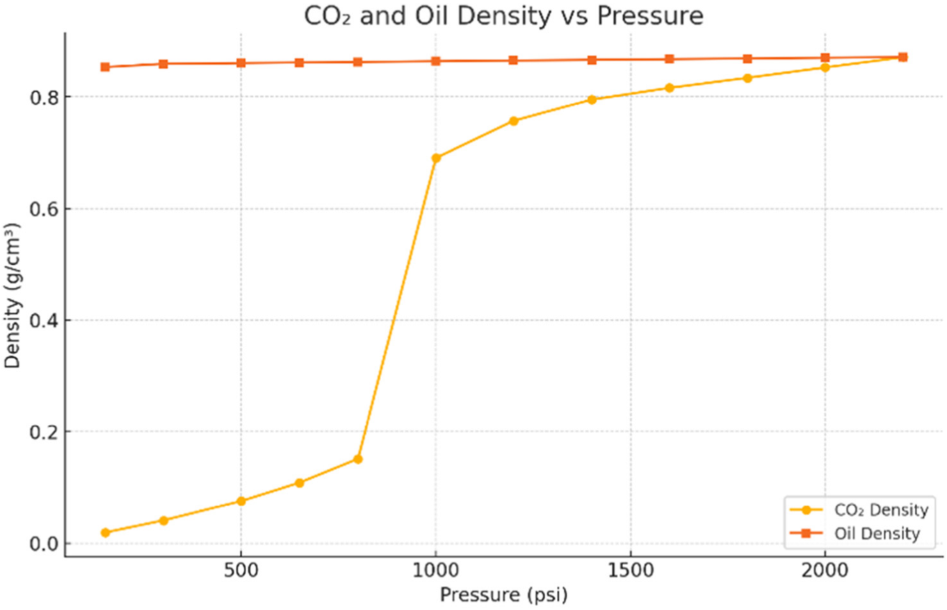


Figure 4. Effect of Pressure on CO_2 and Oil Densities at T = 27 °C

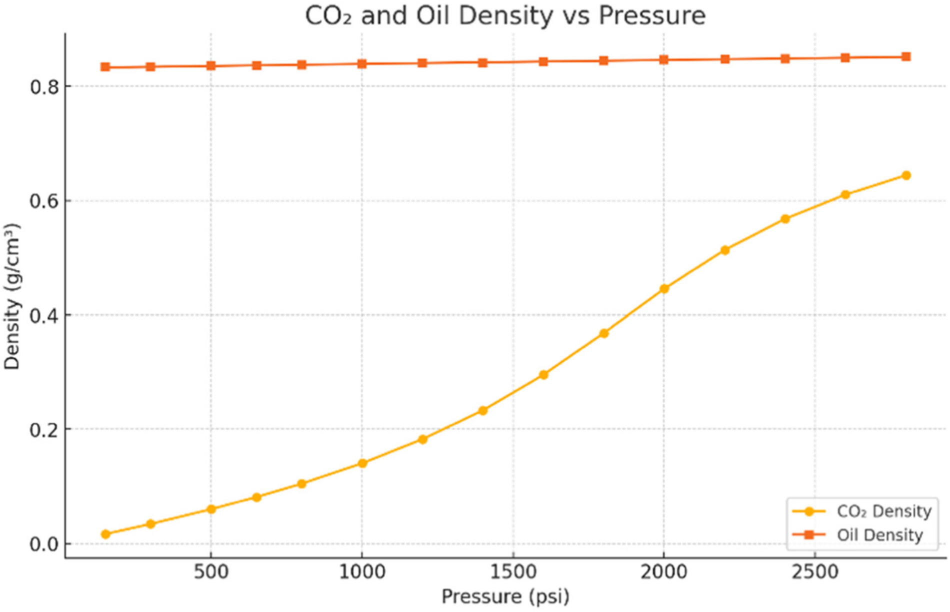


Figure 5. Effect of Pressure on CO_2 and Oil Densities at T = 70°C

The oil samples density was attained by employing the Anton density meter. Figure 4 shows the density measurements at T = 27 °C and at various pressures to investigate the impact of pressure on density measurements. Figure 5 illustrates the density values corresponding to T = 70 °C, which demonstrates density dependency on pressure and temperature.

Figures 4 and 5 show that the oil density tends to decrease as the temperature increases, which is attributed to the thermodynamic behavior of the molecule[19]. When the temperature increases, the molecular kinetic energy increases, which leads to a larger average intermolecular distance. This causes thermal expansion and weakens the intermolecular forces. However, increasing the system pressure results in compression of the fluid (reducing the fluid volume), which is reflected in an increase

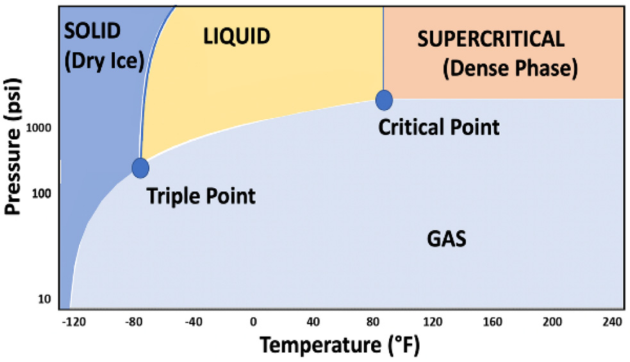


Figure 6. CO_2 phase diagram highlighting the triple point and critical point values (after Greenfelder[21])

in fluid density.[20]. The changes in density for liquids are comparably lower than those for gases because liquids tend to be incompressible (treated as slightly compressible).

3.2. IFT and MMP Measurement. When reaching T = 30.97 °C and pressure = 1070 psi, which are the critical conditions of CO_2 (Figure 6), CO_2 becomes a supercritical fluid, meaning that it does not behave like a liquid or gas. Rather, its density increases and becomes less compressible (like a liquid) and acts as a gas (low viscosity and high diffusivity).

The results from the pendant drop experiment conducted at high pressures and temperatures of 27 °C and 70 °C are shown in Tables 1 and 2, respectively. The tables summarize the measurements of IFT at various pressures and demonstrate that IFT decreases with pressure, indicating enhanced miscibility behavior between the crude oil and CO_2 phases.

The IFT vs pressure data were plotted to estimate the MMP and FCM, as illustrated in Figure 7. The CO_2-oil IFT values decreased as the system pressure increased, which shows that the system gradually approaches the miscibility conditions. The MMP was identified from the plot as the point where the initial near-linear trend of IFT values leveled off to zero IFT (intersection with the X-axis), indicating a region where further pressure increases produced minimal changes in IFT. This stabilization reflects the onset of dynamic miscibility[22]. In contrast, FCM was typically inferred from the extrapolation of a secondary linear trend observed at higher pressures.

Table 1. *Interfacial tension between CO_2 and oil medium at different pressures and 27 °C*

Pressure (psi)	IFT (mN/m)
150	22
300	18.5
500	14.66
650	11.3
800	6.94
1000	3.25
1200	2.63
1400	1.83
1600	1.41
1800	1.28
2000	0.6
2200	No coherent drop

Table 2. *Interfacial tension between CO_2 and the medium oil at various pressures and 70 °C*

Pressure (psi)	IFT (mN/m)
150	20.77
300	18.94
500	16.77
650	15.42
800	13.66
1000	11.62
1200	9.35
1400	7.33
1600	5.67
1800	4.56
2000	3.36
2200	3.11
2400	2.75
2600	2.89
2800	2.73

At T = 27 °C, the IFT showed a sharp exponential decline from 22 mN/m at 150 psi to 0.53 mN/m at 2000 psi. Based on the extrapolated linear trends, the MMP was estimated to be 1175 psi, and the FCM was visualized at 2454 psi.

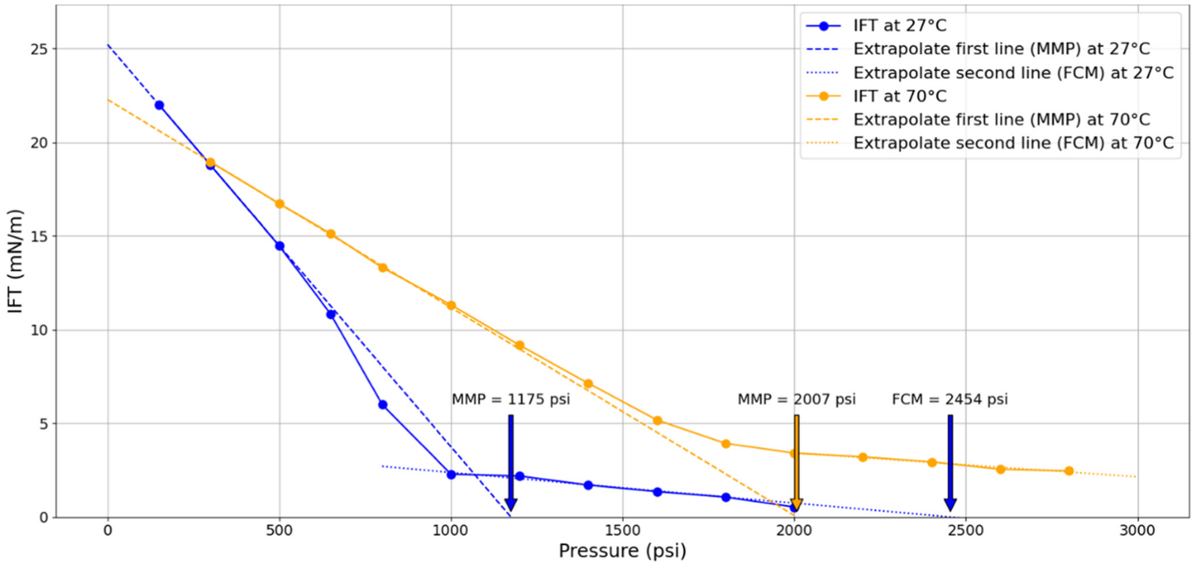


Figure 7. Measured IFT at various pressures at 27 °C and 70 °C

At T = 70 °C, the IFT decreases more gradually and linearly from 18.95 mN/m at 300 psi to 2.46 mN/m at 2800 psi, with no clear stabilization limit observed within the tested pressure range. Thus, the MMP was estimated to be 2007 psi, whereas the FCM could not be determined, indicating that pressures beyond 2800 psi may be required to reach the FCM. The FCM was not obtained at 70 °C due to the difficulty of holding the oil drop at such high pressures.

3.3. Limitations.

There are several limitations that may have influenced the accuracy of the findings, and identifying these limitations will aid in strengthening the credibility of the results. The limitations encountered fall into several categories:

3.3.1. Methodological limitations.

The methodological limitation arises from the approach design of the study. The main assumption in using the pendant drop, where in practice the equilibrium may not have been fully attained due to CO_2 diffusion, which may reflect IFT values from the absolute equilibrium conditions.

3.3.2. Instrumentation & technical limitations.

This limitation is primarily associated with the operational conditions of the laboratory equipment used in this study. The difficulty of holding the oil drop at high pressures limits the IFT data. Additionally, the experimental work was subjected to environmental limitations under which measurements were conducted. One of the limitations is the temperature factor, where small fluctuation occurs that can affect the CO_2-crude oil IFT behavior. In addition, the droplet size, shape stability, and visual clarity were influenced by minor vibrations near the working area.

4. CONCLUSIONS

In conclusion, the miscibility behavior of CO_2-crude oil was investigated through IFT measurements at 27 °C and 70 °C using the pendant drop method. The results demonstrated that at 27 °C, the MMP of CO_2 with oil was relatively low (1175 psi), whereas at 70°C, the MMP increased to 2007 psi using the same CO_2-oil system. This demonstrates that elevated temperatures substantially increase the pressure required to achieve CO_2-oil miscibility.

This study also highlighted the significant contribution of pressure in governing IFT reduction, which is essential for oil recovery during CO_2 injection. These findings provide practical insights for designing CO_2-EOR projects in oil reservoirs, particularly for selecting the optimum reservoir pressures and temperatures. Future research should prioritize the following areas to better integrate and improve the CO_2-EOR method.

It is recommended to use various crude oil samples with different compositions to be tested when measuring IFT. Components such as resins and acids can cause complex IFT behaviors. In addition, such components can cause asphaltene to precipitate and deposit, thereby altering the interfacial properties. Future work should incorporate asphaltene stability to reveal its long-term impact on interfacial phenomena. Finally, evaluating the role of chemical agents and formation brine on IFT interaction between oil components and CO_2 under reservoir conditions can be valuable for optimizing CO_2-EOR.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Noor Alsadig – *College of Petroleum Engineering and Geosciences (CPG), King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0009-0005-7227-3344
Email: S202276780@kfupm.edu.sa

Corresponding Author

Mohamed Gamal Rezk – *Research Mentor, Center for Integrative Petroleum Research (CIPR), and College of Petroleum Engineering and Geosciences (CPG), King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* 0000-0001-9335-0883
Email: mohamed.rezk@kfupm.edu.sa

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support provided by the Department of Petroleum Engineering and the Center for Integrative Petroleum Research (CIPR) at King Fahd University of Petroleum and Minerals (KFUPM). The authors also extend their appreciation to Mr. Ahmad Mahboob for his assistance in running the tests.

REFERENCES

(1) J. Price and B. Smith, "Geologic Storage of Carbon Dioxide–Staying Safely Underground," *International Energy Agency (IEA) Greenhouse Gas R & D Programme,* 2008.

(2) I. E. Agency, "Global Energy Review 2025," ed: International Energy Agency Paris, France, 2025.

(3) X. Zhou, Q. Yuan, X. Peng, F. Zeng, and L. Zhang, "A critical review of the CO_2 huff 'n'puff process for enhanced heavy oil recovery," *Fuel,* vol. 215, pp. 813–824, 2018.

(4) L. Lake, R. T. Johns, W. R. Rossen, and G. A. Pope, *Fundamentals of Enhanced Oil Recovery*. Society of Petroleum Engineers.

(5) P. IEA %J International Energy Agency, France, "Storing CO_2 through enhanced oil recovery," *J International Energy Agency, Paris, France,* 2015.

(6) M. G. Rezk and J. Foroozesh, "Effect of CO_2 mass transfer on rate of oil properties changes: Application to CO_2-EOR projects," *Journal of Petroleum Science and Engineering,* vol. 180, pp. 298-309, 2019.

(7) M. G. Rezk and J. Foroozesh, "Phase behavior and fluid interactions of a CO_2-Light oil system at high pressures and temperatures," *Heliyon,* vol. 5, no. 7, p. e02057, 2019.

(8) S. Iglauer, "Dissolution trapping of carbon dioxide in reservoir formation brine–a carbon storage mechanism," in *Mass Transfer-Advanced Aspects*: InTech, 2011.

(9) E. Secchi. (2024). *Carbon Capture – Is it Really Going to Materialise?* Available: https://www.mdd.com/forensic-accounting-articles/carbon-capture-is-it-really-going-to-materialise/

(10) M. G. Rezk, J. Foroozesh, A. Abdulrahman, and J. Gholinezhad, "CO_2 Diffusion and Dispersion in Porous Media: Review of Advances in Experimental Measurements and Mathematical Models," *Energy & Fuels,* vol. 36, no. 1, pp. 133–155, 2022/01/06 2022.

(11) M. G. Rezk, J. Foroozesh, D. Zivar, and M. Mumtaz, "CO_2 storage potential during CO_2 enhanced oil recovery in sandstone reservoirs," *Journal of Natural Gas Science and Engineering,* vol. 66, pp. 233–243, 2019/06/01/ 2019.

(12) M. Almobarak, Z. Wu, D. Zhou, K. Fan, Y. Liu, and Q. Xie, “A review of chemical-assisted minimum miscibility pressure reduction in CO_2 injection for enhanced oil recovery,” *Petroleum,* vol. 7, no. 3, pp. 245–253, 2021/09/01/ 2021.

(13) S. Taku Ide, K. Jessen, and F. M. Orr, “Storage of CO_2 in saline aquifers: Effects of gravity, viscous, and capillary forces on amount and timing of trapping,” *International Journal of Greenhouse Gas Control,* vol. 1, no. 4, pp. 481–491, 2007/10/01/ 2007.

(14) M. Cao and Y. Gu, “Temperature effects on the phase behaviour, mutual interactions and oil recovery of a light crude oil–CO_2 system,” *J Fluid Phase Equilibria,* vol. 356, pp. 78–89, 2013.

(15) D. Li, S. Xie, X. Li, Y. Zhang, H. Zhang, and S. Yuan, “Determination of Minimum Miscibility Pressure of CO_2–Oil System: A Molecular Dynamics Study,” vol. 26, no. 16, p. 4983, 2021.

(16) C. Huang *et al.*, “Prediction of minimum miscibility pressure (MMP) of the crude oil-CO_2 systems within a unified and consistent machine learning framework,” *Fuel,* vol. 337, p. 127194, 2023/04/01/ 2023.

(17) Peace software. (2007). https://www.peacesoftware.de/einigewerte/co2_e.html.

(18) P. J. Linstrom and W. G. Mallard, “The NIST Chemistry WebBook: A chemical data resource on the internet,” *Journal of Chemical Engineering Data,* vol. 46, no. 5, pp. 1059–1063, 2001.

(19) W. M. Hashim, H. A. Al-Salihi, and F. N. Al Zubaidi, “Effects of temperature on the properties of HL32 oil in the conventional hydraulic actuators,” *Heliyon,* vol. 8, no. 12, p. e11831, 2022/12/01/ 2022.

(20) W. D. McCain Jr, “Properties of petroleum fluids,” 1973.

(21) J. Yang, H. Lian, and L. Li, “Fracturing in coals with different fluids: an experimental comparison between water, liquid CO_2, and supercritical CO_2,” Scientific Reports, vol. 10, no. 1, p. 18681, 2020/10/29 2020.

(22) K. Zhang and Y. Gu, “Two new quantitative technical criteria for determining the minimum miscibility pressures (MMPs) from the vanishing interfacial tension (VIT) technique,” *Fuel,* vol. 184, pp. 136–144, 2016/11/15/ 2016.

KFUPM

Journal of Undergraduate
Research International